UNIVERSITÀ DEGLI STUDI DI FIRENZE
Facoltà di Scienze Matematiche, Fisiche e Naturali

Dottorato di Ricerca in Scienze Chimiche
XXII Ciclo
Settore Disciplinare CHIM/03

Mechanistic studies of oxidative metalloenzymes through X-ray crystallography

Dottoranda: Irene Matera

Tutor:
Prof. Andrea Scozzafava

Coordinatore:
Prof. Gianni Cardini

Contents

Part I

Introduction

1

Bioremediation and biotrasformations

Interest in the microbial biodegradation of pollutants has intensified in recent years as mankind strives to find sustainable ways to cleanup contaminated environments. These bioremediation and biotransformation methods endeavour to harness the naturally occurring microbial catabolic diversity to degrade, transform or accumulate a huge range of compounds including hydrocarbons, polychlorinated biphenyls, polyaromatic hydrocarbons, pharmaceutical substances, radionuclides and metals. Major methodological breakthroughs in recent years have enabled detailed genomic, metagenomic, proteomic, bioinformatic and other high-throughput analyses of environmentally relevant microorganisms providing unprecedented insights into key biodegradative pathways and the ability of organisms to adapt to changing environmental conditions.

The field is recently expanding because the use of chemicals in industry, agriculture and human activities generate huge amounts of hazardous chemical waste that often results in environmental pollution. Microorganisms exhibit enormous capability to degrade both natural and synthetic substances; in fact, they have been forced to evolve to degrade a variety of natural and xenobiotics compounds which are released into the environment. The genetic diversity of such metabolic pathways reflects functional divergences in degradation capabilities. Nature has developed a multitude of processes for the synthesis and the mineralization of aromatic compounds: the synthesis of natural aromatic compounds by plants and bacteria is generally balanced by the mineralization, which is primarily catalyzed by bacteria and fungi. Understanding of the metabolic pathways involve in mineralization of natual compounds should lead to the creation of novel biochemical routes for xenobiotics, that are often persistant chemicals [1].

Currently, aromatic compounds are still mainly obtained from crude oil which is the basis for a multitude of industrial products such as pharmaceuticals, pesticides,

dyes, plastics and other synthetic materials. These aromatic products are either released directly into the environment or are formed during combustion processes and often represent persistant anthropogeneous compounds. Especially halogenated, nitrated or sulfonated compounds cause severe problems in the environment. Although bacteria show short generation times and high genetic flexibility towards a changing surrounding, adaption towards certain new substrates seems to be problematic. This is also reflected by the observation, that often xenobiotics are only unspecifically modified or co-metabolized so that toxic intermediates can accumulate and the cells are seriously damaged or even die. Nevertheless, a wide variety of aromatic xenobiotics are degraded into carbon dioxide, water and inorganic salts by microorganisms [2] [3].

1.1 Dioxygenases

The key step for the breakdown of aromatic xenobiotics is generally the enzymatic cleavage of the thermodynamically very stable aromatic ring system. Under aerobic conditions almost all aromatic compounds are catabolized via a few central intermediates. Those intermediates are usually formed by mono- and dioxygenase activities, which activate the aromatic ring system by the introduction of hydroxyl-groups. In most cases, these central intermediates are (substituted) catechol(s), protocatechuate(s) or gentisate. These activated intermediates serve as substrates for ring-fission dioxygenases which cleave the aromatic compounds into unsaturated aliphatic carboxylic acids [4].

Enzymatic and genetic studies revealed that ring-fission dioxygenases can be separated into functionally different groups [5]. Three main classes of ring-fission dioxygenases have been described. Intradiol and extradiol ring cleaving dioxygenases demand two hydroxy-groups in ortho-orientation to each other (for example catechol or protocatechuate). Intradiol cleaving dioxygenases oxidize the aromatic ring between the two hydroxy groups, whereas an extradiol ring cleavage is characterized by the cleavage of the aromatic ring next to one of the two hydroxy groups [6]. The third known class of ring-cleavage enzymes requires two hydroxy-groups in para-orientation as found in (substituted) gentisates or hydroquinone (1,4-dihydroxybenzene). In the case of gentisate dioxygenases the opening of the aromatic ring occurs between the C-atoms of the aromatic ring which are substituted with a carboxy and a hydroxy group. Similarily, hydroquinones are cleaved next to one of the hydroxy groups carrying C-atoms. Nevertheless, most dioxygenases of this type have been characterized as gentisate dioxygenase [7].

Apparently, the different classes of ring-fission dioxygenases have evolved largely independently since there is no significant sequence homology between the individual groups [5]. However, extradiol dioxygenases and 1,4-dihydroxybenzenes catabolizing dioxygenases share some similarities. For instance, both classes possess catalytically active Fe^{2+} ions in their active center. In contrast, intradiol dioxygenases contain Fe^{3+} ions in their active site [7]. In fact, the cleaving mechanism of extradiol

dioxygenases and 1,4-dihydroxybenzene dioxygenases seems to be similar, and 1,4-diydroxybenzene dioxygenases are often regarded as a peculiar group of extradiol dioxygenases [6].

Ring cleaving dioxygenases are of interest due to their potential utility in the degradation of recalcitrant xenobiotics such as polychlorinated biphenyls, substituted naphtalensulphonates (see Figure 1.1) and other substituted aromatic compounds.

1.2 Laccases

Laccases have received much attention from researchers in last decades due to their ability to oxidise both phenolic and non-phenolic lignin related compounds as well as highly recalcitrant environmental pollutants. Such applications include the detoxification of industrial effluents, mostly from the paper and pulp, textile and petrochemical industries, use as a tool for medical diagnostics and as a bioremediation agent to clean up herbicides, pesticides and certain explosives in soil. Laccases are also used as cleaning agents for certain water purification systems, as catalysts for the manufacture of anti-cancer drugs and even as ingredients in cosmetics. In addition, their capacity to remove xenobiotic substances and produce polymeric products makes them a useful tool for bioremediation purposes [8].

The search for new, efficient and environmentally benign processes for the textile and pulp and paper industries has increased interest towards laccases as biocatalysts, since they work with air and produce water as the only by-product. The direct substrate oxidation of phenol derivatives has been investigated in bioremediation efforts to decontaminate industrial wastewaters. The polymeric polyphenolic derivatives that result from the laccase-catalyzed oxidative couplings are usually insoluble and can be separated, easily, by filtration or sedimentation. At present, one of the main technological applications of laccases is in the textile, dye or printing industries, in processes related to decolourisation of dyes [9].

Laccases have other important applications in the pulp and paper industries for the delignication of woody fibres, particularly during the bleaching process. The structure of wood can be represented as an agglomerate of long, slim fibres (mainly celluloses and hemicelluloses) and of lignin, a complex polymer composed of hydroxyphenyl-propane subunits, linked together by various types of bonds. Lignin is present in the middle lamella, where it acts as a natural glue, and in the secondary cell wall of wood fibres, where it acts as a matrix and forms linkages between the cellulose and hemicellulose chains. To produce paper pulp, lignin must be eliminated, either by mechanical or chemical methods, particularly for the production of the so-called kraft pulp (obtained after cooking with NaOH and Na_2S), which is eventually used to produce paper. Chemical pulping processes use strong chemicals to eliminate the fraction of lignin that is tightly bound to fibres (the so-called bleaching process). Traditionally, chlorine, chlorine oxide and oxygen were also added to nalize the bleaching process but nowadays the use of Cl_2 has been forbidden and the

use of Cl_2O has been restricted; therefore, new approaches for bleaching have been developed: one of them is the use of laccase-based biocatalysts [10].

Laccases could also have uses in green chemistry for organic synthesis that involve typical substrates as phenols and amines, and reaction products as dimers and oligomers derived from the coupling of reactive radical intermediates [11].

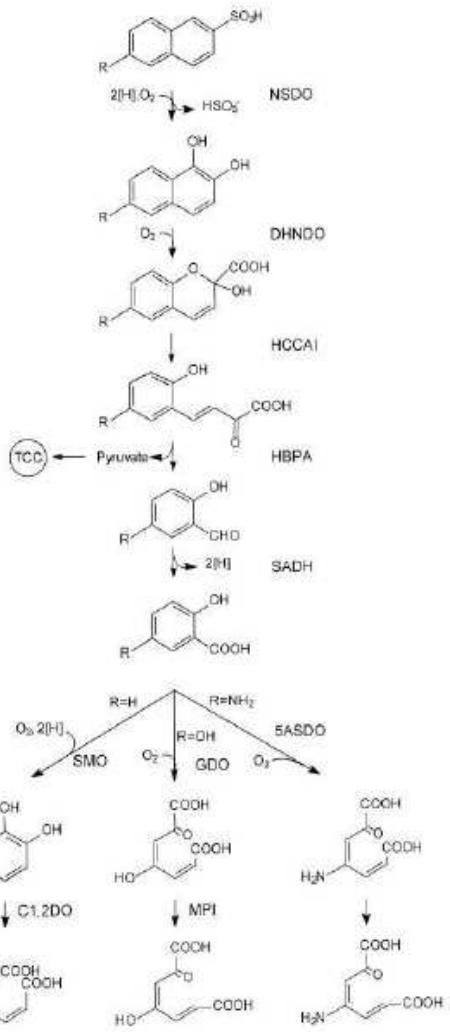


Figure 1.1: Proposed pathways for the degradation of (substituted) naphthalenesulphonates. NSDO, naphthalenesulfonate dioxygenase; DHNDO, 1,2-dihydroxynaphthalene dioxygenase; HCCAI, 2-hydroxychromene-2-carboxylate isomerase; HBPA, 2-hydroxybenzalpyruvate aldolase (hydratase); SADH, salicylaldehyde dehydrogenase; SMO, salicylate 1-monooxygenase; GDO, gentisate 1,2-dioxygenase (an extradiol dioxygenase); 5ASDO, 5-aminosalicylate 1,2-dioxygenase; C1,2DO, catechol 1,2-dioxygenase (an intradiol dioxygenase); MPI, maleylpyruvate isomerase; TCC, tricarboxylic acid cycle.

2

Protein crystallography and kinetics

Protein crystallography is a powerful technique for determining enzyme structure and thus have a tridimensional picture of the spatial position of residues in the protein. This determination should seem to give static informations about the protein, but actually it is really helpful to investigate enzyme mechanism; in fact, the tridimensional model itself, giving the coordinates of the atoms, displays the overall structure but also the arrangement of residues in the active site and the binding of metals and cofactors. Moreover there are several methods in crystallography that are capable to trap adducts of the enzyme binding substrates, inhibitors or substrate omologues: the structures of these kind of adducts often are resolutive in understanding how the protein works. In addition, recently the coupling of crystallography with other techniques, for example Uv-Vis spectroscopy, has allowed to isolate more subtely the steps of reaction, making possible to follow the catalityc mechanism along each state in the reaction coordinate.

2.1 What is protein X-ray crystallography?

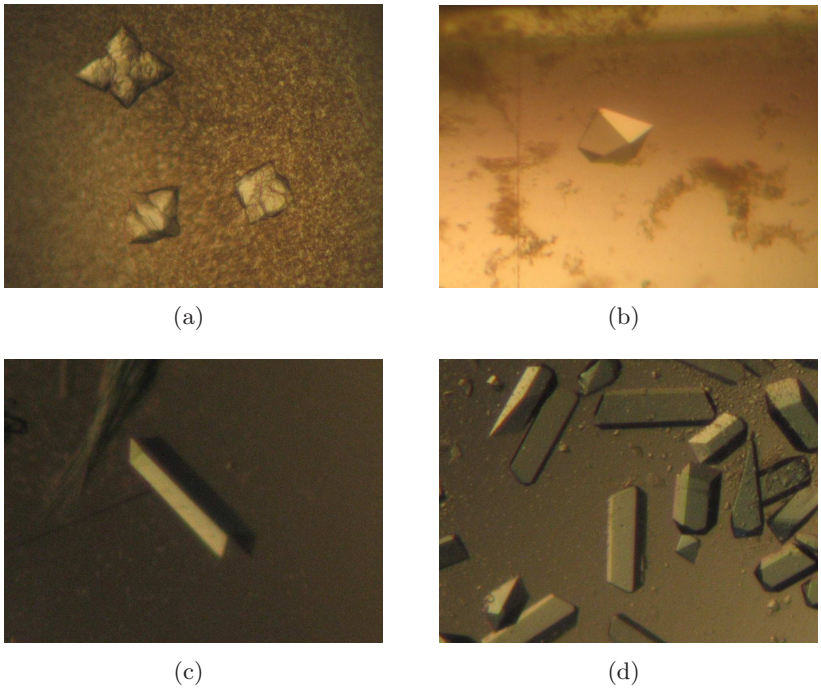
The most common experimental means of obtaining a detailed model of a large molecule, allowing the resolution of individual atoms, is to interpret the diffraction of X-rays from many identical molecules in an ordered array like a crystal. This method is called *single-crystal X-ray crystallography*. As of December 2009, the Protein Data Bank (PDB), the worlds largest repository of macromolecular models obtained from experimental data (called experimental models), contains roughly more than 53,000 protein and nucleic-acid models determined by X-ray crystallography.

In brief, determining the structure of a protein by X-ray crystallography entails growing high-quality crystals of the purified protein, measuring the directions and intensities of X-ray beams diffracted from the crystals, and using computer pro-

grams to produce an image of the crystals contents, that is an electron density map. Finally, the crystallographer must interpret that map, which entails displaying it by computer graphics and building a molecular model that is consistent with the electron density.

2.1.1 Growing protein crystals

In the most common methods of growing protein crystals, purified protein is dissolved in an aqueous buffer containing a precipitant, such as ammonium sulfate or polyethylene glycol, at a concentration just below that necessary to precipitate the protein. Then water is removed by controlled evaporation to raise both protein and precipitant concentrations, resulting in precipitation. Slow precipitation is more likely to produce larger crystals, whereas rapid precipitation may produce many small crystals, or worse, an amorphous solid. In Figure 2.1 few examples of protein crystals are shown.



(a) (b)

(c) (d)

Figure 2.1: Microsnapshots of various protein crystals: a,b) Crystals of Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans*; c) Cystals of Laccase from *Trametes trogii* and d) Laccase from *Steccherinum ochraceum*.

In theory, precipitation should occur when the concentration of protein and precipitant exceeds threshold values, as shown in the phase diagram of Figure 2.2. Crystal formation occurs in two stages, nucleation, and growth. Nucleation, the

initial formation of molecular clusters from which crystals grow, requires protein and/or precipitant concentrations higher than those optimal for slow precipitation (blue region). In addition, nucleation conditions, if they persist, result in the formation of many *nuclei*, and as a result, either an amorphous precipitate or many small crystals instead of a few larger ones. An ideal strategy would be to start with conditions corresponding to the blue region of the phase diagram, and then, when nuclei form, move into the green region, where growth, but not additional nucleation, can occur.

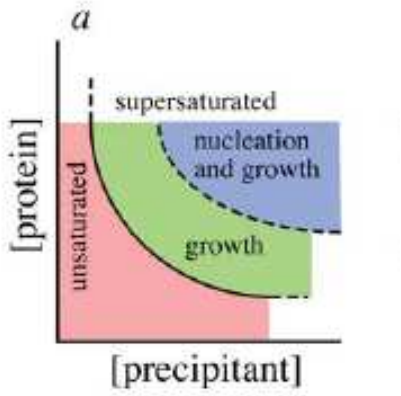


Figure 2.2: Phase diagram for crystallization mediated by a precipitant.

One widely used crystallization technique is vapor diffusion, in which the protein/precipitant solution is allowed to equilibrate in a closed container with a larger aqueous reservoir whose precipitant concentration is optimal for producing crystals. One of many examples of this technique is the sitting-drop method. Figure 2.3 shows the basis of a vapor diffusion sitting-drop method crystallization experiment: a small drop of a mixture of both protein solution and precipitant solution is equilibrated with a reservoir containing only the precipitant solution by sealing the entire environment. During the vapor diffusion the drop, that at the beginning of the experiment is in the pink region of the phase diagram (See Figure 2.2), starts to concentrate and moves in the blue region, were nucleation initiates. Then, depending on the particular phase diagram of that peculiar experiment, we can observe aggregates, amorphous precipitate, needles, small crystals or bigger crystals.

Obviously it is not possible to explore the whole space of phase of every single protein; the precipitants used are many (PEGs of different molecular weights, salts, alcohols), and these are not the only parameter optimizable in a crystallization experiment: we can vary the pH, the kind of buffer, the protein concentration, the temperature and so on. Generally sparse matrix screens are used in the beginning, just to pick out few promising crystallization conditions. Then by the optimization of the parameters in these few conditions regular and hopefully big crystals can be obtained.

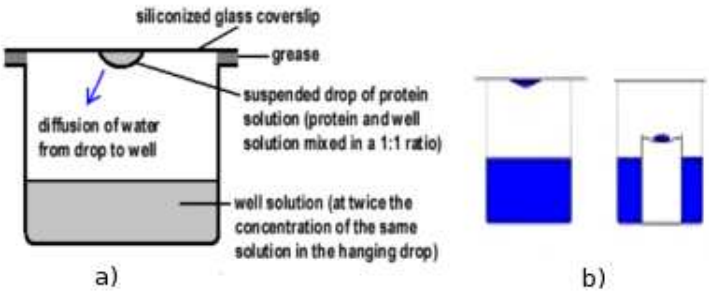


Figure 2.3: The vapor diffusion method: a) the crystallization trial; b) hanging and sitting drop.

In case that it is impossible to grow crystals suitable for diffraction, there are few further tecniques to improve crystallization. The most easy is the use of additives. Additives like salts, alcohols or other organic small molecules, are added directly in the drop, and they really influence the crystallization process. Another useful technique is seeding. From Figure 2.2 it is clear that generally when crystallization begins, the drop is in a region of the phase diagram where nucleation is favoured; thus crystal growth is not the most favourable process, and in few cases crystals stop to grow when their size is not suitable for an X-ray diffraction experiment. Seeding consist on putting a few seeds of crystals, that is few little pieces of an even small crystal already obtained, in an equilibrated drop of both protein and precipitant solutions that is located in a region of the phase diagram where only crystal growth occurs: the seeds act like *nuclei* and they can only grow since nucleation does not occurs. These improving techniques can also be used together; for example in order to crystallize Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans* we first optimized the crystallization conditions with few additives and then we use seeding to improve shape and regularity of the crystals (See Part II).

2.1.2 The observable of X-ray crystallography

Crystallization is often the rate determining step of the whole process of solving a protein structure. Once some crystals are obtained, we can move to the next step, that is irradiate them with an X-ray source and get the diffration pattern: the actual diffractors in the protein are the electron clouds (For a description of how electromagnetic radiation interacts with electrons see Chapter 3). All the informations needed to build the model of the protein are in a set of diffraction images. Roughly, the diffraction pattern is the convolution of a function describing the repetitiveness of the lattice in which the molecules are arranged in the crystal and a function describing the diffraction pattern of all the electrons of a single molecules of protein.

The spots in the diffraction pattern are arranged in a lattice, the reciprocal lattice, that is connected with the real lattice of the crystals by precise mathematical

equations. The intensity of each spot is the results of a sum of the rays diffracted in that particular direction by all the electrons in the protein, thus is connected to the spatial density of the electrons in the crystal. The direct result of crystallographic data collection is a list of intensities for each point in the three-dimensional reciprocal lattice. This set of data is the raw material for determining the structures of molecules in the crystal. We are measuring diffracted X-rays, which are waves, with intensities, that are scalar quantities: obviously we are loosing an information: the phase of the diffracted waves (See Figure 2.4 for a funny representation).

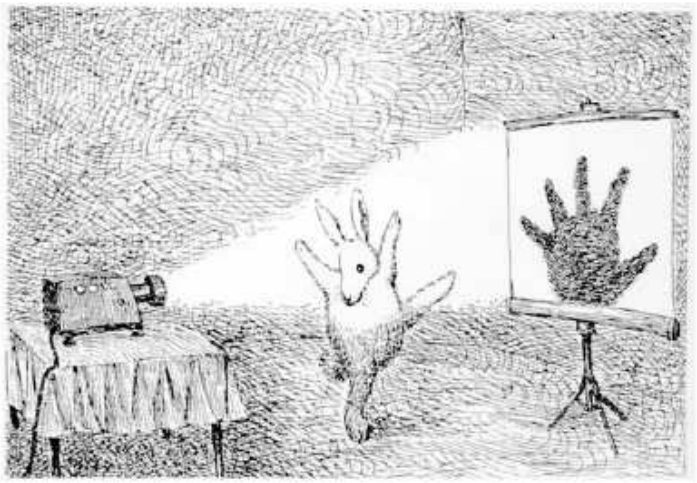


Figure 2.4: Fun in reciprocal space. The New Yorker Collection (1991) by John O'Brien.

Crystals are usually frozen at 100K in a flux of nitrogen while during the diffraction experiment; a cryoprotectant (like a low molecular weight PEGs, glycerol, ethylen glycol) is added to the crystallization solution to avoid icing. The diffraction images are processed with programs that calculate the parameters of the reciprocal lattice (and then those of the real lattice), locate the spots on the image and integrate every single spots to get the intensities.

2.1.3 The phase problem

The X-ray waves diffracted from the crystals (called structure factors) are related to the electron density in the molecule by a Fourier transform. Since we want an image of the electron clouds that surround the molecules in the average unit cell in the crystal, hoping this image will allow us to locate all atoms in the unit cell, we have to find a method to obtain the phases of the structure factor, in order to perform the Fourier transform. From the integration of spots in the diffraction images we get the intensities of each reflection; the amplitude of the structure factor is proportional

to the square root of the intensity of each reflection. Fortunately, there are several ways to recover the lost phases, for example Multiple Isomorphous Replacement (MIR) and Multiwavelenght and Single-wavelenght Anomalous Diffraction (MAD and SAD).

In this study we solved the structures with Molecular Replacement (MR). The use of MR is possible if we know that the structure we want to solve is similar to that of a known protein; the known protein is used as a phasing model. If the new crystals and those of the model are isomorphous, the model phases can be used directly as estimates of the desired phases. If not, we must somehow superimpose the known protein upon the new protein to create the best phasing model. We can do this without knowledge of the structure of the new protein by using Patterson-map comparisons to find the best orientation of the model protein and then using structure-factor comparisons to find the best location of the model protein. A Patterson map is an interatomic vector map created by squaring the structure factor amplitudes and setting all phases to zero. This vector map contains a peak for each atom related to every other atom, with a large peak at 0,0,0, where vectors relating atoms to themselves p̈ile up: If we generate Patterson maps for the data derived from our unknown structure (knowing the intensities only), and from the structure of a previously solved homologue, in the correct orientation and position within the unit cell, the two Patterson maps should be closely correlated. This principle lies at the heart of MR, and can allow us to infer information about the orientation and location of an unknown molecule with its unit cell. An MR search is typically divided into two steps: rotation and translation. Following this, we should have a correctly oriented and translated phasing models, from which we can derive phases which are hopefully accurate enough to derive an interpretable electron density maps.

2.1.4 From electron density to model

In brief, obtaining a detailed molecular model of the unit-cell contents entails calculating the electron density through a Fourier transform using amplitudes of the structure factors computed from measured intensities in the native data set and phases computed from heavy-atom data, anomalous scattering, or molecular replacement. Because the phases are rough estimates, the rst map may be uninformative and disappointing; the map can be improved by an iterative process. The basic principle of this iteration is the following: any features that can be reliably discerned in, or inferred from, the map become part of a phasing model for subsequent maps. Without the input of new information, the map will not improve.

Is is not possible to get any new phase information without modifying the model in some way. The possible modifications include solvent flattening, noncrystallographic symmetry averaging, or introducing a partial atomic model. It is considered best practice to make sure that initial phases are good enough to make the map interpretable. The initial model can the be improved by renement of the atomic coordinates. This method entails adjusting the atomic coordinates to improve the agreement between amplitudes calculated from the current model and the original

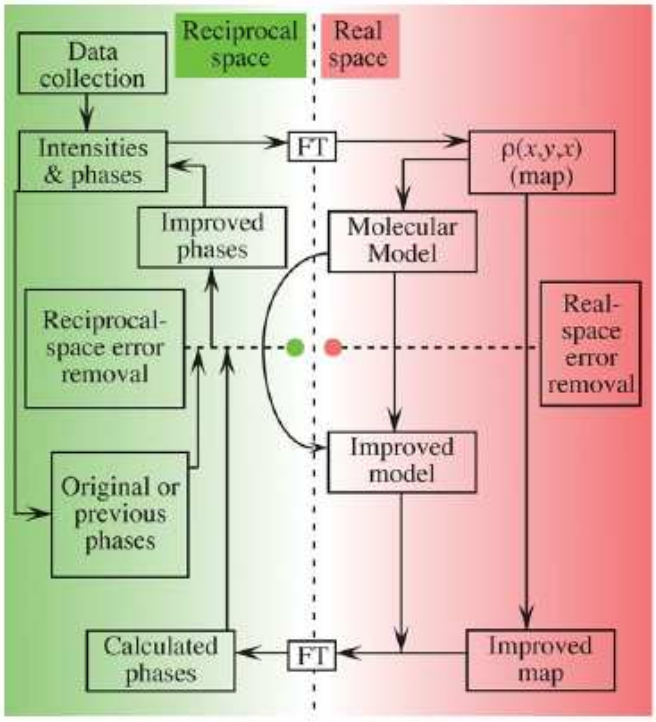


Figure 2.5: Crystallographic structure determination. Operations on the left are based on reciprocal space criteria (improving phases), while those on the right are based in real space (improving the accuracy of atomic coordinates).

measured amplitudes in the native data set. In the latter stages of structure determination, the crystallographer alternates between map interpretation and refinement of the coordinates. At its heart, refinement is an attempt to minimize the differences between measured structure factors and structure factors predicted by the current model, which, in intermediate stages, is incomplete and harbors errors that will eventually be removed. Classical renement algorithms employ the statistical method of least-squares. Figure 2.5 shows how these various methods ultimately produce a molecular model that agrees with the native data.

The minimizing algorithm is based upon several parameters; among these there are obviously the atom positions, because the atom positions in the model determine the calculated stucture factors. But other parameters are included as well. One is the temperature factor B, a measure of how much the atom oscillates around the position specied in the model. Atoms at side-chain termini are expected to exhibit more freedom of movement than main-chain atoms, and this movement amount spread each atom over a small region of space. Diffraction is affected by this variation in atomic position, so it is realistic to assign a temperature factor to each atom and include the factor among parameters to vary in minimizing. From the temperature

factors computed during renement, we learn which atoms in the molecule have the most freedom of movement, and we gain some insight into the dynamics of our largely static model. In addition, adding the effects of motion to our model makes it more realistic and hence more likely to fit the data precisely. Another parameter included in refinement is the occupancy of each atom, a measure of the fraction of molecules in which atoms actually occupy the position specified in the model. If all molecules in the crystal are precisely identical, then occupancies for all atoms are 1.00. Occupancy is included among refinement parameters because occasionally two or more distinct conformations are observed for a small region like a surface side chain, or more often in structure of enzyme-substrate adducts, if the substrate is not bound in all the molecules in the crystal.

The refinement can be followed monitoring two parameters: R_{factor} and R_{free}. The R_{factor} compares overall agreement between the amplitudes of two sets of structure factors, the one observed and the one calculated:

$$R_{factor} = \frac{\Sigma \, ||F_{obs}| - |F_{calc}||}{\Sigma \, |F_{obs}|}$$

Values of R_{factor} range from zero, for perfect agreement of calculated and observed intensities, to about 0.6, a value obtained when a set of measured amplitudes is compared with a set of random amplitudes. A desirable target R_{factor} for a protein model refined with data to 2.5 Åis 0.2. Very rarely, small, well-ordered proteins may refine to values as low as 0.1, whereas small organic molecules commonly refine to values below 0.05.

A more demanding and revealing criterion of model quality and of improvements during refinement is the free R_{factor}, R_{free}. R_{free} is computed with a small set of randomly chosen intensities, the test set, which are set aside from the beginning and not used during refinement. They are used only in cross-validation, a quality control process that entails assessing the agreement between calculated (from the model) and observed data. At any stage in refinement, R_{free} measures how well the current atomic model predicts a subset of the measured intensities that were not included in the refinement, whereas R_{factor} measures how well the current model predicts the entire data set that produced the model. There is a sort of circularity in R_{factor} that is avoided in R_{free}, because R_{free} gives a better and less-biased measure of overall model. In general, during intermediate stages of refinement, R_{free} values are higher than R_{factor}, but in the final stages, the two often become more similar, and the difference is usually around 0.05. Both R are used to guide the refinement, looking for refinement procedures that improve both of them, and proceeding with great caution when the two criteria appear to be in conflict.

In addition to monitoring R_{factor} and R_{free} as indicators of convergence, we can monitor various structural parameters that indicate whether the model is chemically, stereochemically, and conformationally reasonable: bond lengths, bond angles, torsion angles and the Ramachandran plot.

Electron density maps

If the initial phase estimates are sufficiently good, maps will show some of the gross features of unit cell contents. In the rare best cases, with good phases from molecular replacement, first maps are easily interpretable, clearly showing continuous chains of electron density and features like α-helices, perhaps even allowing some amino-acid side chains to be identified. At the worst, the first map is singularly uninformative, signaling the need for additional phasing information (for example another heavy-atom derivative). Usually the minimum result that promises a structure from the existing data is that protein be distinguishable from bulk water, and that dense features like α-helices can be recognized. If the boundary of each molecule, the molecular envelope, shows some evidence of recognizable protein structure, then a full structure is likely to come forth.

There are two maps that are usually used in model building and refinement: the $F_{obs} - F_{calc}$ map and the $2F_{obs} - F_{calc}$ map. The first map emphasizes errors in the current model, but lacks the familiar appearance of the molecular surface found in an F_{obs} map. In addition, if the model still contains many errors, the $F_{obs} - F_{calc}$ map is noisy, full of small positive and negative peaks that are difficult to interpret. The $F_{obs} - F_{calc}$ map is most useful near the end of the structure determination, detecting subtle errors after most of the serious errors are corrected. A more easily interpreted and intuitively satisfying difference map, but one that still allows undue inuence by the model to be detected, is the $2F_{obs} - F_{calc}$ map: this map is everywhere positive, and map contouring at carefully chosen electron densities give something that resembles a molecular surface.

Conversion to a molecular model is usually done as soon as maps reveal recognizable structural features. This procedure, called map fitting or model building, entails interpreting the electron-density map by building a molecular model that fits realistically into the molecular surface implied by the map. Map fitting is done by interactive computer graphics, although automation has eliminated much of the manual labor of this process. At the end of successful refinement, the $2F_{obs} - F_{calc}$ map almost looks like a space-filling model of the protein whereas the $F_{obs} - F_{calc}$ map becomes rather empty.

Usually crystallogaphic models and structure factors are available through the Protein Data Bank (PDB)[12]. Structure can be published and made available by depositing coordinates in the PDB. Depositors are required to run a series of checks for errors and inconsistencies in coordinates and format, and a validation check that produces a report on various model properties. These validation checks are not as extensive as can be done with the online validation tools. Once the files are processed, the PDB makes them available free over the World Wide Web, in a standard text format.

2.1.5 The structure: which informations inside?

The final model consists of the coordinates, the temperature factor and the occupancy of each atom. In analizing a crystallographic structure, there are few important features on that we have to keep in mind in order to understand correctly the physical meaning of the structure itself and derive faultless biological informations:

Resolution

In microscopy, the phrase resolution of 2 Å, implies that we can resolve objects that are 2 Åapart. If this phrase had the same meaning for a crystallographic model of a protein, in which bond distances average about 1.5 Å, we would be unable to distinguish or resolve adjacent atoms in a 2-Åmap. Actually, for a protein structure refined at 2-Åresolution to an R_{factor} near 0.2 the situation is much better than the resolution statement seems to imply. In X-ray crystallography, 2-Åmodel means that analysis included reections out to a distance in the reciprocal lattice of $\frac{1}{2\text{Å}}$ from the center of the diffraction pattern. Presumably, data farther out than the stated resolution was unobtainable or was too weak to be reliable. Although the final 2-Åmap, viewed as an empty contour surface, may indeed not allow us to discern adjacent atoms, prior knowledge in the form of structural constraints on the model greatly increase the precision of atom positions. The main constraint is that we know we can t the map with groups of atoms — amino-acid residues — having known connectivities, bond lengths, bond angles, and stereochemistry.

Thermal motion and Disorder

These are two important physical (as opposed to statistical) reasons for uncertainty in atom positions. Thermal motion refers to vibration of an atom about its rest position. Disorder refers to atoms or groups of atoms that do not occupy the same position in every unit cell, in every asymmetric unit, or in every molecule within an asymmetric unit. The temperature factor B obtained during refinement reflects both the thermal motion and the disorder of atom, making it difficult to sort out these two sources of uncertainty. In some cases, after refinement converges, a few B values fall far outside the average range for the model. This is sometimes an indication of disorder. Careful examination of electron density maps may give evidence for more than one conformation in such a troublesome region. If so, inclusion of multiple conformations followed by refinement of their occupancies may improve the R and the map, revealing the nature of the disorder more clearly. If B were purely a measure of thermal motion at atom (and assuming that occupancies are correct), then in the simplest case of purely harmonic thermal motion of equal magnitude in all directions (called isotropic vibration), B is related to the magnitude of vibration. But the B values obtained for most proteins are too large to be seen as reecting purely thermal motion and must certainly reflect disorder as well.

Disordered regions

Occasionally, portions of the known sequence of a protein are never found in the electron-density maps, presumably because the region is highly disordered or in motion, and thus invisible on the time scale of crystallography. The usual procedure is simply to omit these residues from the deposited model. It is not at all uncommon for residues at termini, especially the N-terminus, to be missing from a model.

Unexplained density

Electron-density maps occasionally show clear, empty density after all known contents of the crystal have been located. Apparent density can appear as an artifac, but this density disappears when a more complete set of data is obtained. Among the possible explanations for density that is not artifactual are ions like phosphate and sulfate from the mother liquor; reagents and detergents used in purification or crystallization; or cofactors, inhibitors, allosteric effectors, or other small molecules that survived the protein purification. Later discovery of previously unknown but important ligands has sometimes resulted in subsequent interpretation of empty density.

2.2 Enzyme activity in crystalline state

Knowing that crystallography study proteins only in the crystalline state, it is relevant to wonder if these molecules are altered when they crystallize and whether the structure revealed by X-rays is pertinent to the molecules action in solution; with a few proteins, it has been found that crystal structures are in conflict with chemical or spectroscopic evidence about the protein in solution. These cases are rare, fortunately, and the large majority of crystal structures appear to be identical to the solution structure.

Probably the most convincing evidence that crystalline structures can safely be used to draw conclusions about molecular activity is the observation that many macromolecules are still functional in the crystalline state. Substrates added to suspensions of crystalline enzymes are converted to product, even if at reduced rates, suggesting that the enzyme catalytic and binding sites are intact. The lower rates of catalysis can be accounted for by the reduced accessibility of active sites within the crystal, in comparison to solution.

In a dramatic demonstration of the persistence of protein function in the crystalline state, crystals of deoxyhemoglobin shatter in the presence of oxygen. Hemoglobin molecules are known to undergo a substantial conformational change when they bind oxygen. The conformation of oxyhemoglobin is apparently incompatible with the constraints on deoxyhemoglobin in crystalline form, and so oxygenation disrupts the crystal. Another evidence of protein crystals retaining its own activity can be the changing colour of the crystals while soaked in a substrate solution; studying the Blue Laccase from *Steccherinum Ochraceum* (see Part III) we found a dramatic

change in the colour of the protein crystal soaked in a solution of acetosyringone (a natural substrate of this enzme): the crystals went from blue to white. Since the first step of the reaction is the reduction of a Cu(II) to Cu(I) we could therefore conclude that the protein retains its function in the crystalline state. Being the crystalline form functional, we can be confident that the X-ray model will show the molecule in its functional form.

Further evidence for the similarity of solution and crystal structures is the compatibility of crystallographic models with the results of chemical studies on proteins. For instance, two reactive groups in a protein might be linked by a cross-linking reagent, demonstrating their nearness. The groups shown to be near each other by such studies are practically always found near each other in the crystallographic model.

In a growing number of cases, both NMR and X-ray methods have been used to determine the structure of the same molecule. Generally the two methods produce similar models; the greatest discrepancies, even though they are not large, lie in the surface loops at the top and bottom of the models. This and other NMR-derived models confirm that protein molecules are very similar in crystals and in solution. In some cases, small differences are seen and can usually be attributed to crystal packing.

Very often the structure of a protein has been obtained from more than one type of crystal form: the resulting models are generally identical, suggesting that the molecular structure was not altered by crystallization. For example, in the second part of this thesis we will discuss the crystal structures of Salicylate 1,2-Dioxygenase from *Pseudaminobacter Salicilatoxidans* obtained from two different crystalline forms: the first belonging to space group $P4_32_12$ and the other to space group I222. The two different crystallographic models are superposable with a r.m.s.d. of 1.546 Å; infact the I222 structure was easily solved by molecular replacement using the $P4_32_12$ structure as starting model.

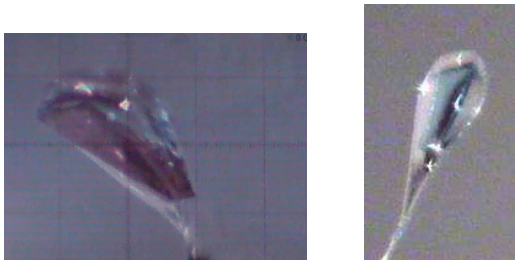
In fact, stable protein crystals contain a large amount of both ordered and disordered water molecules. As a result, the proteins in the crystal are still in the aqueous state, subject to the same solvent effects that stabilize the structure in solution. Viewed in this light, it is less surprising that proteins retain their solution structure in the crystal.

2.3 Isolation of reaction steps

In studing enzyme mechanism it is important to keep in mind that protein structures determined by X-ray diffraction are time-and-space averages. Spatial averaging is over all of the trillions of molecules in the crystal; temporal averaging is over the time required to collect the X-ray diffraction data. Both forms of averaging have an impact on attempts to use the power of X-ray crystallography to visualize discrete steps in an enzyme-catalyzed reaction at atomic resolution. For unambiguous interpretation of the experimentally-observed electron density, that density must come

from a single species at high occupancy (the noise level in an electron density map at typical resolutions is about 25% of its average value). Averaging over all molecules in the crystal means that this condition will only be met if the catalytic reaction is synchronized throughout the entire crystal or if the experimental parameters have been chosen so that a discrete intermediate accumulates. Even if the reaction can be initiated synchronously a distribution of species will quickly develop; a better strategy is, therefore, to adjust temperature, pH and other variables in the experiment so that a particular complex becomes metastable. Time-averaging requires that the lifetime of this complex must exceed the X-ray data collection time or else mixtures of species will form.

Over the past twenty years or so a variety of techniques have been developed to enable normally unstable reaction intermediates to accumulate and be stabilized in protein crystals. The field is clearly moving forward, with the revitalization of ultrafast time-resolved experiments, significant developments in trapping strategies and several biologically important results. Time-resolved and trapping approaches are complementary. Whereas time-resolved experiments minimize perturbations to the sample and are best suited to handling inherent structural heterogeneity, trapping approaches are accessible to a larger variety of projects, provide high-resolution data and may be able to reveal early intermediates inaccessible at present to time-resolved studies [13]. A current trend is the important synergy between kinetic crystallography and *in crystallo* UV/visible, Raman, IR or EPR spectroscopy.



(a) (b)

Figure 2.6: Microsnapshots of frozen protein crystals: a) Crystals of Cathecol 1,2-Dioxygenase from *Rhodococcus opacus*; b) Cystals of Laccase from *Trametes trogii*.

Flash-freezing is probably the most versatile method for stabilizing the desired intermediate, but any set of conditions that prevents its breakdown can be used: working at far from optimum catalytic conditions, changing pH for example, or removing a specie that is required by the enzyme to react with the substrate, like O_2 for oxygen-dependent enzymes.

2.3.1 Cryocrystallography

The beneficial effects of freezing on minimizing X-ray radiation damage are well known and the study of these temperature-dependent effects has become even more important with the use of intense synchrotron X-ray sources. Crystals are soaked for few seconds in a solution in which they are stable (usually the crystallization condition) added up in a cryoprotectant, such as glycerol, and then flash-freeze in a nitrogen stream at 100 K (See Figure 2.6). This enables the surrounding solvent to freeze in a glass state at cryogenic temperatures, thus helps avoid icing and maintain both the integrity of the crystal lattice and its diffraction quality. For a small crystal, cryogenic temperatures can be achieved throughout the crystal within a few milliseconds.

Freeze-trapping

The most commonly used method for stabilization of reaction intermediates in protein crystals is to generate the species at a temperature at which it forms relatively quickly and then rapidly cool the crystal. If the Arrhenius rate equation (which relates, in a logarithmic fashion, the rate of a reaction to the reciprocal of the absolute temperature) holds over all temperature ranges it can be used to calculate how cold the crystal needs to be for any desired reaction time, provided the activation enthalpy of the reaction is known. However, a glass transition exists in the dynamical properties of all proteins at around 150 — 200 K; below this temperature the nonharmonic, collective motions are frozen out. Because such motions appear to be essential for many, if not all, enzyme reactions to proceed from substrate to product, and also for specific substrate binding and product release, it is often only necessary to flash-freeze the crystal to below this temperature in order to trap whatever species was generated, essentially indefinitely. Protein crystals can be soaked in a solution containig the substrate and then flash-frozen to about 100K, a freezing routinely employed to reduce radiation damage so it is often not difficult to try this method of trapping intermediates.

Freeze-triggering

However, although the crystalline state often slows down turnover, rates under 1 s^1 in crystals are common. In such cases, even if synchronous reaction triggering can be achieved, the relatively slow (0.1 s) process of ash cooling may become the limiting factor for catching intermediates. A freeze-trigger strategy can then be attempted, whereby triggering is done at a temperature low enough to lock the reaction (typically under 150 K) and the formation of intermediates results from the limited thermal energy available to the sample.

To trigger the reaction within the crystal at these temperatures, soaking procedures cannot be envisaged and light has to be used. This is why the freeze-trigger strategy is applied mostly to proteins displaying endogenous photosensitivity, like

heme-containing proteins [14] [15]. Recent developments have concerned redox enzymes, whereby X-ray light served as a source of reducing electrons: this is the method employed in our study of the catalityc mechanism of Laccase from *Stec-cherinum Ochrceum* (See Part III). Like the simpler freeze-trapping, also the freeze-trigger strategy relies on the dynamical transition that occurs in the crystal at 150 — 200 K. Upon triggering below this transition temperature, individual molecules may not escape a limited set of low-energy conformational states. Therefore, they need not be synchronized, and triggering can be gentle and slow. However, because only small motions are expected, the structural differences between early states (before and after triggering) are subtle and special care is needed during crystallographic analysis to identify all of them. Whereas some states might hint at the structure of extremely short-lived physiological states (so far not accessed by time-resolved diffraction) and might thus be of extreme interest, others might just be artifacts.

Deconvolution of structural heterogeneity in trapping experiments remains a challenge. Several methods have been used to identify multiple states from a single crystallographic data set, including extrapolation of structure factor amplitudes [16], renement of multiple conformations [17] or difference renement [18]. Strategies in which a series of data sets can be recorded on the same crystal while a parameter is varied (for example UV/visible light dose, X-ray dose or temperature) should be further developed, allowing linear interpolation of the data.

2.3.2 Other tecniques

Working at far from optimum catalytic conditions

The pH-dependence of an enzyme-catalyzed reaction can be such that the turnover time of a particular intermediate is very long at pH values several units away from the optimum. For such enzymes it is sometimes possible simply to incubate the crystal with substrate at this extreme pH: the intermediate will accumulate but its breakdown time will be long compared with the time required to collect the diffraction data. It is often necessary to use low temperatures as well as non optimal pH to ensure that the intermediate does not decompose during the collection time. Unfortunately this strategy cannot be used at the chosen pH value if protein crystals are not stable. If the catalityc process required the presence of oxygen or water, we can try to remove these species from the environment of the crystal while adding the substrate. For hydrolytic enzymes the approach would be to replace water itself, to a species that is not capable of functioning as a nucleophile. The observation that many crystalline enzymes can be transferred to neat organic solvents without disordering opens up the possibility for such substitution [19] [20] [21]. This approach has great promise for the study of many enzymes in view of the widespread role of water as a nucleophile in biology. In this work we isolated the substrate adduct of an oxygen-dependent enzyme, Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans*, by removing oxygen during the soaking process: crystals were soaked in a glove box a then flash-froze in liquid nitrogen directly in

the glove box. The frozen crystals were keeped under liquid nitrogen until mounting them on the diffractometer: defreeze was severely avoided since it can make some oxygen molecules enter the crystals (See Part II).

Direct observation of Michaelis complexes under equilibrium turnover conditions

The strategy makes use of the fact that if an enzyme catalyzes a simple single-substrate/single-product equilibration, all one has to do to observe the Michaelis complex is to soak the crystal in mother liquor containing a substrate concentration significantly greater (about 10 times) than K_m (the Michaelis-Menten equilibrium constant). Substrate will bind to the active site and be converted to product, but product is just the substrate for the back reaction, so the enzyme will settle to equilibrium and a crystal structure determination will show enzyme-substrate, or enzyme-product or a mixture of the two, whatever is the lowest free-energy species. Any isomerase, racemase, mutase, or epimerase can be studied this way, and in some cases other enzymes can also make use of this method if conditions can be found that cause the reaction to be effectively single-substrate/single-product [22].

Establishment of pseudo-steady-state conditions

Another method for trapping intermediates makes use of the fact that, under steady-state conditions, the species prior to the rate-determining transition state in an enzymic reaction will tend to accumulate: if product can be removed and there is enough substrate around to maintain saturation, then the enzyme will be in a substrate-bound form on average. In a protein crystallographic investigation, this can be done most easily by using a flow cell, which allows fresh substrate to flow into the crystal while any product that is formed is washed away [23].

Use of mutant enzymes

Often an enzyme-catalyzed reaction will feature several discrete steps that involve different functional groups in the active site. Mutation of just one of those groups to a non-functional amino acid can cause the reaction to stop at the point at which that group is required, leading to the accumulation of a prior intermediate. This strategy is finding widespread applications [24].

Direct binding of product

Considering that product is often a good competitive inhibitor for many enzymes, it is surprising that more investigators have utilized the simple but very effective strategy of just soaking the enzyme crystals in a concentration of product 10-fold or more higher than its inhibition constant (K_i). For example many phosphate-utilizing enzymes are found to have a phosphate ion bound to their active sites

naturally when their structures are determined, suggesting that it should be easy to do this, and when it has been employed this approach has worked very well [25].

Use of substrate analogs

Finally, a word must be said about the use of inhibitors and substrate analogs. By definition, such molecules cannot show all of the interactions and chemical transformations expected for the actual substrate, but when they function as mechanism-based inactivators they can sometimes be used to trap an intermediate or a close analog of one. Fluorinated substrates are often useful for this kind of study [26]. The danger in such approaches is that analogous compounds do not always bind analogously, so it is imperative that mechanistic conclusions obtained in this way (or any way, for that matter) are consistent with data obtained by other methods such as kinetics, chemical modification, spectroscopy and mutagenesis.

3

Crystallography in four dimensions

When X-ray used in crystallography (12 KeV in energy or about 1 Åin wavelenght) interact with a biological sample, as the protein crystal, only a small fraction of photos are elastically scattered giving rise to the diffraction pattern. The larger fraction of photons deposit energy into the sample and induce a redistribution of the electrons that causes a radiaction damage in the crystal; even though this process has been supposed to have a negative effect on the structural determination, since it causes primarly a loss of diffraction intensity and also strucutal damage in the protein, recently it has been noticed that it is possible to take advantage of it. Berglund, Carlsson and co-workers observed four successive photo-reduction events at the heme center of horseradish peroxidase, that allowed the isolation of the structures of compounds III, I and II of the ferric and finally of the ferrous states of the enzyme in a single set of experiments [27]. Adam and co-workers noticed a subtle expansion of the active site upon a one-electron photo-reduction in superoxide reductase, a non-heme iron protein that scavenges the superoxide radical [28]. In both cases, the structures of the intermediates were obtained from several composite low- to high-dose data sets and their biological relevance was verified by microspectrophotometry.

3.1 Interaction of X-rays with biomaterials

X-rays interact with matter through scattering and absorption. Elastic scattering carries easily accessible information about the structure, and is routinely used for structural studies. It is a relatively rare form of X-ray-matter interactions, and at 12 keV (about 1 Å) photon energy less than one tenth of X-ray interactions with carbon is elastic scattering. During this process, an incoming photon excites an electron. When the electron returns to the ground state, it re-emits the photon with

unchanged photon energy and phase (Figure 3.1).

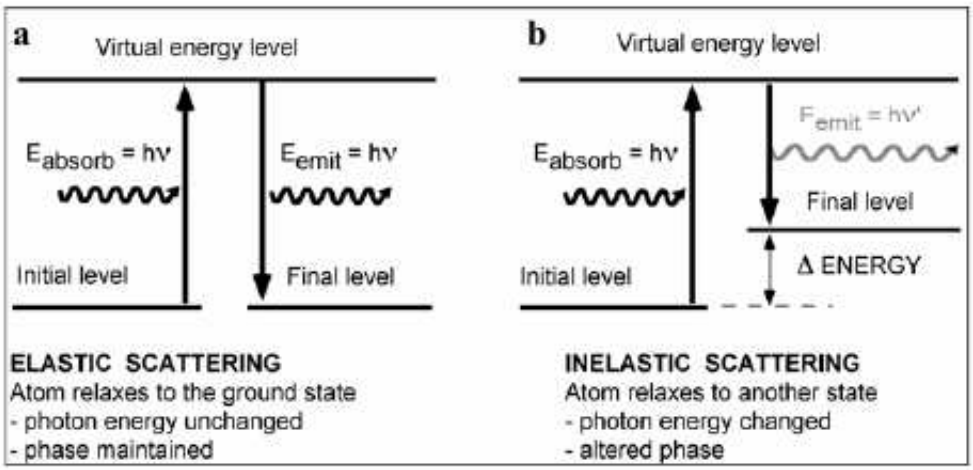


Figure 3.1: Elastic and inelastic scattering: the absorbed photon excites an electron. The atom can relax either through a) re-emitting the photon with unchanged properties (elastic scattering) or b) emitting a photon of different wavelength than the incoming photon (inelastic scattering).

Inelastic scattering, called also Compton scattering, is a relatively rare event at X-ray frequencies although it is the main source of energy deposition with hydrogen, and represents about 3% of all interactions between X-rays and a biological sample at 1 Åwavelength. During inelastic scattering, the incoming photon excites an electron but when the electron relaxes it emits a photon with smaller energy and does not come back to the ground state (Figure 3.1. Alternatively, the excited system may relax to a lower energy level than the initial state, in which case the energy of the photon emitted will be increased.

At 1 Åwavelength, about nine-tenth of the interacting photons will deposit energy into a biological sample, causing damage mainly through the photoelectric effect. The photoelectric effect causes an X-ray photon to be absorbed and a tightly bound electron ejected from an inner shell of the atom (Figure 3.2). Note that weakly bound outer shell electrons may also be ejected in the photoelectric processes, but this happens with a decreasing probability with increasing photon energies. The energy of the departing photoelectron depends on the difference between the energy of the X-ray photon and the binding energy of the electron. The binding energy of an inner shell electron in carbon is about 0.25 keV.

The departing photoelectron leaves a hole in a low lying orbital, and an upper shell electron falls into it. This electron may either emit an X-ray photon to produce X-ray fluorescence (dominant process with heavy elements) or may give up its energy to another electron within the atom, which is then ejected from the ion as an Auger

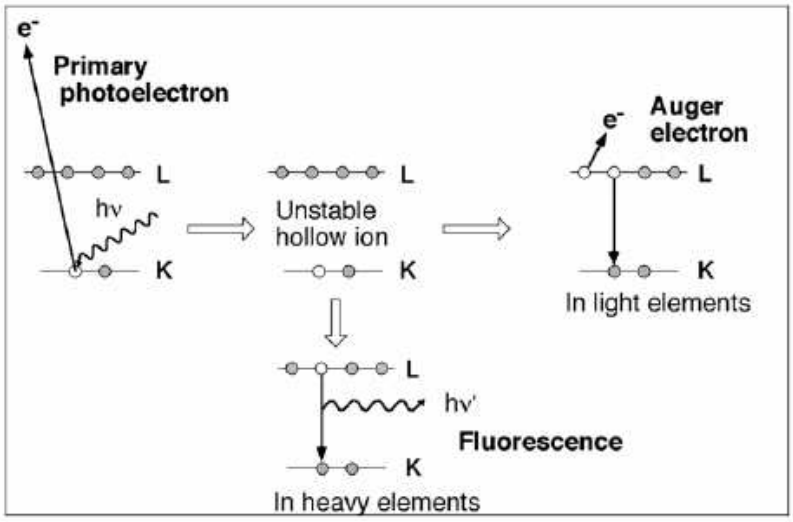


Figure 3.2: The photoelectric effect and atomic relaxations: when an inner shell electron is ejected, a hollow ion is formed in which a K-shell electron is missing; the hollow ion is unstable, ant it can relax by moving an electron from a higher level down to the K-shell to fill the hole. This process can result in the emission of an Auger electron in light elements, or in the emission of a fluorescent X-ray photon in heavy elements.

electron (dominant process with light elements). The probability of fluorescence emission or Auger emission depends on the binding energy of the electron, and through this, on the atomic number of the element. In carbon, nitrogen, oxygen and sulphur practically every photo-emission is accompanied by the emission of a secondary Auger electron, doubling the number of electrons liberated from these atoms for an absorbed photon. The energy of the Auger electrons is usually in the order of about 0.1-0.3 keV, and depends on the difference between the shell binding energies. The departing photoelectron can interact with the electrons left behind, depositing some of it's energy into the atom (a process called shake-up). This extra energy can be added to the energy of the Auger electron. Another possibility is that the excess energy causes ejection of outer shell electrons (a process called shake-off), sometimes also accompanied with the emission of more Auger electrons.

In a macroscopic protein crystal, all emitted electrons become thermalised and trapped in the sample. Thermalisation is mainly caused by inelastic electron-electron interactions. Thermalisation produces secondary cascade electrons, the total number of these is roughly proportional to the energy of the impact electron triggering the cascade. An Auger electron liberates between about 2 and 20 cascade electrons, while a photoelectron of 12keV liberates about 1000 cascade electrons before reaching thermal equilibrium. This process is over within about 100 femtoseconds. The consequence of these processes is that for each elastically scattered photons there

are about 10 000 electrons which will be redistributed in the sample (Figure 3.3).

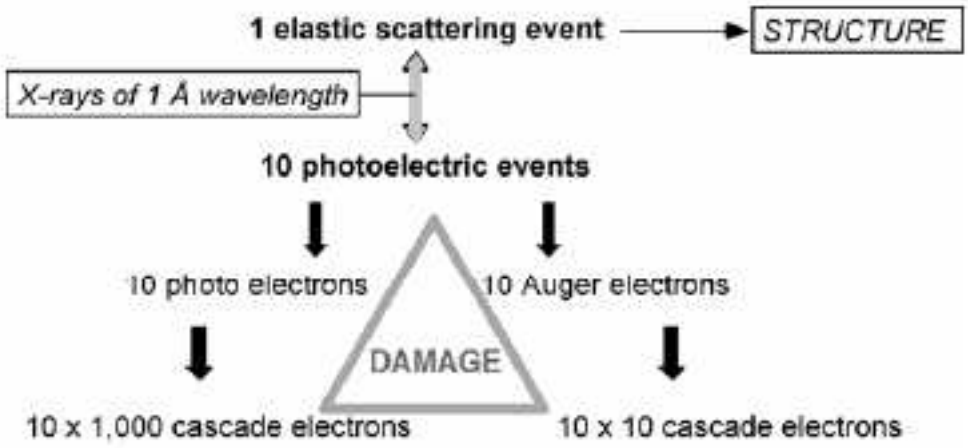


Figure 3.3: Yield of cascade electrons in a in a carbon sample at 12keV photon energy: for each elastically scattered photon 10 000 electrons are released and redistribuited in a protein crystal.

Going to shorter wavelengths improves the ratio between the elastic and the photoelectric effect. This improvement, however, does not result in a corresponding decrease in the ratio between electrons released and elastic scattering events. The discrepancy is due to the fact that at higher X-ray energies, the energy of the photoelectrons will also be higher, and this will then result in proportionally more cascade electrons during the thermalisation process.

3.1.1 X-ray absorption results in a redistribution of electrons in the sample

The thermalisation process plays a dominant role in sample ionisation. Atoms from which electrons are released become oxidised, while electrons liberated from these atoms cause reductions elsewhere in the sample. The process of liberating electrons is very fast (femtoseconds), while the subsequent redox and radical processes, leading to chemical damage caused by the redistribution of electrons takes place on a longer time scale. The slower redox processes can be described in terms of the Marcus theory, which states that the rate of electron transfer between two redox centres (for example an oxidized redox centre and a hydrated electron) is dependent on:

- the temperature
- the distance between the two centres
- the medium between the two centres

- the difference in redox potentials of the two centres
- the reorganisation energy associated with the conformational change needed to adopt the new redox state.

The mechanism of charge redistribution involves hole migration (positive charges) and electron migration (negative charges). In water, the main reactive species formed upon photo-absorption are hydrated electrons and hydroxyl radicals. Hydrated electrons cause reduction while hydroxyl radicals are strongly oxidising. Both electrons and electron holes can migrate along the protein backbone, while the route of electron transfer involving hydrated electrons mainly proceeds through hydrogen bonds. It has been suggested that diffusion processes are involved in X-ray induced redox processes in the sample. In the protein, one can observe changes at disulphide bridges, carboxylates and also at redox centers.

3.2 Following redox processes in protein crystals

The key in taking advantage of the thermalisation effect is to notice that, even though the redistribution of electrons can cause structural changes and disruption (an example are the disulphide bridges), it can also reduce redox centers like metals in metalloproteins and thus simulate the catalityc process: if we are able to control the X-ray dose received by the crystal, we can somehow tune the thermalisation and the reduction of redox centers and thus isolate different reaction step at different redox states of the enzyme. Isolation of each subsequent redox step could results in an *X-ray movie* of the enzymatic reaction.

3.2.1 Redox titrations in protein crystallography

A molecular description of oxygen and peroxide activation in biological systems is problematic because electrons liberated during X-ray data collection easily reduce the active centres of the redox enzymes catalysing these reactions. X-rays leave a fingerprint on structures, especially when high doses are needed to collect a data set, because the strucutre obtained will show the redox centers completelly reduced. This is an increasingly serious problem with high resolution crystal structures where large X-ray doses are needed to reach desired resolutions. Attempts to minimise X-ray-induced interference with redox centres in crystallography include soaking the crystals in excess electron scavengers and working with shorter wavelength X-rays. Cooling can also slow down sample deterioration but it cannot eliminate X-ray-induced ionisation in the sample. However, X-ray data collection at low temperatures can minimize the chemical contribution to X-ray-induced damage but the number of electrons liberated and redistributed in the sample remains practically unaffected upon cooling. The approach of obtainig each redox state of a metalloprotein by collecting different data set at different X-ray dose is thus inappropriate.

Figure 3.4 shows a novel data collection strategy used in our study on Laccase from *Steccherinum ochraceum* to deal with these problems more effectively. The

method is based on damage distribution over many crystals followed by the assembly of composite data sets from small chunks of these data. Composite data sets obtained this way report on structures which received different X-ray doses, and in which a different numbers of electrons were liberated during data collection. This strategy offers the possibility to control and monitor X-ray-induced redox processes at any given X-ray wavelength. The technique permits experiments similar to redox titrations so that redox processes can be followed in the crystal.

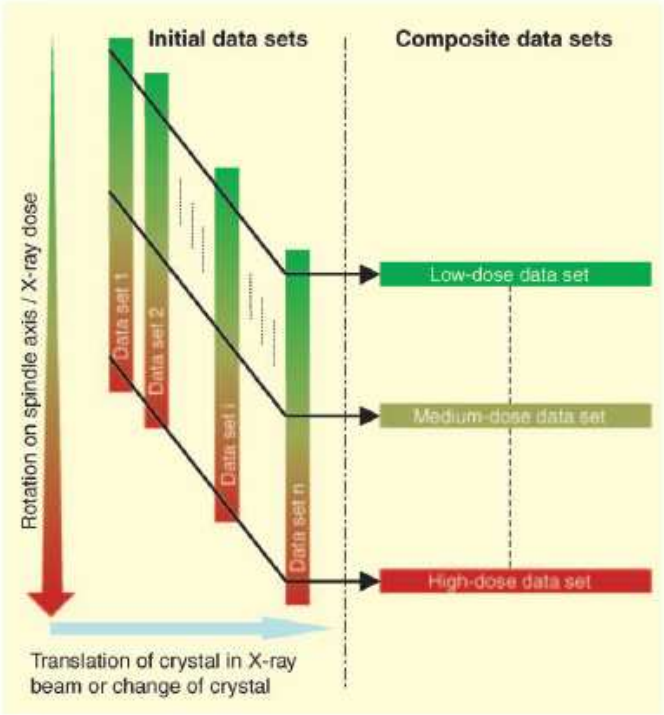


Figure 3.4: The technique: a series of data sets is collected either on one crystal, with adequate rotation and translation between subsequent sets, or on several crystals. Dose-dependent data sets are then reconstructed by properly combining fixed-dose subsets.

3.2.2 Monitoring different redox states in the crystal

The development of single-crystal spectroscopies, primarily single-crystal UV/visible spectrophotometry, has been key to allowing monitoring of reactions in crystals. In this way, the build-up of spectrally distinct reaction intermediates in the crystalline environment can be followed. These studies have shown that, in most cases, the proteins use the same reaction routes as in solution, but the rates and equilibria are often perturbed. In particular, the rates often slow down to such an extent that

intermediates which are only transient in solution can be trapped in the crystal. The use of spectroscopy has enabled monitoring for redox changes in metalloprotein intermediates caused by X-ray exposure during data collection.

The use of single-crystal spectroscopy is of further importance if we perform a redox titration as decribed above. When obtaining a structure for an intermediate to dissect a mechanism it is obviously important to know what is really present in the crystal and in what concentration. In many papers in the literature, only the electron density has been used as an indicator to which species the structure represents. However, it is always preferable to have additional independent information about the nature and the concentration of the intermediate in the crystal. Ideally this should be checked during data collection as the stability of the intermediate is influenced by exposure to X-rays. Chemical reactions involve the shuffling of electrons between ions, atoms and molecules, and the transfer of electrons is accompanied by changes in the electromagnetic spectrum of the reactants. Spectral changes associated with chemical reactions cover a very broad spectral domain. Several spectroscopic methods have been used on single crystals in combination with X-ray crystallography. The most frequently used technique is UV-VIS absorption spectroscopy. Other methods include fluorescence spectroscopy, FTIR spectroscopy, single crystal EPR/ESR, single crystal EXAFS and polarized single crystal XAS, and single-crystal micro-Raman spectroscopy. This list is far from complete, and only indicates the range of methods that can be used to obtain independent information on the nature and concentration of catalytic intermediates in a crystal.

Single crystal microspectrophotometry

In the present study on the mechanism of Laccase, we used UV-VIS microspectrophotometry to obtain independent spectral information on the state and the contents of the crystals.

Microspectrophotometry can be used to correlate electronic transitions with structural transitions in crystals, and the spectral data can be used to follow reactions and to distinguish between different states of the protein and/or the substrates. Note however that spectral work on crystalline proteins is different from spectral study on the same proteins in solution in some important aspects.

Proteins dissolved in dilute aqueous solutions take up independent and random orientations, while proteins in a crystal lattice are fixed relative to each other. Molecules fixed in space absorb plane-polarised light differently from different directions. Linear dichroism refers to differences in the absorption of plane-polarised light as a function of the orientation of the molecule (or condensate) relative to the plane of polarisation of the incident light. Linear dichroism can thus be observed for most single molecules and also for condensed matter with symmetry lower than cubic. This is a feature of optical measurements on crystalline samples. If one wants to get something similar to a solution spectrum, one needs to integrate the spectra over different measurement angles. Even then one may find some differences due to packing interactions within the crystal. Absorption of plane-polarised light with

the electric vector parallel to the principal optical directions of the crystal obeys the Lambert-Beer law and can be used for straightforward spectral measurements.

Another factor is the effect of refraction and reflection on the measured spectrum. When recording a solution spectrum in a spectrophotometer one usually uses a cuvette with its transparent parallel faces normal to the incident beam. In this case there is very little reflection or refraction from the sample. So the best case one can hope for in measurements on crystals is a crystal surface perpendicular to the incident light beam and a parallel crystal face on the exit side. Experience shows that parallel crystal faces mainly exist in text books. Real crystals have odd prism shapes. The crystal is usually also surrounded by additional surfaces (the droplet, the capillary tube), which are usually curved. The combined geometry will give rise to refraction and reflection at all possible interfaces. Reflection and refraction depend on the wavelength and on the angle of incidence. These factors will give us different base lines, depending on the orientation of the crystal (the droplet and the capillary) in the light beam. This effect is the smallest when the crystal has parallel faces, and it is located in the middle of a thin droplet, with faces perpendicular to the light beam. In the case of crystals mounted in capillaries, care should be taken to make sure the light passes through the centre of a capillary and is normal to the axis of the capillary tube.

Changes in the temperature will also influence the spectrum. At low temperatures fewer vibrational states are occupied, and this can lead to a sharpening of certain peaks. Temperature will also influence the occupancy of certain molecular orbitals, and this can produce changes in the spin state of the sample, which will give an altered spectrum. Finally, temperature dependent distortions to the chromophores may occur as the protein is cooled in the lattice, and this can also affect the spectral properties of the sample.

An obvious way around most of these problems is to measure spectral changes in a reaction under identical measurement conditions (fixed geometry, stable crystal orientation, controlled temperature, etc.). In most of the studies described here we used air as a reference. This arrangement removes half of the problems associated with the reproducibility of spectral measurements.

Part II

Salicylate 1,2-Dioxygenase: a study on substrate specificity

4

The structure of Salicylate 1,2-Dioxygenase

The aerobic degradation of aromatic compounds by microorganisms is usually accomplished via the intermediate formation of dihydroxylated aromatics. The typical ring-fission substrates are compounds such as catechol, protocatechuate, homoprotocatechuate, hydroquinone, gentisate or homogentisate (or their substituted derivatives). These ring-fission substrates are subsequently converted to non-aromatic compounds by ring-cleaving dioxygenases, which cleave the carbon-carbon bond of vicinal diols such as catechol(s) and (homo)protocatechuates either between the two phenolic hydroxyl groups (intradiol or ortho-cleavage) or adjacent to them (extradiol or meta-cleavage). These reactions result in the formation of either dicarboxylic acids (substituted muconates) in the case of ortho-cleavage pathways or monocarboxylic acids carrying an additional aldehyde function (or rarely keto-group) in the course of the extradiol cleavage reactions [6].

The para-substituted aromatic diols, such as (substituted) hydroquinones, homogentisate, and gentisate are attacked by ring-fission dioxygenases that cleave the aromatic nucleus between a carbon atom carrying a hydroxyl-group and an adjacent carbon atom. In the case of the hydroquinones carrying an additional carboxy-group, like gentisate and homogentisate, the ring-fission generally occurs in the 1,2-position between the carbon atoms carrying the carboxylated side-chain and the hydroxyl-group in the 2-position. Therefore, these reactions usually also result in the formation of monocarboxylic acids with an additional aldehyde group (during the cleavage of hydroquinones) or dicarboxylic acids (during the cleavage of gentisate or homogentisate). Ring-fission dioxygenases that catalyse these reactions contain, in most cases, catalytically active Fe^{3+} ions (intradiol dioxygenases) or Fe^{2+} ions (extradiol cleaving enzymes and para-substituted diols cleaving enzymes).

There are only a few enzymes known that cleave monohydroxylated aromatic compounds. The majority of these reactions involve aminophenols, such as 3-

hydroxyanthranilate [29], 5-aminosalicylate [30], or 2-aminophenol [31] [32].These ring-fission substrates clearly resemble the conventional dihydroxylated substrates and it is generally assumed that in these cases the amino group fulfils the function of the missing hydroxyl-group in order to activate the aromatic ring for the electrophilic attack of the dioxygen molecule.

More recently, some monohydroxylated compounds without an additional amino group have been described as ring-fission substrates. This was found for compounds such as (substituted) salicylates and 1-hydroxy-2-naphthoate (an intermediate in the bacterial degradation of phenanthrene). These substrates are cleaved between the carbon atoms carrying the carboxy group and the hydroxyl group [33] [34] [35]. Thus, these reactions clearly resemble the ring-fission reactions observed with gentisate 1,2-dioxygenases. The cloning of a 1-hydroxy-2-naphthoate dioxygenase from a *Nocardioides* sp. [36] and a salicylate 1,2-dioxygenase from *Pseudaminobacter salicylatoxidans* BN12 [37] demonstrated significant sequence homologies between gentisate 1,2-dioxygenases, 1-hydroxy-2-naphthoate dioxygenases and the salicylate cleaving enzyme from *P. salicylatoxidans*. The ring-fission dioxygenase from *P. salicylatoxidans* appears to be the most versatile of these enzymes because it converts a wide range of substituted salicylates and 1-hydroxy-2-naphthoate. Since it shows the highest enzyme activities with gentisate as substrate it appears to be a gentisate 1,2-dioxygenase that has been adapted to the conversion of monohydroxylated substrates [37].

In this work we tried to clarify the molecular basis for the extraordinary catalytic competence of salicylate 1,2-dioxygenase from *Pseudaminobacter salicylatoxidans*. We solved the structure of the wild type enzyme and its adducts with substrates salicylate, gentisate and 1-hydroxy-2-naphtoate; in order to understand the peculiar activity of this enzyme we introduced mutations in the residues that we thought involved in substrate recognition on the basis of structural analisys of wild type—substrate adducts. The structures obtained of mutants and mutant—substrate adducts helped us to explain the substrate specificity and the catalytic mechanism of the enzyme.

4.1 Wild type structure

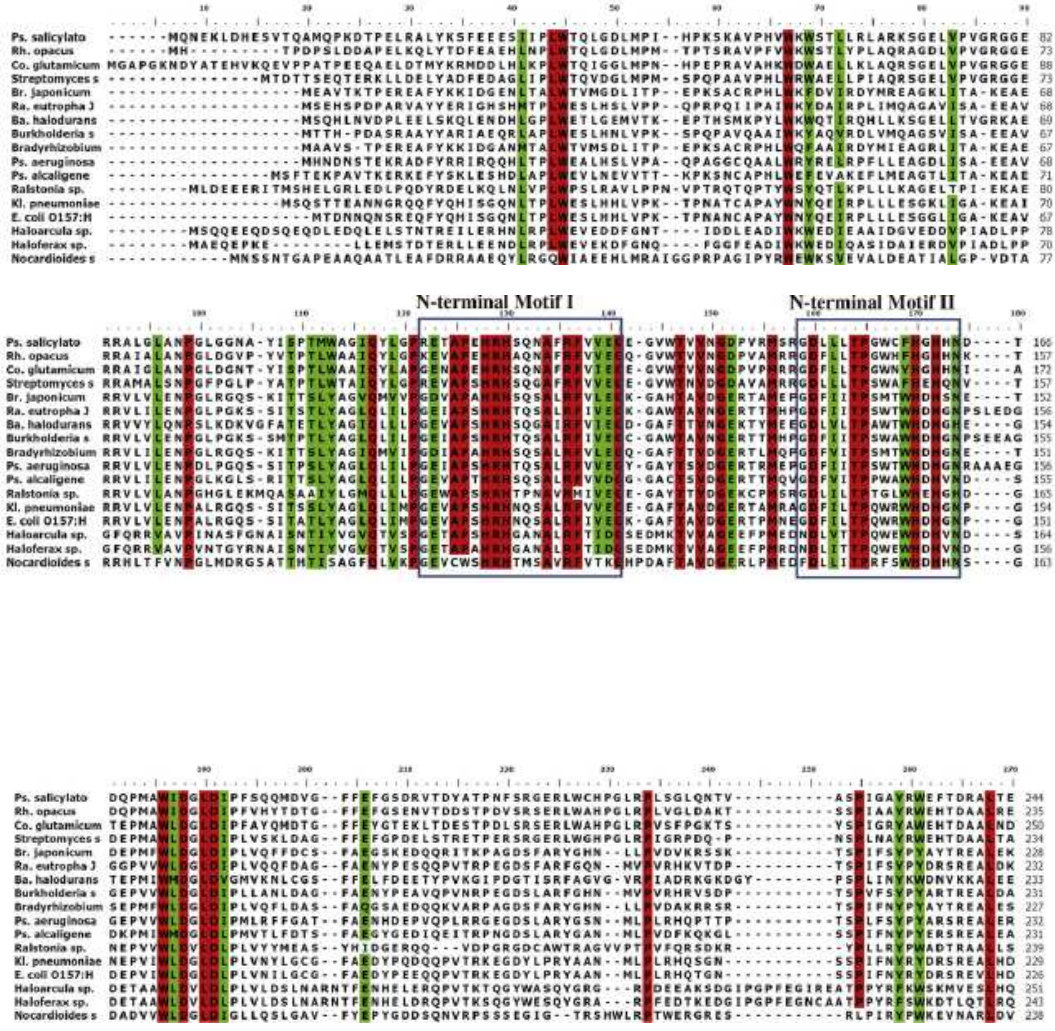
4.1.1 Sequence alignement

Salicylate 1,2 dioxygenase from *P. salicylatoxidans* exhibits sequence homology to several gentisate 1,2-dioxygenases (GDO); among these, the structure of which is known only for the GDO from *Escherichia coli* O157:H7 [38]. ClustalW sequence alignments of salicylate 1,2-dioxygenase with several GDOs as well as a 1-hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7 are reported in Figure 4.1. The highest level of homology of SDO is found with GDO from *Rhodococcus opacus*: 71% identity (80% similarity), whereas the homology with the GDO from *E. coli* O157:H7 (eGDO) is 35% identity and 51% similarity, and that with 1-hydroxy-2-naphthoate dioxygenase is only 29% identity and 44% similarity. Analysis of GDO protein

sequences reveals the presence of a highly conserved $G(X)_5HXH(X)_{3,4}E(X)_6G$ metal-binding motif (motif I), along with a $G(X)_5PXG(x)_2H(x)_3N$ motif (motif II) (Figure 4.1).

These motifs prove that such enzymes are all members of the cupin superfamily, one of the most functionally diverse protein classes named on the basis of a conserved β-barrel fold [39]. The cupin superfamily includes both enzymatic and non-enzymatic members, which have either one or two cupin domains. The cupin fold comprises a motif of six, seven or eight antiparallel β-sheets located within a conserved β-barrel structure. Analysis of cupins has identified at least 18 different functional classes that range from single-domain bacterial enzymes, such as isomerases and epimerases, to two-domain bicupins, like some storage globulins and multidomain transcription factors. The largest subset of the cupin superfamily is the 2-oxoglutarate-Fe^{2+}-dependent dioxygenases. Although the majority of enzymatic cupins contain iron as an active site metal, other members contain either copper, zinc, cobalt, nickel or manganese ions as a cofactor allowing a variety of biochemical functions to occur within the conserved tertiary structure.

As well as the GDO from *E. coli*, SDO is a bicupin with two germin-like β-barrel domains. In each β-barrel, both motifs exhibit mutations with respect to the standard cupins. In particular in the N-terminal β-barrel motif I (residues 113132, see Figure 4.1), glutamic acid in standard cupins has been replaced by alanine (Ala 126) in GDOs/SDO, and the first glycine residue is substituted with arginine (Arg113) in SDO. The glutamic acid to alanine substitution is highly conserved among GDO enzymes. Since other bicupin proteins employ the glutamic acid residue for metal coordination, the alanine substitution in GDOs/SDO may be vital to allow the simultaneous coordination of the substrate and molecular oxygen to the iron ion center and therefore allowing catalysis to occur.


N-terminal Motif I
N-terminal Motif II
Ps. salicylato
Rh. opacus
Co. glutamicum
Streptomyces s
Br. japonicum
Ra. eutropha J
Ba. halodurans
Burkholderia s
Bradyrhizobium
Ps. aeruginosa
Ps. alcaligene
Ralstonia sp.
Kl. pneumoniae
E. coli O157:H
Haloarcula sp.
Haloferax sp.
Nocardioides s

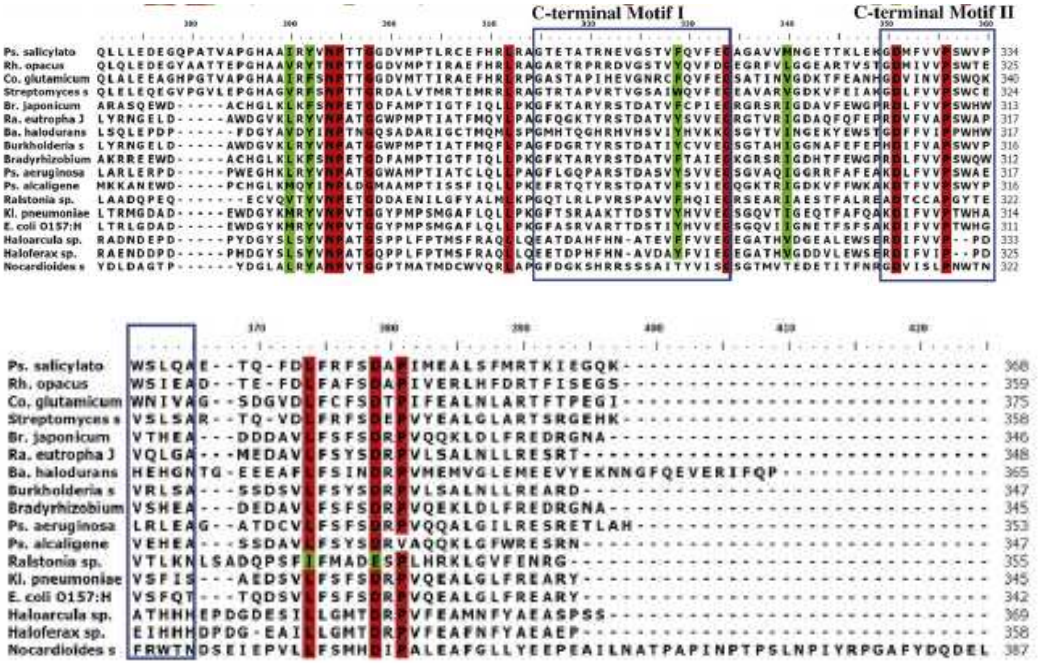


Figure 4.1: Amino acid sequence alignments of salicylate dioxygenase, 1-hydroxy-2-naphthoate dioxygenase, and other known gentisate dioxygenases: (a) *Pseudaminobacter salicylatoxidans* SDO (AAQ91293.1); (b) *Rhodococcus opacus* GDO (ABH01038.1); (c) *Corynebacterium glutamicum* ATCC 13032 GDO (NP_602217.1); (d) *Streptomyces* sp. WA4 GDO (BAC78375.1); (e) *Bradyrhizobium japonicum* USDA 110 GDO (NP_766750.1); (f) *Ralstonia eutropha* JMP134 GDO (AAZ65207.1); (g) *Bacillus halodurans* C-125 GDO (BAB05723.1); (h) *Burkholderia* sp. 383 GDO (YP_372647.1); (i) *Bradyrhizobium* sp. BTAi1 GDO (ZP_00861947.1); (j) *Pseudomonas aeruginosa* PAO1 GDO (NP_251160.1); (k) *Pseudomonas alcaligenes* GDO (AAD49427.1); (l) *Ralstonia* sp. U2 GDO (AAD12619.1); (m) *Klebsiella pneumoniae* GDO (AAW63413.1); (n) *Escherichia coli* O157:H7 GDO (BAB36453.1); (o) *Haloferax* sp. D1227 GDO (AAQ62856.1); (p) *Haloarcula* sp. D1 GDO (AAQ79814.1); (q) 1-hydroxy-2-naphthoate dioxygenase from *Nocardioides* sp. KP7 (BAA31235). In the alignment, positions identical in all sequences are marked with a red background. Regions in which the sequences are similar are marked with a green background. Codes in parentheses refer to NCBI protein accession.

The second mutation, between standard cupins (Gly) and GDOs/SDO, in motif I is observed in only a few known enzymes, such as the GDO from *Streptomyces* sp. (R104), the GDO from *R. opacus* (K104), and in the present SDO (R113). Motif II in the N-terminal β-barrel of the GDO family also exhibits a substitution of glycine by tryptophan in the consensus sequence, a mutation that could contribute to the hydrophobic shielding of the lower part of the active site. The C-terminal β-barrel domain of SDO shows differences from that of the N terminus. In fact, its motifs (motif I, residues 289308; and motif II, residues 324339) are largely mutated and lack the metal-binding residues normally associated with the cupin fold (Figure 4.1).

Extradiol-type dioxygenases have been classified with structure-validated phylogenetic analyses into three evolutionarily independent families [40]: type I extradiol dioxygenases that are part of the vicinal oxygen chelate superfamily, type II extradiol dioxygenases that include multimeric enzymes not classified in a superfamily, and bicupin dioxygenases that compose the type III superfamily [41]. The dioxygenases belonging to the type III superfamily such as 3-hydroxyanthranilate-3,4-dioxygenase (HADO), quercetin 2,3-dioxygenase (QDO), cysteine dioxygenase (CDO), homogentisate dioxygenase (HGDO), and all the GDOs usually contain a single active site in one of the two domains, with the other domain remaining as a non-catalytic vestigial remnant. However, SDO is similar to the other bicupin dioxygenases reported above, in its conservation of the metal co-ordination site within the N-terminal domain. HGDO, on the contrary, exhibits a C-terminal metal co-ordination site [42].

4.1.2 Subunits and quaternary structure

The crystal structure of Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans* [43] (SDO) was determined at 2.9 Åresolution by molecular replacement techniques using the four subunits of gentisate 1,2-dioxygenases from *Escherichia coli* O157:H7 (eGDO) [38] (PDB accession code 2D40) as a starting model. The two enzymes are homo-tetramers and a sequence comparison between the two monomeric subunits shows a relatively low sequence identity (35%).

SDO contains one Fe(II) ion and has a molecular mass of about 41 kDa per subunit [37]. Its overall dimensions are approximately 100 Å× 80 Å× 60 Å. The final model includes a total of 175 water molecules and for each monomer, residues 16368 (the exceptions are reported in Experimentals), and one Fe(II) ion. The asymmetric unit contains one tetramer. The global structure, depicted in Figure 4.2, resembles that of eGDO. The comparison of a single subunit of SDO and eGDO is given in Figure 4.2 as an overall least-squares superposition. Several differences are evident, in particular in the N-terminal region. Whereas the first N-terminal 32 residues are missing from the eGDO structure, probably due to inadequate electron density caused by the mobility of this branch, in SDO, which has a 14 residues longer N-terminal tail, only the electron density of the first 14 or 15 residues is missing. This results in the disclosure of a novel structure fragment in SDO that contains an N-terminal random coil (residues 15-21) followed by an 11 residue α-helix (H1 in Figure 4.3, residues 22-32) and a further 15 residue random coil arm (residues 33-47). A more defined electron density also allowed the identification of the structure in the region comprising residues 57-87 where a β-sheet-α-helix-β-sheet motif is now evident.

The N-terminal cupin domain starts at residue 113 and terminates at residue 164, whereas the C-terminal cupin domain initiates at residue 289 and ends at 339. The two cupin domains are arranged face to face. The interface of the two domains is formed through extensive hydrophobic interactions that permit rigid relative positioning of the N- and C-terminal domains. They are connected through a peptide region composed of a short α-helix (H5, 178-183) a β-sheet (S12, 189-192) followed by a gap of two amino acids missing (196-197) and three α-helices (H6-H8) of different lengths interposed by two short β-sheet (S13-S14) spacers. The C-terminus is then composed of a final α-helix (H9) and a β-sheet (S24). The bicupin folds of each subunit are flanked by extensive regions containing mostly α-helices and random coil. These regions are responsible for the formation and stabilization of the tetrameric organization.

As in eGDO, the structure of SDO exhibits two distinct types of intermolecular interfaces. The first boundary mainly involves symmetrical contacts between molecules A and C, and between molecules B and D (Figure 4.2). Primarily hydrophobic intermolecular interactions are involved at this site, which mainly comprise -helix H5 and -sheet S12 (residues 179-194), and the region encompassing -helix H9, (residues 349-355); the formation of a number of hydrogen bonds also contributes to stabilize the quaternary assembly (i.e. Ser34 with Ala89, Ile36 with

Gly87, Leu38 with Arg83 and Ala85, Asp185 with Ser52 and Ser 331, Phe189 with Ile48 of the facing subunit). The interfaces of molecules A with C, and B with D are further stabilized through swapping of the flexible N-terminal region I (residues 15-52), in the vicinity of the metal co-ordination sites (Figure 4.2). The reciprocal extension of the N-terminus from molecule A to molecule C, and from molecule B to molecule D, respectively, is clearly visible in the present structure. Contrary to what was observed for eGDO, the N-terminus, with the exception of the first 14-15 residues, is ordered and contributes several residues to the active site that possibly influence substrate specificity. The second interface is involved in the interaction of molecules A with D and B with C (Fig. 4.2). The region comprising helices H6, H7, and H8, and the intercalating sheets S13, S14, and S15 is implicated in such interactions that are mainly hydrophobic in nature, although a number of hydrogen bond contacts are observed in this region: Arg209 with Glu244 and Glu251, Gly215 with Val267, Arg217 with Arg265 and Gly272, Ser220 with Ala255. The resulting oligomeric state, similar to the majority of GDOs studied previously, mediated by the two interface types, results in the formation of a large, square ring structure with a central tunnel that exhibits a significant polar character. The quaternary structure is arranged in such a way that the active metal sites are oriented towards the external edge of the tetramer.

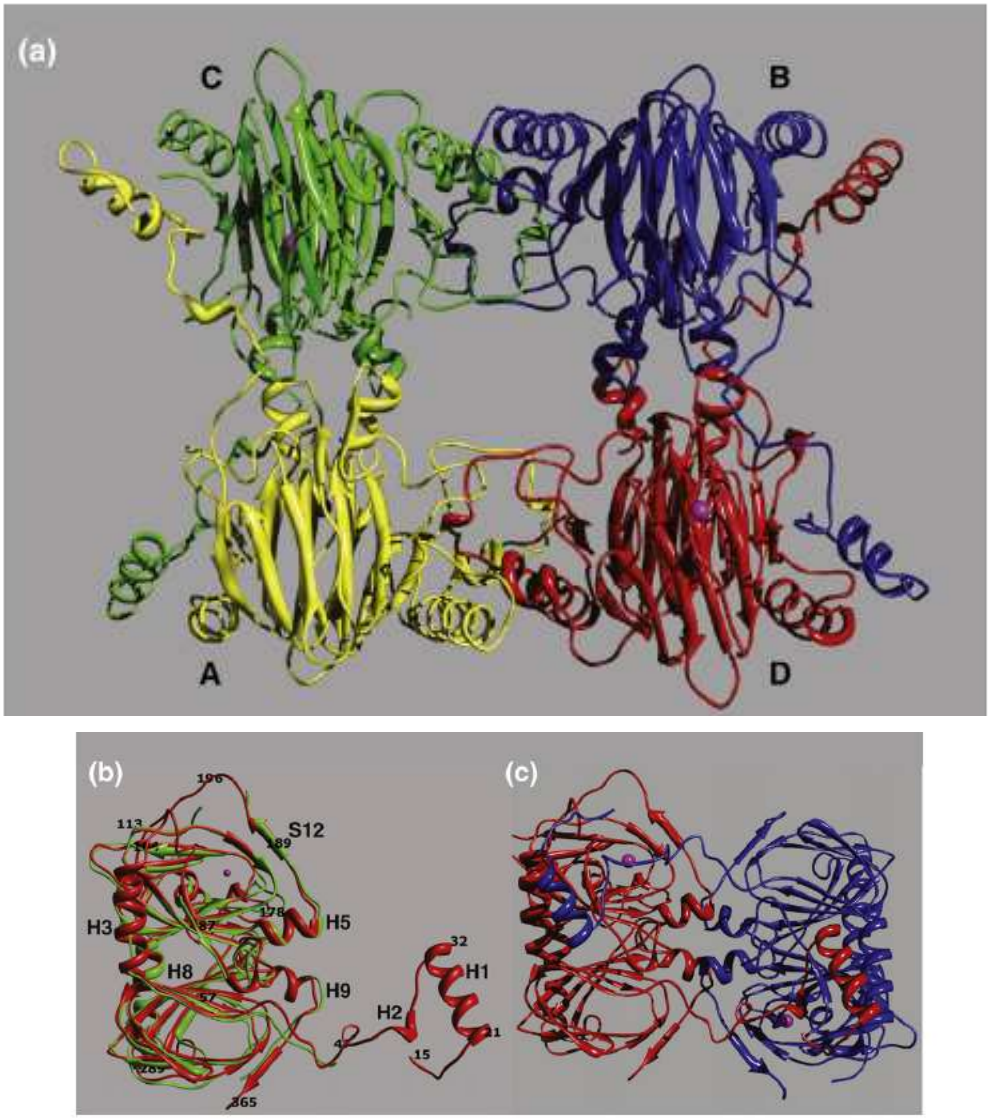


Figure 4.2: a) The tetrameric structure of SDO; the iron ions are represented as magenta spheres; b) superposition between a single subunit of SDO (red) and one of GDO from E. coli (green); c) Dimeric arrangement between subunits D and B of SDO.

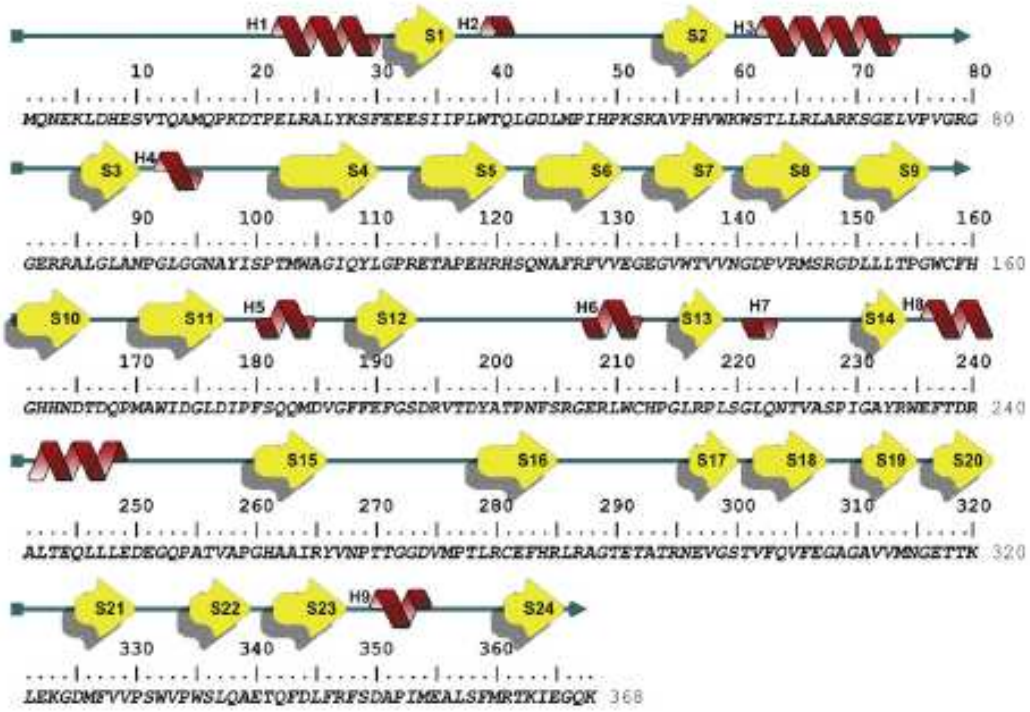


Figure 4.3: Schematic representation of the structure of SDO.

4.1.3 The catalytic domain and the active site metal coordination

A close-up view of the active site of SDO is presented in Figure 4.4. The catalytic center contains a iron(II) ion bound to the enzyme with three histidine residues (His119, His121, and His160) from the N-terminal cupin motifs I and II, a feature recently observed in a number of bicupin dioxygenases.

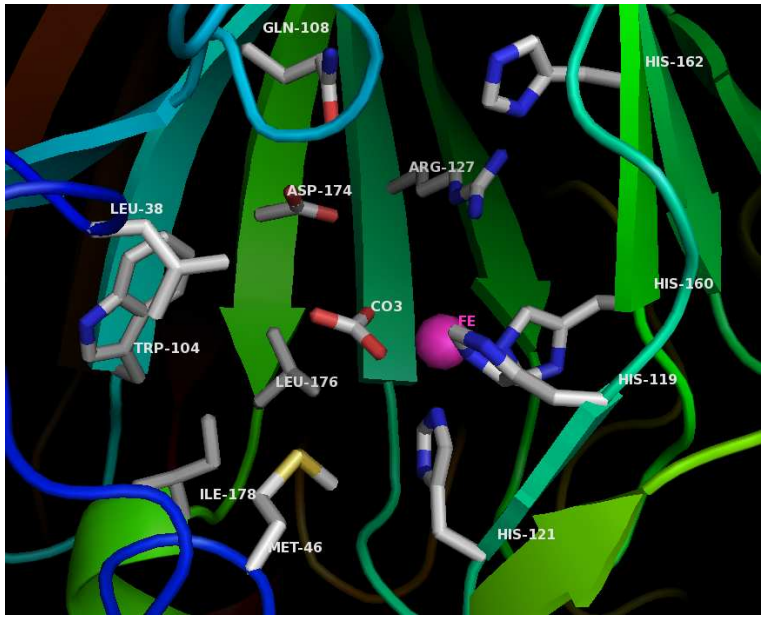


Figure 4.4: Active site of SDO.

Since four subunits (one tetramer) are present in the asymmetric unit of the present enzyme crystal, we analyzed the active site differences among them. The comparison of the bond distances of the catalytic iron ion coordination polyhedron between SDO subunits reveals some heterogeneity in the four different subunits. The three protein ligands to the active metal ion, His119, His121, and His160, exhibit Fe(II)—N bond lengths in the intervals 1.85-2.02 Å, 2.13-2.31 Å, and 1.87-2.17 Å, respectively, depending on the subunit analyzed. Moreover, a single water molecule appears to complete the coordination sphere of the active site iron(II) metal ion in subunits C and D located at 1.93 Åand 2.55 Å, respectively (see Figure 4.5 a) for subunit D). In subunits A and B, a larger electron density was found in the coordination sphere of the iron and was interpreted as an acetate ion (see Figure 4.5 b) for subunit B), probably related to the presence of acetate ions in the crystallizing solution. This moiety is bound in an asymmetric bidentate way with one of the oxygen atoms at 2.53 Åor 2.91 Åand the other at 3.40 Åor 3.37 Ådistance from the iron ion in subunits A and B, respectively; additionally, in subunit B, a water molecule further completes the metal coordination at 2.16 Å.

The coordination of water/hydroxide ions to the active site ferrous iron was

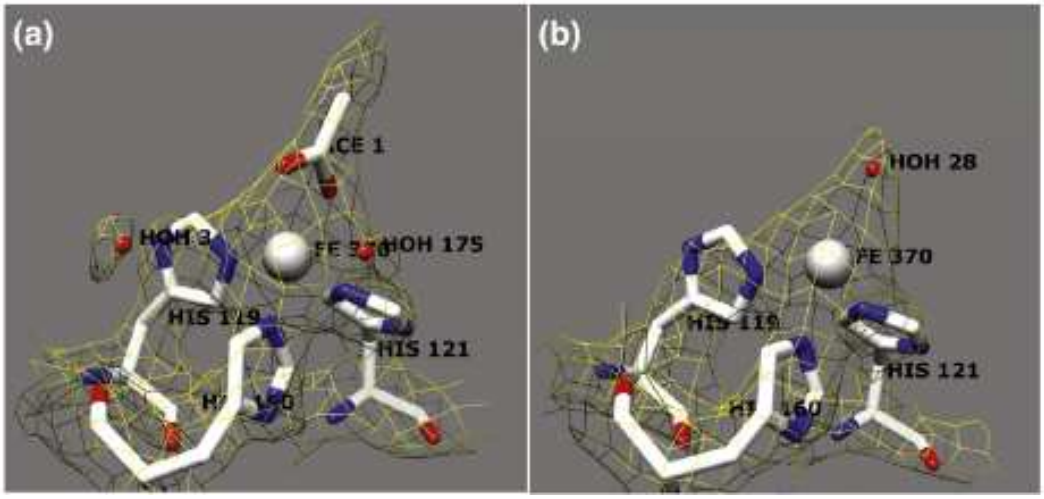


Figure 4.5: Views of the $2F_{obs} - F_{calc}$ electron density maps for the active sites of subunits B and D of SDO. The electron density is contoured at 2σ.

previously suggested from the significant broadening of the electron paramagnetic resonance spectroscopy (EPR) spectrum due to ^{17}O hyperfine coupling for the nitrosyl complex of a GDO in ^{17}O-enriched water, and up to three water molecules were observed bound to the iron ion in the structure of eGDO [38]. In spite of the phylogenetic diversity of extradiol-type dioxygenases, all the members of type I and II classes contain a 2-His-1-carboxylate metal-binding motif [44]. In contrast, among the bicupin dioxygenases (type III) the catalytic metal coordination appears to be heterogeneous. The classical 2-His-1-Glu coordination is observed in 3-hydroxyanthranilate-3,4-dioxygenase [45] and human homogentisate dioxygenase [42]. The crystal structures of two quercetin 2,3-dioxygenases revealed that their metal active centers (containing Fe(II) or Cu(II)) both display two distinct geometries: a distorted tetrahedral coordination, composed by three Histidine residues and one water molecule, or a distorted trigonal bipyramidal environment, that additionally includes a Glu residue [46] [47]. Other type III dioxygenases, such as the rat cysteine dioxygenase and eGDO as well as the present SDO, show a mononuclear iron center involving only three histidine ligands, the iron coordination being completed by water molecules [38] [48].

Regarding the nature of active metal ions, copper [46], manganese [49] and magnesium [50] containing dioxygenases have also been isolated. The structural knowledge on how dioxygenation is carried out by those non-iron centers is not completely clear. Studies of quercetin 2,3-dioxygenase from *Bacillus subtilis* suggest that replacement of the iron at the active site with different metal ions causes the modulation of enzymatic activity in accordance with the Irving-Williams metal ion behavior, further providing clues to the diversification of function in the bicupin family of pro-

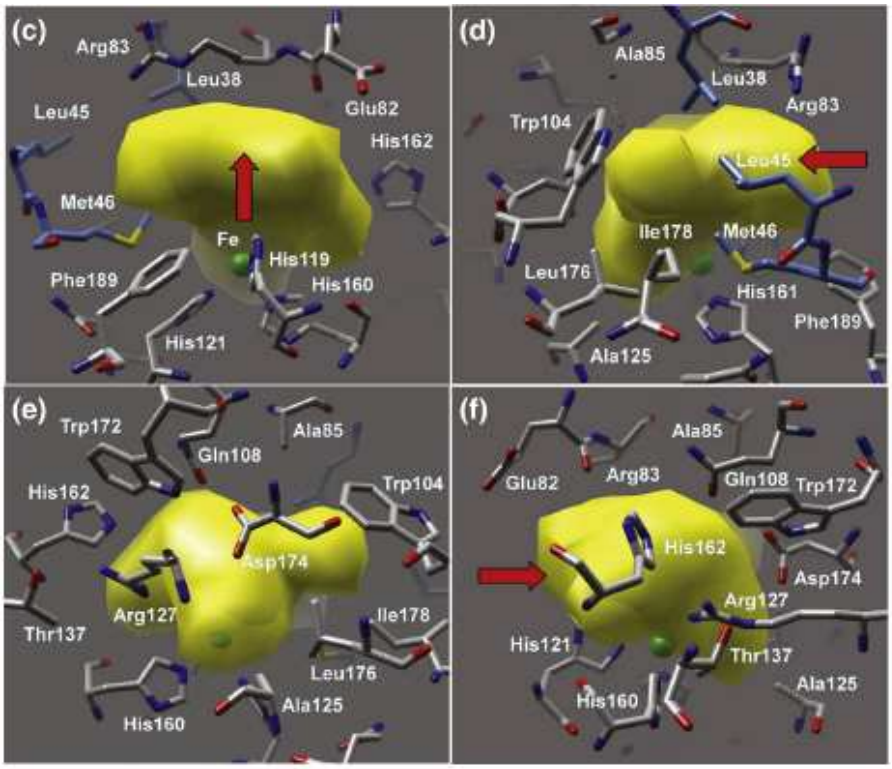


Figure 4.6: Representations of the solvent-accessible surface of the active site cavity of SDO (yellow). The four pictures show 90 counterclockwise rotated (around a vertical axis) views of the cavity c) 0; d) 90; e) 180; and f) 270. The iron ion is represented as a green sphere, and the red arrow indicates the active site entrance.

teins [47]. The metal coordination motif of SDO is comprised in a solvent-accessible pocket depicted in Figure 4.6 and composed of many hydrophilic residues: Met46, Glu82, Arg83, Gln108, Gln123, Arg127, Thr137, His162 and Asp174, and several hydrophobic amino acids: Leu38, Leu45, Ala85, Trp104, Ala125, Trp172, Leu176, Leu178 and Phe189. Many of them are highly conserved among all the known GDOs. Specifically, Leu38, Gln108, Ala125, Arg127, His162, Trp172, Asp174, and Leu176 are strictly conserved in all of the GDO homologues analyzed in Figure 4.1.

Figure 4.7 shows the superposed active sites of SDO and eGDO, the only known structure of gentisate 1,2-dioxygenase. A few amino acid substitutions are observed in the catalytic cavities of SDO/eGDO respectively: Met46/Val32 (not observed in the eGDO structure) Glu82/Val67 (not observed in the eGDO structure), Ala85/Leu71, Trp104/Tyr89, and Phe189/Ala174, the other residues are conserved even if small backbone and side-chain shifts are detected.

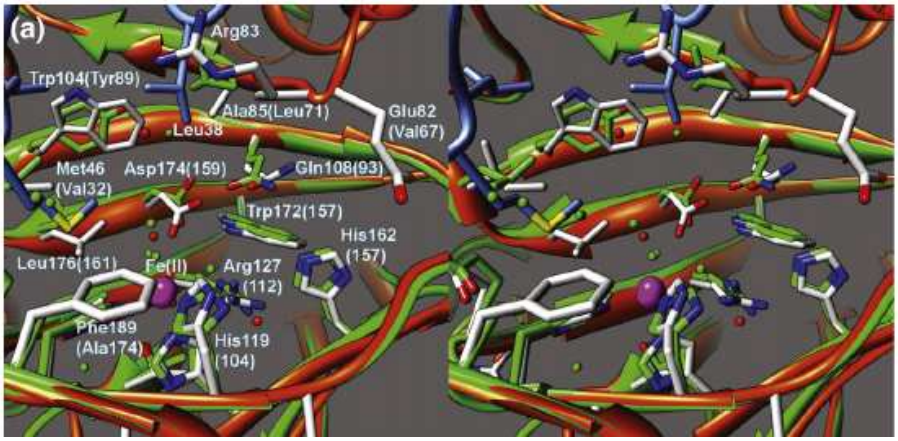


Figure 4.7: Stereoview of the least-squares superposition of the active sites of SDO (subunit D, red-orange; subunit B, light blue) and eGDO (colored green, PDB code 2D40, subunit D). The residue numbers for eGDO are reported in parentheses. The iron ion is represented as a magenta sphere.

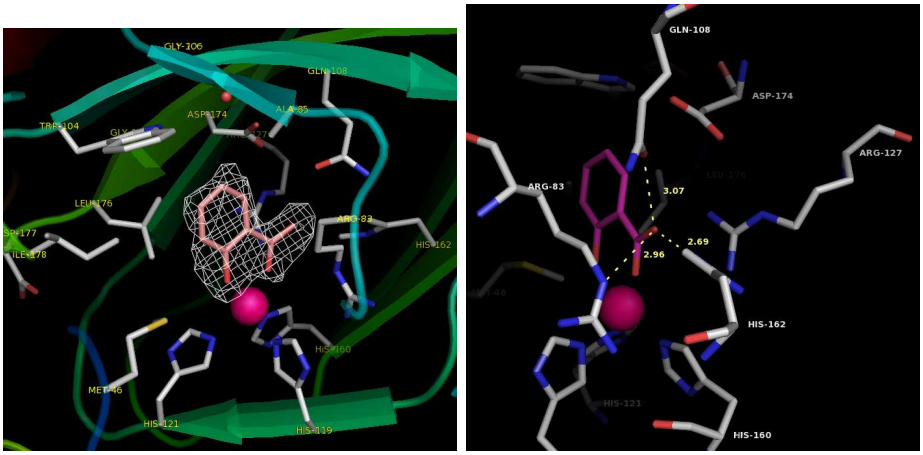
4.2 Adducts with substrates

Soaking experiments under anaerobic conditions were performed in order to isolate the substrate-enzyme complexes. These kind of experiments require good diffraction quality crystals because usually they get damaged after the soaking in the substrates or inhibitor solution. In an attempt to improve the quality of the crystals and the reproducibility of the method we optimized a new crystallization condition. Further the crystals were improved using both additives and seeding techniques. Crystals of this new form are orthorhombic and belong to the space group I222 with unit cell dimension a= 73 Å, b= 86 Å, c= 168 Åand one molecule per asymmetric unit. The overall structure of SDO in this new orthorhombic form strictly resembles that of the previous native structure obtained: the two structure were superposed with an r.m.s.d. of 1.546 Å.

Adducts were isolated by removing oxygen from the environment and freeze-trapping the enzyme-substrate complexes. The anaerobic conditions allow to stop the reaction in its first step, that is the binding of the substrate: the reaction cannot proceed further due to the absence of oxygen. Crystals were introduced in a glove box and soaked in a substrate solution for approximately two hours to ensure the binding of the substrate. Then crystals were frozen inside the glove box and kept under liquid nitrogen in order to hold the reaction in this first step.

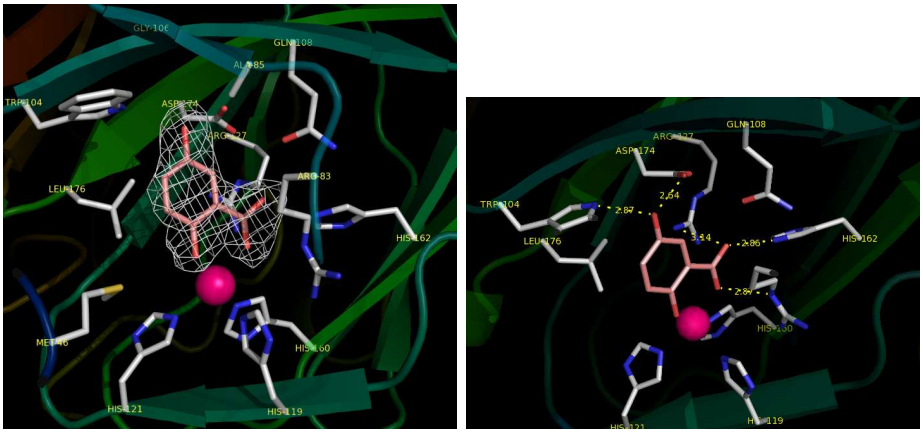
The crystal structures of the complexes with salicylate (Figure 4.8), gentisate (Figure 4.9) and 1-hydroxy-2-naphthoate (Figure 4.10) were obtained under anaerobic conditions by soaking experiments using this orthorhombic crystal form. The substrates complexes models shows that substrates bind to the ferrous ion in the SDO active site in a bidentate manner through the 2-hydroxy group and one of the carboxyl oxygen. A bidentate coordination to the iron was already suggested by spectroscopic studies and docking calculation for this class of enzymes (GDOs) and is common to all the intra and extra-diol dioxygenases so far investigated. The carboxyl oxygen that is coordinate to the iron ion makes also an hydrogen bond with nitrogen NE of Arg83. The other carboxyl oxygen of the three substrates that is not coordinated to the iron makes hydrogen bonds with NE2 of His162, OE1 of Gln 208 and NH2 of Arg127. The enzyme-gentisate complex model shows that the 5-phenolate group of gentisate is hydrogen bound to OD2 of Asp174 and NE1 of Trp104 at a distance of 2.56 Åand 2.88 Årespectively (See Figure 4.9).

Superposing the SDO wild type structure and those of the subtrates adducts, we noticed that the binding of the substrates causes important conformational changes in the enzyme active site: the closure of the loop formed by residues from 78 to 82 in the active site cavity and the stabilization of the loops formed by residues 5—15 and 194—198. These changes will be discussed more precisely in Chapter 6.



(a) (b)

Figure 4.8: The WT — salicylate complex: a) electron density contoured at 3σ and b) interaction of the substrate in the active cavity.



(a) (b)

Figure 4.9: The WT — gentisate complex: a) electron density contoured at 3σ and b) interaction of the substrate in the active cavity.

(a) (b)

Figure 4.10: The WT — 1-hydroxy-2-naphtoate complex: a) electron density contoured at 3σ and b) interaction of the substrate in the active cavity.

5

Structural studies on SDO mutants

SDO activity is unique among the currently known ring fission dioxygenases since, contrary to all the identified GDOs, it is able to catalyze the ring opening of gentisate, as well as salicylate and of a variety of substituted salicylates containing only a single hydroxyl group on carbon-2 and no additional electron-donating substituent. Among them, 5-fluorosalicylate, 3-, 4-,and 5-chlorosalicylate, 3-, 4-, and 5-bromosalicylate, 3-,4-, and 5-methylsalicylate, and 3,5-dichlorosalicylate conversions have been fully characterized [37]. Additionally, SDO is able to cleave 3- and 5-amino-salicylate, as well as 3- and 4-hydro-xysalicylate and 1-hydroxy-2-naphthoate [35] [37].

In this work a structural and mutational study of SDO from *P. salicylatoxidans* was performed in order to understand the unique activity of SDO. Key amino acid residues of SDO from *P. salicylatoxidans* were selected and altered by site-directed mutagenesis, based on amino acid sequence alignments and the results from our crystallographic studies on the enzyme. The major aim of this study was to obtain variant forms with either improved catalytic efficiencies for one of the tested substrates or to generally change the substrate specificity.

The residues that constitute the top of the catalytic cavity in SDO are Trp104, Asp174, Gln108, Ala85; these residues are supposed to interact with the substituent on carbon-5 of the aromatic ring of the substrate, thus these residues should be responsible of the sustrate specificity of the enzyme. Asp174 and Gln108 are very well conserved residues among all the GDOs. The mayority of the GDOs show a Tyr instead of Trp104; therefore a SDO mutant W104Y is expected to have an enhanced activity respect to gentisate. Ala85 is mutated in a His in 1-hydroxy-2-naphtoate dioxygenase; we supposed that the histidine ring could stabilize and orient the naphtoate in the cavity. Therefore the SDO mutant A85H should show an increased selectivity towards 1-hydroxy-2-naphtoate. The last mutant proposed

was a G106A mutant; while most of GDOs have Tyr instead of Trp104 (see Figure 4.1), GDO from *Corynebacterium glutamicum* shows a conservated Trp, but has a Ala instead of Gly106. This mutation could influence the orientation of the Trp104 and thus change the activity of the SDO into a GDO.

5.1 Mutant reactivity

The activity towards salicylate, gentisate and 1-hydroxy-2-naphtoate of wild type SDO and the mutants A85H, W104Y and G106A was measured by spectrophotometric analysis in the laboratory of Prof. Andreas Stolz (Stuttgard, Germany). The results are shown in table 5.1.

Our hypotesis was confirmed for the G106A mutant, that resulted to be a conventional GDO, since it does not show activity towards salicylate and 1-hydroxy-2-naphtoate, but is still able to convert gentisate even with a slighty increased activity. The W104Y mutant is still able to cleave the three substrates, even if the activity is reduced; moreover, the relative activity between te three substrates is similar to the relative activity in the case of wild type enzyme. The A85H mutant shows a reduced activity for 1-hydroxy-2-naphtoate, but is not more able to convert salicylate and gentisate: for this mutant the substrate specificity is as expected, but the decrease in the activity was unexplained on the basis of the sequence alignment.

	k_{cat}/K_M (s^{-1} M^{-1}) $\times$ 10^{-3}			
	WT	A85H	W104Y	G106A
Salicylate	50.06	0.02	7.20	n.d.
Gentisate	800.42	n.d.	219.4	818.68
1-OH-2-naphtoate	168.08	41.52	31.63	n.d.

Table 5.1: Activities of SDO wild-type and mutant variants (n.d. not determined).

	WT		
	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1}) $\times$ 10^{-3}
Salicylate	17.1	0.86	50.06
Gentisate	166.6	133.35	800.42
1-OH-2-naphtoate	140.4	23.60	168.08

Table 5.2: Kinetic parameters and substrate specificities of SDO wild-type enzyme.

	A85H		
	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1}) $\times$ 10^{-3}
Salicylate	398.5	0.01	0.02
Gentisate	n.d.	n.d.	n.d.
1-OH-2-naphtoate	79.66	3.31	41.52

Table 5.3: Kinetic parameters and substrate specificities of SDO A85H mutant (n.d. not determined).

	W104Y		
	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1}) $\times$ 10^{-3}
Salicylate	165.6	1.19	7.20
Gentisate	466.4	102.3	219.4
1-OH-2-naphtoate	144.4	4.57	31.63

Table 5.4: Kinetic parameters and substrate specificities of SDO W104Y mutant.

	G106A		
	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M (s^{-1} M^{-1}) $\times$ 10^{-3}
Salicylate	n.d.	n.d.	n.d.
Gentisate	110.30	90.30	818.68
1-OH-2-naphtoate	n.d.	n.d.	n.d.

Table 5.5: Kinetic parameters and substrate specificities of SDO G106A mutant (n.d. not determined).

5.2 Mutant structures

In order to explain the substrate specificity of SDO and its mutants, the structure of the mutants were solved and compared to the wild type structure. Figure 5.1 shows the superposition of wild type structure to mutant structures. It appears clearly that the mutations do not influence the overall structure of the enzyme: these mutations in fact altered only the binding cavity of the substrate.

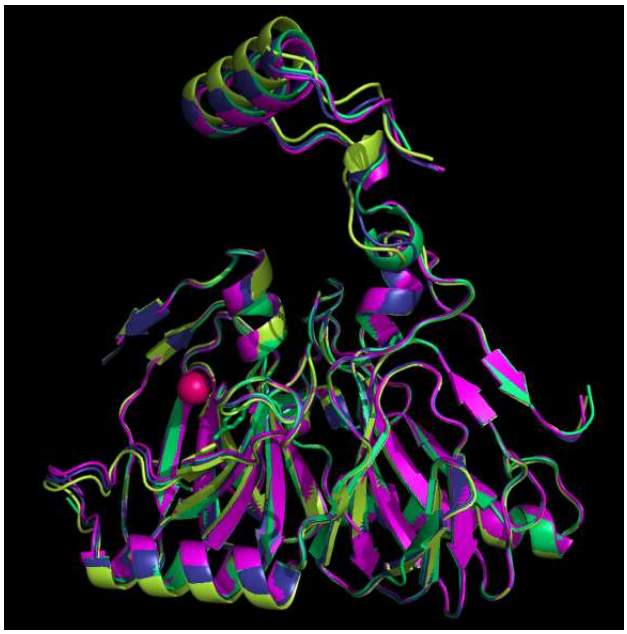
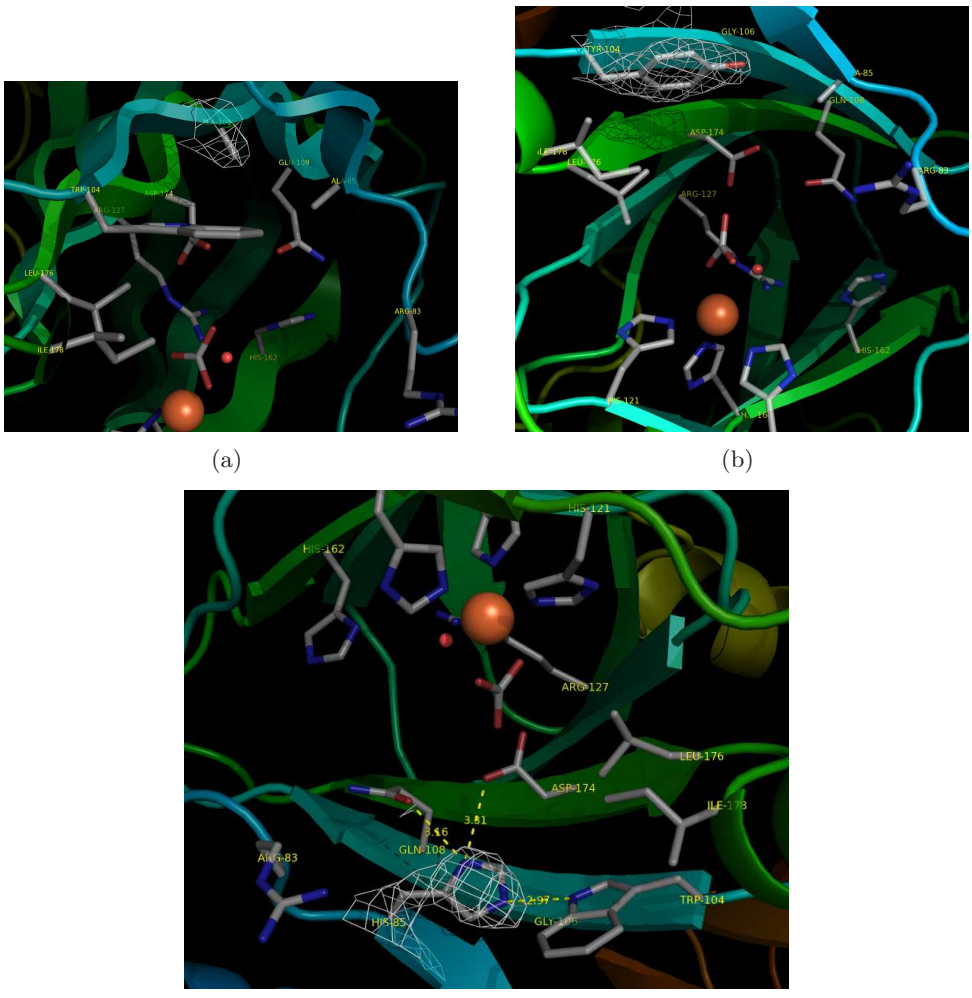


Figure 5.1: Superposition of Wild type (blue) and mutants structures: A85H (yellow), G106A (pink), W104Y (green); in magenta the iron ion.

Figure 5.2 shows the electron density on the mutated residues for A85H, W104Y and G106A mutants. Regarding G106A mutant, the mutated Ala does not seem to influence the orientation of the next Trp104, as expected. The mutant A85H shows an hydrogen bonds, between Ala85 and Trp104, Gln108 and Asp174, that stabilize and orient the histidine side chain towards the top of the binding cavity and thus probably hindering this active site region. In order to share some light on the unique activity of SDO and of the substrate specificity of the enzyme a structural analysis of the SDO mutants in complex with their substrates was performed.



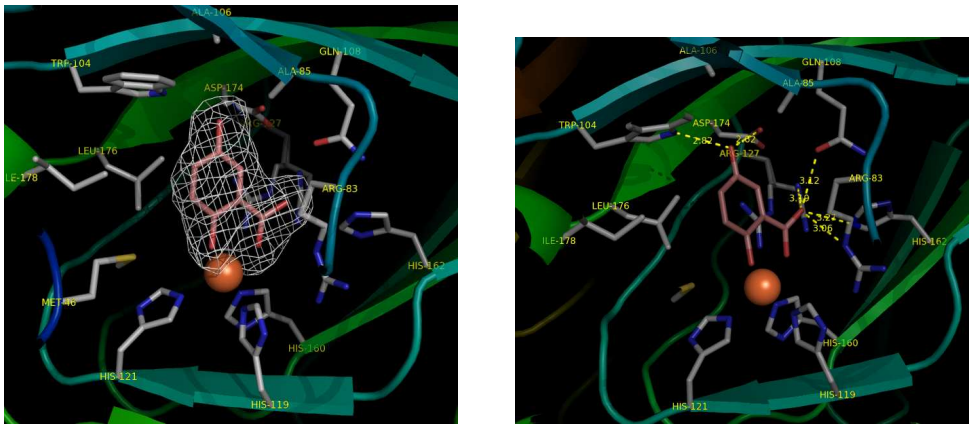
(a) (b)

(c)

Figure 5.2: Electron density (contoured at 2σ) on the mutated residues: a) G106A mutant; b) W104Y mutant; c) A85H mutant

5.3 Description of the obtained adducts

The crystallographic structures of the complexes of G106 mutant with gentisate and salicylate and of the W104Y mutant with gentisate were solved and analized. The substrate binds to the iron ion in a bidentate manner, with the oxygen atom from the 2-hydroxy group and one oxygen atom from the carboxylate group of the substrate. The binding is stabilized by hydrogen bonds between the oxygen atom from the carboxylate that does not bind to the iron and Arg83, Gln108, Arg127 and His162. Since the mutation of Gly106 into Ala106 does not affect the orientation of Trp104, the hydroxy group on C-5 of the substrate is stabilized by hydrogen bonds to Asp174 and Trp104 (See Figure 5.3), as seen previously for the WT — gentisate adduct. For the W104Y adduct the stabilizing hydrogen bonds involve Asp174 and the mutated Tyr104 (See Figure 5.4).

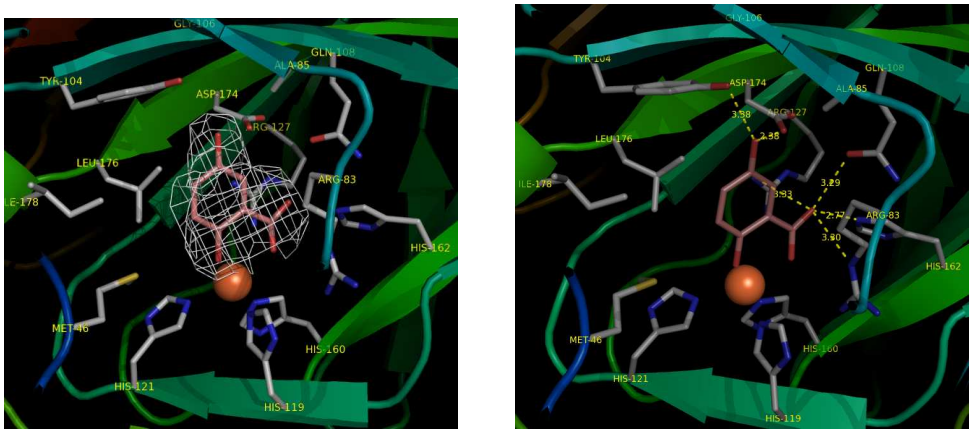


(a) (b)

Figure 5.3: The G106A — gentisate complex: a) electron density (3σ) and b) interaction of the substrate in the active cavity.

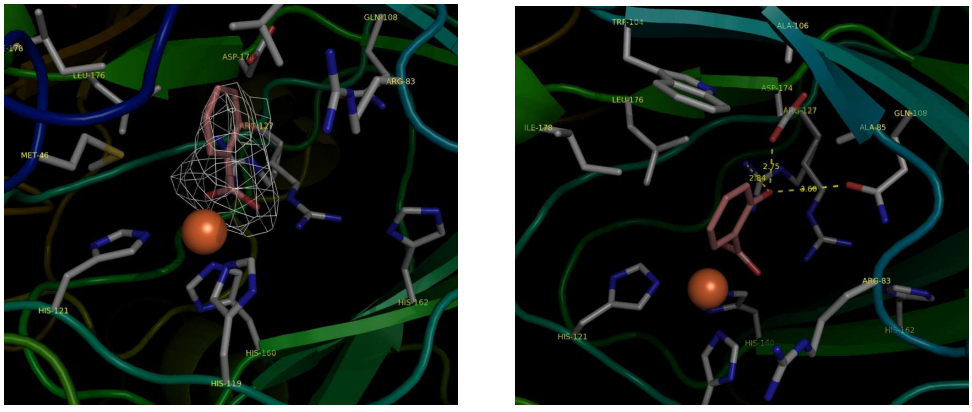
As previously discussed for the Wild type adducts, the binding of the substrate induces a conformational change in the loop formed by residues 78—82 and a stabilization of the loops formed by residues 5—15 and 194—198 (for a more detailed discussion see Chapter 6).

The structure of the G106A mutant with salicylate, that is not more converted by this mutant, shows a different mode of binding of the salicylate to the enzyme, as depicted in Figure 5.5. Salicylate is in fact bound to the iron atom with both the oxygen atoms of the carboxylate, while the hydroxy group remains available to hydrogen bonds towards Gln108 and Arg127. In this case we did not observe the conformational changes in the 78—82, 194—198 and 5—15 loops as in the other complexes structures.



(a) (b)

Figure 5.4: The W104Y — gentisate complex: a) electron density (3σ) and b) interaction of the substrate in the active cavity.



(a) (b)

Figure 5.5: The G106A — salicylate complex: a) electron density (3σ) and b) interaction of the substrate in the active cavity.

5.4 Explaining substrate specificity

Comparison of the adducts of WT SDO and of G106A mutant, both in complex with salicylate and gentisate, allowed us to share some light on the substrate specificity of this enzyme. The adduct of W104Y mutant in complex with gentisate suggested that the activity of this mutant is not due to steric hindrance in the binding cavity, but probabily to changes of the channel required by dioxgen molecules to enter the cavity. Unfortunately we were not able to isolate any adducts of A85H mutant, but the comparison of the A85H structure with the adduct of WT enzyme in complex with 1-hydroxy-2-naphtoate could somehow shed a light on the activity of A85H mutant.

5.4.1 The cleavage of salicylate

In order to explain the peculiar activity of Wild type SDO towards salicylate, WT — gentisate, WT — salicylate, G106A — gentisate and G106A — salicylate complexes were superposed and analized. As already reported Wild type SDO is able to convert both salicylate and gentisate, while G106A mutant only converts gentisate; the G106A mutant was created as a model for the GDO from *Corynebacterium glutamicum*, since this mutation is the unique observable in the active site of the two enzymes.

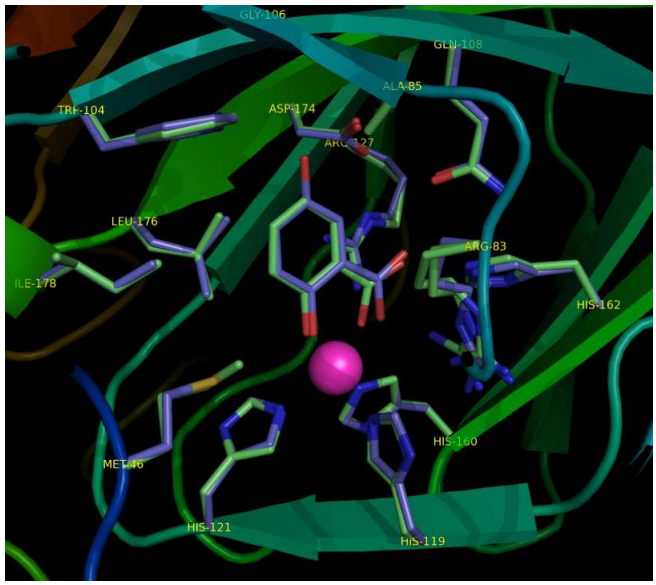


Figure 5.6: WT — salicylate complex (green) superposed to WT — gentisate complex (blue).

Figure 5.6 shows that in Wild type enzyme the binding of both substrates, salicylate and gentisate, is the same: the substrate molecules are completely superposable.

Moreover, gentisate binds to WT SDO and G106A mutant in the same way, as depicted in Figure 5.7. Thus, the absence of activity in cleaving salicylate shown by G106A mutant has to be related to the different coordination of salicylate to the catalytic ferrous ion in the WT — salicylate and G106A — salicylate complexes.

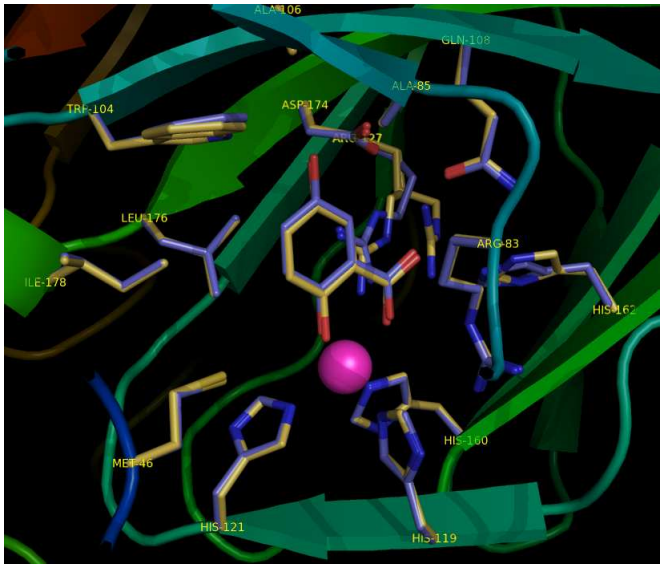


Figure 5.7: Superposition of WT enzyme (blue) and G106A (yellow) both in complex with gentisate.

In fact, we found that salicylate and gentisate bind in different manners to G106A mutant, as shown in Figure 5.8. Since the cleavage required both the hydroxy group and the carboxylate group to be bound to the iron atom, it is clear from the G106 — salicylate complex structure that G106A cannot convert salicylate.

A comparison of the structures of the G106A mutant and WT enzyme both in complex with salicylate revealed the factors that cause the salicylate unproductive binding in the G106A SDO mutant. In figure 5.9 the superposition of these two complexes is depicted; a water molecule near Asp174 and Trp104 is emphasized.

The inactivity of G106A mutant towards salicylate seems to be explained by the absence of this water molecule. In WT enzyme, the presence of this water molecule is possible since the Gly106 residue does not give any steric hindrance: the water molecule thus stabilize the Asp174 side chain in a conformation that make the binding of salicylate possible in the position that let the reaction occur. In the G106A mutant, Ala106 does not allow the presence of this water molecule: Asp174 is then free to assume a different conformation, eventually stabilized by a loose hydrogen bond with Trp104. The different conformation of the side chain of Asp174 reduces the size of the active site cavity thus causing different unproductive binding of the substrate. Although the wrong binding of salicylate prevents the cleavage of the aromatic ring it might explain the inhibitory effect of salicylate (A. Stolz, pers.

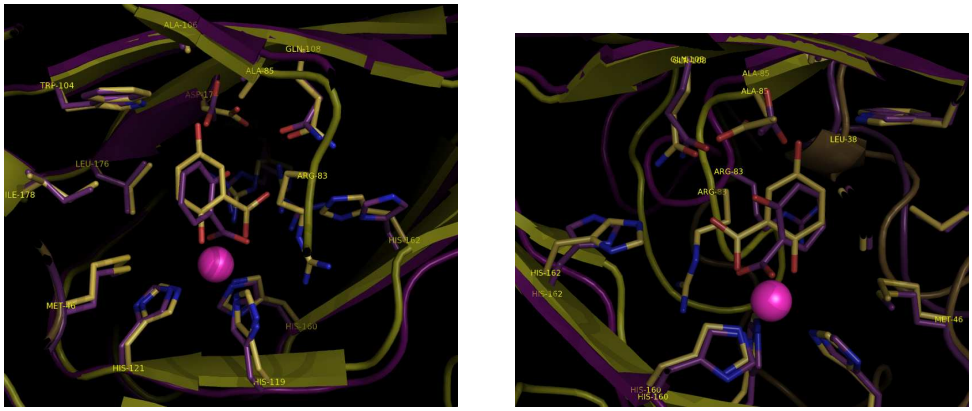


Figure 5.8: Binding of salicylate (purple) and gentisate (yellow) for G106A mutant.

communication).

The presence of this key water molecule does not influence the activity of WT enzyme in cleaving gentisate: the water molecule is displaced when the gentisate binds, probably because the net of hydrogen bonds between Asp174, Trp104 and hydroxy group on carbon-5 on the substrate ring is more favourable than the simple binding of this water molecule. The stabilization effects of the above mentioned net of hydrogen bonds between Asp174, Trp104 and hdroxy group on carbon-5 on substrate ring is probably the explanation of why G106A mutant is still able to convert gentisate: the interactions through hydrogen bond between the hydroxy group and Asp174 and Trp104 stabilize the same side chain conformation of Asp174 as in WT—gentisate complex, allowing the substrate to bind correctly for the conversion.

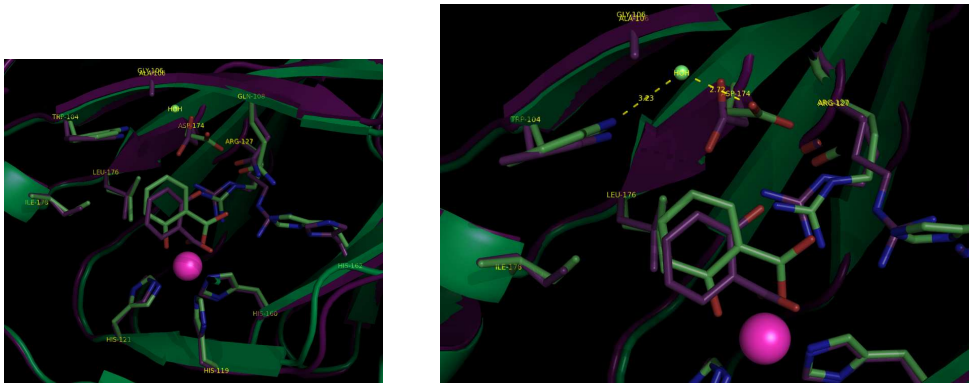


Figure 5.9: Binding of salicylate: superposition of G106A mutant complex (purple) and WT complex (green).

5.4.2 The cleavage of 1-hdroxy-2-naphtoate

A structure of the A85H mutant in complex with 1-hydroxy-2-naphtoate would be important to explain the fact that the wild type enzyme is able to convert 1-hydroxy-2-naphtoate contrarily to the other GDOs. The mutation was performed to obtain a pure 1-hydroxy-2-naphtoate dioxygenase; as shown in Section 5.1, A85H mutant converts only 1-hydroxy-2-naphtoate, but with a dramatic decrease of the activity if compared to the activity of wild type towards 1-hydroxy-2-naphtoate. In Figure 5.10 is depicted the superposition of the A85H mutant structure and WT — 1-hydroxy-2-naphtoate complex. The conformation of the mutated His85 side chain is stabilized by an hydrogen bond to Trp104. This interaction probably gives steric hindrance to the hydroxy group on the C-5 of the substrate aromatic ring, giving the mutant the ability of converting 1-hydroxy-2-naphtoate and salicylate, but not gentisate.

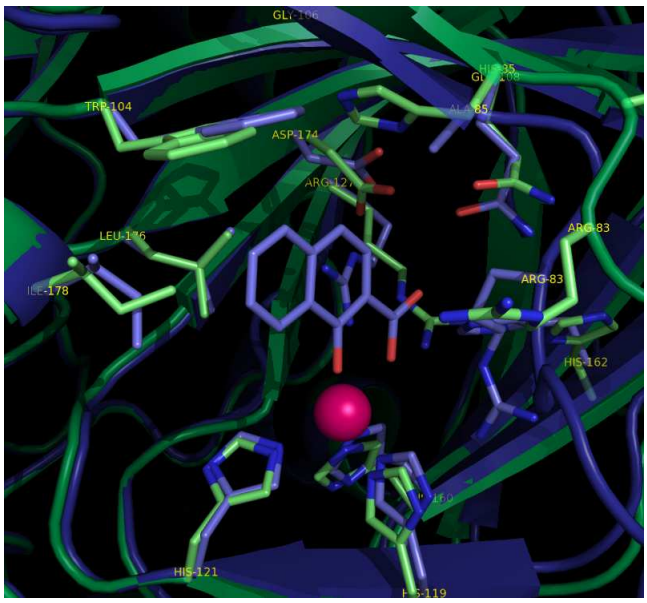


Figure 5.10: Superposition of A85H mutant structure (green) and WT — 1-hydroxy-2-naphtoate complex (blue).

Comparing the sequences of SDO and 1-hydroxy-2-naphtoate dioxygenase from *Nocardioides* sp. KP7 pointed out that Trp104, that in 1-hydroxy-2-naphtoate dioxygenase is a cysteine, could be mutated in order to obtained a efficient 1-hydroxy-2-naphtoate dioxygenase. The double mutant A85H-W104C probably would have a high activity towards 1-hydroxy-2-naphtoate. However, explaining why wild type SDO is able to convert 1-hydroxy-2-naphtoate requires a structural analisys of A85H—1-hydroxy-2-naphtoate complex and a comparison with the WT—1-hydroxy-2-naphtoate complex. This could allow to identificate which are the residues that

play a key role and give SDO the ability of cleaving also this kind of substrate.

5.4.3 Hypothesis on W104Y mutant activity

Figure 5.11 shows the structures of wild type enzyme and W104Y mutant both in complex with gentisate. Trp104 is probably not involved in substrate recognition; recall that the kinetic data obtained from the spectrophotometric assays showed that W104Y had a very similar substrate specificity profile as wild type SDO, although activity is altered and generally showed a decreased efficiency for the mutant. A putative function of Trp104 was discovered during the oxygen uptake measurements performed by A. Stolz, L. Steimer and co-workers (pers. communication). The data obtained from the oxygen uptake measurements showed an inhibition of W104Y by oxygen concentrations ≥ 260 μM.

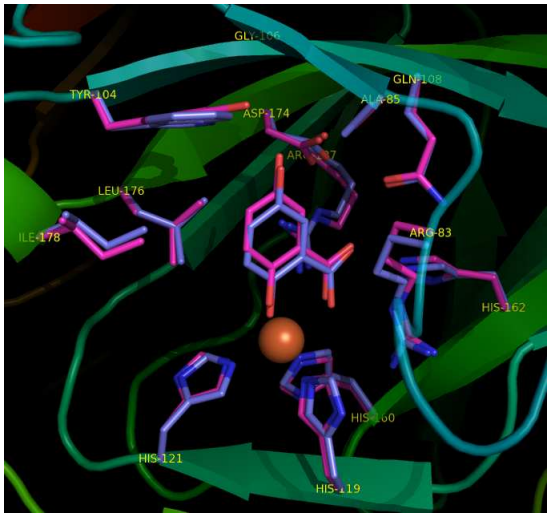


Figure 5.11: Superposition of WT enzyme (blue) and W104Y mutant (pink) both in complex with gentisate.

The binding of molecular oxygen to the catalytic iron occurs after the binding of the substrate (see Chapter 6) and it has a defined route to enter the active site, probably along a hydrophobic surface of the solvent accessible cavity. Therefore, it might be possible that dioxygen has to pass Trp104 and an exchange of this residue to a more hydrophilic amino acid residue might handicap the oxygen entry to its specific binding site and therefore reduce both the enzyme activity and the substrate affinity towards molecular oxygen. Another possibility could be that residue Tyr104 due to its missing aromatic ring compared to Trp104 somehow deregulates the proper access of the dioxygen molecule to the catalytically ferrous ion with the same results as mentioned above.

6

The catalytic mechanism

Figure 6.1 shows the mechanism of GDO proposed by Harpel, Lipscomb and coworkers in 1990 [7]. In these studies the gentisate 1,2-dioxygenases from *P. testosteroni* and *P. acidovorans* were analyzed by a multitude of methods including spectrophotometrical experiments and EPR spectroscopy. These measurements suggested that the iron ions in the active center of gentisate 1,2-dioxygenases possess three sites for the binding of exogenous ligands (two binding sites for the aromatic substrate, and one for molecular oxygen). Furthermore, Harpel and Lipscomb proposed an ordered binding of the organic substrate prior to molecular oxygen without integral changes in the redox state of the metal. In the proposed cleavage mechanism, the binding of molecular oxygen causes a polarization of electron density away from the aromatic ring system of the organic substrate towards the iron bound oxygen. The resulting electron deficiency at the C1-atom of the aromatic ring allows an ionic attack of iron-bound dioxygen at this position. The formed peroxy-intermediate complex is then dissected by a heterolytic cleavage reaction of the O-O bond and one of these oxygens is introduced into the aromatic ring. Harpel and Lipscomb suggested that the cleavage is promoted by the ketonization of the C5-hydroxyl group (or by an imine formation from the amino-group in case of 5-aminosalicylate) and the transfer of electrons into the aromatic ring structure. The subsequent hydrolysis of the anhydride intermediate results in the formation of the product. Since this cleavage mechanism is dependent on an electron-donating substituent at the C5-position, it cannot explain the cleavage of salicylate.

6.1 Trapping the intermediate state for G106A mutant

During the experiments isolating the substrate-enzyme adducts described in Chapter 5, we were able to obtained two more complex structure: the G106A mutant in

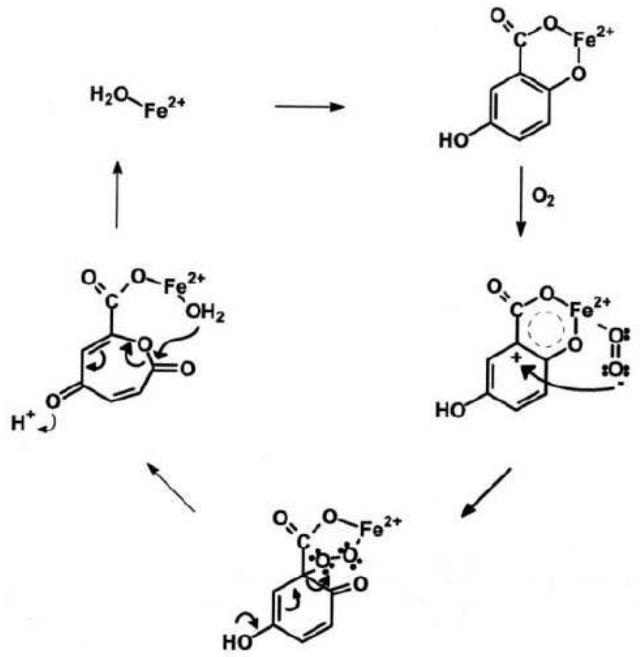


Figure 6.1: Proposed mechanism of gentisate 1,2-dioxygenase action (Harpel and Lipscomb).

complex with both gentisate and dioxygen and the G106A mutant in complex with 2-oxo-5-hydroxyepta-3,5-dien-1,7-dioic acid,the product of the cleavage of gentisate. These two structure complete the whole catalityc circle of G106A mutant, since we obtained also the structure of G106A mutant in the resting state and the structure of G106A mutant in complex with gentisate, as previously discussed.

The isolation of these further steps in the enzymatic reaction was a matter of serendipity; the oxygen concentration in the glove box was keeped around 20—30 ppm during the isolation of adducts. It raised at 60 ppm while crystals of G106A mutants were soaking in the gentisate solution. The result was that the reaction went further to the binding of the dioxygen molecule and more further reaching the formation of the product.

6.1.1 Binding of dioxygen and release of product

Figures 6.2 shows the electron density on substrate and dioxygen molecules obtained in the G106A—gentisate—oxygen adduct. The dioxygen molecule binds on the third open coordination site on the iron atom, the other two sites occupied by the substrate. In addition, the dioxgen molecule is oriented to a region in the binding cavity that is constitued by hydrophobic residues: Leu176 and Ala125.

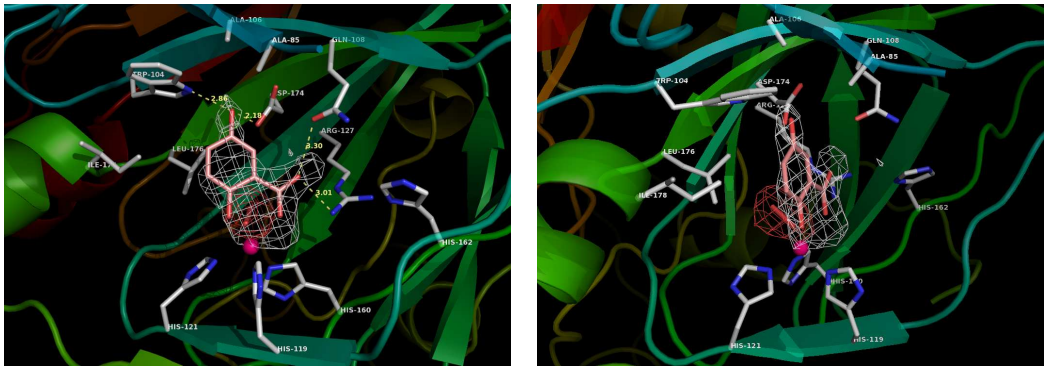


Figure 6.2: Electron density contoured at 3σ for gentisate and dioxygen bound to G106A mutant.

The dioxygen molecule is bound in an end-on manner, at a distance of 2.01 Åfrom the iron atom. This is apparently in contrast with the results obtained for a homoprotocatecuate 2,3-dioxygenase by Kovaleva, Lipscomb and co-workers [51]. In their paper, they report a structure of this enzyme, that is an extradiol dioxygenase, in complex with a 4-nitrocatecol (1,2-dihydroxy-4-nitro-benzene) and dioxygen: the dioxygen molecule is bound side-on, the distances of the two oxygen atoms from iron ion being 2.4 Å; they also observe the alkylperoxo intermediate. In both structures, the orientation of dioxygen molecule is different from the one we observed in our G106 — gentisate — dioxygen structure. This is due to the different states isolated in the structures: in the homoprotocatecutae 2,3-dioxygenase in complex with 4-nitrocatechol and dioxygen, the substrate molecule, expected to be planar, is distinctly buckled with the C-2 hydroxy group moving out of plane, towards the dioxygen molecule. This structure is probably an averaging status between the binding of incoming dioxygen and the attack on the aromatic ring of 4-nitrocatechol. Our structure of G106 — gentisate — dioxygen complex probably isolates a very first step of the reaction, that is the binding of dioxygen, and does not involve the attacking on substrate.

In the structure obtained from a G106A crystal soaked with gentisate in the presence of 60 ppm of oxygen we found an electron density on the iron atom that could easily be modeled with a molecule of 2-oxo-5-hydroxyepta-3,5-dien-1,7-dioate, that is in fact the product of the cleavage of gentisate. The structure shows the product molecule (see Figure 6.3) bound to the iron atom through a oxygen atom from the carboxylate group formed by the reaction on C-2 atom of the substrate: the reaction is frozen during the release of the product.

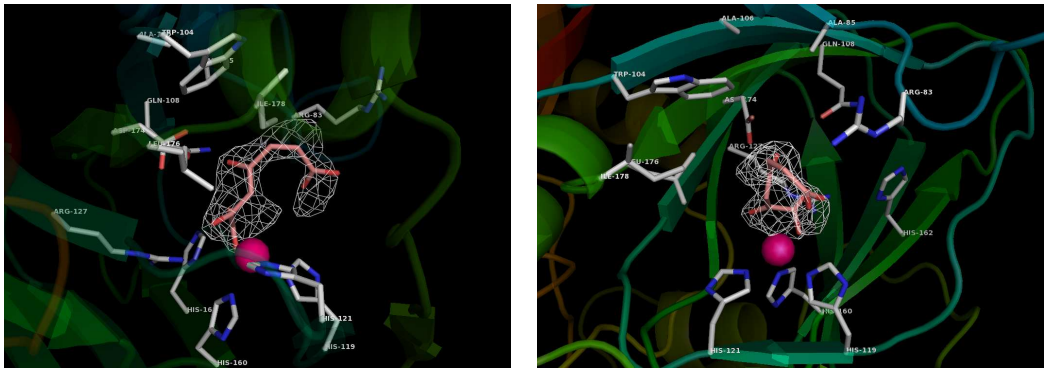


Figure 6.3: Electron density contoured at 3σ for 2-oxo-5-hydroxyepta-3,5-dien-1,7-dioate bound to G106A mutant.

6.1.2 Conformational changes induced by the binding of substrate

The analisys of the structures of G106A mutant and G106A mutant in complex with gentisate and 2-oxo-4-hydroxyepta-3,5-diendioate revealed that the binding of the substrate induced many conformational changes. Figures from 6.4 to 6.6 show the superposition of G106A structure (in yellow) to G106A — gentisate adduct (in green) and G106A — product adduct (in blue) and depict the conformational changes occurring on the loop formed by residues 78—82, the loop formed by residues 194—198 and the loop formed by residues 5—15.

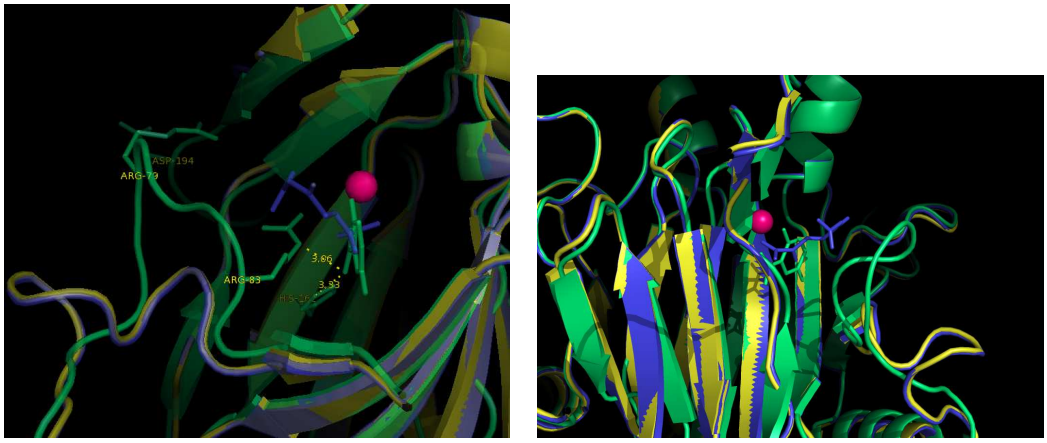
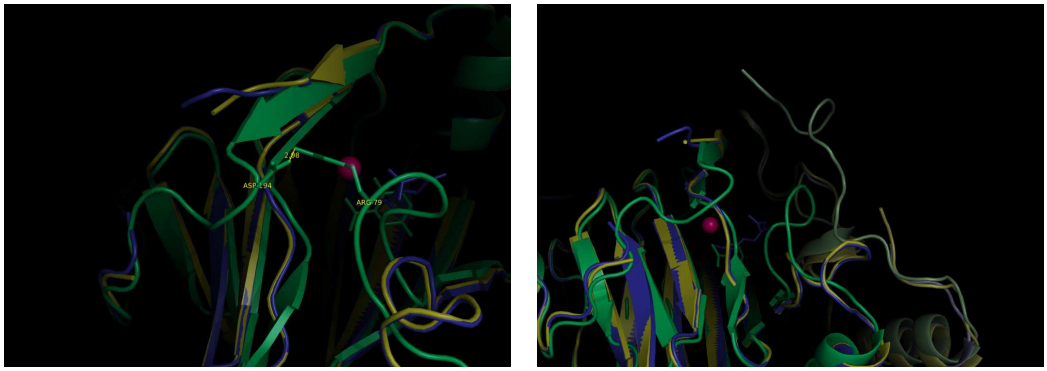


Figure 6.4: Superposition of G106A structure (in yellow) to G106A — gentisate adduct (in green) and G106A — product adduct (in blue): closing of the loop 78—82 upon substrate binding.

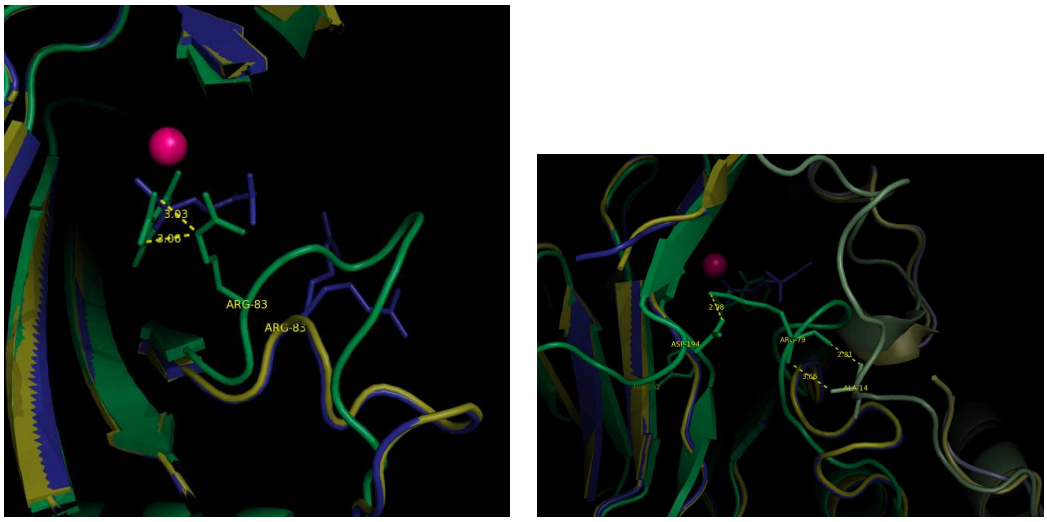
The closing of the loop 78—82 on the active cavity of the enzyme is caused by the interaction of Arg83 (as depicted in Figure 6.6a) directly to the binding substrate, through an hydrogen bond to the carboxylate group of gentisate. Subsequently, the other conformational changes are caused by the hydrogen bonds given by Arg79, the residue at the other end of the moving loop. Figure 6.6b shows the interactions between Arg79 and Asp194 and Ala14 that stabilize loop 194—198 and loop 5—15 respectively. These interactions and the resultant stabilizations allowed us to model the residues in these loops, while they did not give electron density in G106A structure or G106A—product complex. The same effects were observed in every structure of enzyme—substrate complex isolated.

In the G106A—salicylate adduct we did not observe these conformational changes, since the salicylate is differently oriented and, bound to the iron atom with the carboxylate group, cannot interact with Arg83 to induce the closing of the loop on the cavity and consequentely stabilize loop 194—198 and loop 5—15. In fact these residues remained unmodeled in the structure, as in the structures of Wild type enzyme and mutants without any soaking.



(a) (b)

Figure 6.5: Superposition of G106A structure (in yellow) to G106A — gentisate adduct (in green) and G106A — product adduct (in blue): stabilizations due to substrate binding for a) loop 194—198; b) first residues (aa 5—15).



(a) (b)

Figure 6.6: Superposition of G106A structure (in yellow) to G106A — gentisate adduct (in green) and G106A — product adduct (in blue): residues involved in conformational changes due to substrate binding a) role of Arg83 (loop 78—82); b) role of Arg79 (loops 5—15 and 194—198).

6.1.3 Proposed mechanism for Salicylate 1,2-dioxygenase

Former EPR studies by Lipscomb and co-workers on the active site of GDO from *Pseudomonas testosteroni* and *Pseudomonas acidovorans* provided evidence for the simultaneous coordination of substrate and oxygen to the three available sites on the active site iron by eventually displacing bound water molecules and thus completing its six-coordination sphere. The oxidative ring cleaving of gentisate by gentisate-1,2-dioxygenases was suggested to require a direct bidentate ligation to the ferrous iron in the catalytic center as seen in other dioxygenases in order to lower the redox potential of the iron thus favoring the successive binding of molecular oxygen [7].

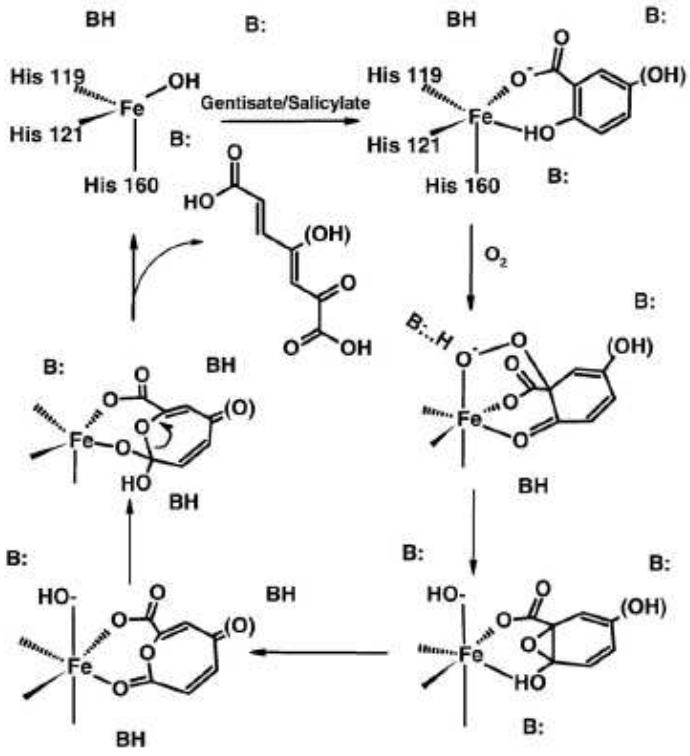


Figure 6.7: Proposed mechanism for SDO.

Figure 6.7 shows the proposed mechanism for SDO. The first step in the reaction is the binding of the substrate, stabilized by a net of hydrogen bonds that involves several residues: Asp174 and Trp104 for the hydroxy group on C5 and Arg127, Arg79, Arg83, Gln108 and His162 on the oxygen atom from the carboxylate group that is not directly bound to the iron ion. These residues assist substrate deprotonation, stabilize its binding to the iron ion and promote the subsequent oxygen attack and ring opening. As seen above, the hydrogen bonds with Arg83 and Arg79 also induce some conformational changes in the tertiary structure of the enzyme;

the most important seems to be the enclosure of the mobile loop wich they belong to, since it close the active cavity.

The heterolytic cleavage of the O—O bond and insertion of one oxygen atom into the ring C1—C2 bond were suggested to be promoted in GDO by ketonization of the 5-hydroxy substituent that would transfer electron density into the ring; this mechanism is not required of SDO, because it cleaves salicylate, but still plays an important role in substrate specificity, making the cleavage of gentisate more favoured than that of salicylate as suggested by the k_{cat} values.

7

Experimentals

All the reagents were of the best purity available from Sigma-Aldrich. The initial trials in crystallization were made using commercial screens: JBScreen Classic from Jena Bioscience and Additive Screen from Hampton Research.

Integration of data was performed with Mosflm [52] or XDS [53]. In refinement some programs belonging to the CCP4 program suite [54] were used: SCALA [55], MOLREP [56], REFMAC5 [57] and ARP [58]. Manual rebuilding of the models was performed using COOT [59]. Stereochemistry was analized with PROCHEK [60] and sequence alignement with ClustalW [61].

7.1 Crystallization

Enzyme purification, activity assay and biochemical characterization on wild type enzyme and mutant variants were carried out by L. Steimer, S. Bürger and A. Stolz from Institut für Mikrobiologie, Universitat Stuttgart (Stuttgart, Germany).

Wild type Salicylate 1,2-Dioxygenase

The enzyme was crystallized at 277 K using the sitting-drop, vapor-diffusion method from a solution containing 12% (w/v) EtOH, 4% (w/v) PEG400, 0.1 M sodium acetate, pH 4.6. Drops were prepared using 1 μl of protein solution mixed with 1 μl of crystallization solution and were equilibrated against 100 μl of crystallization solution [62].

Mutants: A85H, W104Y, G106A

Mutants were crystallized at 277 K using the sitting-drop vapour-diffusion method. The crystallization conditions are different from the wild type condition: we used a

solution containing 8% PEG10000 and 0.1 M tris-HCl pH 8.0. Crystallization was optimize using both seeding techniques and additives. Drops were prepared mixing 1 μl of protein solution with 0.8 μl of crystallization solution; we added 0.2 μl of additives from Hampton Research Additive Screen I. The drops were equilibrated against 100 μl of crystallization solution for 24 hours. Then, since no crystals appeared, we introduced some crystal *nuclei* by adding 0.3 μl of a seeding solution; this solution was prepared by smashing few crystals (previously obtained but not suitable for data collection) in 50 μl of crystallization solution and the diluiting it 1:10, 1:100 and 1:1000. Crystals of good quality appeared after few days; the best seeding diluition was found to be the 1:100, while the best additive were calcium chloride 0.1 M for W104Y and G106A mutants and benzamidine-HCl 2% (w/v) for A85H mutant. Since crystals obtained in this way were far better in shape, dimension and reproducibility, we tried this new way with additives and seeding tecniques on the wild type enzyme. As expected, the wild type crystallization was improved; the best seeding diluition was 1:100 in this case too, and the additive that gave best results is calcium chloride 0.1 M. We therefore used this second condition to grow crystals of the wild type needed for adduct studies.

7.2 Data collection and Refinement

Wild type enzyme in the $P4_32_12$ form

A complete data set at 100 K extending to a maximum resolution of 2.9 Åwas collected at the EMBL beamline BW7B, DESY, Hamburg. Data were collected adding 30% (v/v) glycerol to the mother liquor as cryoprotectant, using an MAR345 image plate detector and a wavelength of 0.8423 Å. Crystals belong to the primitive tetragonal space group $P4_32_12$ with unit cell dimensions a=133.3, c=191.51. Assuming one tetramer per asymmetric unit, the solvent content is 47% (v/v) of the unit cell (V_M =2.3 Å^3/Da). Data processing with Mosflm and SCALA gave 38,218 unique reflections, an R_{symm} of 12.1% and an overall completeness of 99.0%. The statistics for data collection are summarized in Table 7.1.

The structure was solved by molecular replacement techniques using the program MOLREP, and the four subunits of the gentisate 1,2-dioxygenase from *E. coli* O157:H7 [38] (PDB accession code 2D40) as a starting model. The two enzymes are both homotetramers, and a sequence comparison between the two monomeric subunits shows a relatively low sequence identity (33.3%). The model was refined using the program REFMAC5 and applying non crystallographic symmetry constraints. Manual rebuilding of the model was performed using the program Coot. Solvent molecules were introduced automatically using ARP. Refinement resulted in R_{factor} and R_{free} values of 19.3% and 26.1%, respectively. Data refinement statistics are summarized in Table 7.2.

The stereochemistry of the final model was analyzed with PROCHECK. The Ramachandran plot is of good quality; there are 1141 non-glycine and non-proline residues; among these, 972 (85.2%) are in the most favored regions, 151 (13.2%)

Data collection	
Beamline	BW7B, Desy, Hamburg
Space group	$P4_32_12$
Unit cell parameters a(Å); c(Å)	133.02; 190.75
V_M (Å^3)	2.3
Solvent content (%) (v/v)	47
Wavelength (Å)	0.8423
Limiting resolution (Å)	20-2.9 (3.11-2.9)
Unique reflections	38218
R_{symm}	0.12 (0.47)
Multiplicity	4.8 (4.9)
Completeness overall (%)	99.0 (99.6)
I/σ(I)	4.3 (1.7)

Table 7.1: Summary of data collection statistics; values in parentheses are for the highest resolution shell.

Refinement	
Resolution range (Å)	20.0-2.90
Unique reflections, working/free	35989/1894
R_{factor} (%)	19.3
R_{free} (%)	26.1
Non-hydrogen atoms	11033
Water molecules	175
r.m.s.d. from ideal: Bond lengths (Å)	0.015
r.m.s.d. from ideal: Bond angles (∘)	1.717

Table 7.2: Summary of refinement statistics.

are in the additionally allowed regions,12 (1.1%) in the generously allowed regions, and 6 (0.5%) in disallowed regions. The overall mean B factor of the structure after refinement was 44.01 Å^2 for chain A, 43.96 Å^2 for chain B, 43.75 Å^2 for chain C, 42.98 Å^2 for chain D, 30.19 Å^2 for water molecules, and 43.19 Å^2 for all atoms.

The final model is composed of residues 16-368 of subunits A-D, with the exceptions reported below, 175 water molecules, four Fe(II) ions and two acetate molecules. Electron density was missing for residues A1-16, A195-196, A368, B1-15, B195-196, B80-81, C1-14, C78-79, C196-197, C368, D1-14, and D367-368, and were not introduced into the model.

Protein Data Bank accession code

Coordinates and structure factors of wild type structure enzyme have been deposited in the Protein Data Bank with accession code 2PHD.

Mutants and Wild Type in the I222 form

Complete data sets at 100 K were collected for each mutant and wild type enzyme in the I222 form, adding 30% (v/v) glycerol to the mother liquor as cryoprotectant, at EMBL beamline X13, DESY, Hamburg and at Centro Interdipartimentale di Cristallografia Strutturale, Universit di Firenze, Italy (CRIST) using an Oxford Diffraction Xcalibur PX Ultra diffractometer equipped with CCD area detector and monochromated CuKα radiation. The statistics for data collections are summarized in Table 7.3.

Data Collection	WT	A85H	W104Y	G106A
Beamline	CRIST	X13	CRIST	X13
Space group	I222	I222	I222	I222
Unit cell parameters a (Å)	75.75	72.18	73.71	73.99
Unit cell parameters b (Å)	87.16	90.01	87.05	87.85
Unit cell parameters c (Å)	166.89	169.06	167.92	167.86
V_M (Å^3)	3.28	3.29	3.16	3.11
Solvent content (%) (v/v)	62.22	62.27	60.75	60.17
Wavelength (Å)	1.542	0.979	1.542	0.979
Limiting resolution (Å)	47.14-2.20 (2.32-2.20)	56.34-2.50 (2.64-2.50)	47.09-2.40 (2.53-2.40)	47.19-2.10 (2.21-2.10)
Unique reflections	27804	18577	21100	32365
R_{symm} (%)	10.8(49.2)	13.1(51.8)	10.7(50.5)	8.3(39.9)
Multiplicity	5.6(5.1)	3.7(3.8)	4.2(4.2)	4.4(4.4)
Completeness overall (%)	98.3(97.5)	96.3(97.1)	98.4(99.3)	100.0(100.0)
I/σ(I)	13.0(3.2)	9.9(2.7)	12.4(2.8)	12.9(3.9)

Table 7.3: Summary of data collection statistics; values in parentheses are for the highest resolution shell.

The structures were solved by molecular replacement techniques using the program MOLREP using the Wild type $P4_32_12$ structure as a starting model. The model was refined using the program REFMAC5. Manual rebuilding of the model was performed using the program Coot. Solvent molecules were introduced automatically using ARP. Data refinement statistics are summarized in Tables 7.4 and 7.5.

Refinement	WT	A85H	W104Y	G106A
Resolution range (Å)	20.00-2.20	20.00-2.50	20.00-2.40	29.2-2.70
Unique reflections, working/free	26328/1391	17617/946	20011/1073	14709/771
R_{factor} (%)	20.6	17.6	17.1	21.9
R_{free} (%)	26.1	25.2	24.1	28.9
Non-hydrogen atoms	2998	2985	3046	2761
Water molecules	241	241	290	290
r.m.s.d. from ideal: Bond lengths (Å)	0.017	0.015	0.014	0.018
r.m.s.d. from ideal: Bond angles (∘)	1.785	1.576	1.579	1.764
B factor (Å^2)	27.3	25.0	30.7	24.0

Table 7.4: Summary of refinement statistics.

Ramachandran Plot	WT	A85H	W104Y	G106A
Most favoured region	250(88.7%)	249(%88.0)	254(%89.4)	236(85.2%)
Additionally allowed region	30(10.6%)	33(%11.7)	28(9.9%)	34(12.3%)
Generously allowed region	2(0.7%)	1(0.4%)	2(0.7%)	6(2.2%)
Disallowed region	—	—	—	1(0.4%)

Table 7.5: Summary of Ramachandran Plot statistics.

7.3 Isolation of adducts

Adducts were isolated in an environment without oxygen; we used a glove box (Braun — UNIlab). Substrate solutions were prepared dissolving the substrates in the crystallization solution. All of the solution used were degased with nitrogen for few minutes and then equilibrated in the glove box for about two hours, to remove as oxygen as possible. Crystals were transferred in the glove box and then washed in the crystallization solution inside the glove box to remove even residual oxygen.

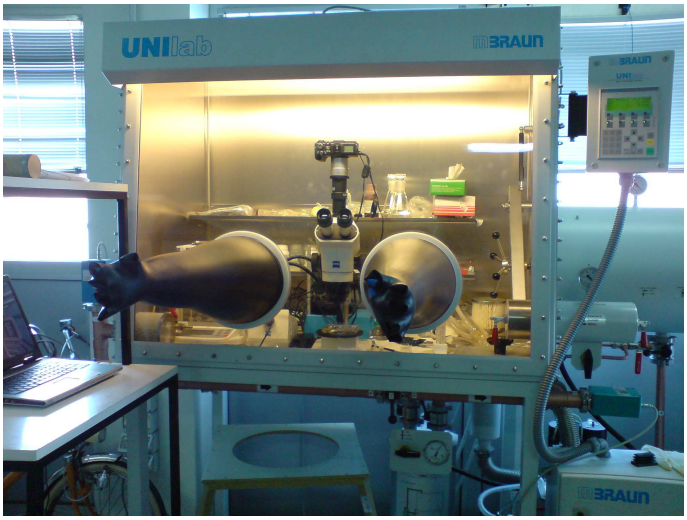


Figure 7.1: The setup for the adduct isolation experiment: the glove box is ready with a microscope connected to an outside laptop through a camera

Adducts were prepaired soaking crystals in a 20 mM solution of salicylate or gentisate, and in a sature solution of naphtoate (less than 20 mM). Crystals were flash-freezed inside the glove box in solution obtained adding 30% (v/v) glycerol to the mother liquor as cryoprotectant. Once transferred outside the glove box, crystals were keeped under liquid nitrogen to prevent the unfreezing that could let some oxygen enter the crystal. The presence of oxygen in the glove box was constantly monitored. The oxygen amount was about 2 and 30 ppm, except the first attempt of G106A mutant soaked with gentisate: in this case the amount rose accidentally to 70-80 ppm.

7.4 Data collection and refinement of enzyme—substrate adducts

Complete data sets at 100 K were collected for each adduct obtained, adding 30% (v/v) glycerol to the mother liquor as cryoprotectant, at EMBL beamline X13, DESY, Hamburg and at beamline XRD1, Elettra, Trieste, Italy. The statistics for data collections are summarized in Tables 7.6, 7.9 and 7.12. Structures were refined using the program REFMAC5. Manual rebuilding of the model was performed using the program Coot. Solvent molecules were introduced automatically using ARP. Data refinement statistics are summarized in Tables 7.7, 7.10 and 7.13. Tables 7.8, 7.11 and 7.14 show Ramachandran plot statistics.

Data Collection	WT — salicylate	WT — gentisate	WT — naphtoate
Beamline	XRD1	X13	X13
Space group	I222	I222	I222
Unit cell parameters: a(Å)	76.65	76.58	76.15
Unit cell parameters: b(Å)	86.96	86.97	86.13
Unit cell parameters: c(Å)	166.47	166.78	166.94
V_M (Å^3)	3.05	3.13	3.01
Solvent content (%) (v/v)	59.42	60.40	58.84
Wavelength (Å)	1.000	0.8123	0.979
Limiting resolution (Å)	47.3-2.10 (2.21-2.10)	20.0-2.15 (2.27-2.15)	46.7-2.35 (2.46-2.35)
Unique reflections	32880	29532	21895
R_{symm} (%)	12.2(42.3)	5.8(30.7)	9.1(45.0)
Multiplicity	4.3(4.3)	4.2(3.8)	3.3(3.2)
Completeness overall (%)	99.9(99.9)	97.2(94.4)	95.3(92.9)
I/σ(I)	11.8(3.2)	16.6(3.8)	11.7(3.1)

Table 7.6: Wild type adducts: summary of data collection statistics; values in parentheses are for the highest resolution shell.

Refinement	WT — salicylate	WT — gentisate	WT — naphtoate
Resolution range (Å)	46.78-2.10	20.00-2.15	46.71-2.35
Unique reflections, working/free	32188/1640	27948/1479	207084/1099
R_{factor} (%)	21.7	18.2	18.1
R_{free} (%)	27.3	22.1	25.2
Non-hydrogen atoms	3243	3168	3246
Water molecules	347	272	333
r.m.s.d. from ideal: Bond lengths (Å)	0.016	0.014	0.015
r.m.s.d. from ideal: Bond angles (∘)	1.923	1.672	1.778
B factor (Å^2)	22.3	27.0	29.5

Table 7.7: Wild type adducts: summary of refinement statistics.

Ramachandran Plot	WT — salicylate	WT — gentisate	WT — naphtoate
Most favoured region	262(87.6%)	268(89.3%)	268(89.3%)
Additionally allowed region	33(11.0%)	29(9.7%)	29(9.7%)
Generously allowed region	2(0.7%)	1(0.3%)	1(0.3%)
Disallowed region	2(0.7%)	2(0.7%)	2(0.7%)

Table 7.8: Wild type adducts: summary of Ramachandran Plot statistics.

Data Collection	G106A — salicylate	G106A — gentisate	G106A — gentisate—O_2	G106A — product
Beamline	X13	X13	XRD1	XRD1
Space group	I222	I222	I222	I222
Unit cell parameters: a (Å)	74.27	76.50	75.25	73.28
Unit cell parameters: b (Å)	86.98	86.89	86.79	87.24
Unit cell parameters: c (Å)	167.63	166.83	167.13	167.90
V_M (Å^3)	3.30	2.84	3.43	3.32
Solvent content (%) (v/v)	62.46	56.32	63.83	62.62
Wavelength (Å)	0.8123	0.8123	1.000	1.000
Limiting resolution (Å)	19.65-2.45 (2.58-2.45)	19.60-2.00 (2.11-2.00)	46.98-2.30 (2.42-2.30)	47.09-2.25 (2.37-2.25)
Unique reflections	20120	37295	24657	25906
R_{symm}	7.9(47.7)	6.5(41.4)	6.6(27.7)	11.7(55.4)
Multiplicity	3.4(2.9)	4.2(3.5)	7.5(7.3)	4.5(4.5)
Completeness overall (%)	99.1(99.5)	98.7(98.0)	99.8(99.6)	99.9(100.0)
I/σ(I)	11.7(2.3)	15.0(3.3)	21.7(6.4)	10.9(2.7)

Table 7.9: G106A mutant adducts: summary of data collection statistics; values in parentheses are for the highest resolution shell.

Refinement	G106A — salicylate	G106A — gentisate	G106A — gentisate—O_2	G106A — product
Resolution range (Å)	19.6-2.45	19.6-2.00	44.8-2.30	46.6-2.25
Unique reflections, working/free	20084/1002	37276/1856	24655/1255	25896/1304
R_{factor} (%)	20.1	16.5	18.9	21.8
R_{free} (%)	28.1	21.2	25.3	27.1
Non-hydrogen atoms	2928	3489	2844	2890
Water molecules	172	532	155	141
r.m.s.d. bond lengths (Å)	0.018	0.017	0.023	0.025
r.m.s.d. bond angles (°)	1.972	1.844	2.173	2.315
B factor (Å^2)	40.0	29.9	35.4	29.7

Table 7.10: G106A mutant adducts: summary of refinement statistics.

Ramachadran Plot	G106A — salicylate	G106A — gentisate	G106A — gentisate—O_2	G106A — product
Most favoured region	239(85.7%)	270(89.7%)	247(88.8%)	243(85.3%)
Additionally allowed region	37(13.3%)	28(9.3%)	29(10.4%)	39(13.7%)
Generously allowed region	2(0.7%)	2(0.7%)	2(0.7%)	2(0.7%)
Disallowed region	1(0.4%)	1(0.3%)	—	1(0.4%)

Table 7.11: G106A mutant adducts: summary of Ramachandran Plot statistics.

Data Collection	W104Y—gentisate
Beamline	X13
Space group	I222
Unit cell parameters a; b; c (Å)	75.35; 87.90; 167.02
V_M ($Å^3$)	3.26
Solvent content (%) (v/v)	61.96
Wavelength (Å)	0.8123
Limiting resolution (Å)	57.17-2.50(2.64-2.50)
Unique reflections	18973
R_{symm} (%)	9.4(39.2)
Multiplicity	3.0(3.0)
Completeness overall (%)	97.5(98.3)
I/σ(I)	9.7(2.2)

Table 7.12: W104Y mutant adduct: summary of data collection statistics; values in parentheses are for the highest resolution shell.

Refinement	W104y—gentisate
Resolution range (Å)	38.9-2.50
Unique reflections, working/free	18004/953
R_{factor} (%)	20.2
R_{free} (%)	27.1
Non-hydrogen atoms	3032
Water molecules	136
r.m.s.d. from ideal: Bond lengths (Å)	0.024
r.m.s.d. from ideal: Bond angles (∘)	2.277
B factor ($Å^2$)	36.4

Table 7.13: W104Y mutant adduct: summary of refinement statistics.

Ramachandran Plot	W104Y—gentisate
Most favoured region	248(82.7%)
Additionally allowed region	45(15.0%)
Generously allowed region	4(1.3%)
Disallowed region	3(1.0%)

Table 7.14: W104Y mutant adduct: summary of Ramachandran Plot statistics.

Part III

Laccase: following the catalytic reaction with a multi-crystal data collection strategy

8

Laccase and its redox states

Laccases (benzenediol oxygen oxidoreductase; EC 1.10.3.2) are glycosilated multi-copper oxidases. Since their discovery more than one century ago in the Japanese tree *Rhus venicifera* [63], laccases have been found to be widely distributed among plants, where they are involved in the wounding response and the synthesis of lignin, and subsequently they were discovered to be present in insects, bacteria and widely diffused in basidiomyceteous and ascomyceteous fungi [64]. The simplest reactions catalyzed by laccases are those in which a vast set of substrates, typically phenols and arylamino compounds, is oxidized to the corresponding radical species by direct interaction with their active site [65] and accompanied by the reduction of molecular oxygen to water. Due to their very broad substrate range they are implicated in an extensive series of functions such as pathogenesis, immunogenesis and morphogenesis of organisms and in the metabolic turnover of complex organic substances such as lignin, humic matter, and toxic xenobiotics [66].

White-rot fungi degrade wood lignin using a combination of specialized intra- and extra-cellular enzymes [67]. Lignin, the most common polymer on earth which provides the structural component of the plant cell wall, is a heterogeneous biopolymer composed by phenyl propanoid units linked by various non-hydrolyzable C-C- and C-O- bonds. The ligninolytic system of white-rot fungi was thought to be mainly composed by lignin-peroxidase and manganese-peroxidase [68]. Laccases, incapable of directly cleaving the non-phenolic bonds of lignin, were not considered as significant components of the ligninolytic system, despite the secretion of large quantities of laccases by the vast majority of white-rot fungi under ligninolytic conditions. More recently it was discovered that white-rot fungi laccases can enlarge their substrate range and are then able to oxidize compounds with a redox potential exceeding their own, such as non-phenolic benzylalcohols [69] [70] [71]. In fact their catalytic competences can be further extended to substrates which are not oxidized directly,

either because they are too large to penetrate into the enzyme active site or because they have a redox potential higher than the laccase itself. Nature overcomes these limitations with the utilization of mediators, which are suitable compounds that act as intermediate substrates for the laccase: the oxidized radical forms of which are sufficiently stable to leave the enzyme site and react with the bulky or high redox potential substrate targets. This finding led to the discovery that laccase-mediator systems effectively play a major role in the biodegradation of lignin and recalcitrant aromatic pollutants [71] [10] [72]. These non enzymatic routes of oxidative polymerizing or depolymerizing reactions are vital in a range of physiological functions such as lignolysis, lignin synthesis, morphogenesis, pathogenesis and detoxification.

Owing to their high and non-specific oxidation capacities, to the lack of a requirement for cofactors and to the use of readily available oxygen as an electron acceptor, laccases are useful biocatalysts with some established and lots of emerging biotechnological applications such as biobleaching, xenobiotics bioremediation, textile dyes decolorization, biosensors, food industry [8] [73] [9].

These multicopper oxidases are generally monomeric glycoproteins containing about 500 amino acids arranged in 3 β-barrel domains assembled to model three spectroscopically distinct copper binding catalytic sites: one type 1 (T1, blue copper, characterized by a strong absorption at 600 nm), one type 2 (T2, normal copper) and one type 3 (T3, EPR-silent antiferromagnetically coupled dinuclear coppers). Four one-electron oxidations of the reducing substrates mentioned above which occur at the T1 site on the protein surface are coupled to the four-electrons reduction of dioxygen to water which occurs at the internal T2/T3 cluster.

Type 1 copper has a trigonal coordination, with two histines and a cysteine as conserved equatorial ligands. A further axial ligand at a longer distance is also observed in several multicopper oxidases. It has been widely argued that this axial position ligand strongly influences the oxidation potential of the enzyme, possibly providing the mechanism for regulating its activity. A mutation from phenylalanine to methionine signicantly lowered the oxidation potential of a fungal laccase from *Trametes villosa* [74]. Type 1 copper confers the typical blue colour to multicopper proteins, which results from the intense electronic absorption caused by the covalent copper-cysteine bond. Due to its high redox potential of about 790 mV, Type 1 copper is the site where substrate oxidation takes place.

Type 2 or normal Cu(II) site is characterized by the lack of strong absorption features in the visible region and reveals usual EPR spectra. It is strategically positioned close to the Type 3 copper, a binuclear center spectroscopically characterized by an electron adsorption at 330 nm (oxidized form) and by the absence of an EPR signal as the result of the anti-ferromagnetic coupling of the copper pair. The Type 3 copper center is also the common feature of another protein superfamily including the tyrosinases and haemocyanins. Type 2 and Type 3 copper form a trinuclear cluster, where reduction of molecular oxygen and release of water takes place. Type 2 copper is coordinated by two histidines and Type 3 copper atoms by six histidines. The strong anti-ferromagnetical coupling between the two Type 3 copper atoms, is related to the presence of a hydroxyl bridge.

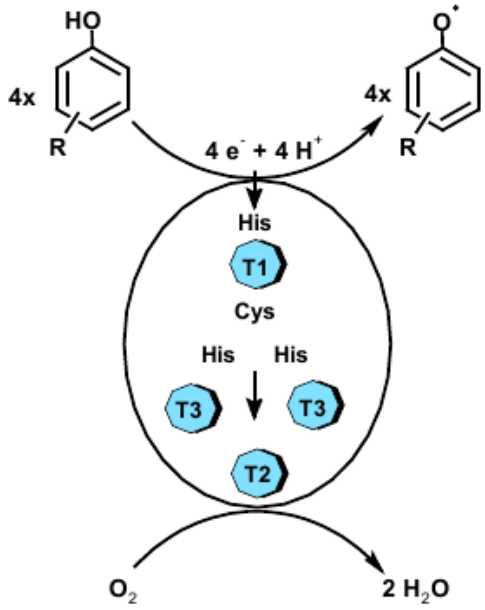


Figure 8.1: Catalytic cycle of laccases.

Regarding fungal laccases at first, the extensive microheterogeneity, presumably caused by variable glycosylation of these enzymes, hindered successful crystallization but deglycosylation performed to obtain high quality diffracting crystals resulted in the loss of copper as in the case of Ducros et al. which reported the crystal structure of a laccase from the fungus *Coprinus cinereus* at 1.68 Åresolution [75] in a form devoid of the type 2 copper and therefore in a catalytically incompetent state. More recently the X-ray structures of laccases from the fungi *Trametes versicolor* (1.9 Åresolution) [76], *Melanocarpus albomyces* (2.4 Åresolution) [77], *Rigidoporus lignosus* (1.7 Åresolution) [78], *Cerrena maxima* (1.9 Åresolution) [79], and the spore-coat laccase from *Bacillus subtilis* (1.7 Åresolution) [80] have been described. Substrate and dioxygen binding was investigated for the *Bacillus subtilis* endospore-coat laccase (substrate: ABTS) [80] and for the *Trametes versicolor* laccase (substrate: 2,5-xylidine) [76].

The unique kinetic and spectral features of copper/dioxygen intermediates in multicopper oxidases have been investigated into details mainly by Reinhammar and Solomon groups [81] [82] [83]. These studies led to several hypotheses of molecular mechanisms for the four-electrons reduction of dioxygen to water by this class of enzymes. X-ray structural studies on ascorbate oxidase, ceruloplasmin and on several fungal and bacterial laccases have further allowed the rationalization of important structural and functional aspects of multicopper oxidases such as the positioning of the T1 site, the electron transfer pathways between T1 and the T2/T3 clus-

ter coppers, the oxygen and water channels, and some of the substrates/inhibitor interactions.

Nevertheless several focal points still need to be elucidated from the structural standpoint including the nature of the intermediates in the oxygen reduction pathway. In particular the biotechnological application of laccases, aiming at the development of various industrial oxidative processes going from the production of fine chemicals and pharmaceuticals to bioremediation of contaminated soil and water, would receive great impulse to improvements from the complete comprehension of the catalytic mechanism of multicopper oxidases, and in particular of their redox potential and substrate selectivity control and a detailed characterization of the high resolution molecular structure of such enzymes will surely help in achieving such aims. Recently in our laboratory, a high resolution structural analysis of a laccase from *Lentinus tigrinus* detected the structure of two sequential intermediates in the molecular oxygen reduction pathway: the peroxide and the native intermediates the latter only hypothesized through spectroscopic, kinetic and molecular mechanics calculations studies [84].

In this study we solved the structure of a blue laccase from the white-rot fungus *Trametes trogii*, showing the binding site of the substrate; the structure allowed to identify two different redox states in the trinuclear copper cluster. Thus we solved the structure of a blue laccase from the white-rot fungus *Steccherinum ochraceum*. Through a peculiar redox titration performed *in cristallo*, we were able to isolate several sequential intermediates in the reduction of dioxygen to water in the trinuclear copper cluster that allow a more detailed comprehension of the redox mechanism of these class of enzymes.

8.1 The Structure of Laccase from *Trametes trogii*

The tridimensional structure of Laccase from *Trametes trogii* (TtL) was solved at 1.58 Åresolution using the molecular replacement technique and employing, as a starting model, the coordinates of the *Lentinus tigrinus* laccase (pdb code: 2qt6) [84]. The final model is composed by a single chain of 496 residues, 595 water molecules, 4 Cu ions, 6 saccharide moieties (4 moleculesof N-acetyl-glucosamide and 2 of mannose), 3 Ca(II) ions, 3 glycerol molecules and a molecule of p-toluate. The schematic structure and folding topology of the monomeric globular glycoprotein TtL, with overall dimensions 70 Å$\times$ 60 Å$\times$ 50 Å, is shown in Figure 8.2. The overall structure is organized in three domains; each of them has a greek key β-barrel topology, strictly related to small copper protein folding like azurin and plastocyanin and common to all the members of the blue multicopper oxidase family, like the ascorbate oxidase and the mammalian ceruloplasmin: domain I is composed by the first 139 residues (1-139), domain II starts at residue 140 and ends at residue 319 and domain III is composed by the last 176 residues (320-496). The structure is stabilized by two disulde bridges: Cys85-Cys485, between domains I and III, and Cys117-Cys204, between domains I and II. TtL is a monomer.

Figure 8.2: Schematic representation of the overall structure of Laccase from *Trametes trogii*; domain I is depicted in blu, domain II in green and domain III in yellow; the copper ions are depicted as magenta spheres.

The density maps of TtL indicate the presence of glycosylation at Asn54 and Asn433 residues. The corresponding electronic densities were modelled to one di(N-acetyl-D-glucosamine), one β-D-mannose, and one α-D-mannose moieties bound to Asn54, and one di(N-acetyl-D-glucosamine) moiety bound to Asn433. Three Ca ions were also found on the enzyme surface: one bound to Thr188 and three water molecules, a second bound to Ser224, Ala103 and four water molecules, the third bound to Asp183 and two water molecules.

8.1.1 The T1 active site

Figure 8.3 shows the coordination of the copper atoms and the electron density obtained for the ions and the surrounding residues. The T1 active site is near to the external surface where organic substrates (mainly phenols or arylamines) are oxidized releasing electrons to the copper ion through a bound His residue (His455). The coordination of the T1 copper ion, which is not solvent exposed, is completed by His394 and Cys450, that belong to domain III. In laccases the T1 copper exhibits a planar triangular coordination with the above mentioned two histidines and one cysteine residue; contrarily to many other multicopper oxidases which show a tetrahedrical coordination of the T1 copper ion with an axial fourth ligand, usually a methionine, to complete the first coordination sphere [85].

A role of redox potential modulation was ascribed to this ligand nature in azurin and recently in a bacterial laccase: when the axial methionine ligand at the T1 copper was substituted by a leucine the redox potential increased of about 100 mV

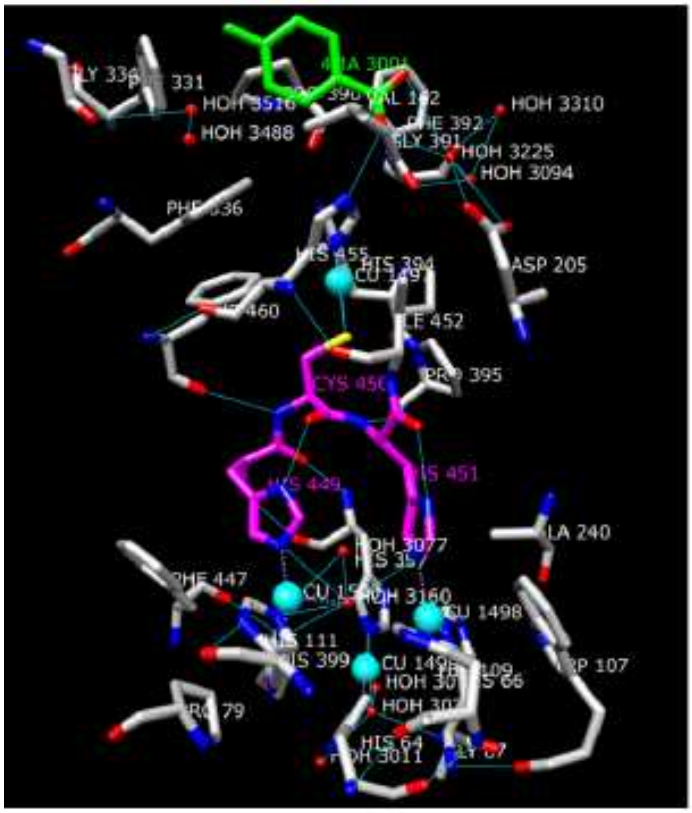


Figure 8.3: Stereoview of the schematic representation of the four copper sites in TtL; the intramolecular electron pathways are shown in magenta.

[86]. Additional experiments of site specific mutagenesis, showed that Leu-Phe substitution at such site resulted in no significant effect [87]. The variations in redox potential of the T1 site observed among laccases as well as in other metalloenzymes cannot be ascribed to a single structural feature but to a sum of factors such as the nature of the direct ligands and to their coordination geometries which seem to be somewhat invariant in many laccases but also to the nature of the second sphere residues present in the structure which influences solvent accessibility, hydrogen bonding, and dielectric anisotropy around the site. In particular in the present structure the occurrence of two hydrophobic residues Phe460 and Ile452 in the near surroundings of the T1 copper ion (3.79 Åand 3.58 Åfrom the copper ion, respectively) would indeed contribute to the high redox potential observed (E_0 = 760 mV). Furthermore, residue Phe460 is additionally surrounded by a large number of hydrophobic residues: Phe336, Phe329, Phe396, Ala389 and Ile338 that would also contribute to rise the redox potential of the copper center. In *Coprinus cinereus* laccase (E_0 = 550 mV) Phe460, Phe329, and Ala389 are substituted by Leu462, Leu333, and Val391 whereas in the low redox potential bacterial *Bacillus subtilis*

laccase (E_0 = 455 mV) Phe460, Phe336, Phe329, Ile338, Ala389 and Phe396 are substituted by Met502 (fourth ligand to the T1 copper), Leu386, Leu374, Leu388, Thr415 and Ile421, respectively. These observations support the fact that several mutations are needed to change the redox potential of the metal site in a 300 mV range.

8.1.2 The T2/T3 trinuclear active site

The electrons taken from the substrates are transferred, through two intramolecular electron transfer pathways, shown in Figure 8.3 in magenta and formed by Cys450, His 449 and His451, from the T1 copper ion to the second active site, where the oxygen molecules bind and are reduced to water molecules. It is composed by a trinuclear copper cluster, with a T2 Cu ion and two T3 Cu ions arranged in a triangular way, embedded between domains I and III and bound to eight histidine residues and two water molecules. T2 Cu is tricoordinated and bound to His64, His397 and one water molecule. The two T3 coppers are tetracoordinated: T3a Cu binds to residues His111, His399 and His449, T3b Cu binds to residues His66, His109 and His451; moreover there is a water molecule asymmetrically bound between the two T3 copper ions (with a distance of 2.98 Åto T3a and 2.23 Åto T3b).

The surface of the two solvent channels which provide access to the trinuclear copper cluster, located in the interior of the protein structure, are shown in Figure 8.4. The first channel points towards the two T3 copper ions on one side of the T2/T3 cluster allowing to the molecular oxygen to enter and to bind to it. The second channel pointing towards the T2 copper ion on the other side of the cluster permits to the water molecules produced in the dioxygen reduction to move to the bulk solvent.

8.1.3 The *p*-toluate ligand

In the current structure a further electron density was found on the protein surface near to the T1 copper site; a molecule of p-toluate (p-methylbenzoate) or p-hydroxybenzoate could be modelled into such density (see Figure 8.5). Neither one of these compounds were added during the crystallization process or during the enzyme purification procedures, therefore we suppose that the bound molecule is a fungal metabolite. This moiety is located in the putative cavity for the reducing substrate: the carboxylate group is hydrogen bonded to His455 (at 2.7 Å) whereas the para-group is hosted in a hydrophobic region composed by Phe331, Phe336, Pro390 and Val162. As p-hydroxybenzoate is known to be a substrate for laccases its binding to the T1 copper center is likely to occur through an hydroxyHis455 hydrogen bond interaction. For this reason the electronic density hosted in the hydrophobic pocket was modeled as a methyl group (p-toluate) rather than as an hydroxyl group (p-hydroxybenzoate).

A weak substrate of laccases, 2,5-xylidine (2,5-dimethylaniline), was observed to bind to the enzyme from *Trametes versicolor* in a similar mode [76]. The ring

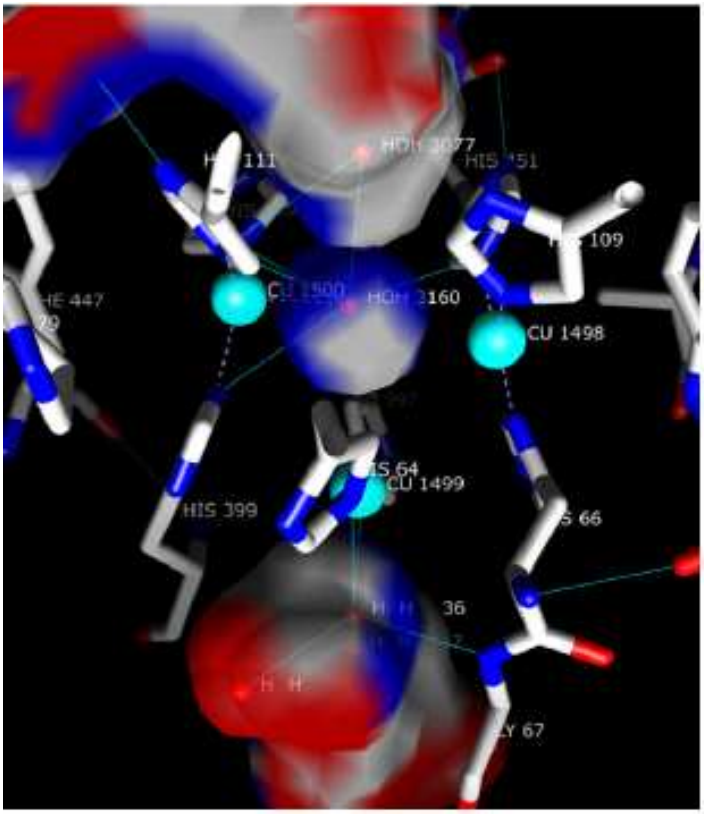


Figure 8.4: T2/T3 trinuclear cluster active site of TtL. The molecular surfaces of the oxygen channel (on top) reaching the trinuclear cluster as well as that of the water channel (on the bottom) are shown.

of the His residue coordinated to the T1 copper is close to the nitrogen of 2,5-xylidine at a distance consistent with a hydrogen bond (3.2 Å). The amino group of 2,5-xylidine was also hydrogen bonded to a terminal oxygen of a vicinal Asp side chain. We tested the properties of p-toluate in order to establish if this compound could have some effect on the enzyme activity and resulted that such molecule does not acts as a substrate and it is only a very weak inhibitor towards ABTS (2,2'-azino-bis(3-ethylbenzothiazoline-6-sulfonic acid)) oxidation enzymatic activity. Anyway the observed adduct may well suggest the mode of binding for a variety of compounds interacting with laccases since to our knowledge the T1 site does not exhibit strong interactions neither with substrates nor with inhibitors.

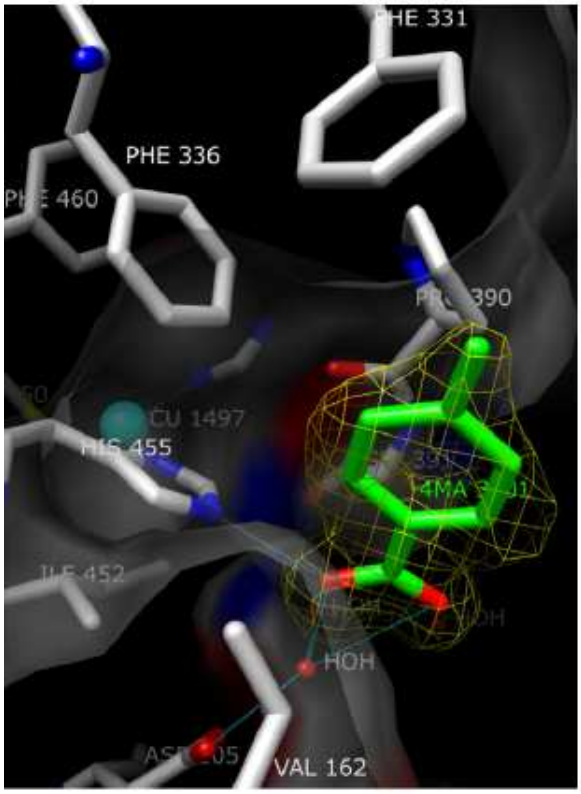


Figure 8.5: Representation of the electron density for the p-toluate moiety bound to the T1 active site of TtL. The electron density is contoured at 2σ. The molecular surface of the site is also shown.

8.2 The structure of Laccase from *Steccherinum ochraceum*

Laccase from *Steccherinum ochraceum* (SoL) shows 63.5% of identity and 74.5% of similarity with laccase from *Trametes trogii*. The SoL structure was solved at 2.00 Åresolution with molecolar replacement using the structure of laccase from *Panus tigrinus* as done for TtL. The final model is composed by three chains of 495 residues each, 1196 water molecules, 12 Cu ions, 6 molecules of N-acetyl-glucosammide, 7 glycerol molecules (added as a cryoprotectant), a Zn(II) ion and 4 sulphate ions (both species added during crystallization). We also determined the sequence of this protein. In the final overall structure we can notice the three domains, typical of blue laccases, as described for TtL: domain I is composed by the first 140 residues (1-140), domain II starts at residue 141 and ends at residue 324 and domain III is composed by the last 170 residues (325-495). The structure is stabilized by two disulfide bridge: Cys86-Cys488, between domains I and III, and Cys118-Cys208, between domains I and II. The secondary structure of these domains is mainly a β-stranded structure, with a β-barrel in domain III.

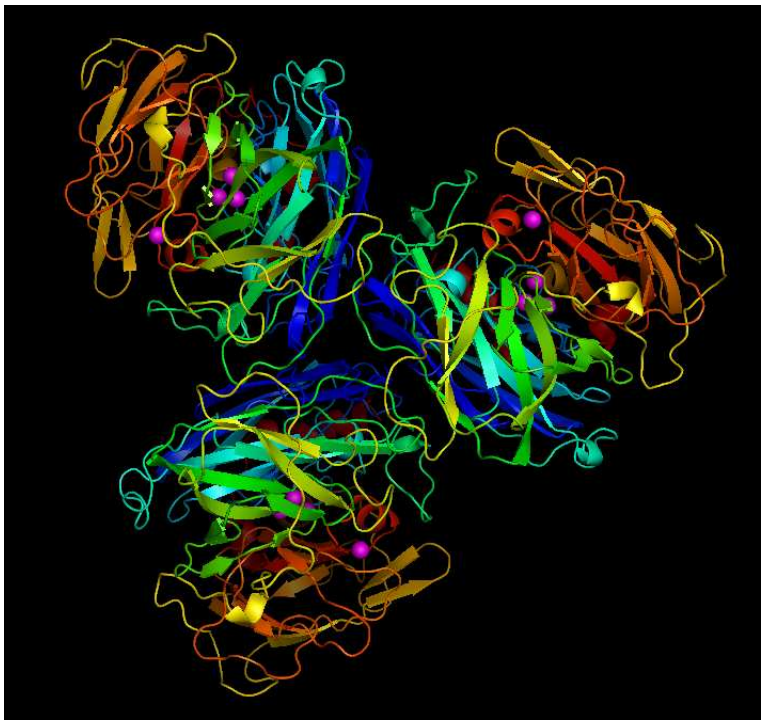


Figure 8.6: The overall structure of laccase from *Steccherinum ochraceum*, showing the three molecules per asymmetric unit.

Figure 8.7 shows the electron density on the copper atoms. The T1 Cu (Cu1) ion coordination involves the three residues His397, Cys453 and His458, which belong to domain III. The trinuclear copper cluster, with the T2 Cu (Cu4) and three two

T3 Cu (Cu2 and Cu3), is embedded between the domains I and III and bound by eight histidine residues and two water molecules. Cu4 is tricoordinated and bound to His65, His400 and a water molecule. Cu2 and Cu3 are tetracoordinated: Cu2 binds the residues His112, His402 and His452, Cu3 binds the residues His67, His110 and His454; moreover in chain A we found a water molecule asymmetrically bound between the two T3 copper ions (with a distance of 2.20 Åto Cu2 and 3.06 Åto Cu3).

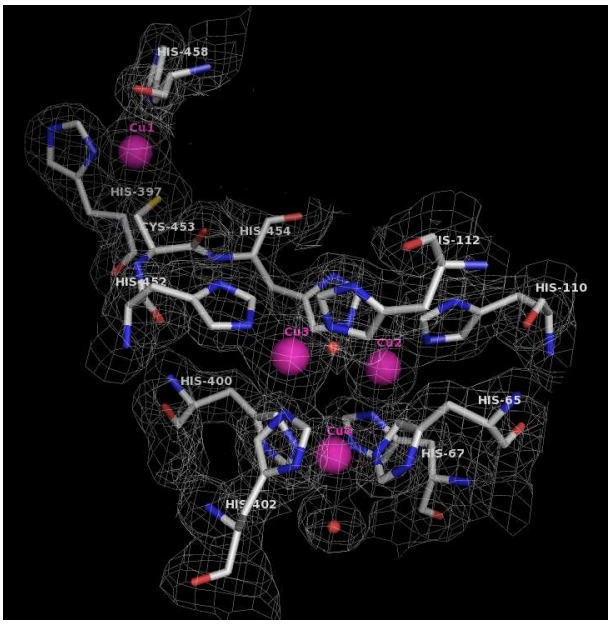


Figure 8.7: T2/T3 copper cluster and T1 copper ion in Blue Laccase from *Steccherinum ochraceum*.

8.3 Evidence of different redox states

As discussed in Chapter 3, the exposure of biological samples to X-ray can cause a reduction of the sample itself; the electrons responsible of such reduction are generated by processes of inelastic scattering of X-ray photons, i.e. photoelectric effect or Auger electrons. The X-ray structures we obtained from an X-ray exposure to an intense radiation source, like a syncrothron, therefore could show a reduction in disulphide bridges, carboxylate groups and above all in the redox center of the enzyme. In fact, comparing two different structures of TtL obtained with different X-ray sources we noticed a varied spatial arragement in the trinuclear copper cluster: distances between copper ions were shorter in the structural model obtained with a Cu-Kα sealed tube, compared to those in the model obtained with a syncrothron radiation source. Figure 8.8 depict the trinuclear copper cluster for the home data set (collected at the CRIST laboratory, Chemistry Dept, University of Florence,

Italy) and the desy data set (collected at the BW7B beamline at the DORIS storage ring, Hamburg, Germany).

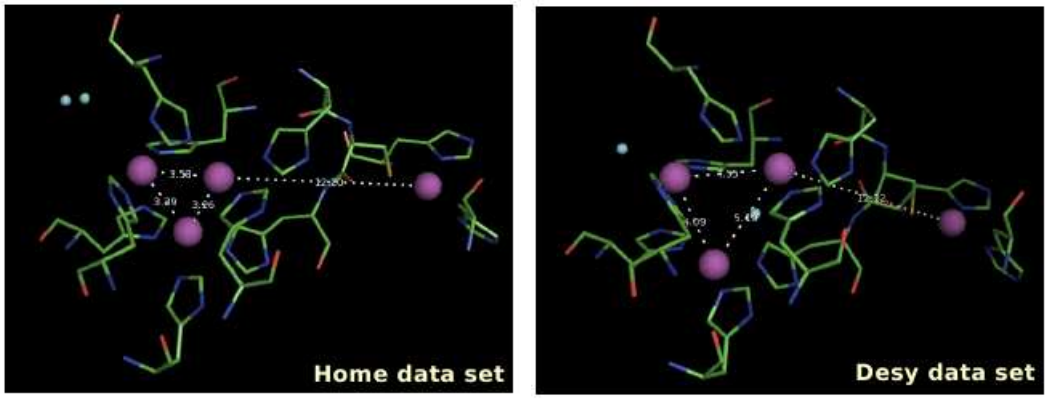


Figure 8.8: Geometry of the metal sites: a) home data set; b) desy data set.

Distance (Å)	home data set	desy data set
T1 Cu — T3a Cu	12.20	12.12
T3a Cu — T3b Cu	3.26	5.19
T3a Cu — T2 Cu	3.58	4.55
T3b Cu — T2 Cu	3.39	4.09

Table 8.1: Distances between T1 Cu and the cluster copper ions in home data set structure and desy data set structure.

The difference in the distances between the copper ions are clearly noticeable: in the desy data set structure they are almost 1 Ålonger than in the home data set structure, as shown in Table 8.1. We associated the different lenght with different redox states of the cluster and thus assumed that the trinuclear site was reduced by intense X-ray exposure during data collection at syncrothron source.

9

The redox titration

The reduction occurring during syncrothron radiation exposure could be an advantage in the mechanistic study of the enzyme, since it could be used to simulate the laccase mechanism, that requires the reduction of the copper ions to oxidise a dioxygen molecule to water. In order to develop further this idea, we collected a set of Uv-Vis absorption spectra *in cristallo* during X-ray exposure to monitoring the redox state of the enzyme. Therefore, a peculiar multi-crystal data collection strategy let us obtain different X-ray structures at increasing radiation doses.

9.1 Coupling Uv-vis tecniques and crystallography

Figure 9.1 shows the Uv-Vis absoption spectra collected *in cristallo* during X-ray exposure at a syncrothron source. Spectra reported were obtained with a SoL laccase. Before exposure we obtained a spectrum similar to those obtained for a laccase in solution: the strong absorption band at 600 nm, that gives the tipycal blue colour to the enzyme, is due to a ligand-metal change transfert between T1 Cu and Cys453. During X-ray exposure, there is a decrease in the absorption at 600 nm, because of the reduction of T1 Cu. Meanwhile we observe the increasing of the absorption at 400 nm, that has been related to the formation of an intermediate in the trinuclear copper cluster, as described by Hakulinen and co-workers in a similar study on laccase from *Melanocarpus albomyces* [88]. Probably the absorption at 400 nm is simply due to the reduction of the trinuclear cluster, or even the reduction of T2 or T3 coppers, because it continues to increase during X-ray exposure until reaching a maximun after 40 minutes; the redox titration described in the next section allowed us to isolate the redox state caused by an X-ray dose corresponding to 40 minutes exposure, which shows the complete reduction of the trinuclear copper cluster.

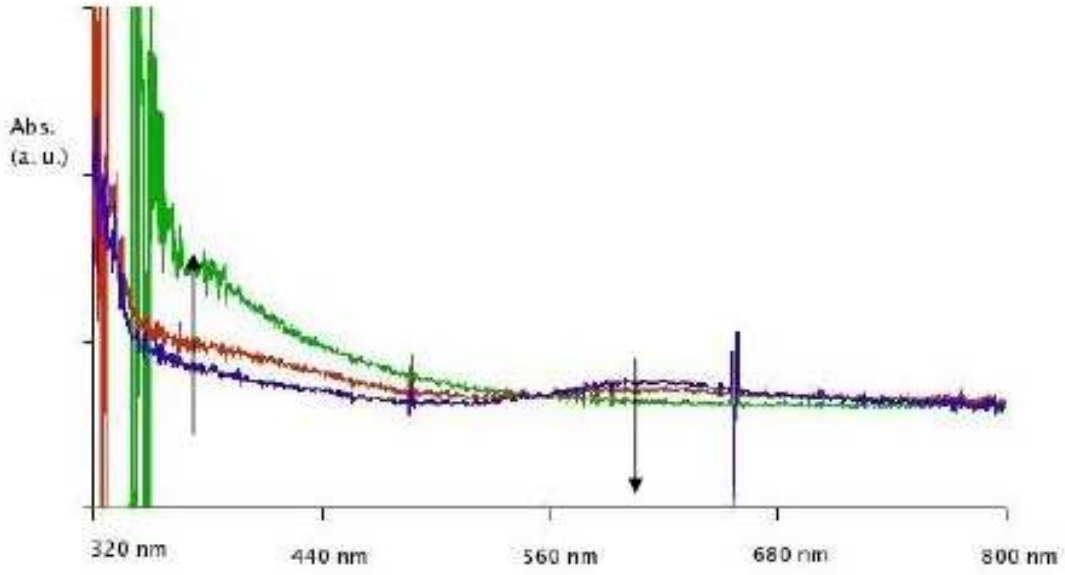


Figure 9.1: Uv-Vis absorption spectra collected during x-ray esposure on a SoL crystal: before exposure (blue); after 40 seconds (red) and after 40 minuter (green) of exposure.

9.2 Isolation of laccase reaction intermediates

The X-ray redox titration was conducted with a peculiar multi-crystal data collection strategy; we obtained different x-ray structures at increasing radiation doses for laccase from *Steccherinum ochraceum*.

The method is based on X-ray damage distribution over many crystals followed by the assembly of composite data sets from small chunks of these data. Composite data sets obtained in this way report on structures which received different X-ray doses, and in which a different numbers of electrons were liberated by inelastic processes during data collection. This strategy offers the possibility to control and monitor X-ray-induced redox mechanism. The technique permits experiments similar to redox titrations so that redox processes can be followed in the crystal.

Figure 9.2 shows how we assembly the composite data sets at increasing X-ray dose. The technique is similar to those used by Berglund and co-workers in their study on horseradish peroxidase [27]. We collected ten complete data collections with ten different SoL crystals; the dose received by each data collection is the same. Then we split every data collection into segments composed by ten images: every segment has received a definite dose, that was calculated with the program RADDOSE [89].

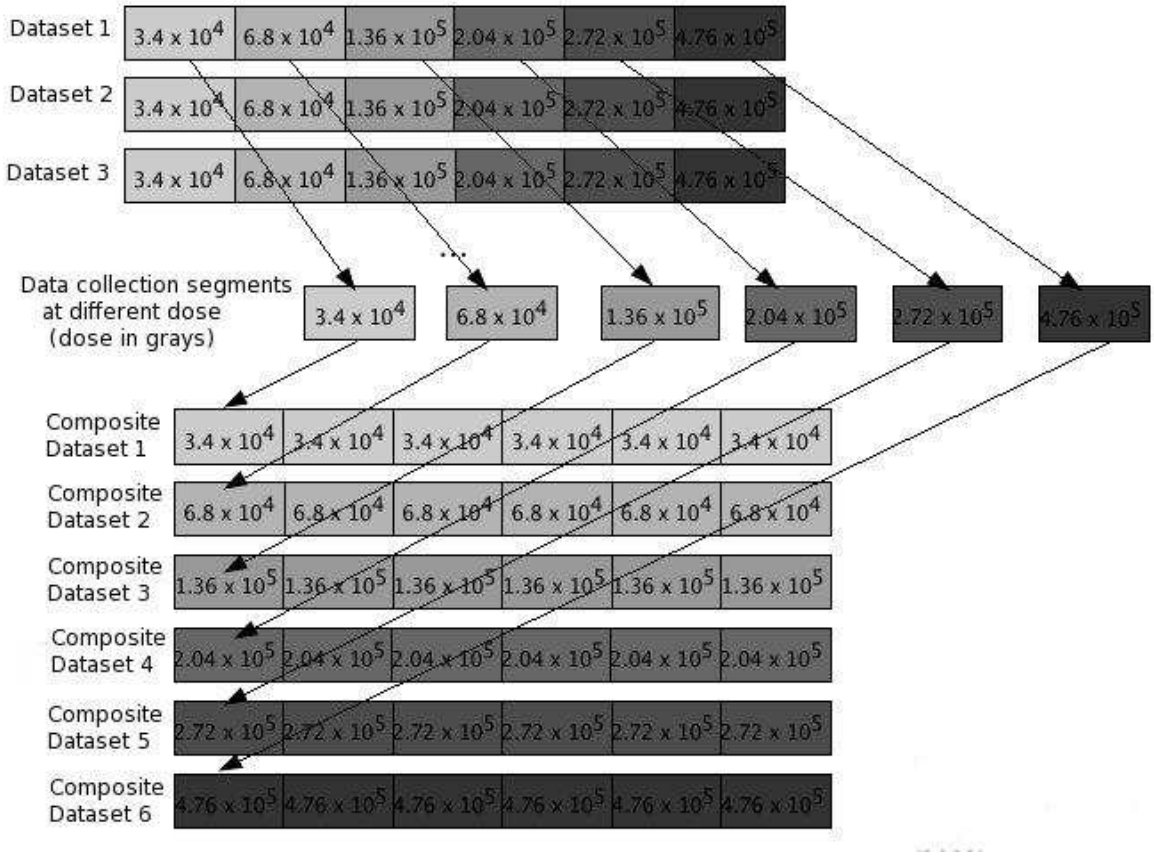


Figure 9.2: The technique: a series of data sets is collected on several crystals. Dose-dependent data sets are then reconstructed by properly combining fixed-dose subsets (dose is reported in grays).

Finally, we merged together the segments that received the same dose: we merged the first segment of each data collection to obtain the composite data collection at lower dose, the second segment of each data collection for a data collection at an increased dose and so on, until the last composite data set that is the one at higher dose.

The structures obtained by the composite data sets show the enzyme from a complete oxidate state (lower dose) to a complete reducted state (higher dose); in between, we isolate the intermediate reaction of the dioxygen reduction to two molecules of water, that is the dioxygen splitting in two atoms of oxygen (remember that usually it is not possible to observe hydrogen atom in x-ray protein crystallography).

Figure 9.3 shows the structure obtained at a dose of 3.4×10^4 grays; the distancs between copper atoms in the trinuclear cluster are similar to those obtained for the home data set of TtL. Therefore we assume that SoL is in its oxidated form. At

a dose of 6.8x10^4 grays, we observe two different states in subunit A e B of the structural model; in subunit B we found a density between T3 copper atoms that was easily modeled with a dioxygen molecule (see Figure 9.4), while in subunit A the density between T3 copper atoms was enlongated compared to subunit B, and thus was modeled with two oxygen atoms splitted at a distance of 2.30 Å(Figure 9.5a).

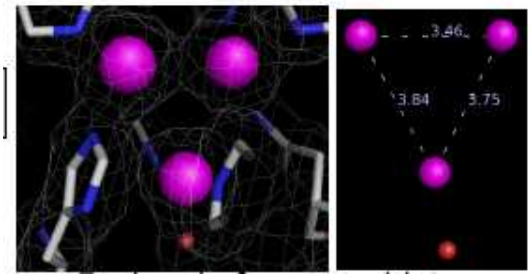


Figure 9.3: Electron density contoured at 2σ and distances in the trinuclear copper cluster (subunit A) for the structure at 3.4x10^4 grays; the enzyme is in its oxidated form.

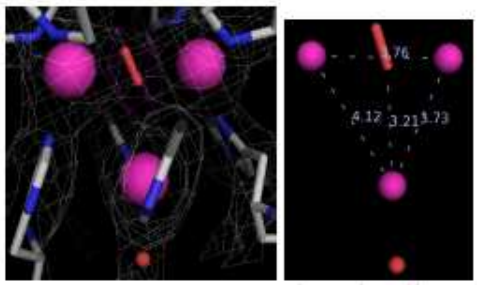
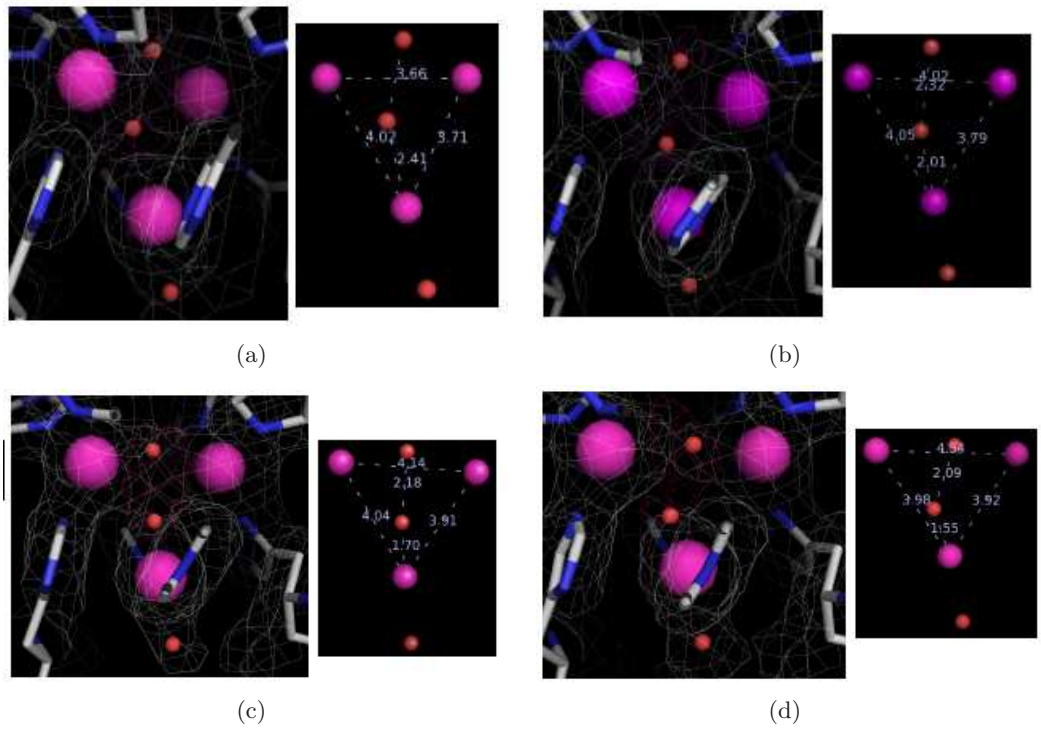


Figure 9.4: Electron density contoured at 2σ and distances in the trinuclear copper cluster (subunit B) for the structure at 6.8x10^4 grays;the dioxygen molecule is bound.

With increasing X-ray doses, we obtained the structures whose trinuclear sites are depicted in Figure 9.5: the splitting of the dioxygen molecule due to reduction to water is progressively more evident, since the distance between the second oxygen atom of the molecule and T2 copper is decreased from 2.41 Åto 1.55 Å.

Finally, at a dose corresponding to 4.8x10^5 grays, we obtained the enzyme completely reduced (Figure 9.6); in this structure, distances between copper ions are in fact close to those in the TtL structure obtained from desy data set.

Table 9.1 summarizes the distances between atoms in the trinuclear cluster for all datasets. It is clearly evident that the exposure to X-ray causes a progressive reduction of the copper atoms, the distances between them increasing constantly



(a) (b)

(c) (d)

Figure 9.5: Electron density contoured at 2σ and distances in the trinuclear copper cluster (subunit A) for the structure at a) 6.8x10^4 grays; b) 1.4x10^5 grays; c) 2.0x10^5 grays; d) 2.7x10^5 grays; the structures show the progressive splitting of dioxygen.

from low-dose data set to high-dose data set. Moreover, the splitting of the dioxygen molecule within the reduction to water is shown by the oxygen atom of the molecule coming closer to T2 copper, from a distance of 2.44 Åto 1.55 Å. The thermalisation effect of X-ray exposure was used to simulate the reduction that occurs normally during the calalityc circle of the enzyme; however, the simulation regards only the very first steps in the reduction, since during the catalisys the electron transfert from Cu atoms to dioxygen gives a newly oxidated form of the cluster and the consequent release of the water molecules. In this redox titration instead, the cluster remains reduced. This process let us to isolate two intermediates of the catalityc reaction, that are the binding of dioxygen and the splitting of dioxygen molecule, as depicted in Fiure 9.7.

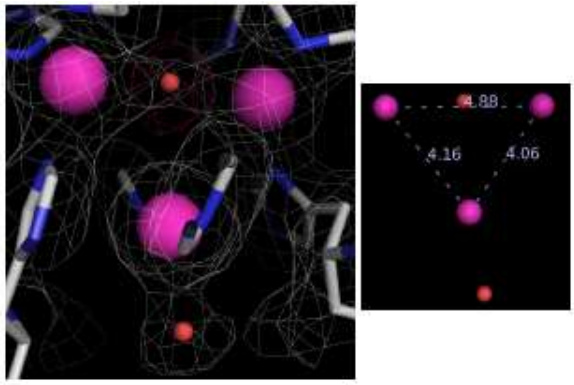


Figure 9.6: Electron density contoured at 2σ and distances in the trinuclear copper cluster (subunit A) for the structure at 4.8x10^5 grays; the enzyme is completely reduced.

Distance (Å) (dose in grays)	data set1 3.4x10^4	data set2 6.8x10^4	data set3 1.4x10^5	data set4 2.0x10^5	data set5 2.7x10^5	data set6 4.8x10^5
T3a Cu — T3b Cu	3.46	3.66	4.02	4.14	4.54	4.88
T3a Cu — T2 Cu	3.84	4.02	4.05	4.04	3.98	4.16
T3b Cu — T2 Cu	3.75	3.71	3.79	3.91	3.92	4.06
T2 Cu — O2	—	2.41	2.01	1.70	1.55	—

Table 9.1: Distances between cluster copper ions and dioxygen atoms in increasing dose structures.

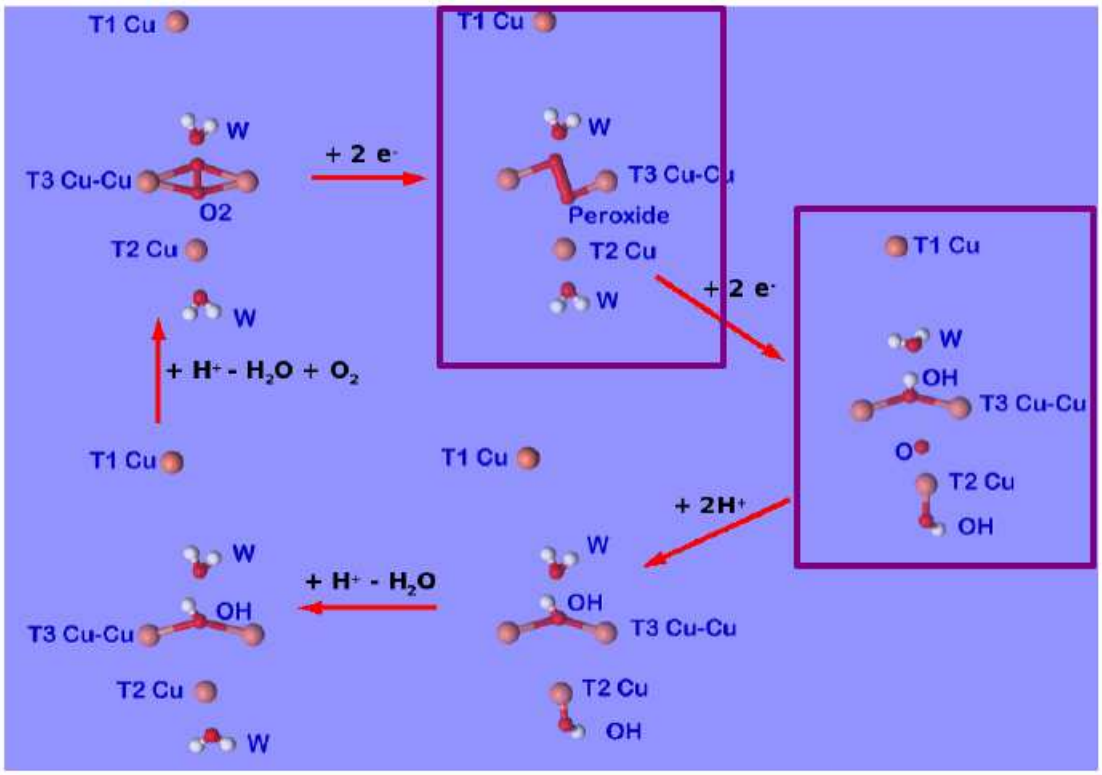


Figure 9.7: Proposed mechanism for Laccase; the two intermediates isolated by the multi-crystal redox titration are highlighted.

10

Experimentals

All the reagents were of the best purity available from Sigma-Aldrich. The initial trials in crystallization were made using commercial screens: JBScreen Classic from Jena Bioscience and Crystal Screen from Hampton Research.

Integration of data was performed with Mosflm [52] or XDS [53]. In refinement some programs belonging to the CCP4 program suite [54] were used: SCALA [55], MOLREP [56], REFMAC5 [57] and ARP [58]. Manual rebuilding of the models was performed using COOT [59]. Stereochemistry was analized with PROCHEK [60] and sequence alignement with ClustalW [61].

10.1 Laccase from *Trametes trogii*

Enzyme purification, activity assay and biochemical characterization were performed in our laboratory by A. Gullotto, S. Tilli and F. Briganti [90].

Crystallization

Crystallization experiments were performed utilizing the sitting-drop vapor diffusion methods in 96-wells plates (Crystal Clear Strips, Douglas Instruments Limited, UK). Preliminary crystallization trials were performed utilizing the Crystal Screen kit I (Hampton Research) and the JBScreens 4 and 5 (Jena Bioscience) at 296 K and the most promising conditions were optimized. One microliter of protein solution were mixed with 1 μl of reservoir solution and equilibrated against 50 μl of precipitant solution. The optimized crystallization conditions: PEG 8000 21%, calcium acetate 0.2 M, PEG 400 3%, Hepes 0.1 M pH 7.5, allowed to obtain diffraction quality crystals which grow after one week at 296 K.

Data Collection

A high-dose data set was collected at the BW7B beamline at the DORIS storage ring, Hamburg, Germany, using a MAR image plate detector and a wavelength of 0.8435 Å. Data were collected at 100 K adding 25% glycerol to the mother liquor as cryoprotectant to a maximum resolution of 1.58 Å. The data were processed and integrated with Mosm and scaled by Scala, from the CCP4 package. Data processing gave 106 982 unique reections, an overall completeness of 99.6% and an Rsym of 0.059. Crystals belong to the primitive orthorhombic space group $P2_12_12_1$ with unit cell dimension a = 84.4, b = 85.1, c = 108.6 Å. Assuming one molecule per asymmetric unit the solvent content is 64% of the unit cell (V_M = 3.5 Å^3/Da). Data collection statistics are summarized in Table 10.1. The low-dose data set was collected at Centro Interdipartimentale di Cristallografia Strutturale, Universit di Firenze, Italy (CRIST), using an Oxford Diffraction Xcalibur PX Ultra diffractometer equipped with CCD area detector and monochromated CuKα radiation; data collection statistics are summarized in Table 10.2.

Data collection	
Beamline	BW7B, Desy, Hamburg
Space group	$P2_12_12_1$
Unit cell parameters a(Å); b (Å); c(Å)	84.4; 85.1; 108.6
V_M (Å^3)	3.5
Solvent content (%) (v/v)	64
Wavelength (Å)	0.8435
Limiting resolution (Å)	85.13-1.58 (1.67-1.58)
Unique reflections	106932
R_{symm}	0.059 (0.362)
Multiplicity	4.1 (3.3)
Completeness overall (%)	99.6 (98.6)
I/σ(I)	15.1 (3.2)

Table 10.1: Summary of data collection statistics of TtL high-dose dataset; values in parentheses are for the highest resolution shell.

Refinement: TtL high-dose structure

The structure was solved with molecular replacement using the structure of a laccase from *Lentinus tigrinus* (PDB code 2QT6) as a startings model and the program MolRep from the CCP4 package. The alignment of the two sequences performed with program Align (http://www.ebi.ac.uk/emboss/align/) shows 70.2% of identity and 80.0% of similarity. The model was rened using the program Refmac 5.1.24 from the CCP4 program suite. Manual rebuilding of the model was performed using the program XTALVIEW. Solvent molecules were introduced automatically using

Data collection	
Beamline	CRIST
Space group	$P2_12_12_1$
Unit cell parameters: a; b; c (Å)	84.93; 84.33; 108.38
V_M (Å^3)	3.41
Solvent content (%) (v/v)	63.62
Wavelength (Å)	1.542
Limiting resolution (Å)	45.69-2.60(2.74-2.60)
Unique reflections	22395
R_{symm}	13.3(34.5)
Multiplicity	2.9(2.8)
Completeness overall (%)	92.4(95.5)
I/σ(I)	8.0(3.3)

Table 10.2: Summary of data collection statistics of TtL low-dose dataset; values in parentheses are for the highest resolution shell.

Refinement	
Resolution range (Å)	22.33-1.58
Unique reflections, working/free	101597/5068
R_{factor} (%)	17.4
R_{free} (%)	19.2
Non-hydrogen atoms	4476
Water molecules	548
r.m.s.d. from ideal: Bond lengths (Å)	0.008
r.m.s.d. from ideal: Bond angles (∘)	1.382

Table 10.3: Summary of refinement statistics of TtL high-dose structure.

ARP. Renement results in R_{factor} and R_{free} values of 17.4% and 19.2%, respectively. Renement statistics are summarized in Table 10.3. The stereochemistry of the nal model was analyzed with PROCHECK. The Ramachandran plot is of good quality; there are 415 non-glycine and non-proline residues; among these, 365 (88.0%) are in the most favored regions, 47 (11.3%) are in the additional allowed regions and 3 (0.7%) in the generously allowed regions. The overall mean B factor of the structure after renement is 17.5 Å^2 for the polypeptide chain, 33.4 Å^2 for water molecules and 19.8 Å^2 for all atoms.

Protein coordinates have been deposited with the Protein Data Bank (Protein Data Bank accession number 2HRG).

Refinement	
Resolution range (Å)	20.0-2.60
Unique reflections, working/free	20765/1111
R_{factor} (%)	17.3
R_{free} (%)	23.1
Non-hydrogen atoms	4068
Water molecules	255
r.m.s.d. from ideal: Bond lengths (Å)	0.021
r.m.s.d. from ideal: Bond angles (∘)	1.942
B factor (Å^2)	33.8

Table 10.4: Summary of refinement statistics of TtL low-dose structure.

Ramachandran Plot	
Most favoured region	350 (84.3%)
Additionally allowed region	59 (14.2%)
Generously allowed region	5 (1.2%)
Disallowed region	1 (0.2%)

Table 10.5: Ttl low-dose structure: summary of Ramachandran Plot statistics.

Refinement: TtL low-dose structure

The structure was solved with molecular replacement using the structure of a laccase from *Lentinus tigrinus* (PDB code 2QT6) as a startings model and the program MolRep from the CCP4 package. The alignment of the two sequences performed with program Align (http://www.ebi.ac.uk/emboss/align/) shows 70.2% of identity and 80.0% of similarity. The model was rened using the program Refmac 5.1.24 from the CCP4 program suite. Manual rebuilding of the model was performed using the program XTALVIEW. Solvent molecules were introduced automatically using ARP. Renement statistics are summarized in Table 10.4 and Table 10.5.

Protein coordinates have been deposited with the Protein Data Bank (Protein Data Bank accession number 2HRH).

10.2 Laccase from *Steccherinum ochraceum*

Enzyme purification, activity assay and biochemical characterization were performed by L. Golovleva and M. Kolomytseva from the Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences (Pushchino, Moscow Region, Russia). [91].

Crystallization

Diffraction quality crystals grow after one week at 296 K using the sitting drop vapour diffusion method: 1 μl of a 20.0 mg/ml protein solution were mixed with 1 μl of a solution containing ammonium sulphate 1.9 M, tris-HCl pH 9.0 0.1 M, Zn chloride 0.01 M and equilibrated against 100 μl of a solution of ammonium sulphate 1.9 M, tris-HCl pH 9.0 0.1 M, Zn chloride 0.01 M.

Data collection

A complete data set was collected at the BW7B beamline at the DORIS storage ring, Hamburg, Germany, using a MAR image plate detector at 100 K adding 20% glycerol to the mother liquor as cryoprotectant. The crystal parameters and data collection statistics are summarized in Table 10.6.

Data collection	
Beamline	BW7B, Desy, Hamburg
Space group	P222
Unit cell parameters a(Å); b (Å); c(Å)	73.54; 140.76 ; 174.18
V_M (Å^3)	2.78
Solvent content (%) (v/v)	55
Wavelength (Å)	0.8423
Limiting resolution (Å)	109.43-2.0 (2.11-2.0)
Unique reflections	120528
R_{symm}	0.079 (0.366)
Multiplicity	3.4 (3.0)
Completeness overall (%)	98.6 (95.8)
I/σ(I)	14.4 (3.3)

Table 10.6: Summary of data collection statistics of SoL dataset; values in parentheses are for the highest resolution shell.

Refinement

The structure was solved with molecular replacement using the structure of a laccase from *Panus tigrinus* (PDB entry 2QT6) as a starting model and the program MolRep from the CCP4 package. The model was refined against 2.00 Åresolution data set, using the program Refmac 5.1.24 from the CCP4 program suite. Manual rebuilding of the model was performed using the programs XtalView and Coot. Solvent molecules were introduced automatically using ARP. Data processing and refinement statistics are summarized in Table 10.7. Refinement results in R_{factor} and R_{free} values of 19.9% and 25.0%, respectively. The stereochemistry of the final model was analyzed with PROCHECK. The Ramachandran plot is of good quality;

Refinement	
Resolution range (Å)	70.3-2.0
Unique reflections, working/free	106056/5650
R_{factor} (%)	19.9
R_{free} (%)	25.0
Non-hydrogen atoms	12690
Water molecules	1196
r.m.s.d. from ideal: Bond lengths (Å)	0.010
r.m.s.d. from ideal: Bond angles (∘)	1.751

Table 10.7: Summary of refinement statistics of SoL structure.

there are 1234 non-glycine and non-proline residues; among these, 1038 (84.1%) are in the most favored regions, 174 (14.1%) are in the additional allowed regions, 13 (1.1%) in the generously allowed regions and 9 (0.7%) are in the disallowed regions.

Sequence

The sequence was obtained from the X-ray structure, and then confirmed by the sequence analysis performed by the group of L. Golovleva.

```
VQIGPVTDLH IVNADIVPDG FVRPAVNAGG TFPGPVIAGN VGDNFQIVTF
NQLIECSMLV HTSIHWHGEF QKGTNWADGP AFITQCPIIV GNSFSYNFNV
PGMAGTYWYH SHLTTQYCDG LRGPFVVYDP NDPDANLYDV DDDSTIITLA
DWYHVLAKEM GAGGAITADS TLIDGLGRTH VNVAAVPLSV ITVEVGKRYR
MRLVSISCDP NYDFSIDGHD MTIIETDGVD SQELTVDEIQ IFAAQRYSFV
LNANQPVGNY WIRANPNSGG EGFDGGINSA ILRYDGATSA DPVTVASTVH
TKCLIETDLH PLSSNGVPGN PHQGGADCNL NLSLGFACGN FVINGVSFTP
PTVPVLLQIC SGANTHADLL PSGSVISLPS NSSIEIALPA GAAGGPHPFH
LHGHDFAVSE SASNSTSNYD DPIWRDVVSI GGVGDNVTIR FCTDNPGPWF
LHCHIDWHLD AGFAIVFAED IPNTASANPV PEAWSNLCPS YDSAH
```

Figure 10.1: SoL sequence obtained from X-ray structure.

10.3 Composition of datasets at increasing X-ray dose

We collected ten complete datasets, each on a different crystal of SoL, at the X12 beamline at the DORIS storage ring, Hamburg, Germany. Table 10.8 summarize the data collection statistics obtained from one of these complete datasets; the statistics are similar to those obtained from the other datasets (data not shown).

Data collection	
Beamline	X12, Desy, Hamburg
Space group	P222
Unit cell parameters a; b; c (Å)	73.50; 140.14; 173.16
V_M (Å^3)	2.49
Solvent content (%) (v/v)	50.63
Wavelength (Å)	1.0028
Limiting resolution (Å)	73.72-2.00 (2.11-2.00)
Unique reflections	99340
R_{symm}	13.2(48.4)
Multiplicity	3.9(4.0)
Completeness overall (%)	92.2(89.4)
I/σ(I)	9.3(2.7)

Table 10.8: Summary of data collection statistics of one of SoL datasets used in the redox titration; values in parentheses are for the highest resolution shell.

On the basis of the Uv-Vis spectra registered during X-ray exposure at beamline XID14-2, ESFR (Grenoble, France) every data collection was composed by 100 images, obtained with an exposure time of 60 seconds; each complete data collection is obtained with an exposure of 100 minutes, that correspond to a dose of 4.8x10^5 grays: at this dose we observed a complete reduction of the enzyme in the Uv-Vis spectra. Each data collection was splitted in ten segment of ten images, then we merged together images from 1 to 10 of each data collection to obtain the low dose composite dataset, images from 11 to 20 for the following composite dataset and so on, until the high dose dataset was composed with images from 91 to 100 of each data collection. Here we report the statistics of the structures obtained for six datasets, since the other structures correspond to higher doses and thus show a complete reduction of the enzyme as observed for dataset #6.

Data Collection	Data set1	Data set2	Data set3	Data set4	Data set5	Data set6
Dose (grays)	3.4x10^4	6.8x10^4	1.4x10^5	2.0x10^5	2.7x10^5	4.8x10^5
Limiting resolution (Å)	73.54-2.00(2.05-2.00)					
Unique reflections	109781	108106	114646	110844	108119	108442
R_{symm}	15.5 (46.7)	15.6 (47.5)	15.2 (40.3)	14.8 (39.1)	14.9 (37.4)	15.1 (39.8)
Multiplicity	3.1(3.0)	3.3(3.2)	3.0(2.9)	3.1(3.0)	3.0(2.9)	3.2(3.0)
Completeness overall (%)	94.0 (90.9)	94.3 (93.9)	94.3 (92.9)	95.8 (93.4)	93.0 (91.9)	94.6 (90.9)
I/σ(I)	6.12 (2.13)	6.26 (2.25)	6.33 (2.34)	6.33 (2.54)	6.21 (2.41)	6.11 (2.55)

Table 10.9: Summary of data collection statistics of the composite datasets; values in parentheses are for the highest resolution shell.

Refinement	Data set1	Data set2	Data set3	Data set4	Data set5	Data set6
Dose (grays)	3.4x10^4	6.8x10^4	1.4x10^5	2.0x10^5	2.7x10^5	4.8x10^5
Resolution range (Å)	20.00-2.00					
Unique reflections working/free	106677 /5334	108601 /5430	108390 /5419	105497 /5275	106841 /5342	107542 /5373
R_{factor} (%)	18.9	18.7	18.6	18.5	18.6	19.0
R_{free} (%)	25.3	24.3	23.4	24.1	24.1	24.7
Non-hydrogen atoms	12288	12291	12301	12278	12279	12272
Water molecules	1097	1106	1111	1093	1092	1084
r.m.s.d.: bond lengths (Å)	0.019	0.019	0.018	0.018	0.018	0.019
r.m.s.d.: bond angles (°)	1.977	1.937	1.908	1.890	1.932	1.984
B factor (Å^2)	18.2	17.3	16.9	17.4	17.0	18.5

Table 10.10: Summary of refinement statistics for the structures obtained from the composite datasets.

Ramachandran Plot	Data set1	Data set2	Data set3	Data set4	Data set5	Data set6
Dose (grays)	$3.4x10^4$	$6.8x10^4$	$1.4x10^5$	$2.0x10^5$	$2.7x10^5$	$4.8x10^5$
Most favoured region (%)	1035 (83.9%)	1035 (83.9%)	1045 (84.8%)	1040 (84.3%)	1038 (84.2%)	1040 (84.3%)
Additionally allowed region (%)	176 (14.3%)	178 (14.4%)	167 (13.5%)	173 (14.0%)	176 (17.3%)	173 (14.0%)
Generously allowed region (%)	11 (0.9%)	7 (0.6%)	10 (0.8%)	8 (0.6%)	8 (0.6%)	8 (0.6%)
Disallowed region (%)	11 (0.9%)	13 (1.1%)	11 (0.9%)	12 (1.0%)	11 (0.9%)	12 (1.0%)

Table 10.11: Structures from composite datasets: summary of Ramachandran Plot statistics.

Part IV

Conclusions

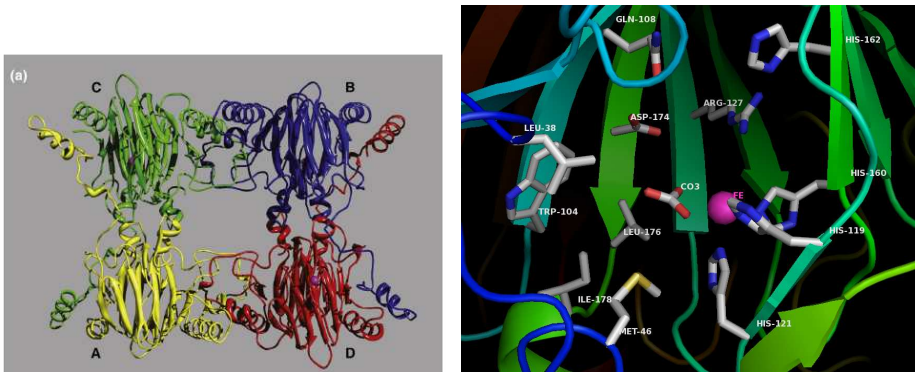
11

Conclusions

Bioremediation, the use of microbial processes to destroy wastes or recalcitrant pollutants, offers significant advantages with respect to chemical and physical treatments: the process is more sustainable and often the cost is lower. Many species of bacteria and fungi are involved in conversion of pollutants through biological processes, since they possess suitable enzymatic systems. Among these systems, bacterial oxygenases and fungal oxidases are key enzymes in biodegradation of toxic pollutants: bacterial oxygenases (such as dioxygenases) catalyze the cleavage of aromatic substituted rings and fungal oxidases (such as laccases) oxidize a broad range of substrates such as polyphenols, methoxy-substituted phenols and aromatic amines. These reactions are involved in the metabolic degradation of recalcitrant substituted aromatic pollutants, like polychlorinated biphenyls or industrial dyes, and inert biopolymeric wastes, such us lignin in the paper industry. Moreover, these oxidative enzymatic systems are involved in developing new and more sustainable bioconversion processes, a key trend in green chemistry. Redox biocatalysts are highly desirable because of the selectivity, controllability and economy of their reactions, in comparison with conventional chemical reactions. Moreover, biocatalyst-based processes require less energy and minimize the amount of waste produced, whilst at the same time being able to improve the quality and functional specifications of products.

In this work we present a study through protein X-ray crystallography on the enzyme mechanism of a Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans* BN12 and two Laccases from different sources, the white-rot fungi *Trametes trogii* and *Steccherinum ochraceum*. The isolation of adducts with substrates of Salicylate 1,2-dioxygenase from *Pseudaminobacter salicylatoxidans* and its mutant variants allowed us to clarify the molecular basis for the substrate specificity and the catalytic mechanism of the enzyme. The characterization of different redox states of a Laccase from *Trametes trogii* allowed us to design a multi-crystal data

collection strategy that lead to the isolation of a key reaction step of a Laccase from *Steccherinum ochraceum*.



(a) (b)

Figure 11.1: a) The tetrameric structure of Salicylate 1,2-Dioxygenase from *Pseudaminobacter salicylatoxidans* BN12; the iron ions are represented as magenta spheres; b) the active site.

The salicylate dioxygenase activity from *Pseudaminobacter salicylatoxidans* BN12 is rather unique among the currently known ring fission dioxygenases because the enzyme is able to cleave various substituted salicylates that carry only a single hydroxyl group and that are not activated by additional electron-donating substituents for a ring fission reaction, while in most cases the cleavage of the aromatic nucleus by bacteria requires two hydroxyl groups. It is able to catalyze the ring opening of gentisate, as well as salicylate and of a variety of substituted salicylates containing only a single hydroxyl group on carbon-2 and no additional electron-donating substituent (1-hydroxy-2-naphthoate). The substrates are cleaved between the carbon atoms carrying the carboxy group and the hydroxyl group.

The structure of this peculiar Salicylate 1,2-Dioxygenase was solved and we pointed out several residues, whose side chains define the binding cavity of the substrate, that could be responsible of the substrate specificity of this enzyme. Several mutations were introduced in this enzyme and we evidenced few mutants that could be useful in understanding more closely the relationship between the mechanistic capabilities of this ring fission dioxygenase and its structural features: the mutant G106A, that converts gentisate (with a slightly increased affinity), but does not convert salicylate, the mutant A85H, that does not convert salicylate and gentisate but converts 1-hydroxy-2-naphtoate (thus its activity is similar to that of a 1-hydroxy-2-naphtoate dioxygenase) and the mutant W104Y, whose activity is decreased compared to the wild type but still converts salicylate, gentisate and 1-hydroxy-2-naphtoate.

	k_{cat}/K_M $(s^{-1}\ M^{-1}) \times 10^{-3}$			
	WT	A85H	W104Y	G106A
Salicylate	50.06	0.02	7.20	n.d.
Gentisate	800.42	n.d.	219.4	818.68
1-OH-2-naphtoate	168.08	41.52	31.63	n.d.

Table 11.1: Activities of SDO wild-type and mutant variants (n.d. not determined).

We solved the structures of the three mutants, and performed soaking experiments to obtain complexes of the wild type enzyme and of the mutants, each one with three substrates: gentisate, salicylate and 1-hydroxy-2-naphtoate; the soaking was performed in a glove box in order to avoid the presence of oxygen and thus freeze the enzyme reaction in its early step of substrate binding. The structures obtained allow us to shed some light on the enzyme substrate specificity and thus understanding the peculiar activity toward salicylate of this dioxygenase; comparing the structures of the wild type enzyme to that of the G106A mutant both in complex with salicylate, we found out how a mutation that involve the addition of a single methyl group in the side chain of a residue in a region far from the site where the substrate actually binds can in fact change dramatically the activity and the specificity of the enzyme. Moreover we obtained the G106A mutant structure in complex with both gentisate and dioxygen and in complex with 2-oxo-5-hydroxyepta-3,5-dien-1,7-dioic acid, the product of the cleavage of gentisate. The isolation of different steps of the catalytic reaction, that are the binding of substrate, the consequent bindign of dioxgen and the release of the product, allowed to understand more closely the enzyme mechanism, revealing the role of residues in the active cavity and of conformational changes due to substrate binding.

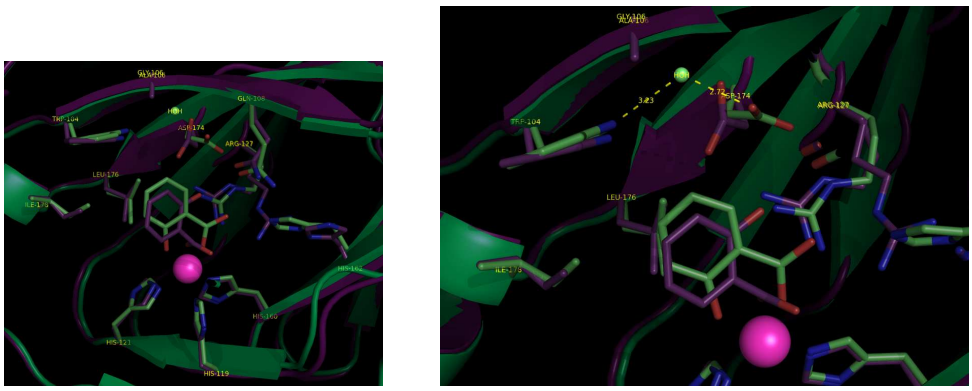
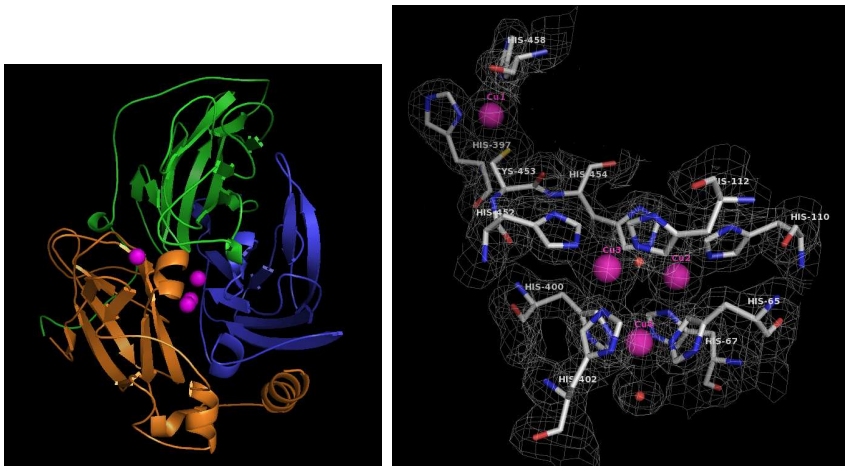


Figure 11.2: Binding of salicylate: superposition of G106A mutant complex (purple) and WT complex (green).

Laccases are blue-copper proteins distributed in plants, fungi and bacteria, that belong to the multicopper protein family, that includes also ascorbate oxidase, ceruloplasmin and bilirubin oxidase. Laccases use molecular oxygen to oxidize various aromatic and non aromatic compounds by a radical catalyzed reaction mechanism; the reduction of dioxygen to water in the trinuclear copper cluster of the enzyme is accompanied by the one-electron oxidation of a substrate, typically a p-hydroxy phenol or other phenolic compounds. The biotechnological application of laccases, aiming at the development of various industrial oxidative processes going from the production of fine chemicals and pharmaceuticals to bioremediation of contaminated soil and water, would receive great impulse to improvements from the complete comprehension of the catalytic mechanism of multicopper oxidases, and in particular of their redox potential and substrate selectivity control and a detailed characterization of the high resolution molecular structure of such enzymes will surely help in achieving such aims.

The first step in the study of the catalityc mechanism was the observation of different redox states in the structures of a Laccase from *Trametes trogii*, depending on the X-ray source used in the data collection: for a data set collected with a Cu-Kα sealed tube, the copper ions in the trinuclear cluster, responsible of dioxygen reduction, appear to be in an oxidated state, since in the structure the distances between them are shorter than 4 Å; instead the structure obtained from a data set collected with a synchrotron source (about two orders of magnitude more intense than the sealed tube) showed distances between thethree copper ions longer than 4 Å. We confirmed that the metal centers were reduced by synchrotron radiation by registering a set of Uv-Vis spectra *in cristallo* during the X-ray exposure on crystals of Laccase from *Steccherinum ochraceum*.

This auto-reduction could seem an inconvenience at first glance, but actually we took advantage of it, since the reduction in the copper cluster is the first step in the reduction of dioxygen to water, and thus the exposure to X-ray could be use to simulate what really happens in the enzyme, without using substrates. We obtained several structures of Laccase from *Steccherinum ochraceum* at increasing doses of sinchrotron radiation: the structures clearly show a progressive reduction of copper ions in the protein, their mutual distances increasing gradually; moreover we isolated different steps of the reduction of dioxygen to water: the binding of the dioxygen molecule and the reduction to peroxide and the further reduction with the splitting of the two oxygen atoms. These structures could be regarded as a three-dimensional movie of the X-ray-driven catalytic conversion of the bound dioxygen species in the active site of laccase. The pictures rising from the present results allow the completion of the structural depiction for an accurate mechanism of dioxygen reduction catalyzed by multicopper oxidases providing general insights into the reductive cleavage of the O-O bonds, a leading problem in many areas of biology.



(a) (b)

Figure 11.3: a) Schematic representation of the overall structure of Laccase from *Trametes trogii*: domain I is depicted in blu, domain II in green and domain III in yellow (the copper ions are depicted as magenta spheres); b) T2/T3 copper cluster and T1 copper ion in Laccase from *Steccherinum ochraceum*.

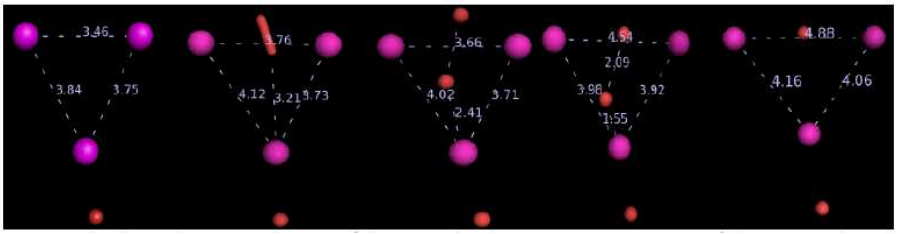


Figure 11.4: The three-dimensional movie of the X-ray-driven enzymatic conversion of dioxygen in the active site of Laccase: copper ions are depicted in pink and oxygen atoms in red.

Bibliography

[1] K. Furukawa. Oxygenases and dehalogenases: molecular approaches to efficient degradation of environmental pollutants. *Biosci. Biotechnol. Biochem.*, 70:2335–2348, 2006.

[2] T. Bosma; H. Harms; A. Zehnder. Biodegradation of xenobiotics in environment and technosphere. In *Biodegradation and Persistance*, pages 163–202. Springer Verlag, Heidelberg/Berlin, 2001.

[3] W. Reineke. Aerobic and anaerobic biodegradation potentials of microorganisms. In *Biodegradation and Persistance*, pages 1–161. Springer Verlag, Heidelberg/Berlin, 2001.

[4] D. T. Gibson; R. E. Parales. Aromatic hydrocarbon dioxygenases in environmental biotechnology. *Curr. Opin. Biotechnol.*, 11:236–243, 2000.

[5] C. H. Chua; Y. Feng; C. C. Yeo; H. E. Khoo; C. L. Poh. Identification of amino acid residues essential for catalytic activity of gentisate 1,2-dioxygenase from pseudomonas alcaligenes ncib 9867. *FEMS Microbiol. Lett.*, 204:141–146, 2001.

[6] F. H. Vaillancourt; J. T. Bolin; L. D. Eltis. The ins and outs of ring-cleaving dioxygenases. *Crit. Rev. Biochem. Mol. Biol.*, 41:241–267, 2006.

[7] M. R. Harpel; J. D. Lipscomb. Gentisate 1,2-dioxygenase from pseudomonas. substrate coordination to active site fe^{2+} and mechanism of turnover. *J. Biol. Chem.*, 265:22187–22196, 1990.

[8] S. R. Couto; J. L. T. Herrera. Industrial and biotechnological applications of laccases: A review. *Biotechnology Advances*, 24:500–513, 2006.

[9] K. Murugesan; P. T. Kalaichelvan. Synthetic dye decolourization by white rot fungi. *Indian J. Exp. Biol.*, 41:1076–1087, 2003.

[10] A. M. Mayer; R. C. Staples. Laccase: new functions for an old enzyme. *Phytochemistry*, 60:551–565, 2002.

[11] S. Riva. Laccases: blue enzymes for green chemistry. *TRENDS in Biotechnology*, 24:219–226, 2006.

[12] H. M. Berman; J. Westbrook; Z. Feng; G. Gilliland; T. N. Bhat; H. Weissig; I. N. Shindyalov; P. E. Bourne. The protein data bank. *Nucleic Acid Res.*, 28:235–242, 2000.

[13] D. Bourgeois; A. Royant. Advances in kinetic protein crystallography. *Curr. Op. Struct. Biol.*, 15:538–547, 2005.

[14] S. Adachi; S. Y. Park; J. R. Tame; Y. Shiro; N. Shibayama. Direct observation of photolysis-induced tertiary structural changes in hemoglobin. *Proc. Natl. Acad. Sci. USA*, 100:7039–7044, 2003.

[15] M. Sugishima; H. Sakamoto; M. Noguchi; K. Fukuyama. Co-trapping site in heme oxygenase revealed by photolysis of its co-bound heme complex: mechanism of escaping from product inhibition. *J. Mol. Biol.*, 341:7–13, 2004.

[16] K. Edman; P. Nollert; A. Royant; H. Belrhali; E. Pebay-Peyroula; J. Hajdu; R. Neutze; E.M. Landau. High-resolution x-ray structure of an early intermediate in the bacteriorhodopsin photocycle. *Nature*, 401:822–826, 1999.

[17] T. Ursby; M. Weik; E. Fioravanti; M. Delarue; M. Goeldner; D. Bourgeois. Cryophotolysis of caged compounds: a technique for trapping intermediate states in protein crystals. *Acta Crystallogr D Biol Crystallogr*, 58:607–614, 2002.

[18] S. Anderson; V. Srajer; K. Moffat. Structural heterogeneity of cryotrapped intermediates in the bacterial blue light photoreceptor, photoactive yellow protein. *Photochem. Photobiol.*, 80:7–14, 2004.

[19] G. K. Farber. Crystallographic analysis of solvent-trapped intermediates of chymotrypsin. *Methods Enzymol.*, 308:201–215, 1999.

[20] N. H. Yennawar; H. P. Yennawar; G. K. Farber. X-ray crystal structure of gamma-chymotrypsin in hexane. *Biochemistry*, 33:7326–7336, 1994.

[21] J. L. Schmitke; L. J. Stern; A. M. Klibanov. Comparison of x-ray crystal structures of an acyl-enzyme intermediate of subtilisin carlsberg formed in anhydrous acetonitrile and in water. *Proc. Natl. Acad. Sci. USA*, 95:12918–12923, 1998.

[22] A. Lavie; K. N. Allen; G. A. Petsko; D. Ringe. X-ray crystallographic structures of d-xylose isomerase-substrate complexes position the substrate and provide evidence for metal movement during catalysis. *Biochemistry*, 33:5469–5480, 1994.

[23] T. R. Schneider; E. Gerhardt; M. Lee; P. H. Liang; K. S. Anderson; I. Schlichting. Loop closure and intersubunit communication in tryptophan synthase. *Biochemistry*, 37:5394–5406, 1998.

[24] K. H. Choi; A. S. Mazurkie; A. J. Morri; D. Utheza; D. R. Tolan; K. N. Allen. Structure of a fructose-1,6-bis(phosphate) aldolase liganded to its natural substrate in a cleavage-defective mutant at 2.3 å. *Biochemistry*, 38:12655–12664, 1999.

[25] E. J. van Asselt; A. M. Thunnissen; B. W. Dijkstra. High resolution crystal structures of the escherichia coli lytic transglycosylase slt70 and its complex with a peptidoglycan fragment. *J. Mol. Biol.*, 291:877–898, 1999.

[26] G. Sulzenbacher; L. F. Mackenzie; K. S. Wilson; S. G. Withers; C. Dupont; G. J. Davies. The crystal structure of a 2-fluorocellotriosyl complex of the streptomyces lividans endoglucanase celb2 at 1.2 åresolution. *Biochemistry*, 38:4826–4833, 1999.

[27] G. I. Berglund; G. H. Carlsson; A. T. Smith; H. Szoke; A. Henriksen; J. Hajdu. The catalytic pathway of horseradish peroxidase at high resolution. *Nature*, 417:463–468, 2002.

[28] V. Adam; A. Royant; V. Niviere; F. P. Molina-Heredia; D. Bourgeois. Structure of superoxide reductase bound to ferrocyanide and active site expansion upon x-ray-induced photo-reduction. *Structure*, 12:1729–1740, 2004.

[29] T. Muraki; M. Taki; Y. Hasegawa; H. Iwaki; P. C. Lau. Prokaryotic homologs of the eukaryotic 3-hydroxyanthranilate 3,4-dioxygenase and 2-amino-3-carboxymuconate-6-semialdehyde decarboxylase in the 2-nitrobenzoate degradation pathway of pseudomonas fluorescens strain ku-7. *Appl. Environ. Microbiol.*, 69:1564–1572, 2003.

[30] H. J. Knackmuss A. Stolz; B. Nortemann. Bacterial metabolism of 5-aminosalicylic acid. initial ring cleavage. *Biochem. J.*, 282:675–680, 1992.

[31] U. Lendenmann; J. C. Spain. 2-aminophenol 1,6-dioxygenase: a novel aromatic ring cleavage enzyme purified from pseudomonas pseudoalcaligenes js45. *J. Bacteriol.*, 178:6227–6232, 1996.

[32] S. Takenaka; S. Murakami; R. Shinke; K. Hatakeyama; H. Yukawa; K. Aoki. Novel genes encoding 2-aminophenol 1,6-dioxygenase from pseudomonas species ap-3 growing on 2-aminophenol and catalytic properties of the purified enzyme. *J. Biol. Chem.*, 272:14727–14732, 1997.

[33] R. L. Crawford; P. E. Olson; T. D. Frick. Catabolism of 5-chlorosalicylate by a bacillus isolated from the mississippi river. *Appl. Environ. Microbiol.*, 38:379–384, 1979.

[34] K. Adachi; T. Iwabuchi; H. Sano; S. Harayama. Structure of the ring cleavage product of 1-hydroxy-2-naphthoate, an intermediate of the phenanthrene-degradative pathway of nocardioides sp. strain kp7. *J. Bacteriol.*, 181:757–763, 1999.

[35] J. P. Hintner; C. Lechner; U. Riegert; A. E. Kuhm; T. Storm; T. Reemtsma; A. Stolz. Direct ring fission of salicylate by a salicylate 1,2-dioxygenase activity from pseudaminobacter salicylatoxidans. *J. Bacteriol.*, 183:6936–6942, 2001.

[36] T. Iwabuchi; S. Harayama. Biochemical and molecular characterization of 1-hydroxy-2-naphthoate dioxygenase from nocardioides sp. kp7. *J. Biol. Chem.*, 273:8332–8336, 1998.

[37] J. P. Hintner; T. Reemtsma; A. Stolz. Biochemical and molecular characterization of a ring fission dioxygenase with the ability to oxidize (substituted) salicylate(s) from pseudaminobacter salicylatoxidans. *J. Biol. Chem.*, 279:37250–37260, 2004.

[38] M. A. Adams; V. K. Singh; B. O. Keller; Z. Jia. Structural and biochemical characterization of gentisate 1,2-dioxygenase from escherichia coli o157:h7. *Mol. Microbiol.*, 61:1469–1484, 2006.

[39] J. M. Dunwell; A. Purvis; S. Khuri. Cupins: the most functionally diverse protein superfamily? *Phytochemistry*, 65:7–17, 2004.

[40] L. D. Eltis; J. T. Bolin. Evolutionary relationships among extradiol dioxygenases. *J. Bacteriol.*, 178:5930–5937, 1996.

[41] J. M. Dunwell; A. Culham; C. E. Carter; C. R. Sosa-Aguirre; P. W. Goodenough. Evolution of functional diversity in the cupin superfamily. *Trends Biochem. Sci.*, 26:740–746, 2001.

[42] G. P. Titus; H. A. Mueller; J. Burgner; C. S. De Rodriguez; M. A. Penalva; D. E. Timm. Crystal structure of human homogentisate dioxygenase. *Nature Struct. Biol.*, 7:542–546, 2000.

[43] I. Matera; M. Ferraroni; S. Burger; A. Scozzafava; A. Stolz; F. Briganti. Salicylate 1,2-dioxygenase from pseudaminobacter salicylatoxidans: Crystal structure of a peculiar ring-cleaving dioxygenase. *J. Mol. Biol.*, 380:856–868, 2008.

[44] K. D. Koehntop; J. P. Emerson; L. Que. The 2-his-1-carboxylate facial triad: a versatile platform for dioxygen activation by mononuclear non-heme iron(ii) enzymes. *J. Biol. Inorg. Chem.*, 10:87–93, 2005.

[45] Y. Zhang; K. L. Colabroy; T. P. Begley; S. E. Ealick. Structural studies on 3-hydroxyanthranilate-3,4-dioxygenase: the catalytic mechanism of a complex oxidation involved in nad biosynthesis. *Biochemistry*, 44:7632–7643, 2005.

[46] F. Fusetti; K. H. Schroter; R. A. Steiner; P. I. van Noort; T. Pijning; H. J. Rozeboom. Crystal structure of the copper-containing quercetin 2,3-dioxygenase from aspergillus japonicus. *Structure. (Camb.)*, 10:259–268, 2002.

[47] B. Gopal; L. L. Madan; S. F. Betz; A. A. Kossiakoff. The crystal structure of a quercetin 2,3-dioxygenase from bacillus subtilis suggests modulation of enzyme activity by a change in the metal ion at the active site(s). *Biochemistry*, 44:193–201, 2005.

[48] C. R. Simmons; Q. Liu; Q. Huang; Q. Hao; T. P. Begley; P. A. Karplus; M. H. Stipanuk. Crystal structure of mammalian cysteine dioxygenase: a novel mononuclear iron center for cysteine thiol oxidation. *J. Biol. Chem.*, 281:18723–18733, 2006.

[49] D. H. Ohlendorf M. W. Vetting; L. P. Wackett; L. Que Jr, J. D. Lipscomb. Crystallographic comparison of manganese- and iron-dependent homoprotocatechuate 2,3-dioxygenases. *J. Bacteriol.*, 186:1945–1958, 2004.

[50] A. Gibello; E. Ferrer; M. Martin; A. Garrido-Pertierra. 3,4-dihydroxyphenylacetate 2,3-dioxygenase from klebsiella pneumoniae, a mg-containing dioxygenase involved in aromatic catabolism. *Biochem. J.*, 301:145–150, 1994.

[51] E. G. Kovaleva; M. B. Neibergall; S. Chakrabarty; J. D. Lipscomb. Finding intermediates in the o_2 activation pathways of non-heme iron oxygenases. *Acc. Chem. Res.*, 40:475–483, 2007.

[52] A. Vagin; A. Teplyakov. Molrep: an automated program for molecular replacement. *J. Appl. Crystallogr.*, 30:1022–1025, 1997.

[53] W. Kabsch. Automatic processing of rotation diffraction data from crystals of initially unknown symmetry and cell constants. *J. Appl. Cryst.*, 26:795–800, 1993.

[54] NUMBER 4 COLLABORATIVE COMPUTATIONAL PROJECT. The ccp4 suite: Programs for protein crystallography. *Acta Cryst. D*, 50:760–763, 1994.

[55] P. R. Evans. Scaling and assessment of data quality. *Acta Cryst.*, D62:71–82, 2005.

[56] A. Vagin; A. Teplyakov. Molrep: an automated program for molecular replacement. *J. Appl. Crystallogr.*, 30:1022–1025, 1997.

[57] G. N. Murshudov; A. Vagin; E. J. Dodson. Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallogr. D*, 53:240–255, 1997.

[58] A. Perrakis; R. Morris; V. S. Lamzin. Automated protein model building combined with iterative structure refinement. *Nature Struct. Biol.*, 6:458–463, 1999.

[59] P. Emsley; K. Cowtan. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, 60:2126–2132, 2004.

[60] R. A. Laskowski; M. W. MacArthur; D. S. Moss; J. M. Thornton. Procheck: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.*, 26:283–291, 1993.

[61] J. D. Thompson; D. G. Higgins; T. J. Gibson. Clustal w: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, 22:4673–4680, 1994.

[62] I. Matera; M. Ferraroni; S. Burger; A. Stolz; F. Briganti. Preliminary crystallographic analysis of salicylate 1,2-dioxygenase from pseudaminobacter salicylatoxidans. *Acta Cryst. F*, 62:553–555, 2006.

[63] H. Yoshida. Chemistry of laquer (hurushi) part 1. *J. Chem. Soc. Tokyo*, 43:472–486, 1883.

[64] H. Claus. Laccases: structure, reactions, distribution. *Micron*, 35:93–96, 2004.

[65] K. Nakamura; N. Go. Function and molecular evolution of multicopper blue proteins. *Cell. Mol. Life Sci.*, 62:2050–2066, 2005.

[66] M. A. McGuirl; D. M. Dooley. Copper-containing oxidases. *Curr. Opin. Chem. Biol.*, 3:138–144, 1999.

[67] A. Leonowicz; A. Matuszewska; J. Luterek; D. Ziegenhagen; M. Wojtas-Wasilewska; N. S. Cho; M. Hofrichter; J. Rogalski. Biodegradation of lignin by white rot fungi. *Fungal Genetics and Biology*, 27:175–185, 1999.

[68] C. S. Evans. Laccase activity in lignin degradation by *Coriolus versicolor*. *FEMS Microbiol. Lett.*, 27:339–343, 1985.

[69] K. Li; F. Xu; K. E. Eriksson. Comparison of fungal laccases and redox mediators in oxidation of a nonphenolic lignin model compound. *Appl. Environ. Microbiol.*, 65:2654–2660, 1999.

[70] C. Johannes; A. Majcherczyk. Natural mediators in the oxidation of polycyclic aromatic hydrocarbons by laccase mediator systems. *Appl. Environ. Microbiol.*, 66:524–528, 2000.

[71] R. Bourbonnais; M. G. Paice. Oxidation of non-phenolic substrates. an expanded role for laccase in lignin biodegradation. *FEBS Lett.*, 267:99–102, 1990.

[72] K. Murugesan. Bioremediation of paper and pulp mill effluents. *Indian J. Exp. Biol.*, 41:1239–1248, 2003.

[73] S. D. Aust; J. T. Benson. The fungus among us: use of white rot fungi to biodegrade environmental pollutants. *Environ. Health Perspect*, 101:232–233, 1993.

[74] S. V. S. Kumar; P. S. Phale; S. Durani; P. P. Wangikar. Combined sequence and structure analysis of the fungal laccase family. *Biotechnol. Bioeng.*, 83:386–394, 2003.

[75] V. Ducros; A. M. Brzozowski; K. S. Wilson; S. H. Brown; P. Ostergaard; P. Schneider; D. S. Yaver; A. H. Pedersen; G. J. Davies. Crystal structure of the type-2 cu depleted laccase from *Coprinus cinereus* at 2.2 åresolution. *Nat. Struct. Biol.*, 5:310–316, 1998.

[76] T. Bertrand; C. Jolivalt; P. Briozzo; E. Caminade; N. Joly; C. Madzak; C. Mougin. Crystal structure of a four-copper laccase complexed with an arylamine: insights into substrate recognition and correlation with kinetics. *Biochemistry*, 41:7325–7333, 2002.

[77] N. Hakulinen; L. L. Kiiskinen; K. Kruus; M. Saloheimo; A. Paananen; A. Koivula; J. Rouvinen. Crystal structure of a laccase from *Melanocarpus albomyces* with an intact trinuclear copper site. *Nat. Struct. Biol.*, 9:601–605, 2002.

[78] S. Garavaglia; M. T. Cambria; M. Miglio; S. Ragusa; V. Iacobazzi; F. Palmieri; C. D'Ambrosio; A. Scaloni; M. Rizzi. The structure of *Rigidoporus lignosus* laccase containing a full complement of copper ions reveals an asymmetrical arrangement for the t3 copper pair. *J. Mol. Biol.*, 342:1519–1531, 2004.

[79] A. V. Lyashenko; I. Bento; V. N. Zaitsev; N. E. Zhukhlistova; Y. N. Zhukova; A. G. Gabdoulkhakov; E. Y. Morgunova; W. Voelter; G. S. Kachalova; E. V. Stepanova; O. G. Koroleva; V. S. Lamzin; V. I. Tishkov; C. Betzel; P. F. Lindley; A. B. Mikhailov. X-ray structural studies of the fungal laccase from *Cerrena maxima*. *J. Biol. Inorg. Chem.*, 2006.

[80] F. J. Enguita; D. Marcal; L. O. Martins; R. Grenha; A. O. Henriques; P. F. Lindley; M. A. Carrondo. Substrate and dioxygen binding to the endospore coat laccase from *Bacillus subtilis*. *J. Biol. Chem.*, 279:23472–23476, 2004.

[81] E. I. Solomon; U. M. Sundaram; T. E. Machonkin. Multicopper oxidases and oxygenases. *Chem. Rev.*, 96:2563–2606, 1996.

[82] E. I. Solomon; P. Chen; M. Metz; S. K. Lee; A. E. Palmer. Oxygen binding, activation, and reduction to water by copper proteins. *Angew. Chem. Int. Ed. Engl.*, 40:4570–4590, 2001.

[83] B. Reinhammar. Kinetic studies on polyporus and tree laccases. In A. Messerschmidt, editor, *Multi-Copper Oxidases*, pages 167–200. World Scientific Publishing, Singapore, 1997.

[84] M. Ferraroni; N. M. Myasoedova; V. Schmatchenko; A. A. Leontievsky; L. A. Golovleva; A. Scozzafava; F. Briganti. Crystal structure of a blue laccase from *Lentinus tigrinus*: evidences for intermediates in the molecular oxygen reductive splitting by multicopper oxidases. *BMC Structural Biology 2007*, 7:60, 2007.

[85] A. Messerschmidt. *Multi-Copper Oxidases*. World Scientic Publishing, Singapore, 1997.

[86] P. Durao; I. Bento; A.T. Fernandes; E.P. Melo; P.F. Lindley; L.O. Martins. Perturbations of the t1 copper site in the cota laccase from *Bacillus subtilis*: structural, biochemical, enzymatic and stability studies. *J. Biol. Inorg. Chem.*, 11:514–526, 2006.

[87] F. Xu; R.M. Berka; J.A. Wahleithner; B.A. Nelson; J.R. Shuster; S.H. Brown; A.E. Palmer; E.I. Solomon. Site-directed mutations in fungal laccase: effect on redox potential, activity and ph profile. *Biochem. J.*, 334:63–70, 1998.

[88] N. Hakulinen; K. Kruus; A. Koivula; J. Rouvinen. A crystallographic and spectroscopic study on the effect of x-ray radiation on the crystal structure of *Melanocarpus albomyces* laccase. *Biochem. Biophys. Res. Commun.*, 350:929–934, 2006.

[89] J. W. Murray; E. F. Garman; R. B. G. Ravelli. X-ray absorption by macromolecular crystals: the effects of wavelenght and crystal composition on absorbed dose. *Appl. Cryst.*, 37:513–522, 2004.

[90] I. Matera; A. Gullotto; S. Tilli; M. Ferraroni; A. Scozzafava; F. Briganti. Crystal structure of the blue multicopper oxidase from the white-rot fungus *Trametes trogii* complexed with p-toluate. *Inorganica Chimica Acta*, 361:4129–4137, 2008.

[91] A. Chernykh; N. Myasoedova; M. Kolomytseva; M. Ferraroni; F. Briganti; A. Scozzafava; L. Golovleva. Laccase isoforms with unusual properties from the basidiomycete *Steccherinum ochraceum* strain 1833. *J. Appl. Microbiol.*, 105:2065–2075, 2008.