We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,800 Open access books available
116,000 International authors and editors
120M Downloads

154 Countries delivered to
TOP 1% Our authors are among the most cited scientists
12.2% Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
1. Introduction

The name for cupin proteins is derived from the Latin term for small barrel, ‘Cupa’. Proteins that belong to the group of cupins adopt a barrel-like structure (Dunwell et al., 2001). According to the database of Structural Classification of Proteins (SCOP) (Murzin et al., 1995), the cupin proteins have been classified as members of ‘RmlC-like Cupins’ superfamily in the ‘Double Stranded Beta Helix’ (DSBH) fold, however, the nomenclature employed in the literature is somewhat ambiguous since JmjC transcription factors (Clissold & Ponting, 2001) display many features typical of the DSBH fold. These common characteristics for the DSBH fold include a pair of four-stranded antiparallel $\beta$-sheets constituting up to eight $\beta$ strands which form the typical $\beta$-sandwich structure. The superfamily comprises of 20 families with members performing diverse functions ranging from enzymatic activities like dioxygenases, hydrolases, decarboxylases, epimerases and isomerases to non-enzymatic functions such as binding to auxin, seed storage, and nuclear transcription factors (Dunwell, 2001, 2004). The nature of substrates used in various enzymatic reactions differs in chemical types, size, and structure. The sequence identity is low among the members of this superfamily. The functional site of members of this superfamily is generally located at the center of a conserved barrel. The cupin domain usually consists of two sequence motifs, each corresponding to two $\beta$-strands. A less conserved region separates these motifs. The conserved motifs, $\text{GX}_{3}$H$\text{HX}_{5}$E$\text{X}_{6}$G and $\text{GX}_{3}$P$\text{GX}_{2}$H$\text{XX}_{6}$N, together contain the residues involved in metal ion binding at the active site, that is known to play a functional role (Dunwell, 2001, 2004). It has been indicated that 10,346 cupin sequences (Finn et al., 2008) have been identified in 843 species that belong to eukaryotes, prokaryotes, archaebacteria and viruses. In some plant species like $\text{O. sativa}$, $\text{V. vinifera}$ and $\text{A. thaliana}$ over 100 cupin sequences have been identified. This indicates the extent to which the cupins have diverged and duplicated in proteomes of various species to perform a variety of functions. Various metal ions, bound at the active site, including Iron, Manganese, Nickel, Copper, Zinc and Cadmium are known to play a functional role in the enzymatic members of cupin superfamily (Table 1). The metal cofactor can influence the chemistry of the catalytic reaction. The metal cofactor typically plays an important role in the function of cupins via an interaction with the substrate. An approach involving structure-based clustering of uncharacterized proteins within a group of proteins of known function can provide clues about their possible functions. It thus appears likely that this method would be a valuable tool for the functional annotation of structural genomic target proteins that are similar in structure despite the lack of sequence similarity.
Table 1. Various proteins possessing the cupin protein fold.
Fig. 1. Selected metal binding sites for cupin proteins. Gab protein (1JR7), Homogentisate 1,2-dioxygenase (1EY2), 3-Hydroxyanthranilate 3.4-dioxygenase (1YFY), Monnose-6-phosphate isomerase (3H1M), Hydroxypyropyl phosphonic acid epoxidase (2BNN), SyrB2 (2FCT), Tetracenomycin polyketide synthesis protein (3H50), Phosphomannose isomerase (1PMI), Transcriptional regulator (1Y9Q), Pirin (1J1L), Hypothetical protein yxaG (1Y3T), Acireductone dioxygenase (1VR3), 5-keto-4-deoxyuronate isomerase (1XRU), E. Coli. sugar isomerase (3KMH), Cysteine dioxygenase (2B5H), Gentiisate 1,2-dioxygenase (3BU7), Bacilysin biosynthesis protein BacB (3H7J), Novel manganese-containing cupin TM1459 (1VJ2), RemF protein (3HT2).
Although metal binding cupins were first noted to possess ligands consisting of 3-His-1-carboxylate, it is now apparent that variations in the consensus sequence consisting of either deletions or substitutions of these ligands can produce alternate metal binding sites (Table 1). Representative examples for metal binding sites are included in Fig. 1. This article will focus on structurally characterized cupin proteins possessing metal binding sites consisting of 2-His, 2-His-1-Gln, 2-His-1-Glu-1-X (X = Gln, Tyr), 3-His, 3-His-1-Glu, 3-His-1-Gln, and 4-His. Efforts to model the active sites for these cupin proteins using synthetic compounds will be discussed and areas which would benefit from future studies will be included.

2. 2-His

2.1 SyrB2 (PDB:2FCT)

The biosynthesis of halogenated products is common in many microorganisms (van Pée, 1996). While incorporation of halogens into aromatic and heteroaromatic moieties is usually catalyzed by FAD-dependent enzymes, the more challenging formation of aliphatic carbon-chlorine bonds is generally accomplished by α-ketoglutarate-dependent dioxygenases (Chang et al., 2002; Guenzi et al., 1998). SyrB2 belongs to this group of oxygenation enzymes (Vaillancourt et al., 2005). The gene encoding SyrB2 is part of the syringomycin synthetase gene cluster of *Pseudomonas syringae* pv. *syringae*. The protein catalyzes a halogenation step in the biosynthesis of the phytotoxic nonribosomal peptide syringomycin (Vaillancourt et al., 2005). Purified SyrB2 did not contain a bound metal. The active enzyme could be partly reconstituted from the apoprotein to levels of 0.4 and 0.9 g equivalents of FeII, depending on the presence of α-ketoglutarate (Blasiak et al., 2006). The physiological reaction performed by SyrB2 is the halogenations at C3 of a protein-linked L-threonine moiety specifically bound to the thiolation domain of another pathway-specific protein, called holo-SyrB1. Free L-threonine is not accepted as the substrate. Holo-SyrB1 binds L-threonine through a thioester linkage involving its phosphopantetheine prosthetic arm. In vitro studies show that SyrB2 rapidly deactivates during catalysis, such that no more than approximately seven turnovers of the enzyme were possible (Blasiak et al., 2006) so that in order to achieve significant conversion of the complex of holo-SyrB1 and L-threonine it was necessary to have SyrB2 present in almost equimolar amounts (Vaillancourt et al., 2006). The dependence of the enzymatic reaction on O2, α-ketoglutarate, and chloride has been established (Blasiak et al., 2006) though it has also been shown that SyrB2 can utilize bromide as a nonnatural reactant in place of chloride (Vaillancourt et al., 2006).

A proposed mechanism for SyrB2 is shown in Scheme 1 (Borowski et al., 2010). In this mechanism, the α-KG binds to the FeII center with its carboxylic and keto groups, while a water molecule occupies the position trans to the two His ligands. According to the scheme the catalytic process starts with A where Cl and α-KG are already bound to FeII. The water molecule is then displaced (A→B) to create an opening for coordination of dioxygen. The following steps involve dioxygen binding (B→C) and oxidative decarboxylation (C→D). The decarboxylated high-spin oxoferyl intermediate D was characterized by Mössbauer and EXAFS methods. This reactive species D is responsible for the C–H bond cleavage (D→E). Then the radical carbon is preferentially subjected to react with chloro ligand to produce the final product F. Since the final step involves a rebound of the OH group (E→F) afforded the hydroxylation product F'.
Synthetic model studies for SyrB2 have been carried out using sterically hindered α-ketocarboxylate 2,6-di(mesityl)benzoylformate (MesBF) with the Fe\(^{II}\) complexes LFeCl\(_2\) (L = \(N,N,N',N'\)-tetramethylpropylenediamine) yielded LFe(Cl)(MesBF). X-ray crystal structures of these complexes showed that they closely model the active site structure of SyrB2 (Friese et al., 2008). Computational studies have been carried out to model the reactivity of SyrB2. From these studies it was concluded that the hydrogen abstraction and radical chlorination steps are strongly coupled (Kulik et al., 2009).

### 3. 3-His

#### 3.1 Cysteine dioxygenase (PDB: 2B5H)

Cysteine Dioxygenase (CDO) initiates the catabolism of cysteine to pyruvate and sulfate, which is essential for the generation of adequate inorganic sulfate and allows pyruvate to enter central pathways of metabolism. Mammalian CDO was assigned to the cupin superfamily. The crystal structure of recombinant rat CDO was determined at 1.5 Å resolution. The active site coordination of CDO comprises a tetrahedrally coordinated Fe\(^{II}\) center involving three histidine nitrogens and a water molecule (Simmons et al., 2006). A structure at 1.75 Å was also reported containing a Ni\(^{II}\) coordinated in distorted octahedral geometry with three histidine nitrogens and three water molecules (McCoy et al., 2006). These two eukaryotic structures revealed the presence of a rare cysteinyl-tyrosine cross-link cofactor (Cys93 and Tyr157) while a structure from a prokaryote (\(R. eutropha\)) did not contain the same modification (PDB: 2GM6). In the prokaryotic structure, the iron adopts a pseudo-octahedral geometry with a sulfate ion hydrogen-bonded to the hydroxyl group of the conserved Tyr. The Cys93 is substituted by glycine, a deviation found in all known prokaryotic CDOs (A previous phylogenetic analysis has shown that the cross-linked cofactor is strictly conserved in all eukaryotes known to possess a CDO homolog (metazoan and fungi) but is absent from most bacterial CDO homologs) (Dominy et al., 2006). It is...
feasible that the eukaryotes, which have a very limited ability to synthesize cysteine de novo and are subject to large fluctuations in dietary-environmental cysteine availability, are under a stronger selective pressure than bacteria to maintain the Cys-Tyr cofactor as an auxillary mechanism for regulating intracellular cysteine concentrations.

Several mechanisms have been proposed for CDO. A mechanism was postulated after the crystal structure of human CDO was available (Scheme 2) (Ye et al., 2007). This proposal starts with FeIII coordinated by His86, His88, and His140 and a water (McCoy et al., 2006). The hydroxyl group of Tyr157 participates in hydrogen bonded to the coordinated water. Upon addition of substrate, the water molecule is displaced by the thiol group of cysteine with the amino group additionally bonded. This binding geometry allows the carboxyl group of cysteine to engage in the hydrogen bonding network formed by the second coordination sphere (i.e. Tyr157, Tyr58, and His155). The dioxygen then binds in an ‘end-on’ fashion. Homolytic scission of the O–O bond occurs in concert with abstraction of a hydrogen atom from the Tyr157, forming a tyrosyl radical. The electron in the O–O bond is used to form a bond with the iron center resulting in an oxoferryl species, Fe(V=O). The radical on the phenoxyl group then abstracts a hydrogen atom from cysteine’s thiol. The ferryl species attacks the lone pair on cysteine’s sulfur, forming a single S–O bond. This intermediate then undergoes reductive elimination to form an S=O bond and FeIII. The sulfenic acid group is deprotonated and finally, CSA is released from the active site. It is likely that the thiol group would become deprotonated once bonded to the metal since its pKa would drop significantly.

Scheme 2. Proposed mechanism for cysteine dioxygenase (McCoy et al., 2006).
Another mechanism that does not employ radical coupling to produce a thioperoxide is shown in Scheme 3 (Joseph & Maroney, 2007). This mechanism proposes that cysteine initially binds to the iron center by the amine and thiolate groups. In the next step, the S-donor is displaced by oxygen with the assistance of Tyr157, initially serving as a stabilizing residue via hydrogen bonding. Oxidation of the S atom then occurs via nucleophilic attack on the $O_2$ ligand. In this mechanism, the iron center is proposed to be principally a site for organizing the reactions and breaking down the spin-forbidden nature of the reaction with $O_2$ rather than playing a major role in activating oxygen by reducing it to superoxide. This is suggested by the fact that Fe$^{II}$ in CDO with all histidine ligation and a positive overall charge is less electron rich than the Fe$^{II}$ centers in many other non-heme Fe$^{II}$ dioxygenases, or for that matter in hemoglobin/myoglobin, and thus is expected to be harder to oxidize to Fe$^{III}$ to produce superoxide.

![Scheme 3. Proposed mechanism for cysteine dioxygenase (Joseph & Maroney, 2007).](image)

Fe$^{II}$-thiolate complexes [[(iPrBIP)Fe$^{II}$](SPh)(Cl)] (1) and [[(iPrBIP)Fe$^{II}$](SPh)(OTf)] (2) [BIP = bis(imino)pyridine] were prepared as models for CDO. The reaction of 1 and 2 with $O_2$ leads to Fe-oxygenation and S-oxygenation, respectively. For 1 + $O_2$, the spectroscopic and reactivity data, including $^{18}O$ isotope studies, are consistent with an assignment of an Fe$^{IV}=O$ complex, [[(iPrBIP)Fe$^{IV}$](O)(Cl)]$^+$, as the product of oxygenation. In contrast, 2 + $O_2$ results in direct S-oxygenation to give a sulfonato product, PhSO$_3^-$. The location of the thiolate group in 1 versus 2 appears to play a crucial role in the outcome of $O_2$ activation. The thiolate groups in 1 and 2 are critical for $O_2$ reactivity and exhibit an important influence over the Fe$^{II/III}$ redox potential (Badiei et al., 2011; Jiang et al., 2010).

3.2 Gentisate 1,2-dioxygenase (PDB: 3BU7)

Gentisate 1,2-dioxygenase catalyze dioxygen incorporation into various organic compounds and plays a key role in the complex degradation pathway of mono- and polycyclic aromatic and hetero-aromatic compounds (Equation 2). The crystal structure of gentisate 1,2-dioxygenase from *Silicibacter pomeroyi* (GDOsp) determined at a resolution of 2.8 Å was reported and this enzyme closely resembles the *E. coli* GDO (Chen et al., 2008). The X-ray structure of GDOsp demonstrated that the enzyme contains two metal-binding centers per
subunit, in which the metal ions were refined as iron. Each Fe\textsuperscript{II} was primarily coordinated by three histidine residues.

The proposed catalytic mechanism for the N domain of GDOsp, is shown in Scheme 4 (Chen et al., 2008). First, gentisate binds to the active site and displaces one water molecules to chelate with the Fe\textsuperscript{II} ion by its deprotonated carboxylate at C1 and phenolic hydroxyl at C2. The Hydroxyl at C5 participates in a hydrogen bond with the Asp175 carboxylic side chain, which, additionally is stabilized by the side chain of Gln108. The chelation and H-bond network position the gentisate substrate at the active site. The substrate dioxygen then coordinates to the Fe\textsuperscript{II} ion forming a Fe-superoxide complex as illustrated in the mechanisms of other extradiol dioxygenases (Kovaleva & Lipscomb, 2007). This complex is stabilized by the imidazole side chain of conserved His162 through electron transfer and hydrogen-bond rearrangement. The attack of the Fe\textsuperscript{III}-bound superoxide at C1 affords an alkylperoxo intermediate. A subsequent Criegee rearrangement of this intermediate results in O–O bond scission and insertion of the first oxygen atom into the aromatic ring to generate an anhydride intermediate. The subsequent transfer of the hydroxyl group from Fe\textsuperscript{II} ion to C2 and resonance rearrangement would give the product maleylpyruvate acid.

![Scheme 4. Proposed mechanism for Gentisate 1,2-dioxygenase (Chen et al., 2008).](image)

### 3.3 Acetylacetone-cleaving enzyme (PDB: 3BAL)
The crystal structure of Dke1 from *Acinetobacter johnsonii* was determined for an inactive zinc-bound variant of the native enzyme that requires Fe\textsuperscript{II} for activity. The enzyme is a dioxygenase that cleaves carbon–carbon bonds in β-diketone substrates via the
incorporation of one atom of molecular oxygen into each site of bond fission (Straganz et al., 2003). Acetylacetone (2,4-pentanedione) is therefore converted into methylglyoxal and acetic acid (Equation 3). The structure reveals coordination of the catalytically inactive zinc via three histidine residues and the conformation of the metal center of Dke1 is very similar to those seen in CDOs.

\[
\text{O}_2 \rightarrow \text{COOH} + \text{CO} \tag{3}
\]

3.4 Bxe_A2876
Bxe_A2876 is an Fe}\text{II}-dependent oxygenase from B. xenovorans that catalyzes the cleavage of carbon–carbon bonds in \( \beta \)-diketone substrates. The enzyme is not inducible by addition of \( \beta \)-diketone to the growth medium. Cell extracts of B. xenovorans contain Bxe_A2876 in the inactive apo-form. Recombinant Bxe_A2876 from E. coli showed mixed fractional occupancy of its 16.1 kDa subunit with metal occupation (0.06 copper; 0.11 iron; 0.17 zinc) (Leitgeb et al., 2009). \( \text{Fe}\text{II} \) elicited enzymatic catalysis of \( \text{O}_2 \)-dependent conversion of various \( \beta \)-diketone substrates. Data from X-ray absorption spectroscopy (XAS) support a five-coordinate or six-coordinate \( \text{Fe}\text{II} \) center where the metal is bound by three imidazole nitrogen atoms at 1.98 Å. In the ‘best-fit’ model, one or two oxygens from water and a carboxylate oxygen (presumably from Glu96) are further ligands of \( \text{Fe}\text{II} \) at estimated distances of 2.04 Å and 2.08 Å, respectively. A proposed catalytic mechanism for this enzyme is shown below (Scheme 5) (Leitgeb et al., 2009). The proposed mechanism involves \( \text{O}_2 \)-dependent formation of a C3 peroxo intermediate. \( \text{Fe}\text{II} \) is proposed to facilitate this reaction. The distal oxygen of the peroxide performs an intramolecular attack on a neighboring carbonyl carbon to yield the dioxetane, from which products are finally generated through concerted C–C and O–O bond cleavage.

![](image)

Scheme 5. Proposed mechanism for Bxe_A2876 (Leitgeb et al., 2009).

4. 3-His-1-Glu

4.1 Quercetin 2,3-dioxygenase (PDB: 1JUH)
Quercetinase is produced by various filamentous fungi when grown on rutin as the sole carbon and energy source. The reaction involves dioxygenation of quercetin to yield the corresponding depside (phenolic ester 2-proto catechuoylphloroglucin carboxylic acid) and carbon monoxide (Equation 4). Crystal structures available for both the eukaryotic copper containing (A. japonicus) and the prokaryotic iron (or manganese) containing (B. subtilis) quercetinases have revealed details on the coordination of the metal and on the organization of the substrate-binding cavity (Fusetti et al., 2002; Gopal et al., 2005). From the X-ray
structures of the enzyme from *A. japonicus* in the absence (PDB code: 1JUH) (Fusetti et al., 2002) and in the presence of substrates (quercetin, PDB code: 1H1M) (Steiner et al., 2008) it is clearly established that: (1) substrates are tightly bound to the catalytic pocket where they coordinate CuII as monodentate ligand through their 3-OH groups, and (2) substrate binding triggers a conformational change of a loop region that holds the substrate in place (Gopal et al., 2005; Steiner et al., 2002).

The dioxygenation mechanism (Scheme 6) that has been proposed from structural studies involves binding of quercetin to the metal copper(II) center followed by reversible formation of a flavonol radical (centered on C2). Reaction of this species with dioxygen leads to an endoperoxide (C2−C4) that decomposes generating the depside and CO (Steiner et al., 2002). In this reaction sequence, dioxygen would access the copper center from the enzyme surface through a predicted channel from molecular simulations (van den Bosch et al., 2004). In addition to these major structural data, biomimetic studies (Grubel et al., 2010; Malkhasian et al., 2007; Pap et al., 2010) and computational investigations (Fiorucci, 2004, 2006, 2007; Siegbahn, 2004) have recently increased our knowledge on these enzymes. However, despite this progress, questions relating to the function of both the prokaryotic and the eukaryotic enzymes, and in particular about the use of different redox metals for the activation of either the substrate or dioxygen, remain to be answered. Recently a quercetinase with a preference for NiII and CoII was discovered (Merkens et al., 2008).

Scheme 6. Proposed mechanism for copper-containing quercetin 2,3-dioxygenase.
4.2 Pirin (PDB: 1J1L)

Pirin is a newly identified nuclear protein that interacts with the oncoprotein B-cell lymphoma 3-encoded (Bcl-3) and nuclear factor I (NFI) (Pang et al., 2004). The crystal structure of human Pirin at 2.1 Å resolution shows it to be a member of the functionally diverse cupin superfamily. The enzymatic role for Pirin is most likely in biological redox reactions involving oxygen, and provides compelling evidence that Pirin requires the participation of the metal ion for its interaction with Bcl-3 to co-regulate the NF-κB transcription pathway and the interaction with NFI in DNA replication. Substitution of Fe by heavy metals thus provides a unique pathway for these metals to directly influence gene transcription. The determined structure suggests a new role for iron in biology and that Pirin may be involved in novel gene regulation mechanisms. The enzyme is widespread in mammals (including humans), plants, fungi and prokaryotes (Wendler et al., 1997). There are crystal structures for human pirin (PDB code 1J1L) and a pirin homologue from E. coli (PDB code 1TQ5). The human protein showed a 3-His-1-Glu metal-binding site (Pang et al., 2004). The structure of the E. coli protein was solved in the presence of CdII. The cadmium was coordinated only by two histidine residues of the potential 3-His metal site (Adams & Jia, 2005). Interestingly, quercetin was reported to act as substrate for pirins in a similar way as it does for QDO (Adams & Jia, 2005).

4.3 Oxalate decarboxylase (PDB: 1UW8)

Oxalate is produced by plants and microbes by the hydrolysis of oxaloacetate or by the oxidation of glyoxylate or ascorbate (Franceschi & Nakata, 2005). Oxalate secreted by fungi promotes the degradation of lignin (Shimada et al., 1997). Accumulation of oxalate in leafy plants such as spinach and Amaranthaceae leads to nutritional stress, as these plants lack the ability to catabolize oxalate. Excess oxalate in the diet of humans may lead to hyperoxaluria which has been implicated in a number of pathological conditions such as formation of calcium oxalate stones in the kidney (urolithiasis), renal failure, cardiomyopathy, and cardiac conductance disorders (Williams & Wandzilak, 1989). Oxalate decarboxylase catalyzes the decarboxylation of oxalic acid to yield formic acid and carbon dioxide (Equation 5). This transformation is chemically interesting because the C–C bond in the substrate is relatively unreactive (Begley & Ealick, 2004; Svedružic et al., 2005). X-ray crystal structures of B. subtilis OxDC and several site-specific mutants (Anand et al., 2002; Just et al., 2007; Just et al., 2004) indicate that the enzyme is a bicupin. Evidence shows that recombinant, wild-type (WT) B. subtilis OxDC contains MnII when expressed in Escherichia coli (Angerhofer et al., 2007; Tanner et al., 2001). The OxDC monomer is composed of two cupin, β-barrel domains, each of which contains a metal-binding site. For samples of recombinant, wild type B. subtilis OxDC with high MnII occupancy, EPR studies have suggested that only a single metal center interacts with acetate or formate (Angerhofer et al., 2007). Three observations support the hypothesis that this MnII center, and hence the OxDC active site, is located in the N-terminal domain. First, this domain contains a ‘substrate channel’, which can be ‘open’ thereby permitting oxalate to diffuse into the MnII-binding site from solution or ‘closed’ to exclude solvent during catalysis. The interconversion of these two ‘states’ is mediated by conformational rearrangement of a tetrapeptide ‘lid’ segment comprising residues 161-165 (Just et al., 2004). Second, a molecule of formate is coordinated to the N-terminal metal ion in one of the OxDC crystal structures (Anand et al., 2002). Third, recent work using OxDC mutants has shown that (i) site-directed mutagenesis of Glu162 (Svedružic et al., 2007) and (ii) modification of the N-terminal active site lid (Burrell et al., 2007).
abolish decarboxylase activity. Although there is general agreement that Mn\textsuperscript{II} in the N-terminal domain mediates OxDC-catalyzed decarboxylation, legitimate questions have been raised concerning the function (if any) of the Mn\textsuperscript{II} bound in the C-terminal cupin domain. This problem has been investigated using a series of OxDC mutants in which Mn\textsuperscript{II} binding is perturbed by mutagenesis of Glu101 and Glu280, which coordinate the metal in the N-terminal and C-terminal domains, respectively (Moomaw et al., 2009). It was demonstrated that decarboxylase activity and total manganese content are sensitive to modification of either metal-binding glutamate residue. OxDC requires molecular oxygen for catalytic turnover, even though there is no net oxidation or reduction for this reaction. A detailed investigation of the catalytic mechanism of recombinant, wild type \textit{B. subtilis} OxDC was reported (Reinhardt et al., 2003). The presence of Mn\textsuperscript{III} during the catalytic cycle has yet to be unambiguously established by EPR (Svedružić et al., 2005). For OxDC from \textit{B. subtilis}, no spectroscopic signature for Mn\textsuperscript{III} or Mn\textsuperscript{IV} was observed by EPR for samples frozen during turnover. This is consistent with either a large zero-field splitting in the oxidized metal center or undetectable levels of these intermediates (Chang et al., 2004). It should also be noted that one theoretical study argues that formation of a Mn\textsuperscript{III} species in the catalytic cycle is not required for catalytic activity (Chang & Richards, 2005). In a recently published mechanism (Scheme 7) (Mathusamy et al., 2006), the first step is the reaction of the resting state Mn\textsuperscript{II} enzyme with oxalate.

\[
\text{HO} \text{C} \text{O} \text{H} \xrightarrow{\text{OxDC}} \text{CO}_2 \ + \ \text{H} \text{C} \text{O} \text{H} \tag{5}
\]

\[
\text{H}_2\text{O} \ + \ \text{O}_2 \ + \ \text{HCO}_3^-
\]

\[
\text{Glu} \text{His} \text{Mn} \text{OH}_2 \text{OH} \xrightarrow{\text{OxH}_2} \text{Glu} \text{His} \text{Mn} \text{O}_2 \text{OH}_2 \xrightarrow{\text{O}_2} \text{Glu} \text{His} \text{Mn} \text{O}_2 \text{O}^-.\]

Next, oxygen binds the Mn\textsuperscript{II} to generate a Mn\textsuperscript{III}-superoxide species. The reversible electron transfer from coordinated oxalate to Mn\textsuperscript{III} superoxide complex is accompanied by a proton...
transfer from the protonated substrate to a nearby residue. Based on the OxDC crystal structure, the Glu333 could function as the general base in the deprotonation reaction due to its close proximity to the bound substrate. Decarboxylation and C–C bond cleavage in a second step, facilitated by the partial positive charge on the carbon that will become formate, then results in formation of a formate radical (Mathusamy et al., 2006). The formate radical then acquires a proton, possibly from residue Glu162, and an electron from the enzyme bound manganese ion to yield the product formate bound Mn$^\text{III}$ superoxide. Loss of formate and dioxygen lead to the resting state of OxDC containing Mn$^\text{II}$ in an octahedral environment.

4.4 Oxalate oxidase (PDB: 1FI2)

Oxalate oxidase (OxOx), also known as germin, is expressed by plants such as wheat and barley and catalyzes the manganese-dependent oxidative decarboxylation of oxalate to carbon dioxide and hydrogen peroxide (Equation 6), and protects plants from the toxic effects of oxalate (Dunwell et al., 2004).

\[
\begin{array}{c}
\text{OxOx} \\
\text{O}_2 \\
\end{array} \quad \text{H} \\
\rightarrow \quad 2\text{CO}_2 + \text{H}_2\text{O}_2
\]

The structure of OxOx (Woo et al., 2000) contains a manganese ion bound to the side chains of conserved glutamate and histidine residues in a site that is located toward the narrow end of the barrel-like domain. Studies by Whittaker and coworkers (Whittaker et al., 2007) have suggested that the catalytically active form of OxOx is likely the Mn$^\text{III}$ form and not the Mn$^\text{II}$ form as was previously thought. In this mechanism (Scheme 8), the active, resting Mn$^\text{III}$ enzyme binds oxalate (step 1) as the monoanion to form an enzyme-substrate complex. Protein side chain residues (Asn75 and Asn85) stabilize the oxalate group via hydrogen bonding interactions. Oxalate has been shown to anaerobically reduce the Mn$^\text{III}$ form of the enzyme to the Mn$^\text{II}$ form (step 2) (Whittaker & Whittaker, 2002). The reduction of the Mn$^\text{III}$ to Mn$^\text{II}$ is closely associated with generation of an oxalyl free radical (seen in Scheme 8 as potentially binding in a chelating manner). The oxalyl radical is very unstable and is known to undergo rapid C–C bond fission nonenzymatically in aqueous solution ($k = 2 \times 10^6 \text{ s}^{-1}$), producing a molecule of carbon dioxide and carbon dioxide radical anion (Hislop & Bolton, 1999). The interception of a carbon dioxide radical anion intermediate by dioxygen during OxOx turnover would generate a second molecule of carbon dioxide and superoxide (step 4). Electron transfer oxidation of Mn$^\text{II}$ to Mn$^\text{III}$ by the protonated superoxide (step 5) involving either an inner-sphere or outer-sphere mechanism would generate hydrogen peroxide (Whittaker et al., 2007). This scheme predicts that one proton is consumed per turnover cycle and that peroxymonocarbonate is not formed as a primary product (Burrell et al., 2007), although the presence of both peroxide nucleophile and carbon dioxide electrophile in the product mixture makes it likely that peroxymonocarbonate (a peracid) will be produced as a secondary product in solution. Formation of peroxymonocarbonate would account for oxidation of Mn$^\text{II}$ OxOx to Mn$^\text{IV}$ OxOx in the turnover-based redox modification of the enzyme, consistent with the oxidation of Mn$^\text{II}$ OxOx by peracetic acid. The role of the metal in this mechanism is oxalate activation through one-electron oxidation of bound substrate by the active site Mn$^\text{III}$ center.
Scheme 8. Proposed mechanism for oxalate oxidase (Whittaker et al., 2007).

An alternative mechanism for OxOx that has also been proposed (Burrell et al., 2007) (Scheme 9) begins with MnII becoming oxidized to MnIII-superoxo after binding to oxalate. Internal redox results in MnII-superoxo bonded to one electron oxidized oxalate. CO₂ is lost resulting in a MnII-superoxo bound carbamion. The next step invokes a percarbonate intermediate which decomposes to release H₂O₂ and CO₂ regenerating the resting state. There has been recent interest in using synthetic compounds to model the active site of OxOx (Fuller et al., 2006; Fuller et al., 2005; Makowska-Grzyska et al., 2003; Scarpellini et al., 2008).

Scheme 9. Proposed mechanism for oxalate oxidase (Burrell et al., 2007).

4.5 Acireductone dioxygenase (PDB: 1VR3)

Acireductone Dioxygenase (ARD) is an enzyme involved in methionine salvage pathway (Scheme 10), which regulates aspects of the cell cycle. The addition of polyamines to cells
Modeling the Metal Binding Site in Cupin Proteins

...accelerates their DNA replication and division, whereas inhibition of polyamine biosynthesis arrests DNA replication and prevents continuation of the cycle. Methylthioadenosine (MTA) is a inhibitor for the biosynthesis of polyamine. The MTA concentration in biological systems is therefore tightly regulated. This control is effected via the methionine salvage pathway, where MTA is recycled through a series of reactions that convert its 5-methylthio-D-ribose group to methionine. This pathway has been found in many organisms including *Homo sapiens*. In this pathway, the γ-thiomethyl group of MTA is returned to methionine. MTA is derived from S-adenosylmethionine (SAM), which in turn is formed by S-alkylation of methionine at the 5'-ribose carbon of ATP. Acireductone is the penultimate intermediate in the methionine salvage pathway. This compound reacts readily with O₂ to yield a keto acid which is the precursor of methionine and formate. Overexpression of ARD in *E. coli* results in the formation of two enzymes that share a common amino acid sequence. These enzymes differ in the metal ion present (Ni⁺⁺ or Fe⁺⁺) at the active site (Dai et al., 1999). The structure of the nickel-containing version have revealed that the Ni-ARD possesses Ni⁺⁺ coordinated to three histidine nitrogens, one carboxylate oxygen, and two water molecules (Pochapsky et al., 2002; Xu et al., 2006), however, a variant enzyme bearing replacement of Glu69 by glutamine displayed similar properties to those of the wild-type enzyme (Straganz et al., 2006) suggesting that the conserved glutamate residue is not important for enzyme function.

![Scheme 10. Methionine Salvage Pathway from *Klebsiella pneumonia*.](image1.png)

Depending on the metal bound at the active site, ARD catalyzes two different reactions. ARD with ferrous iron bound promotes the cleavage of 1,2-dihydroxy-5-methylthiopent-1-en-3-one to a precursor of methionine and formate. The Ni⁺⁺-bound variant catalyzes an off-pathway reaction, leading to products methylthiopropionate, CO and formate thus preventing the recycling to methionine (Dai, 1999, 2001). The proposed catalytic pathway for ARD in its Fe⁺⁺ and Ni⁺⁺ forms (Pochapsky et al., 2002) involves initial bidentate coordination...
of the deprotonated substrate to the metal center (Scheme 11, A). Dioxygen reduction by the metal-substrate complex involving either inner or outer-sphere mechanism (B), facilitated by the reactive substrate. Following formation of the peroxide intermediate (C), the reaction branches into two paths depending on the metal involved. The Fe$^{II}$-dependent pathway I begins (D) with nucleophilic attack on the adjacent carbonyl carbon to yield a four-membered ring (E), which then decomposes into formate acid and α-oxo-acid products (F).

Scheme 11. Proposed mechanisms for Fe$^{II}$ and Ni$^{II}$ ARD (Pochapsky et al., 2002).

For pathway II, the reaction of the Ni$^{II}$ enzyme (D') is suggested to involve intramolecular rotation of the metal-bound substrate across its C2-C3 bond such that the C3 carbonyl group is now coordinating (E') and the peroxide attack is directed towards the C3 carbonyl carbon (F'), with opening of the five-membered dioxolane ring to yield formate, carbon monoxide, and the corresponding carboxylate products (G'). From EXAFS studies (Al-Mjeni et al., 2002), substrate binding is believed to cause dissociation of one water and one histidine ligand. Efforts to model the Ni$^{II}$ form of the enzyme have most notably been carried out by Berreau and coworkers (Berreau et al., 2011; Grubel et al., 2010; Rudzka et al., 2010; Szajna-Fuller, 2007a, 2007b; Szajna, 2004, 2005).

5. 3-His-1-Gln

5.1 Bacilysin biosynthesis protein BacB (PDB: 3H7J)

Bacilysin is a non-ribosomally synthesized dipeptide antibiotic that is active against a wide range of bacteria and some fungi. Synthesis of bacilysin (L-alanine-[2,3-epoxycyclohexano-
4]-L-alanine) is achieved by proteins in the bac operon, also referred to as the bacABCDEF (yuvBCDEF) gene cluster in B. subtilis. The enzyme BacB, an oxidase, catalyzes the synthesis of 2-oxo-3-(4-oxocyclohexa-2,5-dienyl)propanoic acid which is a precursor to L-anticapsin. This protein is a bicupin containing bound Co\textsuperscript{II} and Fe\textsuperscript{II} (Rajavel et al., 2009). The residues coordinating the Co\textsuperscript{II} ion in the N-terminal domain are His50, His52, His91, and Gln56, whereas those in the C-terminal domain include His162, His164, His202, and Gln168. In the N-terminal domain, the remaining coordination sphere is completed by two water molecules, whereas the C-terminal domain revealed additional density for a bound substrate ligand. The Fe\textsuperscript{II} is attached to the protein via three protein-derived carboxylate groups. The reaction catalyzed by BacA and BacB is shown in Scheme 12. In these reactions prephenate is first converted into an intermediate which is followed by the formation of 2-oxo-3-(4-oxocyclohexa-2,5-dienyl) propanoic acid by BacB. A mechanism for BacB has not yet been proposed.

Scheme 12. Reaction pathway for the formation of Anticapsin (Rajavel et al., 2009).

6. 4-His

6.1 RemF protein

RemF is a polyketide cyclase involved in the biosynthesis of the aromatic pentacyclic metabolite resistomycin (Scheme 13) in Streptomyces resistomycificus. The enzyme is a member of a structurally hitherto uncharacterized class of polyketide cyclases. The X-ray structure of RemF was determined and refined to 1.2 Å resolution (Silvennoinen et al., 2009). The enzyme subunit exhibits a β-sandwich structure with a structure characteristic for the cupin fold. RemF possesses a metal binding site located inside a predominantly hydrophobic cavity. A zinc ion is coordinated to four histidine side chains, and two water molecules in distorted octahedral ligand geometry, highly unusual for zinc binding sites in proteins. Based on sequence comparisons the three predicted cyclases essential for resistomycin biosynthesis, namely RemI, RemF and RemL have been designated to three of the ten sequence families (Fritzsche et al., 2008). Reml is related to the cyclases represented by tetracenomycin aromatase/cyclase with the characteristic ‘helix-grip’ fold (Ames et al.,...
RemL and RemF belong to two different cyclase families that have not yet been structurally elucidated. The Zn$^{II}$ located at the active site of RemF could act as a Lewis acid in the aldol condensation reaction catalyzed by RemF, reminiscent of class II aldolases.

Scheme 13. Reaction pathway for the biosynthesis of resistomycin.

### 6.2 Carotenoid oxygenase

Carotenoid oxygenases (CarOs) are involved in the biosynthesis of retinal and structurally related pigments and have been reported in animals, plants, and more recently cyanobacteria (Ruch et al., 2005). They are described as Fe$^{II}$-dependent enzymes and catalyze the 15–15' bond cleavage of a carotene precursor, as shown in Scheme 14. CarOs are very different from the other enzymes discussed with regard to both structure and catalytic mechanism. The structure of CarO from the cyanobacterium *Synechocystis* sp. PCC6803 (Kloer et al., 2005) was solved in the presence of substrate. The proposed reaction pathway...
for CarO involves initial substrate binding, reaction of the Fe\(^{II}\) center with dioxygen followed by oxidative cleavage of the double bond resulting in products (Kloer et al., 2005).

### 7. Conclusion

Proteins possessing the cupin DSBH fold are present in all realms of life and can participate in a diverse set of enzymatic reactions. In many cases, these enzymes are assisted by various first row transition metal cofactors. The number of metalloproteins discovered within this class appears to be increasing at a rapid rate. High throughput structural methods and utilization of expression systems has assisted in the increased discovery of cupin enzymes.

The catalytically active metal binding sites consist of a variation of Histidine, Glutamine, Aspartate/Glutamate, and/or Tyrosine ligands. This chapter was focused on proteins which had been structurally elucidated and with new potentials for investigations regarding their mechanisms. The mechanisms employed by these enzymes are diverse and their detailed investigation can be probed using synthetic model systems. Since synthetic models can be systematically altered and subjected to conditions not be suitable to biological investigation, there is much promise in slowing down catalytic reactions using low temperature methods and potentially isolating intermediates which can then be investigated by various physical methods and in some cases, by crystallography. Furthermore, physical studies on model complexes along with computational investigations can potentially provide information regarding correlations between electronic and structural parameters.

### 8. Acknowledgement

Financial support from Oakland University and NSF Grant No. CHE-0748607 is gratefully acknowledged. The National Science Foundation (NSF) award (CHE-0821487) is also acknowledged.

### 9. References

JCSG (Joint Center for Structural Genomics) (2008). Crystal Structure of Domain of Unknown Function with a Cupin Fold (YP\_752209.1) from *Shewanella frigidimarina* NCIMB 400 at 2.05 Å resolution.


Modeling the Metal Binding Site in Cupin Proteins


Bio-mimicry is a fundamental idea of how to mimic the Nature™ by various methodologies as well as new ideas or suggestions on the creation of novel materials and functions. This book comprises seven sections on various perspectives of bio-mimicry in our life: Section 1 gives an overview of modeling of biomimetic materials; Section 2 presents a processing and design of biomaterials; Section 3 presents various aspects of design and application of biomimetic polymers and composites are discussed; Section 4 presents a general characterization of biomaterials; Section 5 proposes new examples for biomimetic systems; Section 6 summarizes chapters, concerning cells behavior through mimicry; Section 7 presents various applications of biomimetic materials are presented. Aimed at physicists, chemists and biologists interested in biomineralization, biochemistry, kinetics, solution chemistry. This book is also relevant to engineers and doctors interested in research and construction of biomimetic systems.

How to reference
In order to correctly reference this scholarly work, feel free to copy and paste the following:
