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)†

B. Gopal,[‡,§] Lalima L. Madan,[‡] Stephen F. Betz,[||,⊥] and Anthony A. Kossiakoff*[,§,@]

Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560 012, India, Department of Biochemistry and Molecular Biology, University of Chicago, Chicago, Illinois 60637, Institute for Biophysical Dynamics, University of Chicago, Chicago, Illinois 60637, and GeneFormatics, Inc., 5830 Oberlin Drive, San Diego, California 92121

Received July 21, 2004; Revised Manuscript Received October 15, 2004

ABSTRACT: Common structural motifs, such as the cupin domains, are found in enzymes performing different biochemical functions while retaining a similar active site configuration and structural scaffold. The soil bacterium *Bacillus subtilis* has 20 cupin genes (0.5% of the total genome) with up to 14% of its genes in the form of doublets, thus making it an attractive system for studying the effects of gene duplication. There are four bicupins in *B. subtilis* encoded by the genes yvrK, yoaN, yxaG, and ywfC. The gene products of yvrK and yoaN function as oxalate decarboxylases with a manganese ion at the active site(s), whereas YwfC is a bacitracin synthetase. Here we present the crystal structure of YxaG, a novel iron-containing quercetin 2,3-dioxygenase with one active site in each cupin domain. Yxag is a dimer, both in solution and in the crystal. The crystal structure shows that the coordination geometry of the Fe ion is different in the two active sites of YxaG. Replacement of the iron at the active site with other metal ions suggests modulation of enzymatic activity in accordance with the Irving−Williams observation on the stability of metal ion complexes. This observation, along with a comparison with the crystal structure of YvrK determined recently, has allowed for a detailed structure−function analysis of the active site, providing clues to the diversification of function in the bicupin family of proteins.

The double-stranded β-helix (cupin) domain has been extensively explored in "reaction space", resulting in the evolution of numerous, diverse catalytic activities supported by the same structural scaffold (*1*). Thus, while the tertiary structure is conserved, variety in biochemical function is provided by variations in the residues of the active site and the identity of the bound metal ion (*2*, *3*). The genome of *Bacillus subtilis* shows four bicupin sequences encoding the proteins YwfC, YvrK, YoaN, and YxaG. Considerable sequence identity has established that YwfC is similar to the protein BacB. BacB is a part of the bacitracin synthetase operon and is involved in peptide synthesis. Earlier work on the functional characterization of YvrK, YoaN, and YxaG (*4*) had established that YvrK and YoaN were oxalate decarboxylases, containing a manganese ion at the active site. The crystal structure determination of *B. subtilis* YvrK (*5*) has now allowed for a detailed structural analysis of this oxalate decarboxylase while also revealing that YvrK is a hexamer both in solution and in the crystal. Recent studies (*6*, *7*) have characterized the protein YxaG as an iron-containing quercetin 2,3-dioxygenase. Here we report the crystal structure of YxaG (Figure 1) that shows YxaG to be a dimeric enzyme with two active sites per monomer, each containing an Fe^{2+} ion.

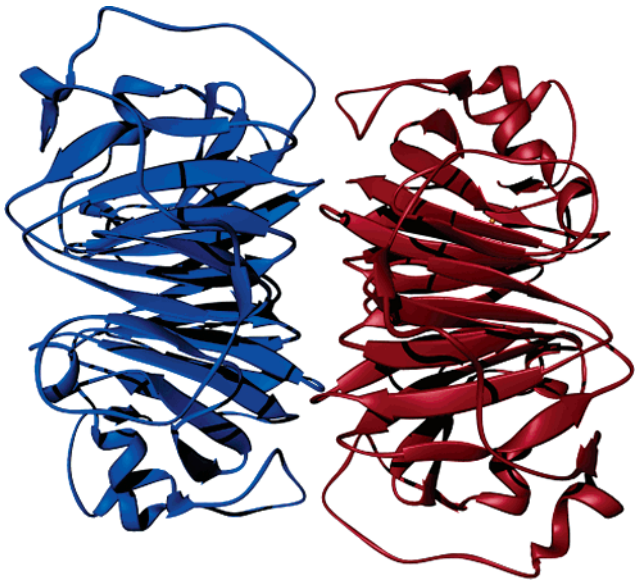

FIGURE 1: Ribbon diagram of a YxaG dimer. The two monomers, shown in red and blue, are related by an approximate 2-fold axis of symmetry.

† Financial support from the start-up grant provided by the Indian Institute of Science is gratefully acknowledged.

* To whom correspondence should be addressed: Cummings Life Sciences Center, 920 E. 58th St., Chicago, IL 60637. Telephone: (773) 702-9257. Fax: (773) 834-2777. E-mail: koss@cummings.uchicago.edu.

‡ Indian Institute of Science.

§ Department of Biochemistry and Molecular Biology, University of Chicago.

|| GeneFormatics, Inc.

⊥ Current address: Neurocrine Biosciences, Inc., 10555 Science Center Dr., San Diego, CA 92121.

@ Institute for Biophysical Dynamics, University of Chicago.

10.1021/bi0484421 CCC: $30.25 © 2005 American Chemical Society
Published on Web 12/13/2004

The structural scaffold which characterizes a member of the cupin superfamily of proteins is that of a conserved β-barrel containing two sequence motifs, $GX_5HXHX_{3,4}EX_6G$ and $GX_5PXGX_2HX_3N$ (*2*, *4*). These proteins are involved in several metabolic pathways and can perform very different enzymatic reactions; they could be carboxylases, dioxygenases, germins or germin-like proteins from higher plants, phosphomannose isomerases, and polyketide synthases, among others. This diversity in function exhibited by members of this family has been suggested to be a consequence of the ancient nature of this protein fold, having been classified as a member of the "small molecule binding domains" which spread across the archea, eubacteria, and eukaryota (*1*). A characteristic feature of this class of proteins is their high thermal stability and resistance to proteases, a feature variously ascribed to an increased number of subunit contacts, hydrophobic interactions and hydrogen bonding, efficient packing, short loops, fewer cavities, or the presence of ions, cofactors, metabolites, compatible solutes, and covalent conjugates (*3*). The thermostable nature of the cupin domain agrees well with observations that these proteins are induced or highly expressed under conditions of abiotic stress, including high temperatures in microbes and eukaryotes (*2*). The active site, located in the center of the β-barrel and flanked by residues from motifs 1 and 2, can bind different metal ions. These metal ions can influence the activity of these proteins, although the presence of a metal ion is not required for some members of this family, prominent examples of this feature being the dTDP-4-dehydrorhamnose 3,5-epimerase and transcription factor AraC (*2*). The manganese ion, on the other hand, serves as a cofactor in the generation of hydrogen peroxide from superoxide and oxalate ions in superoxide dismutase or oxalate decarboxylase. Dioxygenases have copper or iron at the active site, whereas type 1 or 2 phosphomannose isomerases have zinc at their active site. An interesting feature of a change in activity upon variation of the identity of the metal ion has been reported in the case of the enolase from *Klebsiella pneumoniae*, where a 1,3-oxygenalytic reaction or 1,2-oxygenalytic reaction has been observed depending on whether the metal ion is either Ni or Co, or replaced with Fe (*8*). The diversity seen in biochemical activity is further enhanced in the case of bicupins, proteins having two cupin domains fused to each other. The bicupins, often suggested as products of gene duplication events, thus have the potential to function as multienzymes, with each domain geared to adopt a functional role.

B. subtilis is an excellent model system for analyzing gene duplication or gene fusion events. It is estimated that *B. subtilis* has 568 (14% of the total) of its genes in the form of doublets and 273 (7% of the total) in the form of triplets. While 58% of all genes have a counterpart with a known function, 26% of the proteins are not similar to any other proteins in the protein databases. Only 1200 genes (~30%) have been experimentally identified in *B. subtilis*, thus necessitating the experimental confirmation of the predicted functions of the unannotated proteins in the genome that encoded genes with an identifier having a "y". Among other features, it has a grouping of 20 cupin genes (0.5% of the total of 4100) with evidence of duplication or fusion to produce the genes encoding two-domain bicupin proteins. Among the four bicupin proteins encoded by the *B. subtilis* genome, sequence features clearly set YwfC apart from the other three, YvrK, YoaN, and YxaG. YvrK and YoaN share the maximum level of sequence similarity and, not surprisingly, share a common function. YvrK and YoaN are oxalate decarboxylases, which convert oxalate to formamate and CO_2. This enzymatic activity is characterized by a critical requirement for manganese salts. The protein YxaG does not cluster with the oxalate decarboxylases from a phylogenetic perspective (*2*). Another interesting feature is the length of the linker sequence between the two cupin sequence motifs: 15 residues in YxaG, 16 in YwfC, and 20 in YvrK and YoaN. Phylogenetic analysis (*2*, *3*) suggests YjlB (18 residues) as a probable progenitor for YxaG which implies a decrease in intermotif spacing after gene duplication. Sequence comparison suggests that the most likely progenitor of YvrK and YoaN (20-residue spacing) is YkrZ. It is more similar to YvrK than to YoaN, which suggests an evolutionary sequence of events [(YkrZ)×2 → YvrK → YoaN (*2*)]. These sequence variations, especially the length between the two cupin motifs, are probably exploited by the bicupin proteins to explore "reaction space" by providing changes in the residues which confer substrate specificity (*2*, *3*), while still maintaining the robust structural framework provided by the cupin domain. The role of the bound metal ion in the control of enzyme activity is less clear, as is cooperativity in substrate binding, or other interactions between the two active sites.

Quercetin 2,3-dioxygenases from various species of *Aspergillus*, *Aspergillus japonicus* (*9*), *Aspergillus flavus* (*10*), and *Aspergillus niger* DSM 821 (*11*), have been characterized and were found to contain a type II copper ion at the active site. In the case of the *A. japonicus* quercetin 2,3-dioxygenase, an analysis of the EXAFS region suggested that the active site can be equally well described by either four or five ligands [3N(His) + 1−2O] at an average distance of ~2 Å. However, when the enzyme is complexed with quercetin, the copper environment undergoes a transition to a five-coordinate cage (*12*, *13*). Of the two potential active sites, only the one located in the N-terminal domain was found to have a bound copper ion. Studies of the *A. niger* DSM 821 enzyme, however, show that 1 mol of enzyme contained 1−1.6 mol of Cu, suggesting a partial occupancy at the second site (*11*). EPR[1] studies on the *A. japonicus* enzyme reveal several differences in the copper coordination between this and the *A. niger* quercetin 2,3-dioxygenase. The crystallographic analysis of the *A. japonicus* enzyme reveals that the copper center has two alternative conformations, the main form (~70%) being pseudotetrahedral and the minor form (~30%) having a mixed trigonal bipyramidal/square pyramidal geometry (*9*). The crystal structure of YxaG reported here shows a similar scenario of different conformations for the active site Fe ion. The reaction mechanism for the copper-containing quercetin 2,3-dioxygenase has been proposed on the basis of that of the intradiol dioxygenases, which utilize high-spin Fe(III) in place of copper. To examine the effect of different metal ions on enzymatic activity, we report our observations on the replacement of the Fe^{2+} ions at the active site.

[1] Abbreviations: HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; PEG, polyethylene glycol; EDTA, ethylenediaminetetraacetic acid; Tris, tris(hydroxymethyl)aminomethane; DDTC, diethyl dithiocarbamate; EPR, electron paramagnetic resonance.

MATERIALS AND METHODS

Protein Purification. The gene encoding YxaG was cloned between the NheI and XhoI restriction sites of bacterial expression vector pET-15b (incorporating an N-terminal polyhistidine tag) and pET-23b (incorporating a C-terminal polyhistidine tag), to simplify protein purification. After the plasmids containing the appropriate insert had been transformed into BL21(DE3) cells (Novagen, Inc.), the cells were grown to an optical density at 600 nm of 0.4 when the cells were induced with 1 mM IPTG (final concentration). Following this, the temperature for growth was reduced to 25 °C and cells were grown for a further 6 h before they were spun down and stored at −80 °C until they were used. The cells were resuspended in lysis buffer [50 mM Tris-HCl and 200 mM NaCl (pH 7.5)]. After sonication for 4 min on ice, the cell debris was separated from the crude cell lysate by centrifugation for 25 min at 15 000 rpm on a Sorvall centrifuge. After equilibration with the TALON resin (4 mL of resin/10 g of cell paste) and a washing step with 50 mM Tris, 200 mM NaCl, and 5 mM imidazole (pH 7.5), the fusion protein is eluted from the column in the elution buffer [50 mM Tris-HCl, 200 mM NaCl, and 300 mM imidazole (pH 7.5)]. After thrombin cleavage and removal of the N-terminal polyhistidine tag, the purified protein was further subjected to size exclusion chromatography on a Sephacryl Hiprep 16/60 S-200 HR column (Amersham Biosciences Inc.). On the basis of the elution volume of YxaG in the size exclusion chromatography experiment, we infer that this protein is a dimer in solution. Of the two buffers that were examined [50 mM Tris, 250 mM NaCl, and 10 mM DTT (pH 7.5)] or phosphate buffer [50 mM sodium phosphate, 250 mM NaCl, and 10 mM DTT (pH 7.5)], the protein in the Tris buffer was more suitable for crystallization and biochemical experiments.

Crystallization and Determination of the Structure of C-Terminally His-Tagged YxaG. The purified protein YxaG was crystallized in 100 mM Na-HEPES (pH 7.5), 2 M ammonium sulfate, and 2% PEG 400 by the hanging drop method with 2 μL of protein (at 8 mg/mL) mixed with 2 μL of the crystallization solution. Diffraction data were collected on beamline ID19 of the Argonne Advanced Photon Source (University of Chicago) on a Quantum ADSC CCD detector using vitrified crystals maintained at a temperature of 100 K using the Oxford Cryosystems Cryostream device. Data were processed using the HKL suite of processing data (*14*). Molecular replacement trails were performed with the *A. japonicus* dioxygenase (PDB entry 1GQG) using molrep (*15*) using data to 3.5 Å resolution providing a correlation coefficient of 22% and an *R*-factor of 54%. The sequence of YxaG is 17% identical and 39% similar to that of the *A. japonicus* copper-containing dioxygenase that was used as a model for molecular replacement. Improvement of the phasing was performed by iterated cycles of NCSREF (*16*) and warp (*17*) which led to an improvement of the electron density map. Iterated cycles of model building using O (*18*) and refinement using refmac (*19*) led to the final model with the R_{cryst} of 22.6% and the R_{free} of 27.4%. The final model consists of 658 residues, four Fe^{2+} ions, and 298 water molecules with excellent stereochemistry. More than 89% of all the residues fall within the most favored region of the Ramachandran plot. Four residues are missing from the N-terminus and three from the C-terminus in both monomers. The coordinates and structure factors have been deposited in the Protein Data Bank (PDB code: 1Y3T).

Activity Assay. Quercetin 2,3-dioxygenase activity for YxaG was determined spectrophotometrically using a method that has been described previously (*6*, *7*). One unit of enzymatic activity as described by this method corresponds to the amount of enzyme required to convert 1 μmol of quercetin per minute at 25 °C in the reaction buffer [25 mM MES (pH 6.0)]. The reaction buffer was modified to 25 mM HEPES (pH 7.5) due to the pI of this protein (pI ∼5.8). A combination of MES and Tris buffers has been employed by Barney et al. (*7*) to monitor the rate dependence of pH on this enzyme. The results that we obtain are comparable to those published earlier at pH 6.0(6) and pH 7.5(7) for this enzyme. The final reaction mixture contained 0.025 μM YxaG and 50 μM quercetin in 25 mM HEPES buffer (pH 7.5). YxaG hydrolyzes quercetin, leading to a decrease in absorbance at 367 nm which is then followed as a time course measurement (the absorption maximum of quercetin is at 367 nm; $\epsilon_{367} = 17\,200\ M^{-1}\ cm^{-1}$ under these reaction conditions). The UV−visible spectra and time course measurements were obtained using a V-530 UV−visible spectrophotometer (Jasco) using a cell having a path length of 1 cm.

Studies to evaluate the effect of various transition metal ions on the activity of YxaG involved the preparation of metal ion-free apo-YxaG. The chelation conditions were optimized to provide the highest yield of apo-YxaG. The metal ion chelation process involved incubation of the protein with a 10-fold excess of diethyl dithiocarbamate (DDTC) and ethylenediaminetetraacetic acid (EDTA) in 50 mM Tris-HCl buffer (pH 7.5) at 4 °C for 30 min. Following this incubation, the protein was extensively dialyzed against the reaction buffer [25 mM Hepes (pH 7.5)] to remove the chelating agents. The chelation was determined to have reached completion when the enzyme activity was observed to be less than 2% of the holoenzyme. The apoenzyme was then incubated with a 100-fold molar excess of the transition metal ion(s) examined to obtain holo-YxaG. As quercetin itself chelates several transition metal ions, a control experiment was performed using bovine serum albumin (BSA) as a replacement for YxaG. The control experiment, performed in triplicate, allowed for the estimation of the errors involved in the measurement of enzymatic activity due to metal ion chelation by the substrate quercetin.

RESULTS AND DISCUSSION

The crystal structure of YxaG from *B. subtilis* reveals a compact arrangement of two cupin domains. Two active sites are seen in the N- and C-terminal domains, both of which have a bound metal ion. The N- and C-terminal domains of YxaG are joined by a flexible linker 18 residues in length. This linker also forms the lid of the active site of the N-terminal domain. Domain 1 (residues 1−148) and domain 2 (residues 177−333) can be superimposed with a rmsd of 1.11 Å. The two domains are related by a 2-fold axis of symmetry. The dimeric arrangement of the two bicupin monomers seen in Figure 1 is similar to that of the *A. japonicus* dioxygenase (*9*). The interface between the two domains is substantial, involving a buried surface area of

~3100 Å^2, a major component (72%) of it being contributed by the nonpolar residues, as calculated by naccess (*20*). The buried surface area between the two cupin domains is similar to those seen between the domains of the *A. japonicus* Cu^{2+} containing quercetin 2,3-dioxygenase and the Mn^{2+} binding oxalate decarboxylate YvrK protein from *B. subtilis* (*5*).

Comparison between the Active Sites. There are two active sites, defined by the Fe^{2+} ion at their respective centers, separated by a distance of 27.8 Å. As seen in active sites of the other bicupin enzyme structures reported so far, these sites lie buried in the core of the β-barrel of the cupin domain. The enzymatic activity in proteins having a bicupin scaffold appears to be dictated by three main factors: the loop that governs the movement of the substrate to the active site, the nature of the residues that line the cavity of the active site, and the nature of the metal ion at the site (*3*). The stretch of the polypeptide comprising residues 149−177 forms a loop at the active site of domain 1. In the case of the Cu^{2+}-containing dioxygenase from *A. japonicus*, this loop is totally disordered in the free form and regains some order upon binding quercetin (*12*, *21*). Almost all of this region could be traced in YxaG in the apo form, albeit some residues have been modeled with partial occupancy, reflecting the poor electron density map in this region. This loop conformation lies between the "distant" and "closed" conformation of *A. japonicus* dioxygenase. This feature may partly be due to the rigidity provided by a stretch of a 3_{10}-helix at the base of the active site loop adjoining the cavity. As seen in Figure 2, a quercetin moiety was modeled into the active site of the N-terminal domain, based on the coordinates of Steiner et al. (*12*, *13*). This occupies the active site cavity without any major steric hindrance, although the packing inside the cavity appears to be more flexible than that of the *A. japonicus* structure. The residues lining the active site show differences between the two dioxygenases (*A. japonicus* dioxygenase/YxaG): Tyr35 versus Ala, His66 versus His, Met51 versus Val, Thr53 versus Leu, Glu73 versus Glu, Phe75 versus Ile, Phe114 versus Tyr, Met123 versus Leu, and Ile127 versus Thr. In general, bulky residues from *A. japonicus* are replaced by less bulky ones in YxaG, allowing for a widening of the active site cavity. The crystal structures of the enzyme−substrate complexes of kaempferol and quercetin with the *A. japonicus* dioxygenase show some clear differences with that of the native enzyme (*12*). The main difference is in the linker region comprising residues 146−205 of the *A. japonicus* enzyme. In particular, the Cα atom of Ser166 shows a positional difference of 7.5 Å, whereas the main cupin scaffolds were seen to superpose with a rmsd of 0.85 Å. Although the packing of the modeled quercetin in the YxaG structure appears to be less tight than in the *A. japonicus* enzyme, it appears that the enzymatic mechanism wherein a molecular oxygen attacks at an activated flavonol C2 atom would be maintained.

Metal Ion Binding Geometry. The active site residues in the cupin motifs comprising His62, His64, Glu69, and His103 in domain 1 and His234, His236, Glu241, and His275 in domain 2 coordinate the Fe^{2+} ion(s) at the two active sites. The main distinguishing feature between the metal ion coordination at the N-terminal active site and the C-terminal site is the glutamate residue (Glu69) drawn closer to the active site iron with concomitant changes in the rotamers of the histidine residues and the position of the water molecule.

Table 1: Summary of Data Collection, Processing, and Refinement Statistics[a]

wavelength (Å)	0.98
resolution (Å)	20−2.38
no. of reflections	200413
no. of unique reflections	33292
redundancy	6.01
completeness (%)	98.2 (94.3)
R_{sym} (%)[b]	8.9 (59.2)
I/σ	20.37 (2.01)
R_{cryst}[c]	22.6 (34.3)
R_{free}[d]	27.4 (37.9)

[a] Values for outer shells are given in parentheses. [b] $R_{sym} = \sum_j|\langle I\rangle - I_j|/\sum\langle I\rangle$, where I_j is the intensity of the jth reflection and $\langle I\rangle$ is the average intensity. [c] $R_{cryst} = \sum_{hkl}|F_o - F_c|/\sum_{hkl}|F_o|$. [d] R_{free} was calculated like R_{cryst} but on 5% of the data excluded from the refinement calculation.

Table 2: Quaternary Structures of Enzymes Having a Bicupin Scaffold

protein	quaternary structure	metal (domain I/II) and coordination number	ref
B. subtilis YvrK oxalate decarboxylase	hexamer	Mn/Mn; 6, 6	*5*
phosphomannose isomerase	monomer	Zn/none; 6	*30*
A. japonicus quercetin 2,3-dioxygenase	dimer	Cu/none; 5	*9*
B. subtilis YxaG quercetin 2,3-dioxygenase	dimer	Fe/Fe; 5, 5	this work

Table 3: Close and Distant Conformations of the Active Site(s)

close conformation (site 1)	distant conformation (site 2)
Fe−OE1 of Glu69, 2.1 Å	Fe−OE1 of Glu241, 2.44 Å
Fe−NE2 of His62, 2.32 Å	Fe−NE2 of His234, 2.15 Å
Fe−NE2 of His64, 2.16 Å	Fe−NE2 of His236, 2.23 Å
Fe−NE2 of His103, 2.29 Å	Fe−NE2 of His275, 2.16 Å
Fe−Wat136, 2.21 Å	Fe−Wat91, 2.44 Å

In the case of the *A. japonicus* dioxygenase, the ligand coordination was reported to exhibit two different forms: a trigonal bipyramidal arrangement with three histidines, one glutamate, and a water molecule which was estimated to exist for 70% of the time and a square pyramidal arrangement comprising the same ligand donors (*9*, *21*). A distorted version of the first arrangement is seen in the N-terminal active site of YxaG, while the second form is seen in the C-terminal domain (Table 3 and Figure 2a). The significance of this arrangement in determining the nature of the metal ion at the active site is not clear. The electron density surrounding the coordinating water molecule in the N-terminal domain (as in the case of the *A. japonicus* dioxygenase) was elongated, prompting us to examine the possibility of placing another water molecule in that position, thus offering a coordination number of six for the reactive Fe^{2+} ion. Subsequent refinement gave a high-temperature factor to this additional water, leading us to interpret the elongated density to be a feature of the mobility of a single water molecule.

We note that a bound dioxygen species at the active site could also lead to an elongated density. This scenario seemed to be less probable as trapping an active dioxygen species in the crystal requires specific cryotrapping procedures to

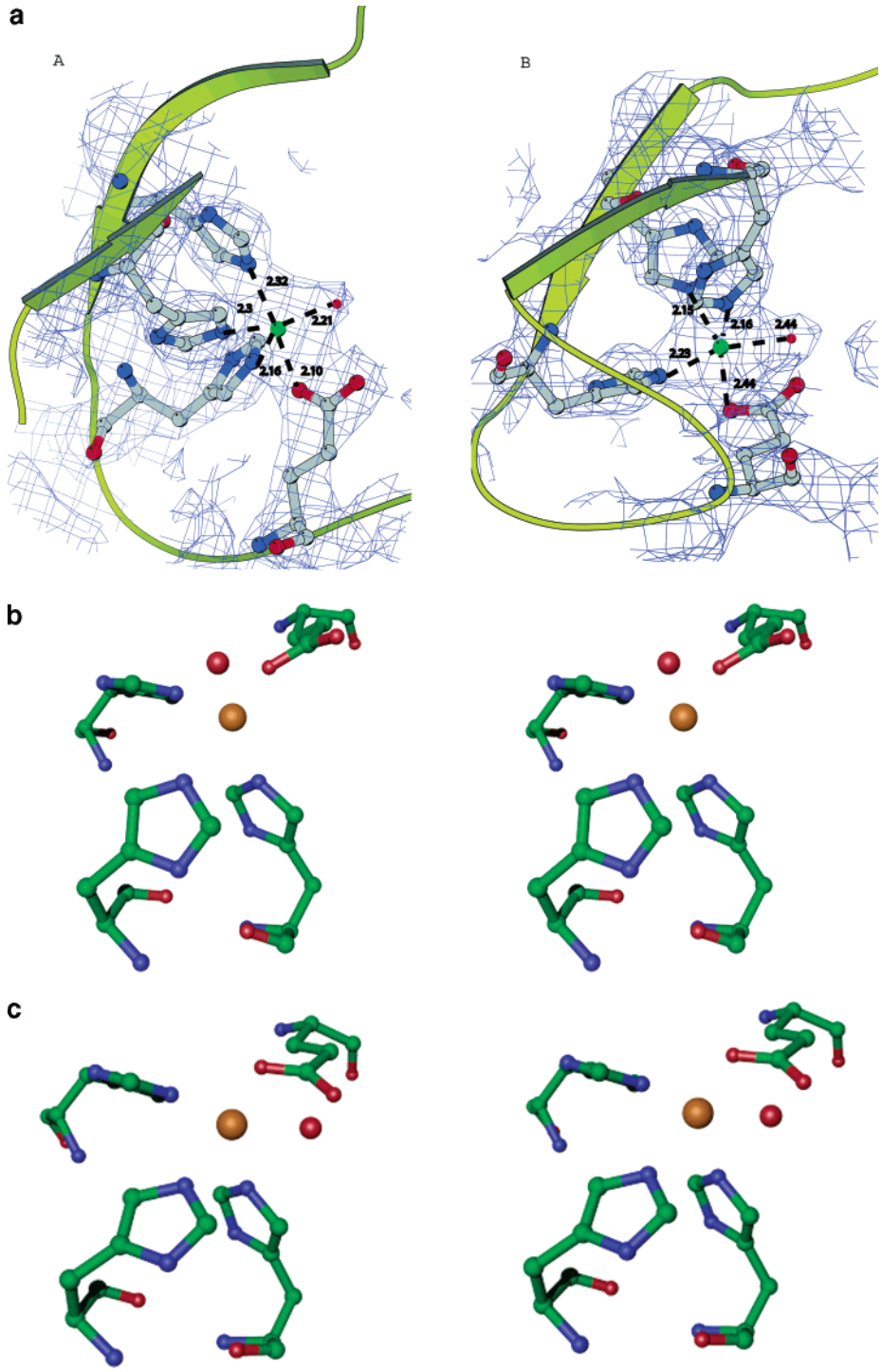

FIGURE 2: (a) Snapshots (A and B) of the two active sites. The $2F_o - F_c$ map is shown here at a contour level of 1.5σ. The Fe^{2+} ion is pentacoordinated in both sites, with a slight distortion of geometry. The geometric distortion is mainly due to the close and distant conformations of the glutamate residue which coordinates the metal ion. (b) Stereoview of the Fe ion coordination at site 1 (the close conformation). (c) Stereoview of the Fe ion coordination at site 2 (the distant conformation).

be in place (*22*). We did not attempt this procedure. An attempt to model a dioxygen ion at this site was unsuccessful as seen by the high *B* factors for the dioxygen during refinement, suggesting that the modeled water was more appropriate.

The crystal structure suggests that the Fe−water distances are 2.21 Å in the close and 2.44 Å in the distant conformations. The ideal Fe−O distance, in the case of a dioxygen, would be 1.8 Å. Indeed, the paper by Sjogren and Hajdu reports the breaking of the bound dioxygen complex after 100° of exposure on a synchrotron beamline. All these facts lead us to believe that the elongated density belongs to a water molecule. A similar interpretation of the elongated electron density near the metal ion was offered by Fusetti et al. for the *A. japonicus* dioxygenase, and subsequently, the quercetin-bound structure (*12*) showed this water molecule is replaced with a hydroxy group of quercetin, allowing the copper ion to be pentacoordinate while maintaining its formal Cu^{2+} coordination state. The catalytic relevance of the close and distant conformation of the active site glutamate remains

FIGURE 3: Model of a quercetin moiety at the active site of YxaG based on the crystal structure of the substrate-bound quercetin 2,3-dioxygenase structure as reported by Steiner et al. There is more room for this substrate in YxaG than the *A. japonicus* dioxygenase as some bulky residues lining this cavity are replaced with lighter ones in YxaG (as noted in the Results and Discussion). The active site residues (His62, His64, Glu69, and His103) have been labeled.

unresolved: in the close conformation, the side chain of the glutamate is close enough to abstract a proton from the substrate's 3-OH group, while in the distant conformation, it would be easier to substitute for the metal ion. The electron density map surrounding the two active sites of YxaG is seen in Figure 2a. The differences in conformational geometry are apparent from the Fe−ligand distances (Figure 2b and Table 3).

The average distances from the coordinating ligands to the metal ion are larger than those seen in either the *A. japonicus* site binding Cu^{2+} or the YvrK sites which coordinate Mn^{2+}. Superposition of the coordinates of the oxalate decarboxylase YvrK onto YxaG shows that the N-terminal Fe^{2+} superposes almost exactly on the Mn^{2+} ion in YvrK, while the C-terminal Mn^{2+} ion appears to have shifted by 1.8 Å (overall superposition shows a rms deviation of 1.7 Å). It is relevant, in this context, to note that the N-terminal domains of these three proteins appear to be more similar (~46% similar) than the overall degree of similarity of ~39%. It thus appears that not only the bicupin scaffold but also the active site geometry appears to have been maintained in the two enzymes from *B. subtilis*. The recent report on the characterization of YxaG mentions that this protein is sensitive to Fe(II) chelators, as in the extensively studied catechol dioxygenases (*6*, *7*). EPR studies using nitric oxide as a label confirmed the presence of an Fe^{2+} ion at the active site. Two *g* resonance peaks at 3.99 and 2.03 were used to arrive at this conclusion. Binding of quercetin resulted in the disappearance of the *g* resonance peak at 3.99, an indicator that the substrate blocks the nitric oxide label from binding to the Fe^{2+} ion. The Fe ion coordination number of 5 as seen in the two active sites of YxaG has been reported to be the most common conformational arrangement seen in the structures deposited in the Protein Data Bank (*23*).

The structure of YvrK (*5*) suggests a secondary hydrophobic shell to contribute toward substrate specificity. The only hydrophilic residues in either of the two domain active site lining regions in YvrK are Arg92, Arg270, and Glu333. These have been replaced with Phe, Ile, and Arg in YxaG, a reflection of the different substrates that interact with these two proteins. A relevant comparison in this regard is with phosphomannose isomerase, which binds mannose 6-phosphate and has an active site lined by hydrophilic residues. The surface potentials (*24*) of these reactive centers are compared in Figure 4. Four bicupins, YxaG, *A. japonicus* dioxygenase, YvrK, and phosphomannose isomerase, were examined for gross surface charge features that can dictate substrate interaction. The orientation of these four proteins seen in Figure 4 is that of looking down the two β-barrels of these bicupin structures. The charge distribution seen at this surface is that which is presented to an incoming substrate moiety. The most perceptible difference, while

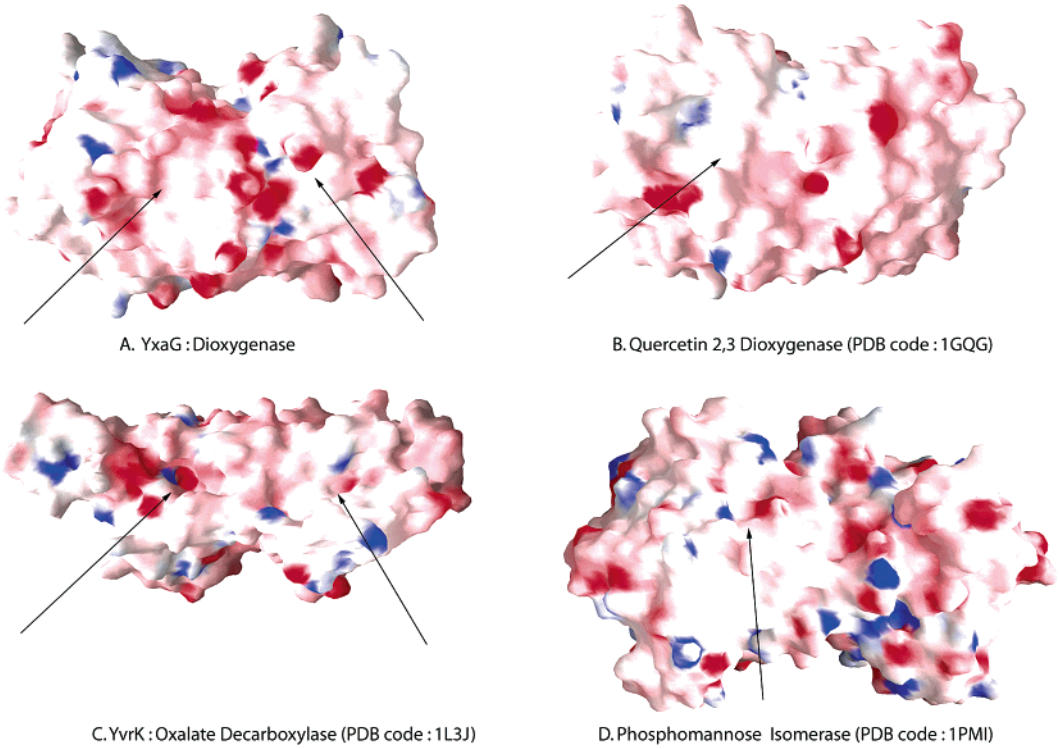

FIGURE 4: Surface charge differences may help in determining substrate specificity. Comparison of the surface-charged potential presented to the substrate moiety in four bicupins: YxaG, *A. japonicus* dioxygenase, *B. subtilis* YvrK, and phosphomannose isomerase. Substrate binding sites have been denoted with an arrow. Quercetin 2,3-dioxygenase (PDB entry 1GQG) and phosphomannose isomerase (PDB entry 1PMI) have only one active site located in the N-terminal cupin domain.

comparing these surfaces, is seen in Yxag which has a positively charged ridge that runs between the two active sites. This feature is not seen in the other bicupin structures that were examined, although other charge differences can be seen. The electrostatic surface representation thus points to features, away from the active site cavity, that may play a role in allowing for interaction with a range of substrates (oxalate in YvrK, quercetin in quercetin dioxygenase or mannose 6-phosphate in phosphomannose isomerase). Thus, alongside the shorter loops connecting the strands which make up the immediate scaffold of the active site (a feature which arises due to the different sequence spacing between the two cupin motifs), point variants among residues further away also contribute to substrate specificity.

Quaternary Organization. Table 2 summarizes the quaternary structures of the enzymes having a bicupin scaffold for which crystal structures are available to date. The two dimeric bicupins reported so far are the dioxygenases (*A. japonicus* quercetin 2,3-dioxygenase and YxaG). The major distinguishing feature in the dimerization of these two proteins is the presence of carbohydrates on the dimer interface in the *A. japonicus* protein, which the authors suggest is an important stabilizing component for the dimeric arrangement (*9*). YxaG dimerizes in the same orientation as does the *A. japonicus* protein, involving a substantial buried surface area (~2900 Å^2, ~66% of which is contributed by nonpolar residues) 500 Å^2 greater than that of the *A. japonicus* homologue (~2400 Å^2, ~63% of which is contributed by nonpolar residues). The increase in the amount of surface area buried in YxaG is probably a reflection of the lack of carbohydrate interactions in this protein. An interesting feature is that, even among the monocupins, apart from the exception of isopenicillin *N*-synthase (*25*), all other reported structures with the monocupin scaffold (*5*) are either dimers or hexamers, suggesting that the genomic events involving the formation of bicupins from monocupins may have been driven by the oligomerization potential of their protein products.

The major structural feature that appears to stabilize the hexameric organization of YvrK is the presence of interlocking clawlike α-helical structures that form the trimer. This feature is observed in the seed-storage proteins such as phaseolin, canavalin, and proglycinin, all of which are trimeric bicupins (*3*). YxaG lacks the clawlike protrusion; on the other hand, it has a large stretch of exposed hydrophobic residues that become buried upon dimer formation.

Kinetic Measurements and Influence of Different Metal Ions on Enzyme Activity. *A. japonicus* quercetin 2,3-dioxygenase, the first quercetinase to be biochemically and structurally characterized, was seen to possess one active site, bound to a Cu^{2+} ion. Biochemical reports characterizing *B. subtilis* YxaG have established it as an Fe^{2+}-containing quercetin 2,3-dioxygenase, albeit traces of other transition metal ions, namely, Cu^{2+} and Mn^{2+}, were also reported to be found (*6*, *7*). The completely conserved nature of the active site metal ion coordinating residues and the structural similarity between the N-terminal active site of *A. japonicus* quercetinase and the two sites of YxaG prompted us to examine the effect that a replacement of the Fe^{2+} ion at the active sites would have on enzyme activity. Various metal ions were examined (Figure 5) as potential replacements for the active site Fe^{2+} ions. The increase or decrease in enzyme activity when compared to the native enzyme (purified recombinant YxaG with no externally added metal ions) suggests an activity profile that has an approximate match

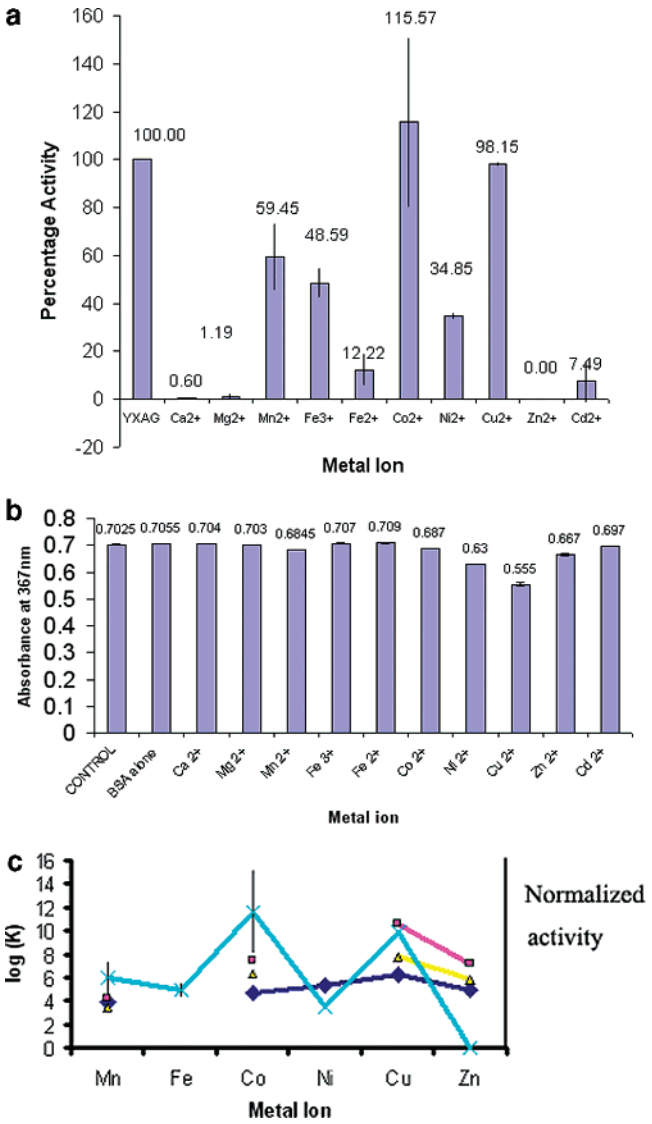

FIGURE 5: Activity profile for YxaG. (a) Activity assays were performed as described in Materials and Methods. The activity of freshly purified recombinant YxaG was treated as 100% in this experiment as no significant variation in this value was observed for different protein preparations. Metal ion-free YxaG was obtained as reported in Materials and Methods. The percentage reactivation of metal ion-free YxaG upon incubation with different metal ions is shown. These data suggest a higher activity with cobalt and copper ions at the active site as opposed to the Fe^{2+} ions located in the crystal structure and confirmed by EPR methods by other groups (*7*, *8*). (b) In a control experiment, quercetin itself acts as a chelator of metal ions. To account for this effect, a control experiment was performed where BSA was used instead of YxaG. This experiment shows the extent of variation in absorbance at 367 nm due to quercetin−metal ion interaction. This control experiment thus shows that the variation in YxaG activity upon its incubation with different metal ions is appreciably greater than the effects that can be ascribed to quercetin−metal ion interaction. (c) Normalized activity (with respect to the native enzyme) is overlaid on a plot depicting the Irving−Williams series. YxaG (cyan), His (log K_2) (yellow), His (log K_1) (red), and glutamate (log K) (black) are shown in this graph. The variation in enzyme activity appears to be correlated with the log K values of the His and Glu residues reported by Irving and Williams (*25*). The active site of YxaG has three His residues and one Glu residue.

to the Irving−Williams metal ions series based on the stability of metal ion complexes (*26*). The most noteworthy feature of this observation is that of Cu^{2+}, which appears to be most suited for the dioxygenase activity. These studies thus highlight the observations made previously (*13*, *27*), based on sequence and structural analysis, that Fe^{2+} preceded Cu^{2+} in the active sites of enzymes in the evolutionary ladder. The correlation also reinforces the observations made on carbonic anhydrase II (*28*) where the metal ion specificity was seen to be governed by a preferred coordination number and geometry of the metal ion.

The kinetic measurements on YxaG performed under the reaction conditions mentioned in Materials and Methods suggest a V_{max} value of 1.84 units/mg of protein and a K_m value of 1.5 μM. These results compare favorably with the values reported recently (*6*, *7*). One feature of the enzyme activity profile is that of a decrease in enzyme activity with an increase in substrate concentration. This profile can be explained as an artifact caused by the chelation of the active site Fe^{2+} ion by quercetin (*29*), as cooperative effects can be excluded given a normal Michaelis−Menten curve for the enzyme reaction (*6*). It is pertinent to note in this context that the "lid" of the active site cavity in YxaG is seen to be more rigid when compared to the *A. japonicus* quercetinase.

Conclusion. The crystal structure of *B. subtilis* YxaG reveals an Fe^{2+}-containing quercetin 2,3-dioxygenase with both sites showing full occupancy for the metal ion. Although biochemical analysis has shown quercetin to be the most likely substrate, it is possible that another natural substrate exists for this enzyme. Studies on enzymatic activity as a function of metal ion variation approximate a trend in the activity profile to the Irving−Williams series of metal ions. The structure also holds forth the tantalizing prospect of YxaG being a multifunctional enzyme, based on the variation seen in the metal ion coordination in the two active sites. The crystal structure and activity measurements for YxaG thus suggest a potential mode of regulation of enzyme activity by varying the metal ion on the active site.

ACKNOWLEDGMENT

We thank Dr. Andrezj Joachimiak for his helpful suggestions and access to beamline ID19 at the Advanced Photon Source.

REFERENCES

1. Anantharaman, V., Aravind, L., and Koonin, E. V. (2003) Emergence of diverse biochemical activities in evolutionarily conserved structural scaffolds of proteins, *Curr. Opin. Chem. Biol. 7*, 12−20.
2. Dunwell, J. M., Khuri, S., and Gane, P. J. (2000) Microbial relatives of the seed storage proteins of higher plants: Conservation of structure and diversification of function during evolution of the cupin superfamily, *Microbiol. Mol. Biol. Rev. 64*, 153−179.
3. Dunwell, J. M., Culham, A., Carter, C. E., Sosa-Aguirre, C. R., and Goodenough, P. W. (2001) Evolution of functional diversity in the cupin superfamily, *Trends Biochem. Sci. 26*, 740−746.
4. Tanner, A., Bowater, L., Fairhurst, S. A., and Bornemann, S. (2001) Oxalate decarboxylase requires manganese and dioxygen for activity, *J. Biol. Chem. 276*, 43627−43634.
5. Anand, R., Dorrestein, P. C., Kinsland, C., Begley, T. P., and Ealick, S. E. (2002) Structure of oxalate decarboxylase from *Bacillus subtilis* at 1.75 Å resolution, *Biochemistry 41*, 7659−7669.
6. Bowater, L., Fairhurst, S. A., Just, V. J., and Bornemann, S. (2004) *Bacillus subtilis* YxaG is a novel Fe-containing quercetin 2,3-dioxygenase, *FEBS Lett. 557*, 45−48.
7. Barney, B. M., Schaab, M. R., LoBrutto, R., and Francisco, W. A. (2004) Evidence for a new metal in a known active site: Purification and characterization of an iron-containing quercetin

2,3-dioxygenase from *Bacillus subtilis*, *Protein Expression Purif. 35*, 131−141.
8. Dai, Y., Wensink, P. C., and Abeles, R. H. (1999) One protein, two enzymes, *J. Biol. Chem. 274*, 1193−1195.
9. Fusetti, F., Schroter, K. H., Steier, R. A., van Noort, P. I., Pijning, T., Rozeboom, H. J., Kalk, K. H., Egmond, M. R., and Dijkstra, B. W. (2002) Crystal structure of the copper-containing quercetin 2,3-dioxygenase from *Aspergillus japonicus*, *Structure 10*, 259−268.
10. Oka, T., and Simpson, F. J. (1971) Quercetinase, a Dioxygenase containing copper, *Biochem. Biophys. Res. Commun. 43*, 1−5.
11. Hund, H.-K., Breuer, J., Lingens, F., Huttermann, J., Kappl, R., and Fetzner, S. (1999) Flavonol 2,4-dioxygenase from *Aspergillus niger* DSM 821, a type 2 Cu^{II}-containing glycoprotein, *Eur. J. Biochem. 263*, 871−878.
12. Steiner, R. A., Kalk, K. H., and Dijkstra, B. W. (2002) Anaerobic enzyme·substrate structures provide insight into the reaction mechanism of the copper-dependent quercetin 2,3-dioxygenase, *Proc. Natl. Acad. Sci. U.S.A. 99*, 16625−16630.
13. Steiner, R. A., Meyer-Klaucke, W., and Dijkstra, B. W. (2002) Functional analysis of the copper-dependent quercetin 2,3-dioxygenase. 2. X-ray absorption studies of native enzyme and anaerobic complexes with the substrates quercetin and myricetin, *Biochemistry 41*, 7963−7968.
14. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode, *Methods Enzymol. 276*, 307−326.
15. Vagin, A., and Teplykov, A. (2000) An approach to multi-copy search in molecular replacement, *Acta Crystallogr. D56*, 1622−1624.
16. Murshudov, G. N., Vagin, A. A., and Dodson, E. J. (1997) Refinement of molecular structures by the maximum-likelihood method, *Acta Crystallogr. D53*, 240−255.
17. Perrakis, A., Harkiolaki, M., Wilson, K. S., and Lamzin, V. S. (2001) ARP/wARP and molecular replacement, *Acta Crystallogr. D57*, 1445−1450.
18. Jones, T. A., Zou, J. Y., Cowan, S. W., and Kjeldgaard, M. (1991) Improved methods for the building of protein models in electron density maps and the location of errors in these models, *Acta Crystallogr. A47*, 110−119.
19. Murshudov, G. N., Lebedev, A., Vagin, A. A., Wilson, K. S., and Dodson, E. J. (1999) Efficient anisotropic refinement of molecular structures using FFT, *Acta Crystallogr. D55*, 247−255.
20. Hubbard, S. J., Campbell, S. F., and Thornton, J. M. (1991) Molecular recognition. Conformational analysis of limited proteolytic sites and serine proteinase protein inhibitors, *J. Mol. Biol. 220*, 507−530.
21. Steiner, R. A., Kooter, I. M., and Dijkstra, B. W. (2002) Functional analysis of the copper-dependent quercetin 2,3-dioxygenase. 1. Ligand-induced coordination changes probed by X-ray crystallography: Inhibition, ordering effect, and mechanistic insights, *Biochemistry 41*, 7955−7962.
22. Sjogren, T., and Hajdu, J. (2001) Structure of the bound dioxygen species in the cytochrome oxidase reaction of cytochrome cd1 nitrite reductase, *J. Biol. Chem. 276*, 13072−13076.
23. Harding, M. M. (2001) Geometry of metal−ligand interactions in proteins, *Acta Crystallogr. D57*, 401−411.
24. Nicholls, A., Sharp, K., and Honig, B. (1991) Protein folding and association: Insights from the interfacial and thermodynamic properties of hydrocarbons, *Proteins 11*, 281−290.
25. Roach, P., Clifton, I. J., Fulop, V., Harlos, K., Barton, G. J., Hajdu, J., Andersson, I., Schofield, C. J., and Baldwin, J. E. (1995) Crystal structure of isopenicillin N-synthase is the first from a new structural family of enzymes, *Nature 375*, 700−704.
26. Irving, H., and Williams, R. J. P. (1953) Stability of transition metal complexes, *J. Chem. Soc.*, 3192−3210.
27. Kooter, I. M., Steiner, R. A., Dijkstra, B. W., van Noort, P. I., Egmond, M. R., and Huber, M. (2002) EPR characterization of the mononuclear Cu-containing *Aspergillus japonicus* quercetin 2,3-dioxygenase reveals dramatic changes upon anaerobic binding of substrates, *Eur. J. Biochem. 269*, 2971−2979.
28. McCall, K. A., and Fierke, C. A. (2004) Probing determinants of the metal ion selectivity in carbonic anhydrase using mutagenesis, *Biochemistry 43*, 3979−3986.
29. Cornard, J. P., and Merlin, J. C. (2002) Spectroscopic and structural study of complexes of quercetin with Al(III), *J. Inorg. Biochem. 92*, 19−27.
30. Cleasby, A., Wonacott, A., Skarzynski, T., Hubbard, R. E., Davies, G. J., Proudfoot, A. E., Bernard, A. R., Payton, M. A., and Wells, T. N. (1996) The X-ray crystal structure of phosphomannose isomerase from *Candida albicans* at 1.7 Å resolution, *Nat. Struct. Biol. 5*, 470−479.

BI0484421