JBC Papers in Press. Published on April 1, 2004 as Manuscript M401692200

Crystal Structure of 4-chlorocatechol 1,2-dioxygenase from the Chlorophenol-utilizing Gram-positive *Rhodococcus opacus* 1CP

Marta Ferraroni[1], Inna P. Solyanikova[2], Marina P. Kolomytseva[2], Andrea Scozzafava[1], Ludmila Golovleva[2], Fabrizio Briganti[1*]

[1] Dipartimento di Chimica, Università di Firenze, Via della Lastruccia 3, I-50019 Sesto Fiorentino (FI), Italy

[2] Skryabin Institute of Biochemistry and Physiology of Microorganisms, Russian Academy of Sciences, 142290 Pushchino Moscow region, Russia.

***Corresponding Author:**
Prof. Fabrizio Briganti
Dipartimento di Chimica,
Università di Firenze,
Via della Lastruccia 3
I-50019 Sesto Fiorentino FI, Italy
e-mail: fabrizio.briganti@unifi.it

Running title: Crystal structure of 4-Chlorocatechol 1,2-dioxygenase

Summary

The crystal structure of the 4-chlorocatechol 1,2-dioxygenase from the Gram-positive bacterium *Rhodococcus opacus (erythropolis)* 1CP, a Fe(III) ion containing enzyme involved in the aerobic biodegradation of chloroaromatic compounds, has been solved by multiple wavelength anomalous dispersion using the weak anomalous signal of the two catalytic irons (1 Fe/257 amino acids) and refined at 2.5 Å resolution (Rfree 28.7%; R factor 21.4%).

The analysis of the structure and its comparison with the structure of catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus* ADP1 (*Ac* 1,2-CTD), highlights significant differences between these enzymes. The general topology of the present enzyme comprises two catalytic domains (one for each subunit) related by a non-crystallographic twofold axis and separated by a common 'α-helical zipper' motif which consists of five N-terminal helices from each subunit; furthermore the C-terminal tail is significantly shortened with respect to the known *Ac* 1,2-CTD. The presence of two phospholipids binding in a hydrophobic tunnel along the dimer axis is shown here to be a common feature for this class of enzymes.

The active site cavity presents several dissimilarities with respect to the known catechol cleaving enzyme. The catalytic non-heme iron(III) ion is bound to the sidechains of Tyr134, Tyr169, His194 and His196 and a co-crystallized benzoate ion, bound to the metal center, reveals details on a novel mode of binding of bidentate inhibitors and a distinctive hydrogen bond network with the surrounding ligands. Among the amino acid residues expected to interact with substrates, several are different from the corresponding analogs of *Ac* 1,2-CTD: Asp52, Ala53, Gly76, Phe78 and Cys224; in addition, regions of largely conserved amino acid residues in the catalytic cleft show different shapes resulting from several substantial backbone and side-chain shifts. The present structure is the first of intradiol dioxygenases which specifically catalyze the cleavage of chlorocatechols, key intermediates in the aerobic catabolism of toxic chloroaromatics.

Introduction

Halogenated aliphatic and aromatic hydrocarbons are a major environmental concern since large amounts of such toxic chemicals, resistant to both chemical oxidation and biological degradation, have been released into the ecosphere in the last decades as pesticides, fire retardants, solvents, etc. (1,2). The substitution for hydrogen with xenobiotic halogens in readily biodegradable hydrocarbons, has in fact led to substances with significantly decreased microbial decomposition and, consequently, to their environmental accumulation (3).

Specialized strains of aerobic and anaerobic bacteria have been discovered to use selected halogenated hydrocarbons as sole carbon and energy sources (1,4,5).

The aerobic catabolism of chloroaromatics usually occurs through two different pathways (5,6). Compounds containing one or two chlorine atoms are usually converted to chlorocatechols and then catabolized through the modified ortho-cleavage pathway (10,11). Aromatic compounds containing more than two chlorine atoms are converted to hydroxyquinol or chlorohydroxyquinols and then cleaved by specific intradiol dioxygenases (see Scheme 1) (12-18). A few interesting exceptions have been observed (7-9).

The modified *ortho* cleavage pathway is a central oxidative bacterial pathway that channels chlorocatechols, derived from the degradation of chlorinated benzoic acids, phenoxyacetic acids, phenols, benzenes, and other aromatics into the energy-generating tricarboxylic acid pathway (19-21) (see Scheme 1). These chlorocatechols are further degraded by the addition of molecular oxygen and the subsequent cleavage between two adjacent hydroxyl groups catalyzed by non-heme Fe(III)-dependent metalloenzymes, classified as intradiol dioxygenases (19).

Dioxygenases that cleave catechols in an intradiol fashion can be divided into two structurally different families: protocatechuate 3,4-dioxygenases (3,4-PCDs) and catechol 1,2-dioxygenases (1,2-CTDs). 3,4-PCDs, specific for hydroxybenzoates, are typically composed of two homologous subunits in large oligomeric complexes $[\alpha\beta Fe^{3+}]_{2-12}$ (22), catechol 1,2-dioxygenases are dimers $[\alpha Fe^{3+}]_2$ with identical or very similar subunits and generally comprise enzymes with markedly different substrate specificities: catechol, alkylated catechols, hydroxyquinols (type I enzymes) (23-27), and chlorocatechols (type II enzymes) (5, 28-41). Many of these enzymes have been extensively characterized in terms of biochemical, spectroscopic and kinetic properties (23-41). Mechanistic and structural details have also been gathered for one catechol 1,2-dioxygenase (42) and several protocatechuate 3,4-dioxygenases (43-47). In particular the crystal structures of protocatechuate 3,4-dioxygenases revealed that the catalytic ferric ion is ligated by two histidine and two tyrosine residues, plus a hydroxyl ion in a trigonal-bipyramidal geometry (43). These residues are conserved in all investigated intradiol dioxygenases. Further exhaustive studies by Lipscomb et al. described the mode of binding of substrates, substrate analogues, and inhibitors to 3,4-PCDs, and allowed to speculate on the subsequent interaction with dioxygen (45-47). Recently, structural information on the 1,2-CTD family was obtained, although limited to a single enzyme, the catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus* ADP1 (*Ac* 1,2-CTD hereafter). The structure of this enzyme markedly differs from those of 3,4-PCDs; it is characterized by the presence of a novel structural hydrophobic helical zipper motif as a subunit linker, and shows that only one subunit is used to fashion each active-site cavity (42). Furthermore 1,2-CTD anaerobic adducts with catechol and 4-methylcatechol confirmed several observations seen in 3,4-PCDs such as iron tyrosinate displacement upon substrate binding (42, 45-47).

Since the substrate specificities of till now characterized chlorocatechol 1,2-dioxygenases (type II enzymes) considerably differ from 1,2-CTDs (type I enzymes) (28-41), we believe that structural details on chlorocatechol cleaving dioxygenases are needed to further understand the mechanism of substrate selection.

X-ray absorption measurements on 4-chlorocatechol 1,2-dioxygenase from *Rhodococcus opacus (erythropolis)* 1CP (*Rho* 1,2-CCD hereafter) and other intradiol cleaving dioxygenases have provided the first insights into the recognition of different chemical structures by *ortho*-cleaving dioxygenases (48).

Diffraction quality crystals have been obtained and X-ray diffraction studies are in progress for other intradiol dioxygenases with different substrate specificities (49-54).

In this paper we report the crystal structure determination for *Rho* 1,2-CCD solved using a MAD experiment and refined at 2.5 Å. This is the first dioxygenase that catalyzes the degradation of chloro-catechols for which the three-dimensional structure has been determined providing details on the conformation of the active site and providing information on which particular residues are involved in substrate selection for enzymes of the 1,2-CCD family.

Experimental Procedures

Protein Preparation, Crystallization and Data Collection

Rho 1,2-CCD was purified from *Rhodococcus opacus* 1CP as previously reported (41). The enzyme was crystallized at 293 K using the sitting drop vapor diffusion method from a solution containing 1.6 M ammonium sulphate, 0.1 M sodium chloride, 100 mM Tris-HCl, pH 7.5 and 5-15% glycerol (53). All the reagents were of the best purity available from Sigma-Aldrich (Milano, IT).

Diffraction data extending to a maximum resolution of 2.5 Å was collected at the X-ray diffraction beamline at Elettra, Trieste. Data were collected using a MAR image plate detector and a wavelength of 1.0 Å. Crystals belong to the primitive hexagonal space group $P6_322$ with unit cell dimensions a=b=89.33, c=313.39. Assuming one molecule per asymmetric unit, the solvent content is about 61% of the unit cell (V_M= 3.13 $Å^3$/Da) (73). For all the data collections, crystals of the native enzyme were cooled at 100 K adding 30% glycerol to the mother liquor solution as cryoprotectant. Due to the large unit cell dimensions an image plate detector was utilized in order to reach the maximal resolution of the crystal without having too many overlapping reflections.

Structure Determination and Refinement

All molecular replacements attempts, using coordinates of known intradiol dioxygenases structures as a model, failed to provide a solution for *Rho* 1,2-CCD.

The structure of the enzyme *Rho* 1,2-CCD has been solved by multiple wavelength anomalous dispersion (MAD) using the weak anomalous signal of the two catalytic irons (1 Fe/257 amino acids). MAD data were collected at the BW7A beamline, EMBL, DESY, Hamburg. The data collected at three wavelengths (inflection, peak, remote) were processed

and integrated with DENZO and scaled by SCALEPACK, from the HKL program suite (55). Data processing statistics are summarized in Table 1.

The program SOLVE (56) was used to identify the two iron sites and for phases calculation. The 3.5 Å MAD phases were improved and extended to 2.5 Å by solvent flattening and histogram mapping using the program DM from the CCP4 program suite (57). The resulting electron density map was of high quality and allowed the manual tracing of 440 out of 514 amino acid residues using the program QUANTA (58). Refinement of the model, performed at 2.5 Å resolution, using the program Refmac 5.1.24 from the CCP4 program suite (57), resulted in R-factor and R-free values of 21.4 and 28.7 %, respectively.

The electron densities corresponding to residues Met A1, Met B1 and to the side chains of Lys A89, Lys B66, Lys B173, His B256, and Gln B255 were missing and therefore the corresponding atoms and bonds were not modeled. Electron density resembling one benzoate molecule bound to the iron ions was found in both active sites. Since no benzoate like compounds were added during the crystallization procedures a possible explanation could be that such compound is formed during cell growth. Inhibition tests revealed that benzoate is a weak competitive inhibitor for the present *Rho* 1,2-CCD. Electronic density corresponding to a Tris buffer molecule was also found in a surface crevice interacting with residues Thr28, Tyr31, Glu32, and Met35 (from subunit B, helix H2) and Pro 48, Asp52, Glu56 (subunit A, helices H3 and H4) Gln75, Gly76 and Pro77 (subunit A, random coil).

The final coordinates have been deposited in the Protein Data Bank (Protein Data Bank accession number 1S9A).

Structure Analysis

The stereochemical quality of the models was assessed using the program PROCHECK (59).

The Ramachandran plot is of good quality; there are 430 non-glycine and non-proline residues; among these, 374 (87%) are in the most favoured regions, 50 (11.6%) are in the additional allowed regions, 5 (1.2%) in the generously allowed regions and 1 (His B256) (0.2%) in disallowed regions.

The secondary structure was defined utilizing the DSSP database and program (60).

Multiple sequence alignments were performed using the ClustalX program (61).

Global structures superimpositions were carried out by utilizing the matching algorithm implemented into the HEX 4.1 program (62). Least squares fit of the active site regions were performed using the McLachlan algorithm as implemented in the program ProFit 2.2 (Martin, A.C.R., www.bioinf.org.uk/software/profit/) specifying as the fitting subset the four amino acid ligands to the catalytic iron ions (63).

Electrostatic Potentials were estimated first transforming the pdb coordinate file into a pqr file containing partial charges and radii for each atom by using the PDB2PQR web service and then solving the second order differential Poisson-Boltzmann equation which relates the electrostatic potential in a dielectric to the charge density using the macroscopic electrostatics with atomic detail (MEAD) program package (64,65).

3- and 4-substituted catechols were manually docked into the active site by first simulating the dissociation of Tyr169 from the iron center. As previously observed, it was assumed that the substrates bind to the iron in a bidentate fashion and with orientations of their aromatic ring similar to those observed for catechol or 4-methyl catechol in *Ac* 1,2-CTD or for the benzoate ion in *Rho* 1,2-CCD (42). Slight rotations and/or tilts of the iron bound substrate molecules did not result in changes in the amino acid residues interacting with the substrates ring substituents.

PyMol, UCSF Chimera, and MSMS were used to produce ribbon diagrams, electrostatic potential surfaces, electron density and other representations (66-68).

Results and Discussion

Overall structure and Linker Domain

The 1,2-CCD from the Gram-positive bacterium *Rhodococcus opacus (erythropolis)* 1CP is a homodimeric protein with overall dimensions 105 x 40 x 40 Å. The statistics for data collection, phasing and structure refinement are summarized in Table 1. The final model includes residues 2-257 for each monomer, two Fe(III) ions, two benzoate ions, two phospholipids, 284 water molecules and one Tris buffer molecule.

The global structure resembles that of catechol 1,2-dioxygenase from *Acinetobacter calcoaceticus* ADP1, the only known structure for this class of dioxygenases (see Figure 1)(42).

The overall fold of the dimer comprises two catalytic domains made up by a number of β-sheets and several random coils, related by a non crystallographic two-fold axis, and a linker domain composed by several α-helices coming from both monomers, and located at the center of the molecule. Two phospholipids are bound inside a channel formed by the two protein monomers at the center of the linker domain (see Figure1C). Each subunit entirely hosts a catalytic pocket containing an iron(III) ion accessible to the substrate from the hollow side of the dimer.

The linker domain is mainly composed by two long (H1 and H2) and three short (H3-H5) α-helices supplied by each subunit: four helices from the N-terminus of each monomer form a hook which is interacting with the equivalent motif from the other subunit and to the fifth helix which elongates from its catalytic domain (Figures 1 and 2).

ClustalX sequence alignments of *Rho* 1,2-CCD, with a variety of other chlorocatechol 1,2-dioxygenases and the representative of 1,2-CTDs: *Ac* 1,2-CTD are reported in Figure 3. The sequence identity between *Rho* 1,2-CCD and *Ac* 1,2-CTD is about 30%. Significant

differences are observable in the helical linker domain and in the C-terminal region by comparing 1,2-CCDs to *Ac* 1,2-CTD sequences.

Figure 4 shows the 3D structural least squares superposition of a single subunit of *Rho* 1,2-CCD and *Ac* 1,2-CTD. The first C-terminal α-helix and the following random coil region present in the conventional catechol 1,2 dioxygenases, are totally missing in *Rho* 1,2-CCD. The two long H1 and H2 helices in *Rho* 1,2-CCD (see Figure 2) closely match the second and third helices of *Ac* 1,2-CTD whereas the short H3 helix which provides several residues to the active site cavity (Leu49, Asp 52 and Ala 53) is partly shifted and about half of the fourth helix (H4) is missing with respect to the corresponding ones from *Ac* 1,2-CTD. On the contrary, the secondary structure of the central section of *Rho* 1,2-CCD thoroughly resembles that of the 1,2-CTD family. Finally, the N-terminal region of *Rho* 1,2-CCD misses the seventh helix, the last long random coil and the final β-sheet present in *Ac* 1,2-CTD.

The presence of two bound phospholipids seems to be a common feature of this class of enzymes (42). Their precise identity could not be determined due to the absence of the electron density of the head groups and the length of each tail was based on the length of the electron density and the stereochemistry of known phospholipids. In the present case they were modeled utilizing a phosphatidylcholine molecule with two C14/C15 hydrophobic tails. They are placed at the interface between the two subunits, one at each end of a large hydrophobic channel made up by helices H1 and H2 from both monomers, with the head group directed outwards into the solvent and the tail moieties pointing inwards towards each other (see Figures 1B and 1C). Various hypothesis have been made on the role of the hydrophobic tunnel and of the bound phospholipids but additional studies are needed to clarify their functional role (42).

The Catalytic domain

The core of the catalytic domain is made up by a series of β-sheets arranged in a β-sandwich conformation and by several random coils positioned between the linker domain and the β-sheets assemblies (see Figures 1 and 4). The active site metal center is located in a random coils region flanked on one side by the β-sheets assemblies of each monomer and on the other side by the α-helices of the linker domain.

The catalytic pocket of *Rho* 1,2-CCD is formed by several residues: Leu49, Asp52, and Ala53 from helix H3, Ile74, Gln75, Gly76, Pro77, Phe78, Phe79 from a single random coil, Trp126 from sheet S3, the iron ligand Tyr134 from a second random coil, Tyr169 from sheet S6, Ile171 from a third random coil, Arg191 from sheet S7, His194 and His196 from sheet S8, Gln210 from sheet S9, and Cys224 from another random coil.

Residues Leu49, Pro77, Phe78, Ile171, (Pro172), and the backbones of the iron ligands Tyr134 and Tyr169 compose the hydrophobic active site entrance.

Figure 5 shows the electrostatic potential mapped onto the molecular surfaces of *Rho* 1,2-CCD (A) and *Ac* 1,2-CTD (B) in the surroundings of the catalytic site entrances. Contrarily to what observed for 3,4-PCDs the groups neighboring the opening of the active site for both enzymes do not provide any marked positive electrostatic potential which is supposed to guide the substrate into the catalytic cleft for 3,4-PCDs (43). In fact in catechol cleaving dioxygenases the substrate binding driving forces are mainly of hydrophobic nature since the substrate catechol is mostly undissociated at physiological pH.

A close view of the active site residues, with the corresponding Fo–Fc density overlaid is represented in Figure 6A. The catalytic center contains a mononuclear iron(III) ion, bound to four amino acid residues Tyr 134, Tyr 169, His 194, and His196. The His_2Tyr_2 coordination is typical of all intradiol ring cleaving dioxygenases (26). X-ray absorption spectroscopy (XAS) data on the same chlorocatechol 1,2-dioxygenase indicates that the

native enzyme is pentacoordinated with two spheres of atoms: either two at 1.9 Å and three at 2.1 Å, or three at 1.9 Å and two at 2.1 Å (48). In the present crystal structure, a benzoate ion is found to bind to the iron ion in a bidentate asymmetric mode in place of the metal bound water molecule/hydroxide ion, generally found in the native enzymes, expanding the iron coordination number to six (Figure 6A). For this reason the iron ion bond distances observed in the present crystal structure are longer than those determined by the extended X-ray absorption fine structure (EXAFS) as well as those reported for the native *Ac* 1,2-CTD structure (see Table 2 for comparison). His194 exhibit the shortest bond (~2.0 Å); Tyr134, His196, and Tyr169 have longer bond distances (~2.2 Å) and the benzoate molecule asymmetrically bound with O1 (~2.5 Å) and O2 (~2.1 Å) completes the iron coordination sphere (see also Table 2).

Extensive studies on inhibitors, substrates and substrate analogs adducts of 3,4-PCAs and Ac 1,2-CTD have revealed several important features of the mechanism of exogenous ligands binding to their active site: phenolate competitive inhibitors bind to iron in a monodentate way; those showing low affinity generate distorted trigonal bipiramidal iron coordination geometries with the phenolate displacing the water/hydroxide ligand whereas the highest affinity inhibitors bind to iron concomitantly to the solvent molecule generating six-coordinated distorted octahedral metal coordination geometries (45,47). It has also been observed that substrates (in anaerobic conditions) or substrate-analogs bind to the iron center in a asymmetric bidentate mode resulting in distorted trigonal bipiramidal iron geometry due to the simultaneous dissociation of the axial tyrosinate and of the water/hydroxide iron ligands, both supposed to function as proton acceptors during substrate deprotonation and binding. Therefore, in the anaerobic substrate adducts, the sixth coordination site is unoccupied whereas in substrate-analogs adducts it is occupied by water or cyanide (if added to the buffer)(42, 45-47). Octahedral iron coordination is also observed in 3,4-PCDs at low

pH where a sulfate ion completes the coordination sphere of the active metal center (47). In the present structure we notice a different mode of binding of the iron chelating benzoate which causes the dissociation of the water/hydroxide iron ligand but it does not trigger the detachment of the tyrosinate residue thus resulting in a distorted octahedral iron coordination.

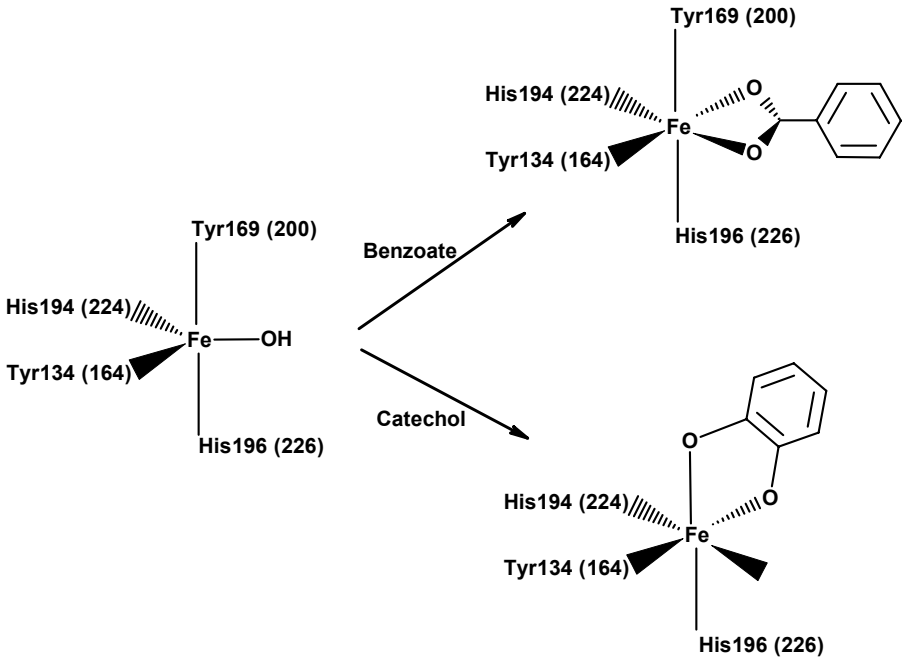
In Figures 6B-D are shown different views of the solvent-exposed surface in the active site region of Rho 1,2-CCD color coded for the calculated electrostatic potential. By positioning the enzyme molecule with the iron ion at the bottom of the catalytic pocket, we observe a region, made up by the iron ligands His194, His196, by the backbone nitrogen of Phe78, by the ring nitrogen of Trp126 and, on the opposite side, by Arg191 and Gln210, with positive electrostatic potential surrounding the metal ion, which should support the substrate/inhibitor deprotonation and binding. The active site zone opposite to the iron ion is, in contrast, made up by the negatively charged residue Asp52 and a negative electrostatic potential is also surrounding Glu56, Thr177, and Met181. This negative potential area could have implications for the correct orientation of the catechol substrates. The intermediate active site region is made up by residues Ala53, Ile74, by the backbones of Gly76 and Pro77 and particularly by Cys224 which has been previously suggested to interact with the chlorine substituents of the aromatic ring of the substrate molecules (42).

Figure 7A reports the superposed active sites of *Rho* 1,2-CCD and native *Ac* 1,2-CTD obtained by minimizing onto the four amino acid iron ligand residues (Tyr 134/164, Tyr 169/200, His 194/224, and His196/226 for *Rho* 1,2-CCD/*Ac* 1,2-CTD)(62). Several amino acid substitutions are observed into the catalytic cavity, the most evident being Asp52/Pro76, Ala53/Gly77, Phe78/Leu109, and Cys224/Ala254 (*Rho* 1,2-CCD/*Ac* 1,2-CTD), whereas the other residues are basically conserved although significant backbone and side chain shifts are detected. As reported above, in *Rho* 1,2-CCD the benzoate ion is asymmetrically chelated to iron and it is stabilized by an hydrogen-bonding network which connects the benzoate O2

atom to Arg191 NH1 (hydrogen-bonded further to Gln210, Asp222 and Tyr212) and the benzoate O1 atom to a well ordered W10 active site water molecule (hydrogen-bonded further to Gln75 and Trp126). The binding of benzoate also results in the absence of a second active site water molecule (W572 in *Ac* 1,2-CTD) caused by the conformational orientations of Arg191 and of the hydrogen-bonded Gln210 which resulted to be very similar to those observed when catechols bind to *Ac* 1,2-CTD (although benzoate is not able to trigger the dissociation of Tyr169). Arg191 and Gln210, conserved across all 1,2-CTDs and 3,4-PCDs, are supposed to promote the substrate positioning and deprotonation, providing significant van der Waals interactions with the ring to prevent its rotation into two equatorial positions and/or providing an electrostatic countercharge to the build up of electron density on a carbon atom of the ring (42, 44-47).

The structural alignment of the active site residues of the catecholate complex of *Ac* 1,2-CTD and *Rho* 1,2-CCD is shown in Figure 7B. The main interactions of *Ac* 1,2-CTD with the substrate involve residues Leu73, Pro76, Ile105, Pro108, Leu109, Arg221, Phe253 and Ala254. The catechol molecule in *Ac* 1,2-CD binds to the iron ion with one hydroxyl group in an axial position *trans* to His226 and the other in an equatorial position *trans* to Tyr164.

The benzoate observed in *Rho* 1,2-CCD binds to the metal center occupying the ligand position *trans* to His194 which is left empty in the productive catechol complex (in *Ac* 1,2-CTD) for the subsequent binding of the oxygen molecule (see scheme 2, the numbering in parenthesis refers to *Ac* 1,2-CTD). Although the mode of benzoate binding to *Rho* 1,2-CCD is radically different from that of catechol in *Ac* 1,2-CTD the aromatic ring orientation is the same (see Figure 7B).



Scheme 2

The catalytic mechanism of intradiol cleavage dioxygenases has been proposed to proceed via activation of the catechol substrate by iron(III) to give an iron(II) semiquinone, which reacts directly with dioxygen to give a hydroperoxide intermediate, which then undergoes Criegee rearrangement via acyl migration to give muconic anhydride (69). An alternative mechanism for migration of the electron-deficient acyl group, via a benzene-oxide–oxepin interconversion, has also been proposed (70). A comparison of *Rho* 1,2-CCD with 3,4-PCDs and *Ac* 1,2-CTD indicates that residues surrounding the postulated oxygen-binding pocket are partially conserved, in particular Pro77 (Pro108 and Pro15 in *Ac* 1,2-CTD and *Pseudomonas putida* 3,4-PCDs respectively) is retained, whereas Phe 78 is substituted by Leu 109 and by Tyr16 in *Ac* 1,2-CTD and *Pseudomonas putida* 3,4-PCDs respectively, nevertheless the side-chains of such residues point towards the external part of the cavity and the backbone portion which is near to the proposed oxygen-binding location is mainly unaffected by such substitutions suggesting that the oxygen molecule has similar interactions in all intradiol ring cleaving dioxygenases (45).

The residues surrounding the ring of the benzoate molecule and supposed to be involved in the correct positioning of the aromatic substrate are Leu49, Asp52, Ala53, Ile74, Gly76, Pro77, Phe78, Gln210 and Cys224. The present enzyme is able to catalyze the ring opening of a variety of substituted catechols (41). We attempted to dock 3- or 4-chloro-catechols into the active site. In the case of the 3-chloro-derivative, two different orientations are possible: if the substituent is oriented towards the internal part of the cavity it would settle into a pocket formed by residues Ile74 and Gly76 (both conserved in 1,2-CTDs), but if the substituent is oriented outwards it would essentially interact with residue Phe78 (substituted for Leu109 in *Ac* 1.2-CTD). Nevertheless, the preferred substrates for *Rho* 1,2-CCD are those bearing a substituent in position 4. If the 4-substituted catecholate binds with the substituent oriented towards the inner part of the catalytic site, as observed in the 4-methylcatechol adduct of *Ac* 1,2-CTD, it would be hosted in a cleft modeled by Asp52, Ala53 (replaced with Pro76 in *Ac* 1,2-CTD), Gly 76, and Cys224 (Ala254 in *Ac* 1,2-CTD), whereas if it prefers the outwards orientation it would mainly interact with Leu49 (conserved in *Ac* 1,2-CTD).

From the ClustalX sequence alignments of *Rho* 1,2-CCD, with a variety of other chlorocatechol 1,2-dioxygenases (Figure 3) it can be noticed that several regions contributing residues to the active-site cavity show significant amino acid substitutions. In particular, the 1,2-CCDs from *P. putida* pAC27, *R.. euthropha* pJP4, *P. chlororaphis* RW71 and the 3-chlorocatechol specific 1,2-CCD from *Rh. Opacus* CP1 (same strain expressing the present enzyme) exhibit higher activities towards 3-chlorocatechol than *Rho* 1,2-CCD. Possible reasons for the different observed specificities could be the replacement of Phe 78, a residue expected to interact with substituents in position 3, with Tyr and in the most internal side of the cavity Ala 53 which is exchanged with Val in all the other 1,2-CCDs described (35,36,41,71). Furthermore, Cys 223 and 224, suggested to be important for interacting with chlorine substituents are generally conserved in CCDs, as shown in Figure 3, but in the

present enzyme only Cys 224 is present whereas residue 223 is substituted for Ser (42). It has also to be noted that in the recently crystallized 3-chlorocatechol preferring 1,2-CCD from *Rh. opacus* CP1 both cysteines 223 and 224 are replaced by Val and Ala respectively, suggesting that the selection for chlorine substituents is a complex issue (54).

Conclusions

The present crystal structure evidences significant differences between *Rho* 1,2-CCD and 1,2-CTDs. Several secondary structure deletions are observed but in particular a number of changed residues inside the active cleft are believed to be responsible for substrate selection. The most significant should be Asp52, Ala53, Gly76, Phe78 and Cys224 which seem to be directly involved in interactions with the differently substituted substrates. Few amino acid differences (Phe78Tyr and/or Ala53Val) between the present enzyme which exhibits a marked specificity for 4-substituted substrates and a series of other CCDs more active with 3-substituted catechols, are believed to be responsible for the observed substrate selectivity differences among chlorocatechol 1,2-dioxygenases.

Acknowledgements:

We gratefully acknowledge the technical skills of Dr. Samuele Ciattini. This work was supported by the European Commission RTD Programme Copernicus grant ICA2-CT-2000-10006. We further acknowledge the "European Community - Access to Research Infrastructure Action of the Improving Human Potential Programme to the EMBL Hamburg Outstation, contract number: HPRI-CT-1999-00017". Finally, the contribution of the Italian Ministero Università e Ricerca Scientifica, Cofin 2002 funding and the support of the Gruppo Nazionale di Ricerca per la Difesa dai Rischi Chimico-Industriali ed Ecologici - C.N.R. are also acknowledged.

References

1. Ghosal D., You, I.-S., Chatterjee, D.K., Chakrabarty, A.M. (1985) *Science* **228**, 135-142.
2. Alexander M. (1981). *Science* **211**, 132-138.
3. Haggblom M.,Valo R.,(1995) in *Microbial Transformation and Degradation of Toxic Chemicals*, Willey-Liss. Inc., 389-439.
4. Frantz, B. & Chakrabarty, A.M. (1986) in *The bacteria*, vol. 10, *Ac*ademic Press, New York.
5. Reineke, W. & Knackmuss, H.-J. (1988). *Annu. Rev. Microbiol.* **42,** 263-287.
6. Chaudhry, G.O. & Chapalamadugu, S. (1991). *Microbiol. Rev.* **55**, 59-79.
7. van der Meer, J. R., van Neerven A. R. W., de Vries E. J., de Vos W. M., and Zehnder A. J. B. (1991) *J. Bacteriol*. **173**, 6–15.
8. Potrawfke, T., Armengaud, J., Wittich, R.-M. (2001). *J. Bacteriol.* **183**, 997-1011.
9. Kaschabek, S.R., Kasberg, T., Muller, D., Mars, A.E., Janssen, D.B. & Reineke, W. (1998). *J. Bacteriol.* **180**, 296-302.
10. Tiedje, J.M., Duxbury, J.M., Alexander, M. & Dawson J.E. (1969). *J. Agric. Food Chem.* **17**, 1021-1026.
11. Duxbury, J.M., Tiedje, J.M., Alexander, M. & Dawson J.E. (1970). *J. Agric. Food Chem.* **18**, 199-201.
12. Apajalahti, J.H.A., Salkinoja-Salonen M.S. (1987). *J.Bacteriol.* **169**, 5125-5130.
13. Joshi, D.K., Gold, M.H. (1993). *Appl. Environ. Microbiol.* **59**, 1779-1785.
14. Kozyreva, L.P., Shurukhin, Y.U., Finkel'shtein, Z.I., Baskunov, B.P. & Golovleva, L.A. (1993). *Mikrobiologiya* **62,** 78-85.
15. Li, D.-Y., Eberspächer, J., Wagner, B., Kuntzer, J. & Lingens, F. (1991). *Appl. Environ.* 57, 1920-1928.

16. Sangodkar, M., Chapman, P.J., Chakrabarty, A.M. (1988). *Gene* **71**, 267-277.

17. Sze, I.S. & Dagley, S. (1984). *J. Bacteriol.* **159**, 353-359.

18. Travkin, V.M., Jadan, A.P., Briganti, F., Scozzafava, A. & Golovleva, L.A. (1997). *FEBS Lett.* **407**, 69-72.

19. Bollag, J.-M., Briggs, G.G., Dawson, J.E. & Alexander, M. (1968). *J. Agric. Food Chem.* **16**, 829-833.

20. Duxbury, J.M., Tiedje, J.M., Alexander, M. & Dawson J.E. (1970). *J. Agric. Food Chem.* **18**, 199-201.

21. Tiedje, J.M., Duxbury, J.M., Alexander, M. & Dawson J.E. (1969). *J. Agric. Food Chem.* **17**, 1021-1026.

22. Que L. Jr., Ho R.Y.N. (1996) Chem. Rev. *96*, 2607-2624.

23. Hammer A., Stolz A., Knackmuss H. (1996) *Arch.Microbiol.* 166, 92-100.

24. Nakai C., Horiike K., Kuramitsu S., Kagamiyama H., Nozaki M.. **(**1990) *J. Biol. Chem*. **265,** 660–665.

25. Maltseva O.V., Solyanikova I.P., Golovleva L.A.. (1991) *Biokhimiya* **56**, 2188–2197. (*Biochemistry*(Moskow) **56**, 1548–1555.)

26. Briganti F., Pessione E., Giunta C., Scozzafava A. (1997) *FEBS Letters* **416**, 61-64.

27. Briganti F., Pessione E., Giunta C., Mazzoli R., Scozzafava A. (2000) *J. Protein Chem.* **19**, 709-716.

28. Moisseeva O.V., Linko E.V., Baskunov B.P. & Golovleva L.A. (1999) *Microbiology* 68:400-405 Translation of Mikrobiologiya (Russia) **68,** 461-466

29. Moisseeva, O.V., Belova, O.V., Solyanikova, I.P., Schlomann, M., Golovleva, L.A. (2001). Biochemistry (Mosc) **66**, 548-555.

30. Solyanikova I.P., Maltseva O.V., Vollmer M.D., Golovleva L.A. & Schlömann M. (1995) *J. Bacteriol*. **177:** 2821-2826.

31. Gorlatov S.N., Maltseva O.V., Shevchenko V.I. & Golovleva L.A. (1989) *Microbiology,* 58: 647-651. Translation of Mikrobiologiya (Russia) **58**, 802-806

32. Dorn, E. & Knackmuss, H.-J. (1978). *Biochem. J.* **174**, 73-84.

33. Dorn, E. & Knackmuss, H.-J. (1978). *Biochem. J.* **174**, 85-94.

34. Ngai, K.-L. & Ornston, L.N. (1988). *J. Bacteriol.* **170,** 2412-2413.

35. Pieper, D., Reineke, W., Engesser, K.-H. & Knackmuss, H.-J. (1988). *Arch. Microbiol.* **150**, 95-102.

36. Broderick, J.B. & O'Halloran, T.V. (1991). *Biochemistry* **30**, 7349-7358.

37. Solyanikova, I.P., Maltseva, O.V. & Golovleva, L.A. (1992). *Biochemistry (Mosc)* **57**, 1310-1316.

38. Hinteregger, C., Loidl, M. & Streichsbier, F. (1992). *FEMS Microbiol. Lett.* **97**, 261-266.

39. Bhat, M.A., Ishida, T., Horiike, K., Vaidyanathan, C.S. & Nozaki, M. (1993). *Arch. Biochem. Biophys.* **300**, 738-746.

40. Miguez, C.B., Greer, C.W. & Ingram J.M. (1993). *Can. J. Microbiol.* **39**, 1-5.

41. Maltseva, O.V., Solyanikova, I.P., Golovleva, L.A. (1994). *Eur. J. Biochem.* **226**, 1053-1061.

42. Vetting M.W & Ohlendorf D.H. (2000). *Structure Fold Des.* **39**, 429-440.

43. Ohlendorf, D.H., Lipscomb, J.D., Weber, P.C. (1988). *Nature(London)* **336**, 403-405.

44. Orville, A.M, Elango N., Lipscomb, J.D., Ohlendorf, D.H. (1997) *Biochemistry* **36**, 10039-10051.

45. Orville, A.M., Lipscomb, J.D., Ohlendorf, D.H. (1997). *Biochemistry* **36**, 10052-10066.

46. Elgren, T.E., Orville, A.M., Kelly, K.A., Lipscomb, J.D., Ohlendorf, D.H. & Que, Jr. L. (1997). *Biochemistry* **36**, 11504-11513.

47. Vetting M.W., D'Argenio D.A., Ornston L.N. & Ohlendorf D.H. (2000). *Biochemistry* **39**, 7943-7955.

48. Briganti, F., Mangani, S., Pedocchi, L., Scozzafava, A., Golovleva L.A., Jadan, A.P. & Solyanikova, I.P. (1998). *FEBS Lett.* **433,** 58-62.

49. Benvenuti, M., Briganti, F., Scozzafava, A., Golovleva, L., Travkin, V.M., Mangani, S. (1999). *Acta Crystallogr. D Biol. Crystallogr.* **55**, 901-903.

50. Earhart, C.A., Radhakrishnan, R., Orville, A.M., Lipscomb, J.D. & Ohlendorf, D.H. (1994a). *J. Mol. Biol.* **236**, 374-376.

51. Earhart, C.A., Hall, M.D., Michaud-Soret, I., Que, Jr. L. & Ohlendorf, D.H. (1994b). *J. Mol. Biol.* **236**, 377-378.

52. Ludwig, M.L., Weber, L.D. & Ballou D.P. (1984). *J. Biol. Chem.* **259**, 14840-14842.

53. Ferraroni M., Ruiz Tarifa M.Y., Briganti F., Scozzafava A., Mangani S., Solyanikova I.P., Kolomytseva M.P., Golovleva L.A., (2002) *Acta Crystallogr.* **D 58**, 1074.

54. Ferraroni M., Ruiz Tarifa M. Y., Scozzafava A., Solyanikova I. P., Kolomytseva M. P.,. Golovleva L, Briganti F. (2003) *Acta Crystallogr* **D 59**, 188-190.

55. Terwillinger, T.C. & Berendzen, J. (1999) *Acta Crystallogr.* **D 55**, 849-861.

56. Otwinowski, Z. & Minor, W. (1997) *Methods Enzymol.* **276**, 307-326 .

57. Collaborative Computational Project, Number 4. (1994) *Acta Crystallogr.* **D 50**, 760-763.

58. Quanta Simulation, Search, and Analysis, July1997, San Diego Molecular Simulation Inc.

59. Laskowski R A, MacArthur M W, Moss D S & Thornton J M (1993). *J. Appl. Cryst.*, **26**, 283-291.

60. Kabsch,W. and Sander,C. (1983) *Biopolymers* **22**, 2577-2637

61. Thompson,J.D., Gibson,T.J., Plewniak,F., Jeanmougin,F. and Higgins,D.G. (1997) Nucleic *Ac*ids Research, **24**, 4876-4882.

62. Ritchie D.W. and Kemp G.J.L. (1999) *J. Comp. Chem.* **20**, 383-395.

63. McLachlan A.D. *Acta Cryst*. **A38**, 871-873.

64. Nielsen JE, Andersen KV, Honig B, Hooft RW, Klebe G, Vriend G, Wade RC (1999) *Protein Eng.* **12,** 657-62.

65. Bashford, D. (1997) in Yutaka Ishikawa, Rodney R. Oldehoeft, John V.W. Reynders, and Marydell Tholburn, editors, Scientific Computing in Object-Oriented Parallel Environments, volume 1343 of Lecture Notes in Computer Science, pages 233-240, Berlin,. ISCOPE97, Springer.

66. DeLano, W.L. The PyMOL Molecular Graphics System (2002) DeLano Scientific, San Carlos, CA, USA.

67. Huang, C.C., Couch, G.S., Pettersen, E.F., and Ferrin, T.E. (1996) *Pacific Symposium on Biocomputing* **1**, 724.

68. Sanner, M.F., Olson, A.J., and Spehner, J.C. (1996) *Biopolymers* **38**, 305-320.

69. Jang HG, Cox DD, Que L Jr (1991) *J Am Chem Soc* **113**, 9200-9204.

70. Eley KL, Crowley PJ, Bugg TDH (2001) *J Org Chem* **66**, 2091-2097.

71. Potrawfke T., Timmis K.N., Wittich R.-M. (1998) *Appl. Environ. Microbiol.* **64**, 3798-3806.

72. Luzzati V. (1952) *Acta Crystallogr* **5**, 802-810.

73. Mattews B.W. (1968) *J.Mol. Biol.* **33**, 491-497.

Keywords

dioxygenase, chlorocatechol, intradiol cleavage, 4-chlorocatechol, X-ray crystallography, aromatic catabolism

Abbreviations

1,2-CCD, Chlorocatechol 1,2-dioxygenase

1,2-CTD, Catechol 1,2-dioxygenase

3,4-PCD, Protocatechuate 3,4-dioxygenase

MAD, Multiple wavelength anomalous dispersion

TRIS, tris(hydroxymethyl)aminomethane

Figure Legends

Figure 1 Schematic representation of the overall structures of (A) *Ac* 1,2-CTD and (B) *Rho* 1,2-CCD, (C) 90° rotated (B) around the horizontal axis. The iron ions are represented as pink spheres; the phospholipid molecules are also shown.

Figure 2 Schematic Representation of the secondary structure of *Rho* 1,2-CCD assigned using the program DSSP (60).

Figure 3 Amino acid sequence alignments of Chlorocatechol dioxygenases and *Ac* 1,2-CTD (61). (a) *Rhodococcus opacus* 1CP 3-Chlorocatechol specific CCD (NCBI protein accession O67987), (b) *Rhodococcus opacus* 1CP CCD (NCBI protein accession CAD28142, present enzyme), (c) *Achromobacter xylosoxidans* A8 CCD (NCBI protein accession CAD56206), (d) *Ralstonia eutropha* NH9 CCD (NCBI protein accession BAA74530), (e) *Delftia acidovorans* P4a CCD (NCBI protein accession AAC35836), (f) *Pseudomonas chlororaphis* RW71 CCD (NCBI protein accession CAB52140), (g) *Pseudomonas sp.* P51 CCD (NCBI protein accession P27098), (h) *Burkolderia cepacia* 2a CCD (NCBI protein accession AAK81678), (i) *Variovorax paradoxus* TV1 CCD (NCBI protein accession AAC01745), (j) *Ralstonia eutropha* plasmid pJP4 CCD (NCBI protein accession P05403), (k) *Pseudomonas putida* plasmid pAC27 CCD (NCBI protein accession A27058), (l) *Ralstonia sp.* JS705 CCD (NCBI protein accession CAA06968), (m) *Pseudomonas aeruginosa* CCD (NCBI protein accession T44616), (n) *Pseudomonas aeruginosa* JB2 CCD (NCBI protein accession AAC69474), (o) *Pseudomonas sp.* GT241-1 CCD (NCBI protein accession AAR88247), (p) *Ralstonia eutropha* JMP134 CCD (NCBI protein accession AAC44730), (q) *Burkolderia sp.* NK8 CCD (NCBI protein accession BAB56009), (r) *Burkolderia sp*. R172 CCD (NCBI protein accession AAP92136),

(s) *Defluvibacter lusatiensis* CCD (NCBI protein accession CAD60254), (t) *Acinetobacter calcoaceticus* ADP1 CTD (NCBI protein accession P07773). In the line above the alignment the strongly conserved positions are indicated in the following ways: '*****' positions which have a single, fully conserved residue; '**:**' positions with fully conserved 'strong' groups; '**.**' positions with fully conserved 'weaker' groups, according to ClustalX (61).

Figure 4 (A) Least squares superimposition of a single subunit of *Ac* 1,2-CTD (green) and *Rho* 1,2-CCD (red). (B) 90° horizontally rotated (A).

Figure 5 Electrostatic Potential Surfaces of (**A**) *Rho* 1,2-CCD and (**B**) *Ac* 1,2-CTD. The active site entrances are indicated by yellow circles. The blue color corresponds to positive and red to negative potentials.

Figure 6 Fo-Fc Electron density map for the active site of *Rho* 1,2-CCD (**A**). The electron density is contoured at 3σ. (**B-D**): representations of the solvent-accessible surface in the active site region of *Rho* 1,2-CCD. The surface is color coded on the basis of the calculated electrostatic potential. The blue color corresponds to positive and red to negative potentials. The iron ion is represented as a green sphere. The three pictures show 90° vertically rotated views of the cavity, the active site entrance is indicated by a yellow arrow.

Figure 7 Superposition of the active sites structure of *Rho* 1,2-CCD to native *Ac* 1,2-CTD (orange, pdb code: 1dlm) (**A**) and to the catechol complex of *Ac* 1,2-CTD (magenta, pdb code: 1dlt) (**B**).

Table 1.

Summary of Data Collection and Atomic Model Refinement Statistics

	native	Fe peak	Fe inflection	Fe remote
Data Collection				
Wavelength (Å)	1.0	1.7436	1.7439	1.5806
Limiting resolution (Å)	2.5	3.0	3.0	3.0
Unique reflections	26170	15072	14647	14376
R_{sym} (%)[a]	9.8 (67.9)	7.1 (30.7)	6.8 (31.6)	8.4 (38.1)
Multiplicity	4.4	13.2	12.6	12.7
Completeness overall (%)	92.23(99.7)	95.5 (95.4)	92.4 (93.7)	94.5 (95.2)
$\langle I/\sigma(I)\rangle$	13.2(2.8)	32.7 (5.7)	33.3 (6.3)	30.5 (5.3)
Overall figure of merit				
Before density modification	0.56			
After density modification	0.84			
Refinement				
Resolution range (Å)	20-2.5			
Unique reflections, working/free	24226/1288			
R_{factor} (%)[b]	21.4			
R_{free}(%)	28.7			
Non-hydrogen atoms	4147			
Water molecules	284			
r.m.s.d. bonds(Å)	0.010			
r.m.s.d. angles (°)	1.298			

Numbers in parentheses are for the highest resolution shell.

[a] $R_{sym} = \Sigma |I - \langle I\rangle| / \Sigma I$

[b] $R_{factor} = \Sigma |F_{obs} - F_{calc}| / \Sigma F_{obs}$

Table 2.

Bond Distances and Angles of the Catalytic Iron Ion Coordination Polyhedron for Rho 1,2-CCD*

	Distance X-Fe (Å) *Rho* 1,2-CCD *(Ac 1,2-CTD)*	Angle X-Fe-Y134	Angle X-Fe-Y169	Angle X-Fe-H194	Angle X-Fe-H196
Tyr 134/ (*164*)	**2.2 (*1.85*)**	**-**	**92.2**	**88.9**	**84.8**
Tyr 169 (*200*)	**2.2 *(2.20)***		**-**	**93.1**	**170.3**
His 194 (*224*)	**2.0 (*2.15*)**			**-**	**96.0**
His 196 (*226*)	**2.2 (*2.0*)**				**-**
Benz O1	**2.5**				
Benz O2	**2.1**				

* The bond distances in parentheses are those observed in *Ac* 1,2CTD.

Estimated bond distance errors for *Rho* 1,2-CCD structure are ~0.35 Å, based on Luzzati analysis (72).

Figure 1

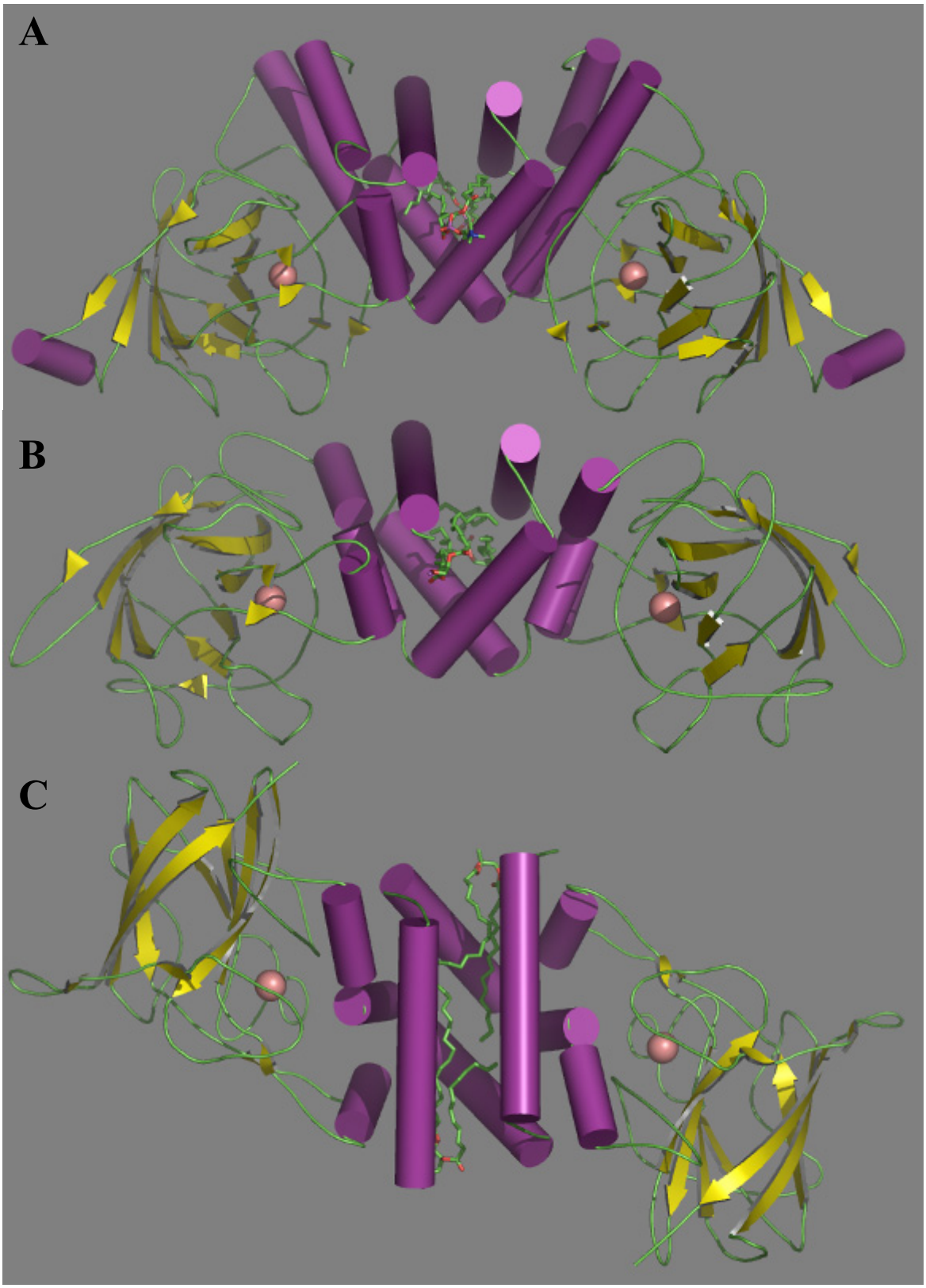


Figure 2

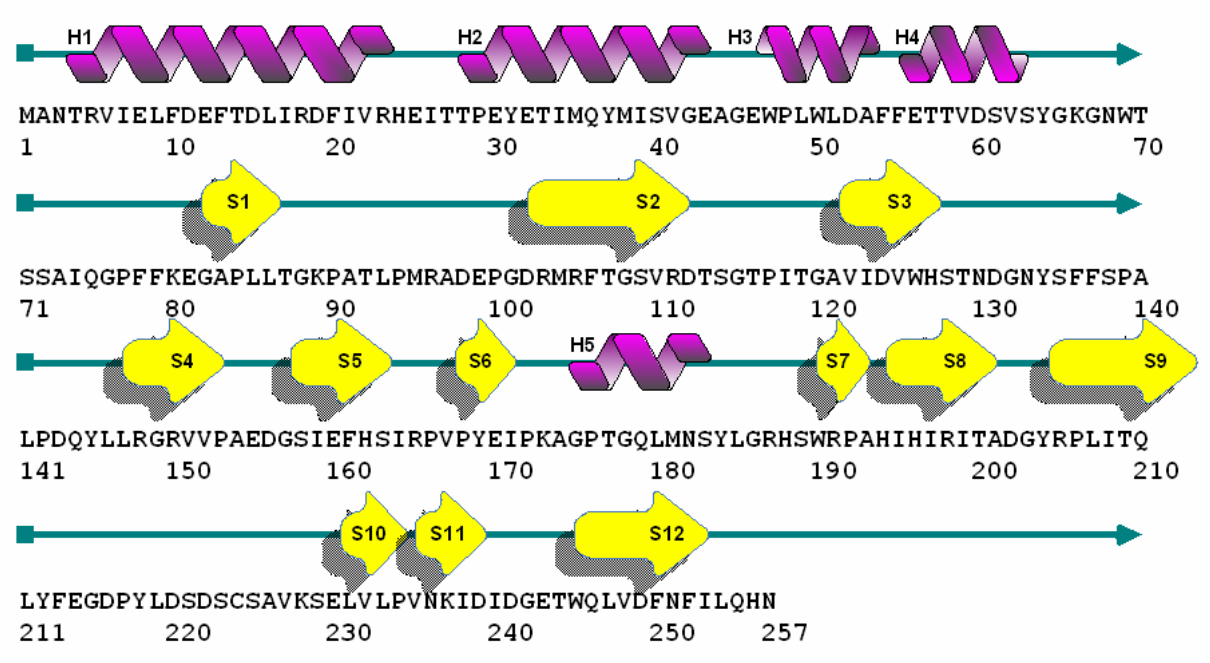


Figure 3

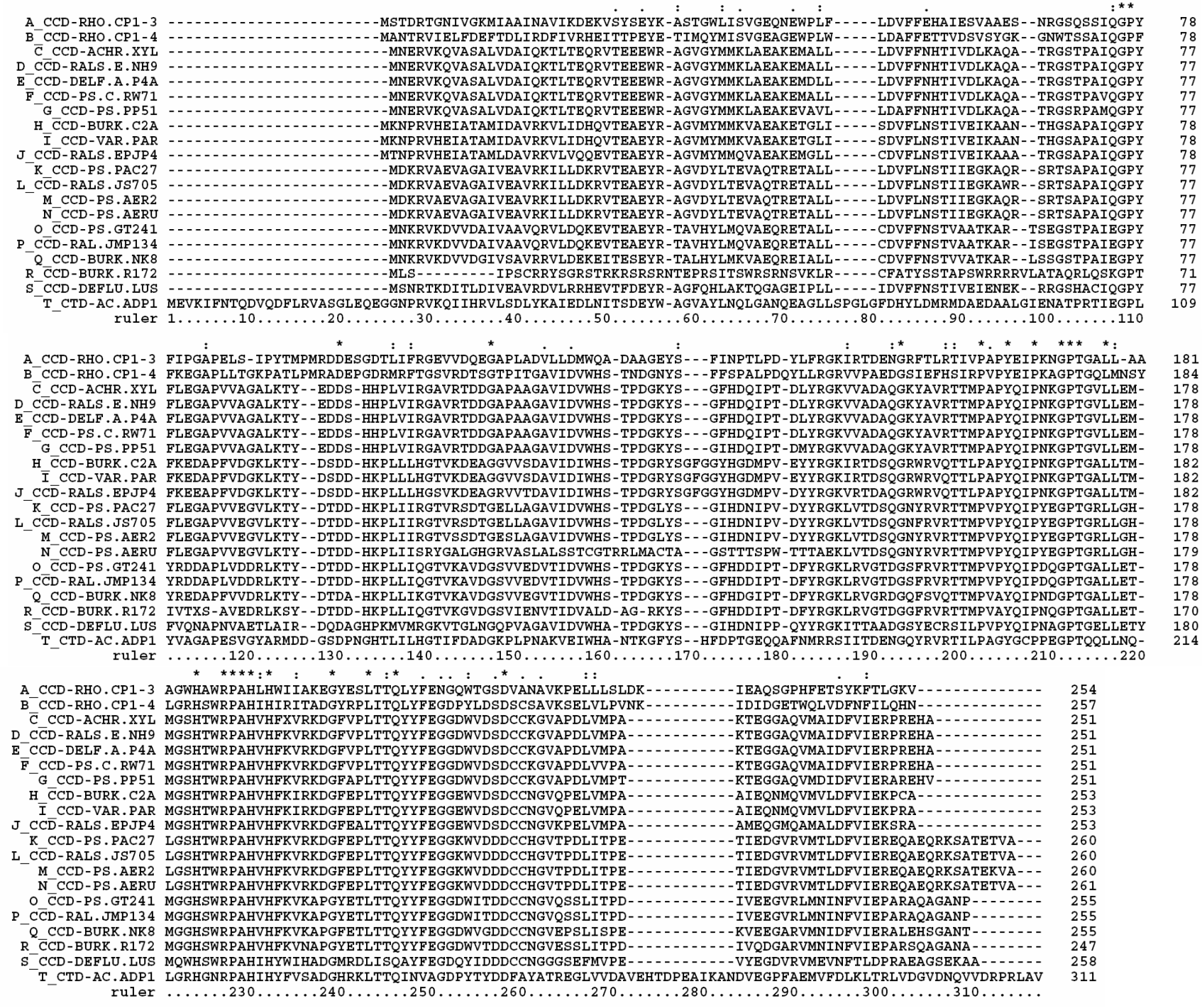


Figure 4

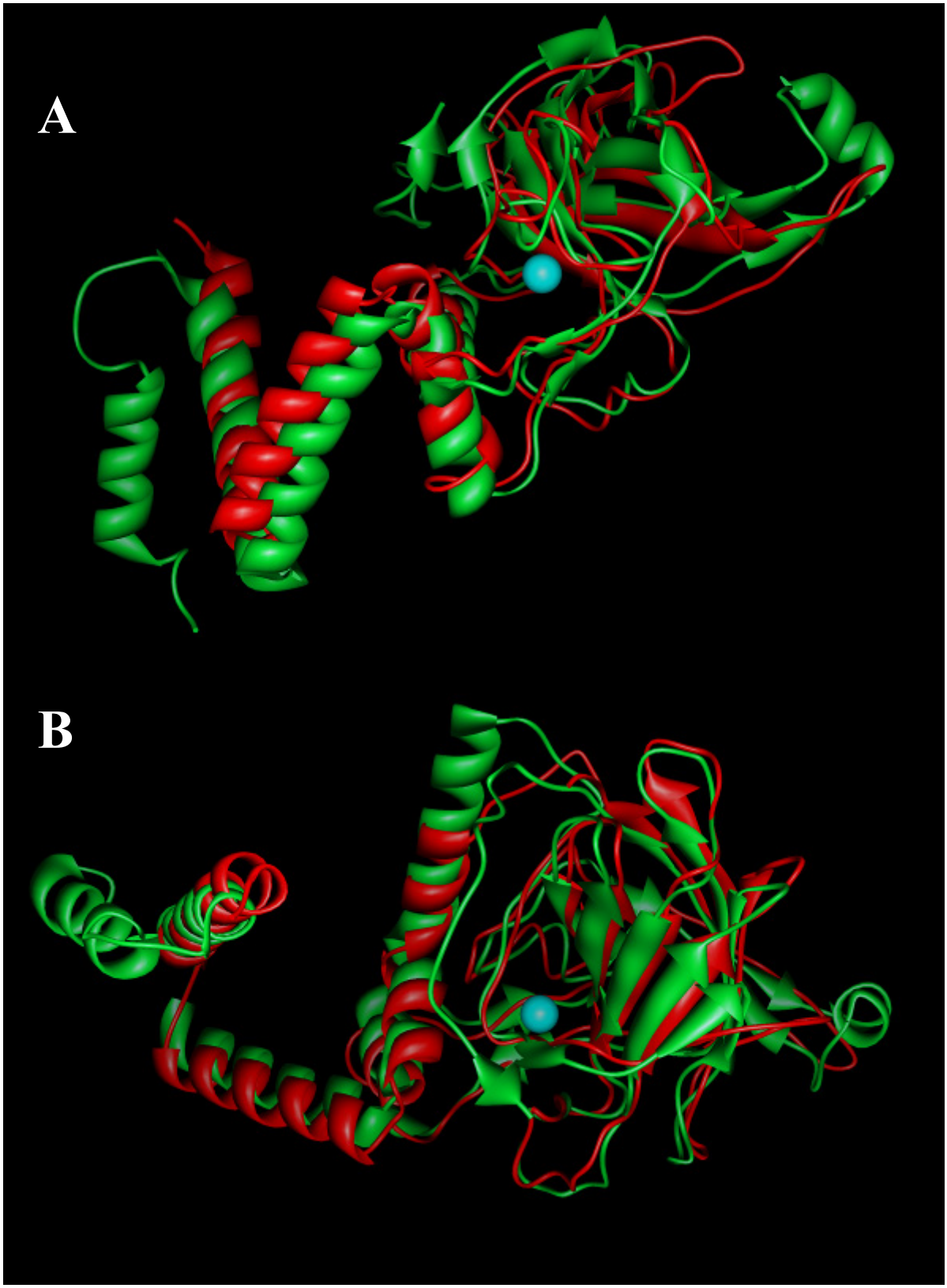


Figure 5

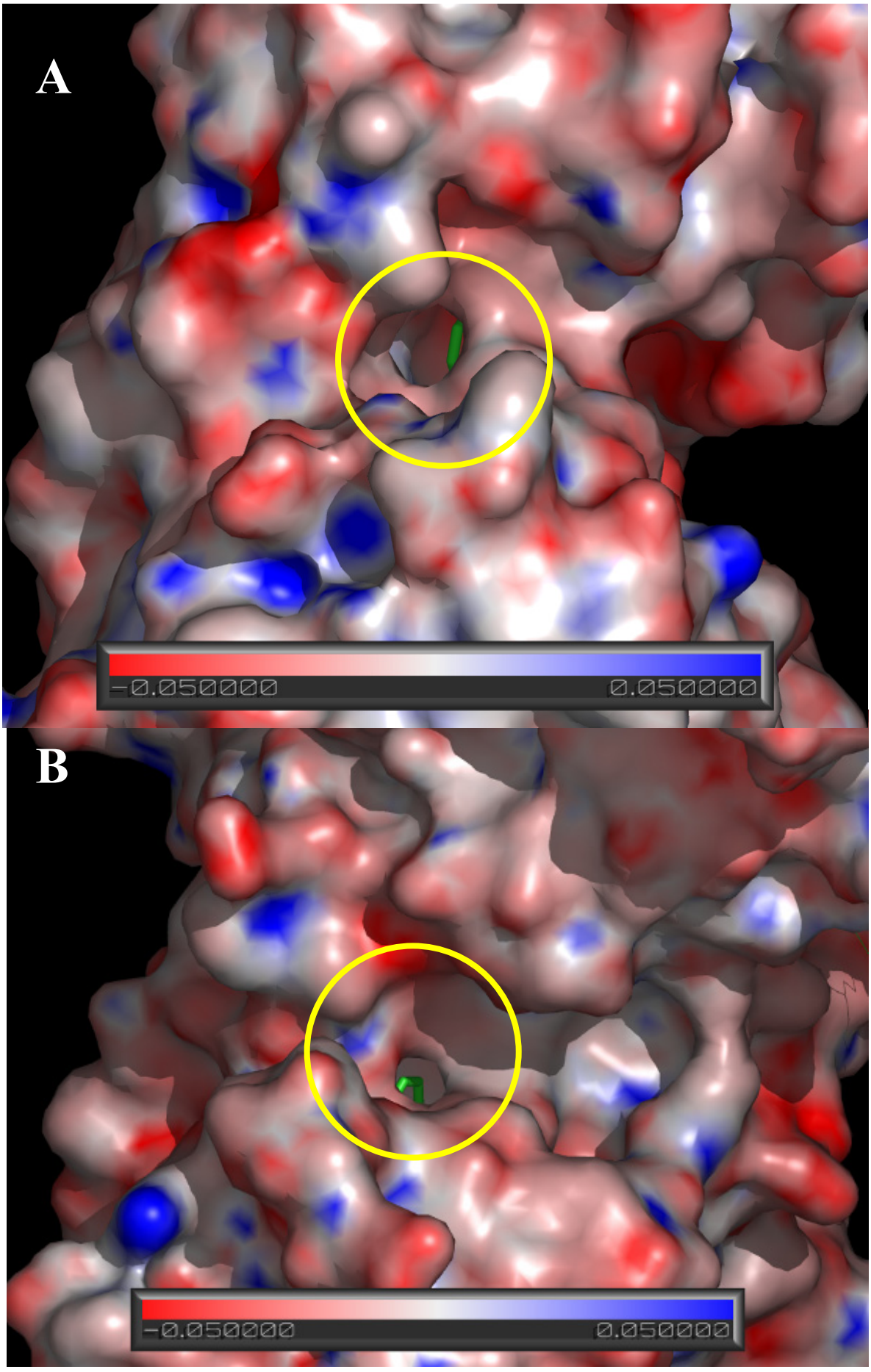


Figure 6

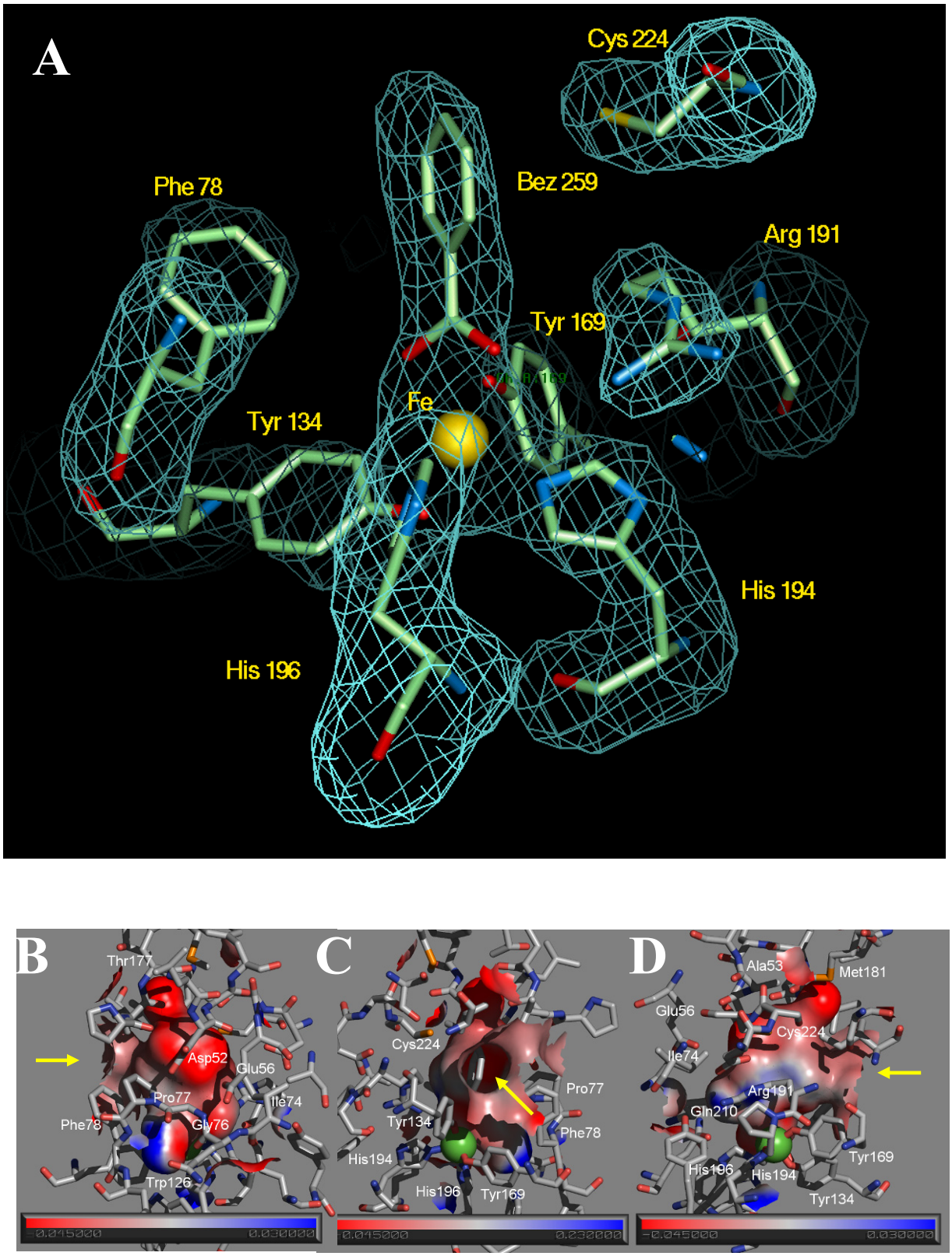
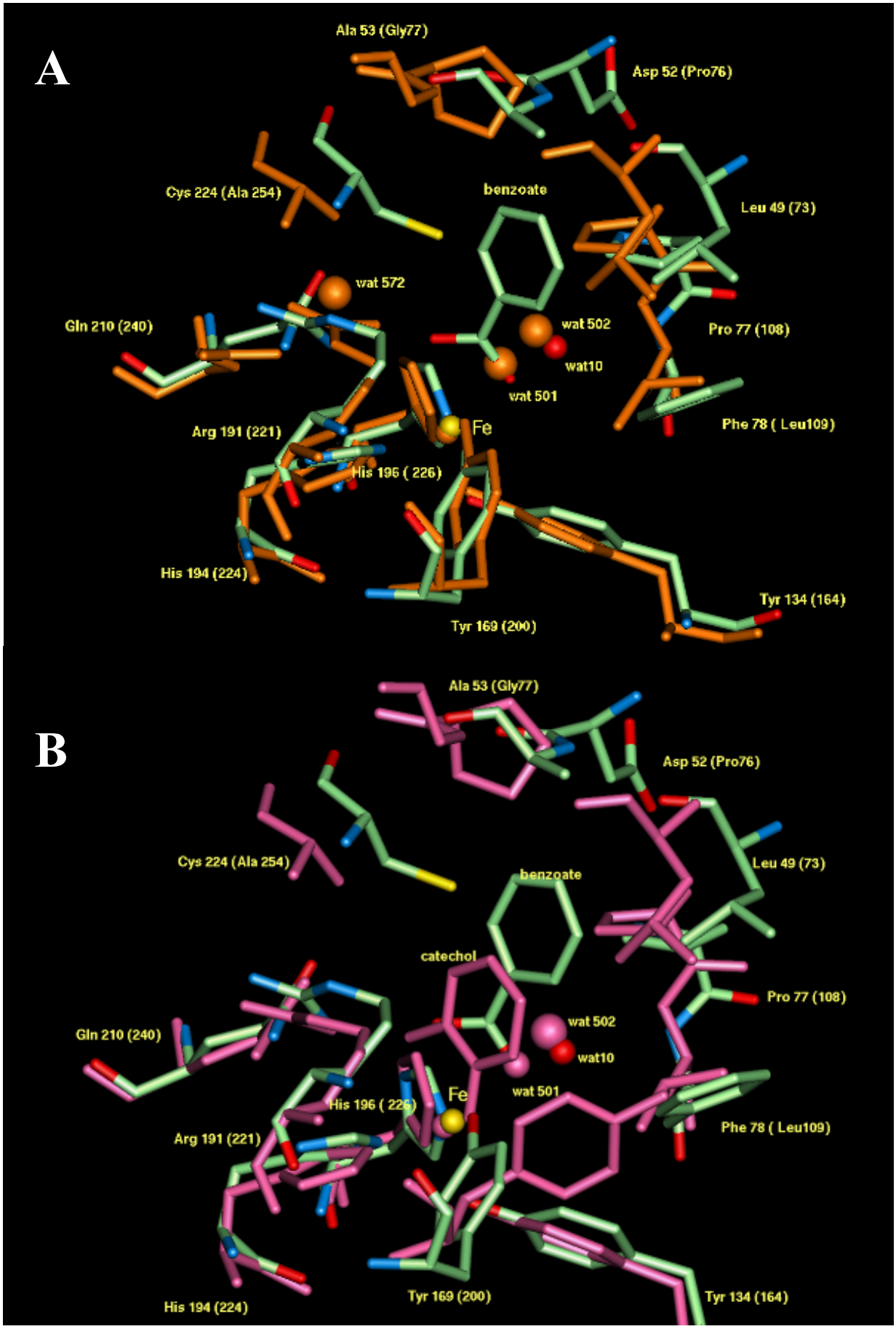
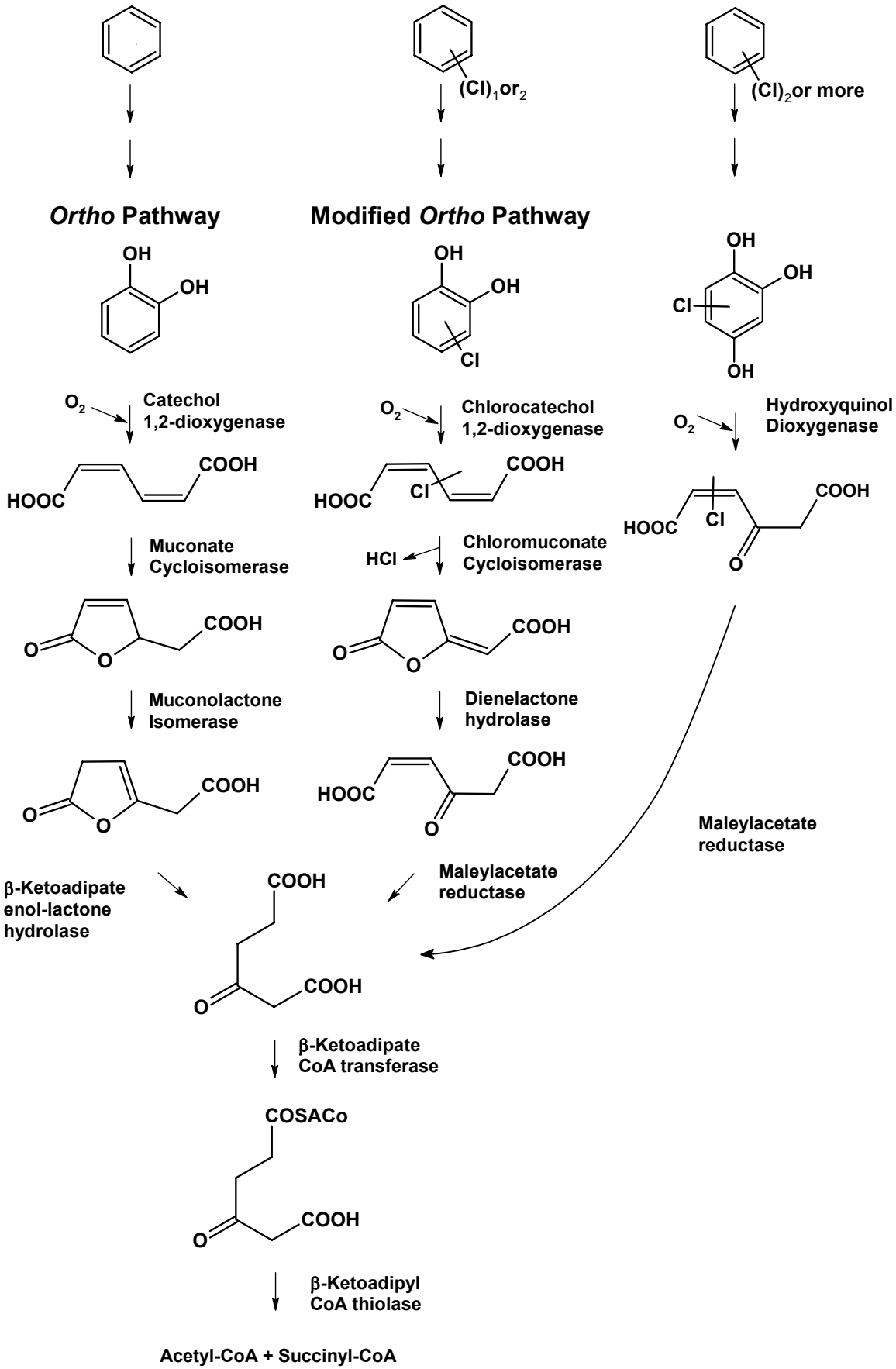


Figure 7





Scheme 1

Crystal structure of 4-chlorocatechol 1,2-dioxygenase from the chlorophenol-utilizing gram-positive rhodococcus opacus 1CP

Marta Ferraroni, Inna P. Solyanikova, Marina P. Kolomytseva, Andrea Scozzafava, Ludmila A. Golovleva and Fabrizio Briganti

J. Biol. Chem. published online April 1, 2004
