

QM/MM study of the reaction mechanism of *Mt*CYP121

QM/MM study of the C-C coupling reaction mechanism of CYP121, an essential

Cytochrome p450 of *Mycobacterium tuberculosis*

Victoria G. Dumas^{1,2}, Lucas A. Defelipe^{1,2}, Ariel A. Petruk², Adrian G. Turjanski^{1,2} and Marcelo A. Marti^{1,2,*}

¹Departamento de Química Biológica, Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Intendente Güiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina.

²Instituto de Química Física de los Materiales Medio Ambiente y Energía (INQUIMAE), UBA-CONICET. Facultad de Ciencias Exactas y Naturales, Universidad de Buenos Aires. Ciudad Universitaria, Intendente Güiraldes 2160, C1428EGA, Ciudad Autónoma de Buenos Aires, Argentina.

*To whom correspondence should be sent: marcelo@qi.fcen.uba.ar

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process which may lead to differences between this version and the Version of Record. Please cite this article as an 'Accepted Article', doi: 10.1002/prot.24474

© 2013 Wiley Periodicals, Inc.

Received: Jul 18, 2013; Revised: Oct 15, 2013; Accepted: Nov 04, 2013

Abstract:

Among 20 p450s of *Mycobacterium tuberculosis* (*Mt*), CYP121 has received an outstanding interest, not only due to its essentiality for bacterial viability, but also because it catalyzes an unusual carbon-carbon coupling reaction. Based on the structure of the substrate bound enzyme, several reaction mechanisms were proposed involving first Tyr radical formation, second Tyr radical formation and C-C coupling. Key and unknown features, being the nature of the species that generate the first and second radicals, and the role played by the protein scaffold each step.

In the present work we have used classical and quantum based computer simulation methods to study in detail its reaction mechanism. Our results show that substrate binding promotes formation of the initial oxy complex, Compound I is the responsible for first Tyr radical formation, and that the second Tyr radical is formed subsequently, through a PCET reaction, promoted by the presence of key residue Arg386. The final C-C coupling reaction possibly occurs in bulk solution, thus yielding the product in one oxygen reduction cycle. Our results thus contribute to a better comprehension of *Mt*CYP121 reaction mechanism, with direct implications for inhibitor design, and also contribute to our general understanding of these type of enzymes.

Keywords: CYP121, Cytochrome p450, *Mycobacterium tuberculosis*, QM/MM, Cyclo-di-Tyrosine, Reaction Mechanism, Molecular Dynamics, Electron Transfer.

Introduction

Tuberculosis has been a long standing major health problem specially in the developing world, causing over two millions deaths annually (WHO reports). The major causes for this impact are related to the need for long, several times incomplete, treatment with current first line drugs (isoniazid and rifampicin), the emergence of drug resistance strains, and the negative drug-drug interactions with HIV positive patients. This scenario has thus created an urgent need for novel anti tubercular drugs ^[1]. A great progress in our understanding of *Mycobacterium tuberculosis* (*Mt*) is derived from its complete genome sequence, which among other unexpected findings, revealed the presence of 20 different CYP genes, encoding Cytochromes of the p450 type of enzymes (or simply CYPs) ^[2,3].

The CYPs are ubiquitously distributed heme based enzymes, typically acting as monooxygenases that transfer a single oxygen atom to the substrate, among other oxidative reactions. To perform the catalysis, CYPs must bind their substrates into the active site pocket, which is located on top of the distal side of the heme, where oxygen binds. Comparison of many available CYP structures, shows that the shape and size of the substrate binding pockets are diverse and are the key for defining the substrate and reaction specificity ^[4]. The CYPs are involved in a variety of metabolic processes, and while eukaryotic ones are membrane bound, those from bacteria are cytosolic enzymes that interact with a soluble redox partner. In the last decades, microbial CYPs have emerged as potential drug targets, in particular since the early discovery that some of them, like the CYP51 (sterol 14 alfa demethylase) in yeast ^[5], were been shown to be essential. Also, the success of azole based inhibitors ^[6], currently used to treat fungal infections has reinforced the search for inhibitors to other members of this very important protein family ^[1,7].

Among the 20 *Mt* CYPs, CYP121 has received an outstanding interest, not only due to its essentiality for bacterial viability [8], but also because it has been recently shown to catalyze an unusual (for a CYP enzyme), carbon-carbon coupling reaction. *Mt*CYP121 converts its substrate, cyclo-di-tyrosine (cYY), to a product called mycocyclosine (mcy) (Scheme 1A), where the two Tyr aromatic rings are covalently joined through a C-C bond in the *ortho* position with respect to each Tyr OH group [9]. The structure of *Mt*CYP121 was solved in the absence [10] and, in the presence of its substrate [9], as well as in complex with differentazole compounds [10-12], and several single point mutant structures were obtained [8]. The enzyme was also extensively characterized in terms of its biochemistry by spectroscopic methods and kinetic measurements [13]. The X-ray structure of the ferric cYY bound enzyme (Figure 1), shows that one of the substrate Tyr lies closer to the heme (we will call this the *close*-Tyr), with its OH group interacting with two structural water molecules (W1 and W2 in Figure 1), one of which is also coordinated to the ferric iron, while the other Tyr (which we will call the *far*-Tyr) is pointing towards the protein surface, i.e the substrate binding site entry. Adjacent to the cYY, there are two water-filled pockets suggesting an easy solvent access to the active site. Interestingly, comparison with the substrate free structure shows that no major conformational changes occur upon cYY binding (C- α RMSD between both structures is only 0.16 Å) [9]. The iron atom is coordinated to water W1, which also shows hydrogen bond interactions with Ser237. This residue possibly replaces the conserved Thr residue found in other p450s which has been shown to be involved in oxygen binding and acting as proton donor to the iron bound superoxide [14,15]. A unique characteristic of CYP121 is the presence of an arginine residue (Arg386) close to the iron, which has been described, to strongly affect the heme redox potential [8,10]. Last but not least, a recent work analyzed the reactivity of *Mt*CYP121 towards different substrate analogues, showing the importance of both the diketopiperazine ring and the

correct positioning of the Tyr hydroxyl for efficient conversion of cYY^[16]. It is important to note at this point, that although much has been learned, the precise mechanism and the structural reasons that allow *Mt*CYP121 to perform this uncommon reaction, remain largely unknown.

Understanding a CYP reaction mechanism requires detailed knowledge of the nature and electronic structure of its key oxidant species, compound I (CpdI), and its associated reaction cycle (Shown adapted for *Mt*CYP121 reaction in Scheme 1). In the archetypical CYP reaction cycle, the resting state is usually a ferric heme hexacoordinated (6c) state, with the proximal cysteine and a water molecule (or hydroxide anion) bound at the distal site. After substrate binding, heme reduction to the ferrous state and water release, the enzyme is prepared to bind dioxygen in order to form the Fe^{II}-O₂ complex, which is further reduced to yield the ferrous-superoxo species (Fe^{II}-O₂⁻). Both electrons are in prokariotes derived from CYP redox partners (CYP reductases), which usually use NADH/FADH as the final electron source. Interestingly, considerable diversity has been observed in bacterial CYP redox partners^[4]. The ferrous-superoxo species has an intrinsic unstable nature, and after taking two protons from its environment, the cleavage of the O-O bond is produced, with the concomitant release of a water molecule, yielding the formation of the high valence oxo-ferryl species, called compound I (Cpd I). Cpd I electronic structure is usually described as an [P^{*+}-Fe^{IV}=O²⁻] state, thus an iron(IV) bound to a di-anionic oxygen atom, and a cationic porphyrin. Cpd I is a very strong oxidant and is usually able to abstract a single hydrogen atom from aliphatic carbons. In the common hydroxylation mechanism, the OH ligand from the resulting Cpd II (usually described as either [P^{*+}-Fe^{III}OH⁻ or P-Fe^{IV}OH⁻]) is rebound to the aliphatic radical leading to the corresponding alcohol and the enzyme ferric resting state. However, Cpd II, has also been shown experimentally to be able to catalyze hydroxylation and other oxidative reactions^[4].

SCHEME 1

Determining how a given CYP is able to perform C-C coupling reactions in the context of the presented scheme is not straight-forward. Interestingly, aromatic C-C coupling reactions have been also reported for other CYPs, for example CYP56 (*DIT2*) of *Candida albicans*, which catalyzes the conversion of *N*-formyl tyrosine into *N,N'*-bisformyl dityrosine [17]; CYP245A1 or *Stap* from *Streptomyces sp.* [18] which catalyzes the intramolecular C-C bond formation in cromopyrrolic acid (a key step in the synthesis of staurosporine) and CYP158A2 from *Streptomyces coelicor* [19] that catalyzes the coupling of two or three flaviolin molecules, leading to the corresponding dimers, among others [20–22]. Common proposal for these C-C coupling reaction mechanism involves the consecutive generation of two aromatic ring radicals, that subsequently recombine establishing the new C-C bond. Key in this proposal are the nature of the species that generate the first and second radicals, and the role played by the protein scaffold in these and the subsequent recombination step, as shown in Scheme 1B.

FIGURE 1

As already mentioned in CYP121, cYY is bound with one tyrosine ring close to the heme iron (the *close-Tyr*), while the other is far from it (the *far-Tyr*) (As shown Figure 1). In the present case, for the Tyr-Tyr coupling reaction to happen, the two (one in each) Tyr radicals have to be produced, which would then combine to yield the cyclic product. First radical formation is proposed to occur in the *close-Tyr*, due to hydrogen atom transfer to Cpd I (Reaction steps a-b-c in Scheme 1B). The second radical formation, must be formed now in the *far-Tyr*, and several mechanisms have been proposed. The first one assumes that mono radical cyclo-dityrosine (cY^*Y) may leave the protein active site to the solution and re-enter in the inverted orientation (Reaction step d' in Scheme 1B), yielding the second radical either in a new complete enzyme cycle (i.e thorough Cpd I) or in the

same cycle using Cpd II as the oxidant. Another interesting possibility is that internal electron transfer (IET) occurs from the far to the *close* tyrosine, thus resulting in first radical localization in the *far*-Tyr (Reaction step d in Scheme 1B), and allowing the second radical to be produced by the heme, again either using Cpd I or Cpd II as the oxidant. Finally, C-C coupling can occur with or without the aid of the protein matrix (Reaction steps f and f' in Scheme 1B). Discerning between these possibilities and understanding the role played by the protein matrix and active site in order to catalyze the reaction, is a very important issue not only due the relevance of CYP121 as a potential drug target but also for the general understanding of how CYPs selectively catalyze a given reaction. In this context we decided to tackle this problem using computational methods, namely Quantum Mechanics (QM), Molecular Dynamics (MD) and hybrid QM/MM based methods.

Since their introduction by Warshel and Levitt ^[23] almost forty years ago, QM/MM methods have been extensively used to study enzyme reaction mechanism, granting an efficient way of incorporating the effects of the protein environment on the electronic structure of enzyme active sites ^[23–29]. In particular, Density Functional Theory (DFT) based QM/MM methods have allowed the detailed study of the electronic structure and reaction mechanism of many metalloproteins ^[30–32], particularly heme proteins and including several works on cytochrome p450 reaction mechanisms, although, to our knowledge none of these works analyzed *Mt*CYP121 reaction. ^[30,31,33–35] Our group has a long standing tradition in the use of QM/MM methodologies to study enzymatic reactions using restrained energy minimizations, particularly in heme proteins, which allowed to understand the role played by the heme group and the surrounding protein framework along the process ^[36–40].

In the present work we have thoroughly studied the reaction mechanism of *Mt*CYP121 using the above mentioned computer simulation tools. Our result show that substrate binding is key for

initiating the reaction, that two Tyr radicals could be generated in only one heme reduction cycle, and that final C-C coupling reaction possibly occurs in bulk solution.

Materials and Methods

Classical Molecular Dynamic simulations. The starting structure for the present calculations was the crystallographic structure of CYP121 complexed with cyclo Tyr-Tyr (PDB ID: 3G5H). The desired heme ligand (oxo, peroxy, dioxygen) was added to the heme iron *in-silico* to generate the desired ternary complex. Given the importance of water molecules in the active site ^[9], the five of the water molecules that were present inside MtCYP121 active site in the starting crystal structure, were conserved.

All classical simulations were performed using the PMEMD module of the Amber11 package of programs ^[41], with the ff99SB force field parameters ^[42] for all aminoacids. The heme parameters used in all different coordination states studied were developed and thoroughly tested in previous works ^[43–49]. Partial charges for the different CYY sites were determined using the suggested protocol for amber force-field compatibility. Briefly, the corresponding system electronic structures were determined at the HF/6-31G* level, and charges were derived using the Restraint Electrostatic Potential (RESP) method. ^[50] All other parameters were assigned through analogy to the Tyr parameters in the ff99SB force field, or from the GAFF force field when required. These parameters are available under request. These methodology and both Tyr radical and Tyrosinate parameters were thoroughly tested in our previously work ^[51]. All the starting structures were immersed in a pre-equilibrated octahedral box of TIP3P water molecules up to a distance of 10 Å from the protein surface. The resulting system consists of ca. 32900 atoms. All simulations were performed at 1 atm and 300K, and maintained with the Berendsen barostat and thermostat, respectively ^[52]. Periodic

boundary conditions and Ewald sums (grid spacing of 1 Å) were used to treat long-range electrostatic interactions; a 10 Å cut-off was used for computing direct interactions. Production simulation times are described in the main text.

Thermodynamic Integration Free Energy Calculations. Thermodynamic integration (TI) scheme^[53] as implemented in the AMBER package^[54] was used for computing the Classical Free energy differences for the following steps: (i) Tyrosine protonation/deprotonation $cY^*Y-OH \leftrightarrow cY^*Y-O^-$ reaction (Figure 3, steps 1 and 3) and (ii) Charge transfer/reorganization in mono-radical negatively charged substrate, i.e the $cY^*Y^- \leftrightarrow cYY^*$ reaction (Figure 3, step 2). The method consists in slightly changing the system (in this case the di-tyrosine) as defined by the corresponding classical force field parameters from those corresponding to the initial, to the final condition. For example, for step 2, we start with close-Tyr with those parameters corresponding to a radical Tyr (charge 0) and far-Tyr with those parameters corresponding to a Tyrosinate (charge -1). Then, we change them to those of a Tyrosinate in the close-Tyr and a Tyr radical in the Far-Tyr. The change is performed by representing the system as a linear combination of the initial and final state according to the following Hamiltonian, and varying the coupling parameter λ in eleven discrete steps, from $\lambda = 0$ to $\lambda = 1$.

$$H = H_i \lambda + (1 - \lambda) H_f$$

Here H corresponds to the actual Hamiltonian (i.e the force field using to perform the MD), H_i corresponds to the classical force field with parameters corresponding to the initial state, and H_f corresponds to the classical force field with parameters corresponding to the final state.

For each reaction, 11 windows corresponding to λ values of 0.01, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9 and 0.99 were simulated, and for each one at least a 2 ns long MD was performed, or until

convergence of the corresponding $\langle dE/d\lambda \rangle$ value for each window. All the other MD simulations conditions were the same as above. The Gibbs free energy was then obtained by numerically integrating the $\langle dE/d\lambda \rangle$ vs λ curve. To have an error estimate of the obtained free energy values two independent estimations were performed (each using a separate set of 11 windows). The reported values correspond to the average \pm Standard Deviation (SD) of the two values, obtained with the different simulations sets.

Quantum and QM/MM methods.

Starting structures for QM/MM simulations were taken from the corresponding classical MD simulations, after slowly (200ps) cooling of selected snapshots to 0K. All QM/MM calculations were performed at the DFT level using the SIESTA code with our own QM/MM implementation called Hybrid^[55]. In this method, the quantum and the molecular mechanics subsystem are combined through a hybrid Hamiltonian introducing a modification of the Hartree potential and a QM/MM coupling term. The protein (or classical) environment affects the electronic density in a self-consistent fashion due to the addition of the classical-point charge potential to the Hartree potential. The coupling term has two main contributions representing the electrostatic interaction between the electrons and nuclei, defining the QM charge density with the classical point charge and an additional term corresponding to the van der Waals interactions between the atoms in the quantum and classical regions through a 6-12 Lennard-Jones potential^[55]. For all atoms, basis sets of double zeta plus polarization quality were employed and all calculations were performed by using the generalized gradient approximation functional proposed by Perdew, Burke, and Ernzerhof (PBE)^[56]. For all systems studied, the spin-unrestricted approximation was used unless explicitly noted. The different QM subsystems included the heme group (without side chains), the heme ligand (depending on the species to be simulated in each case) the close tyrosine side chain of the

substrate (for first and second radical formations) and the complete substrate (for the analysis of the internal ET step), and the thiol group of the proximal cysteine. For the study of the C-C coupling reaction only the substrate was treated at the QM level. The rest of the protein unit and the water molecules were treated classically. QM/MM methods have been successfully applied for the study of enzyme reactions including heme proteins. Particularly, the Hybrid method showed excellent performance for medium and large systems, and was proven to be appropriate for biomolecules, specifically heme proteins, as shown by several works from our group.^[43,49,57]

Using QM methods, thorough analysis of reaction mechanisms requires the determination of accurate free energy profiles. There are several strategies that allow proper and reliable determination of a reaction free energy profile, like umbrella sampling^[58], Multiple Steered MD combined with Jarzynski's equation^[55,59], Metadynamics^[60] or the recently developed paradynamics^[25-27], however all of them require extensive sampling, which is computationally very expensive and difficult to achieve at the DFT QM/MM level for large systems containing transition metals, such as heme proteins. Therefore for the present case, we resorted to compute potential energy profiles by using restrained energy minimizations along the reaction path that connects reactant and product states^[55,57,61]. For this approach, an additional term, $V(\xi) = k (\xi - \xi_0)^2$, was added to the potential energy, where k is an adjustable force constant (set to be 200 kcal/mol·Å² here), ξ_0 is a reference value, which was varied stepwise with an interval of 0.1 Å, along the reaction coordinate. By varying ξ_0 , the system is forced to follow the energy minimum reaction path along the given coordinate ξ . To avoid possible hysteresis problems in the RC scans due to accidental changes in the MM part of the system, a distance cut-off of 10Å was used, that only allows MM atoms that close to the QM active site to move during the RC scan. The energy convergence error for all QM/MM calculations was below 0.2 kcal/mol for each optimized structure and for the energy

barriers defined by the corresponding energy minimum reaction path. This methodology has been extensively used in several works and proven to successfully avoid the mentioned problem [37,39,55,62,63].

Results

In this work we analyze in detail the reaction mechanism performed by *MtCYP121* by means of QM/MM and MD simulations. Given the complexity of the reaction performed by this protein we examined separately the following three key steps: i) First Tyr^{*} or cY^{*}Y generation, ii) Second Tyr^{*} or cY^{*}Y^{*} generation, and iii) C-C bond formation or mcyC, as depicted in Scheme 1B. The results are presented separately for each step and are then discussed in the context of the whole reaction mechanism.

1. First Tyr^{*} generation

1.1 MD simulation of cYY bound to MtCYP121. The most potent and common oxidant in p450 types of cytochromes is, as already mentioned, Cpd I. Once generated in the enzymatic reaction cycle, Cpd I is capable of abstracting a hydrogen atom and inserting an oxygen atom in variety of substrates. The most common way of determining the initial oxidation product in any given CYP is therefore, to look at the substrate atoms that are close to the iron [35]. In order to determine the possible mechanisms for first Tyr^{*} generation, starting from the crystal structure of substrate bound *MtCYP121*, we performed 30 ns MD simulation of cYY bound to *MtCYP121*, with the heme in the Oxy, Cpd I, and Cpd II states, and analyzed the corresponding relevant geometrical parameters as shown in Table I (See Methods for details on the set-up of the system and simulation parameters). Cpd II and Oxy-heme were analyzed not only for comparative purposes, but also due to the fact

that Cpd II has been shown to be able to perform oxidative reactions, while the oxy complex is known to be able to abstract protons from the surroundings. Analysis of the corresponding RMSD and RMSF plots (not shown) using the X-ray structure as the reference, show that protein structure remains stable during the simulation time scale.

The results show that for Cpd I, TyrO does not establish a permanent hydrogen bond with the ferryl oxygen, since the presence of several (up to five) water molecules in the active site compete as hydrogen bond acceptors, and usually a water bridges the interaction between TyrOH and the ferryl oxygen. However, an hydrogen bond interaction between the phenol group and Cpd I can be established transiently (closest FeO1-TyrH_{OH} distance is 1.57 Å) due to the high mobility of the *close-Tyr* and the surrounding water molecules. A similar picture is observed for Cpd II. For the oxy complex, on the other hand, a permanent hydrogen bond interaction is established between TyrOH and the bound ligand with the phenol hydrogen thus remaining close to both oxygen atoms. In all cases, the phenol hydrogen is closer than the ortho carbon hydrogen to the reactive heme bound oxygen atom. Thus, it seems that the CYP121 scaffolds positions the *close-Tyr* phenol group not permanently at reactive distance from Cpd I and Cpd II oxygen atoms, but allows to establish a tight hydrogen bond with the bound dioxygen.

TABLE I

1.2 Hydrogen abstraction from Cpd I. Starting from a selected representative structures from the MD simulations of cYY bound to *MtCYP121* in the Cpd I state where an hydrogen bond is present between the ferryl oxygen and TyrH_{OH}, we analyzed the hydrogen abstraction reaction using QM/MM simulations. The phenolic hydrogen is more acidic and closer to the heme than any of the aromatic hydrogens, thus we only considered abstraction of this hydrogen. Finally, and due to the known Multi State Reactivity^[4] of Cpd I we analyzed both the doublet (D) and quartet (Q) spin

states. The corresponding energy profile for hydrogen atom abstraction results are shown in Figure 2A, while selected structural and electronic structure parameters are shown in Table II.

TABLE II

The obtained energy barriers for the D and Q spin states are similar and small, less than 6 kcal/mol and smaller in the Q state. The Q state corresponds to the ground state along the explored range of the reaction coordinate. The reaction results in the formation of a Tyr radical and a Cpd II type heme, as evidenced by the Mulliken and spin population analysis, which show a close to neutral Tyr with an unpaired electron (Tyr total spin is close to unity). Cpd II displays, as expected, a triplet ground state with two unpaired electrons and a mixture of $[\text{P-Fe}^{\text{III}}\text{-OH}^* \rightleftharpoons \text{P-Fe}^{\text{IV}}\text{-OH}]$ character (as evidenced by the charge and spin of the resulting OH group). Thus we confirm that Cpd I is able to abstract the first hydrogen from the *close*-Tyr in cYY and with a low activation barrier.

1.3 Alternative oxygenated species as oxidants. As described in the introduction, during the CYP-oxygen reaction cycle, dioxygen first binds to ferrous iron, and is then reduced to form iron-superoxo species ($\text{Fe}^{\text{III}}\text{-O}_2$). For comparative purposes and as a possible alternative, we also evaluated whether the iron-dioxygen or iron-superoxo species are able of abstracting hydrogen from the *close*-Tyr. We used the same protocol and reaction coordinate as for the Cpd I reaction described above, and found that the oxy complex is unable to abstract a proton/hydrogen from the *close*-Tyr (data not shown). However, the energy profile for the transfer reaction with the iron-superoxo species, presented in SI (Figure S1), suggests that a in this cases proton/hydrogen can be transferred with a very small barrier. The resulting Mulliken and Spin atomic populations (shown in SI), however, show that the reaction corresponds to a proton transfer, since the Tyr product is left with a net negative charge, close to unity. Thus –the iron-superoxo species is not capable of

generating the first Tyr radical, although *close*-Tyr can participate as proton source in the first steps of the CYP reaction cycle to yield Cpd I.

● FIGURE 2

1.4 Evaluation of OH rebound step as a competing reaction. The first step in the commonly observed hydroxylation reactions in CYPs, usually consists as in this case, in the generation of an alkyl-radical and Cpd II. This state, usually proceeds through the so called rebound step, consisting in Cpd II OH ligand radical addition to the radical substrate, resulting in the desired hydroxylated product. In MtCYP121 the rebound mechanism would lead to formation of a catechol (ortho-biphenol) in the *close*-Tyr, in one CYP cycle. To analyze the likelihood for the occurrence of this competing reaction, we computed the energy profile for the rebound step in both the doublet and quartet spin states. The results presented in Figure 2C, show that the rebound step is exergonic, but has a moderately high barrier in both studied spin states (16-20 kcal/mol). Thus, in order to avoid this unwanted reaction, the second Tyr radical generation should have a lower barrier compared to the reported rebound step energy barrier.

2. Second Tyr* generation

2.1 Hydrogen atoms abstraction by Cpd II. The second Tyrosine radical formation in MtCYP121 is, as previously described, the less understood step and there are several mechanistic possibilities (see reaction steps d, d' and e in Scheme 1B). The simplest proposal assumes that mono radical cYY*, where the *close*-Tyr bears the radical, may leave the protein active site and re-enter in the inverted orientation, thus leaving the radical in the *far*-Tyr and a neutral *close*-Tyr. The second radical, must now be generated again in the *close*-Tyr, and could occur either in a new complete enzyme cycle (i.e through Cpd I) or in the same cycle using Cpd II as the oxidant. We have shown above that

Cpd I can perform the corresponding H abstraction, so now we will analyze the likelihood that Cpd II generates a Tyr radical, in both the T and S states. For this sake we selected a snapshot from the cYY Cpd II MD simulation and performed the corresponding QM/MM optimizations and reaction energy profiles. The results are shown in Figure 2B and Table III.

TABLE III

In Cpd II the ground state is the T state, which lies ca. 3.5 kcal/mol below the S state, displays a 3.5 kcal/mol barrier for hydrogen abstraction, and the reaction is endergonic in a similar amount. (Figure 2B red line). In the S state the barrier is very small (ca. 1.5 kcal/mol) and also slightly endergonic. The final step for the reaction includes water release upon spin transition of the resulting ferric heme. Again, analysis of the Mulliken and spin populations show that the product corresponds to a radical Tyrosine. The charge and spin of the resulting heme moiety show that it corresponds to a ferric water bound heme. Thus, although with some energetic cost, Cpd II is also capable of producing a cYY* *close*-Tyr radical in CYP121.

2.2 Internal Tyr-Tyr electron transfer. Since both Cpd I and Cpd II are able to perform *close*-Tyr radical formation in *Mt*CYP121, both radicals could in principle be produced in one heme oxidation cycle. However, if this has to occur efficiently, a reorganization of the first radical product has to take place, implicating either the release and re-entering of the cYY* intermediate, or the migration of an electron from the *far*-Tyr to the *close*-Tyr in an intramolecular electron transfer (IET) reaction (reaction steps d' and d in Scheme 1B). Given the already mentioned potential complications with a mechanism involving cY*Y release, we thought about a possible mechanism that would generate the second Tyr radical but retaining the cY*Y moiety inside CYP121. We propose that this could occur through an IET, that would result in electronic reorganization that stabilizes a state with the

spin density located in the *far*-Tyr, and the electron density reorientates to the *close*-Tyr, or even directly to the heme group. As will be shown below, the process may occur through several elemental steps, that could or could not involve protonation/deprotonation process of the involved species (*vide infra*). Moreover, recent results from our group^[51] for Tyr* containing peptides in solution, show that IET steps can be tightly coupled to proton uptake and release, leading indeed to a Proton Coupled Electron Transfer (PCET) processes. In this case PCET is defined in its broadest sense following Hammes-Shiffer and Soudackov's work,^[64] i.e. involving movement of a proton and an electron through different bonds in either a time-concerted or sequential manner^[51]. We now turn to the analysis of this last possibility.

2.2.1 Classical analysis of Tyr to Tyr* ET. We begin our analysis by computing the free energy change (ΔG) associated with the Tyr to Tyr* ET (or PCET) process in MtCYP121, using a TI scheme as described in methods. We divided the process in three elemental steps consisting of deprotonation of the *far*-Tyr (Figure 3, step 1), moving the electron from the *far*-TyrO⁻ to the *close*-TyrO*, which corresponds to radical migration to the *far*-Tyr process (Figure 3, step 2), and *close*-TyrO⁻ protonation (Figure 3, step 3).

FIGURE 3

Since it is well known that absolute free energy changes of protonation equilibrium computed with classical force fields using TI are significantly shifted from experimental values, several works show that relative free energies of a given residue in a protein environment with respect to the reference value in solution, are able to correctly predict the shift or trend in the pKa values^[65,66]. Thus, in the present case, each step was evaluated for the substrate inside MtCYP121 and in water solution. We report the values corresponding to the difference between the protein and water free energies for each step, and a negative value means that the step is more favorable to occur inside

MtCYP121 compared to bulk solution. The results presented in Table IV show that removing the proton from the *far*-Tyr is unlikely to occur spontaneously. Tyr pKa in solution is about 10.2^[67], but since the deprotonation free energy inside *MtCYP121* is smaller than in bulk solution ($\Delta\Delta G = 17.3 \pm 1.0$ kcal/mol), its pKa must be even higher. This is not the case for the *close*-Tyr where deprotonation free energy is close to the bulk solvent value and thus a similar to bulk pKa is expected, that could result in protonation of the *close*-Tyr after IET. Most importantly, the radical transfer step, is highly exergonic, and compensates the free energy of the first deprotonation, resulting in an overall spontaneous PCET reaction ($\Delta\Delta G = -14 \pm 3.2$ kcal/mol). The PCET nature of the IET step is clearly evidenced by the fact that only considering the first two (or all three steps) together, a spontaneous reaction is obtained. Thus the *MtCYP121* enzyme framework strongly favors a neutral radical *far*-Tyr and a negatively charged *close*-Tyr.

TABLE IV

To get an insight into the molecular basis for the radical localization in the *far* Tyr, we looked for those residues that could either promote the presence of a neutral *far*-Tyr and/or a charged *close*-Tyr. Arg386 is a key residue located in close proximity to the *close*-Tyr and given its natural positive charge at physiological pH (Figure 1), we decided to analyze its influence in the charge transfer process. For this sake we computed the free energy of the IET in the R386A mutant *MtCYP121* (Figure 5, step 2). Analysis of the structure and dynamics of the R386A mutant protein show that no major changes in the protein structure occur (data not shown), a fact that is consistent with the X-ray structure of the corresponding mutant^[8]. Interestingly, and consistent with a predominant role of this arginine in guiding the radical transfer process, the resulting $\Delta\Delta G = -10.5 \pm 0.6$ kcal/mol for the charge transfer step (step 2 in Figure 3) is significantly smaller than in the *wt* protein of -28.8 ± 1.6 kcal/mol. Thus Arg386 plays a key role for the proposed IET process.

It is important to note that the above presented data only measures the effect of the protein on the process of moving the negative charge, as defined by the classical parameters, from one Tyr to the other. Thus it represents an ideal or extreme case where tyrosine residues are either neutral or have a one electron charge. A more accurate calculation would require the use of QM/MM free energy methods as performed and reviewed by Kamerlin et. al.^[25] which, as already mentioned, is computationally too demanding for the present system.

2.2.2 QM/MM analysis of Tyr to Tyr ET.* To get further and more detailed insight into how the protein could modulate or promote the IET we performed QM/MM optimizations of mono-radical and negatively charged substrate (cYY^{*-}) directly related with the IET process and analyzed the resulting charge and spin distributions (Table V).

We first computed the electronic structure of the cYY^{*-} species inside the protein, considering the heme moiety classically (i.e when charge transfer to the heme group cannot be produced). In agreement with the above described classical results, the Mulliken and Spin population results show that the protein framework favors radical (spin) localization in the *far*-Tyr and negative charge in the *close*-Tyr. Moreover, we also computed the proton affinity of the *far*-Tyr inside *MtCYP121* considering a neutral *close*-Tyr, or a radical *close*-Tyr (i.e after the first hydrogen abstraction step). The results show that the presence of radical Tyr lowers the proton affinity by 6 kcal/mol, thus favoring *far* Tyr proton release.

Even more revealing are the results obtained when the heme moiety in the Cpd II state is added to the QM system. As shown in Table V, for both spin states considered (Q and D), the *close* and *far*-Tyr display a radical character (small charge and $>0.5e^-$ unpaired spin density) while the heme displays ferric hydroxy like state ($Fe^{III}-OH^-$), which corresponds to the *MtCYP121* heme resting

state. Moreover, the difference between both spin states is that, while the D state presents both Tyr residues coupled antiferromagnetically and the only unpaired electron remains in the heme, in the Q state cYY displays the expected bi-radical triplet state character (see below). Thus, these results suggest that once the first radical is formed in the *close*-Tyr, when the *far*-Tyr releases its phenolic proton the electron is taken by Cpd II, resulting in a biradical cYY and a ferric hydroxo bound heme. Thus an IET seems to operate in MtCYP121 to generate efficiently and fast the second Tyr radical. It is important to note that in the present work we did not evaluate the coupling and reorganization energy parameters that both together, with the corresponding change in free energy, control the ET rate. However, we note that IET in adjacent and/or covalently bound residues (as the two Tyrs and the heme in the present substrate) is very fast if the reaction is exergonic, and thus is not expected that the coupling and reorganization energy contribute to the overall reaction rate.^[68,69]

TABLE V

3. cY*Y* to mcyc: C-C bond formation occurs in solution

3.1 MD simulations of the cY*Y* in solution and in MtCYP121. As mentioned in the introduction, the C-C coupling reaction after bi-radical formation, can occur either inside the protein, or in solution after a possible di-radical substrate (cY*Y*) release. In any case, for the reaction to take place, the two possible ortho carbons must come close enough to allow for bond formation to occur (*vide infra*). In order to analyze the probability of this step, we firstly performed 30 ns long MD simulations of cY*Y* in solution and inside MtCYP1212 and measured all possible four relevant C-C distances. The resulting histogram for cY*Y* free (Figure 4) shows all C-C pairs have a bimodal distribution with peaks at 7Å and 10.5Å (which correspond to the distance of the closest and furthest ortho carbons in the most stable conformation). Interestingly, a significant population of conformations where both carbon atoms are at 5Å is found, with a minimum distance of ca. 4.5 Å.

The results obtained from the MD simulation of cY^*Y^* inside protein show that in this case the closest ortho-carbons pair are nearer than others during most of simulation time, and arriving at the same closest distance as those found for the cY^*Y^* in solution. Thus similar closest distance of ca. 4.5 Å between two potential reactive carbons can be found in solution or inside *MtCYP121*, with the protein framework increasing the population of the closest conformation.

3.2. Mycocyclusine (mcy) product formation in vacuum. We begin the study of the Tyr^*-Tyr^* (cY^*Y^*) for the isolated bi-radical in vacuum to have it as reference reaction. The corresponding reaction energy profile was determined using the $Tyr1Ce-Tyr2Ce$ distance as a reaction coordinate.

The results and key structures along the reaction are shown in Figure 5 and Table VI. As expected, the plot shows that for long distances the ground state corresponds to the triplet (T) spin state. Mulliken and Spin population analysis shows one unpaired electron localized in each Tyr, and both Tyr being neutral. The singlet (S) state lies ~11 kcal/mol above the T state in vacuum in the reactive species, with both Tyr still being neutral. The closest distance between two ortho carbons in the optimized structure in vacuum is 4.6 Å, which is similar to the closest possible value in solution. When the C-C distance starts to decrease the energy of the T state raises up continuously while the S state energy also increases but with a smaller slope. At a C-C distance of ca. 2.4 Å both surfaces cross (we will call this structure TS_A). After the crossing the S state curve drops to a first minimum corresponding to a first intermediate state (IS_A). The effective barrier starting from the ground T state is 16 kcal/mol. Interestingly, the TS_A structure displays a value for the dihedral angle around the $C\alpha-C\beta$ bond of ~137°, nearly the most populated value between both possible tyrosine rotamers found in solution for cYY (64° and 183°). The IS_A state, which lies 6 kcal/mol below cYY in the T state, displays a well established C-C bond (1.58 Å), with the concomitant conversion of both ortho

carbons from pure aromatic sp² to an sp³ configuration (Figure 5, IS_A structure). In the IS_A T state in both unpaired spins are localized in the same Tyr showing clearly that a significant change in the electronic structure of the molecule has occurred.

FIGURE 4

To proceed from IS_A to the final product, Ce hydrogens must be transferred to the phenol oxygens. Structural analysis of IS_A shows that each Ce hydrogen is closer to the other Tyr oxygen atom (2.43Å) than to the same Tyr phenolic oxygen (2.73Å). We therefore analyzed the sequential and cross hydrogen transfer using the corresponding difference in Ce-He and He-Oz distances as the reaction coordinate. The corresponding profile shows that the first hydrogen is transferred with a ~20 kcal/mol barrier passing through TS_B and leading to IS_B. In IS_B one Tyr has a planar ortho C while the other one has a tetrahedral ortho C (Figure 5, IS_B structure). Transfer of the second hydrogen has very small barrier (~2 kcal/mol) and finally leads to the product in which both ortho carbons adopt a planar orientation (Figure 5, mcyc structure). As expected from the structural analysis trying to transfer the hydrogen to the same Tyr oxygen results in much higher barrier of ~45 kcal/mol than to the other Tyr oxygen residue (Figure S2).

Since the obtained barrier for the C-C coupling is quite high, we decided to explore also an alternative mechanism, where first the proton/hydrogen migrates from the ortho-carbon to the phenolic oxygen, and the C-C bond formation occurs for the corresponding ortho-carbon located Tyr radicals. The obtained barrier for this proton/hydrogen transfer is over 50 kcal/mol (See Figure 3 in SI), thus, significantly higher than the barrier observed for C-C coupling in the phenolic oxygen located radical Tyrosines.

TABLE VI

3.3 *Tyr*·Tyr* C-C bond formation in solution.* To study the Tyr coupling reaction in solution we started from a snapshot taken from the corresponding MD simulation with the C-C atoms at *ca.* 5 Å.

The corresponding energy plot, presented in Figure 5 solid line, has the same shape as that previously described for the reaction in vacuum. In the reactant structures the ground state corresponds to the T state, with two Tyr radicals, and an S state lying about ~19 kcal/mol above it. When the C-C distance starts to decrease, the energy of both states raises up continuously. In the T state surface, a TS (with a barrier of 23 kcal/mol) is found at C-C distance of *ca.* 2 Å. However, at a distance of *ca.* 2.5 Å the T state and S state profiles cross, yielding an effective transition state structure (TS_A), which lies about 16 kcal/mol above the ground state energy, a value which is similar to that obtained in vacuum. After crossing this TS_A, the energy drops along the S to the first IS_A. The reaction then proceeds similarly to what is observed in vacuum, with a first hydrogen/proton transfer with a high barrier, and a very small barrier for the second hydrogen/proton (data not shown).

FIGURE 5

3.4 *Tyr*·Tyr* C-C bond formation inside the MtCYP121.* To study the Tyr coupling reaction inside MtCYP121 we used the same strategy described above for the reaction in vacuum and in solution.

The corresponding energy plot, presented in Figure 5B dotted line, has the same shape as those previously described. In the reactant structures the ground state corresponds to the T state, with two Tyr radicals and an S state lying about ~11 kcal/mol above it. When the C-C distance starts to decrease the energy of both states raises up continuously. In the T state surface a TS (with a barrier of 40 kcal/mol) is found at C-C distance of *ca.* 2 Å, but at a distance of *ca.* 2.2 Å the both S and T state surfaces cross (TS_A), yielding an effective 32 kcal/mol barrier with respect to the ground state energy. After crossing the TS_A the energy drops along the S state to the first IS_A, lying about 5

kcal/mol above the reactants. The reaction then proceeds similarly to what is observed in vacuum and in solution, with a first hydrogen/proton transfer with a high barrier, and a very small barrier for the second hydrogen/proton transfer.

Thus the results show that in all cases vacuum-solution-protein the C-C coupling reaction occurs through the same elemental steps, with formation of the C-C bond and spin transition, depicting the higher barrier and thus assigned as the rate limiting step. The most important finding, comparing all three environments is that inside *MtCYP121*, the energy barrier for this first step is significantly higher than in vacuum/solution, possibly reflecting the fact that the protein binding site is prepared to host the cYY and not the mcyc product, and thus additional energy is required to allow the two tyrosines to come close enough to establish the C-C bond.

Discussion

In the present work we have analyzed in detail the reaction mechanism performed by *MtCYP121*, which involves the C-C bond coupling reaction between two ortho carbon atoms of the two phenol groups of cyclo-di-Tyrosine (cYY), to yield mycocyclosine (mcyc), by means of QM/MM and molecular dynamics simulations. For this sake we divided the reaction in three elemental steps: 1- first Tyr* generation, 2- second Tyr* generation and 4- the cY*Y* C-C coupling reaction. A complete detailed proposal of the reaction mechanism derived from our results highlighting the key role of *MtCYP121* protein matrix is shown in Scheme 2, with key energetic values reported in Table VII, and discussed below.

TABLE VII

Our results for the first step show that most likely Cpd I is responsible for generation the first Tyr* in the cYY *close*-Tyr. The hydrogen abstraction reaction (from the phenolic oxygen) has a

relatively small barrier, for example when compared to hydrogen abstraction barriers in alkyl-hydroxylating CYPs displaying barriers in the 10-20 kcal/mol range ^[35] and is iso- or slightly exergonic. What seems interesting is the fact that the hydrogen bond interaction between Cpd I oxo and the Tyr-OH is established only transiently, while a permanent hydrogen bond is formed when dioxygen is bound to the heme. We also showed that neither dioxygen, nor peroxide are capable of producing the first Tyr^{*}, the former reaction is thermodynamically unfeasible, while in the latter case a proton (not an hydrogen) is abstracted leading to a negative Tyr (that would be reprotonated by the water molecules in the active site, see below), and eventually participating as a proton source for the formation of Cpd I. Thus, it seems that the *close*-Tyr of the cYY substrate is positioned in a sub optimal position with respect to the heme iron. Although it is tempting to pay minor attention to this observation, proteins hardly ever do something without a specific purpose.

The question then arises, as to why the phenol group is positioned in a suboptimal position. To answer this issue we propose that the phenol positioning has the following two key consequences. First, once the first Tyr^{*} and Cpd II are produced, two competing reactions can occur, the proposed PCET that allows the *Mt*CYP121 reaction to proceed, but also the typical OH rebound step observed for many CYPs performing simple hydroxylating reactions. Hydroxylation reactions in phenols typically occurs in the ortho carbon, but in order to take place the carbon atom must be close to the Cpd II -OH, a fact that in *Mt*CYP121 occurs only transiently. Even for this most reactive structure we computed the energy barrier of the rebinding reaction yielding an energy barrier between 16-20 kcal/mol (Figure 2B). Barriers obtained for the OH rebound step in other CYPs are much smaller, in the 3-4 kcal/mol range, and can be even smaller for some spin states. This difference can be simply explained by the reactive carbon relative position which in cYY lies at ca. 5 Å from the Cpd II oxygen, while in hydroxylating CYPs optimal distance is less than 3 Å ^[35].

Another possible explanation, for the high barrier observed for the rebound step, is the spin delocalization of the Tyr radical on the whole phenol group. In alkyl radicals, which are the usual substrates of CYP hydroxylations, spin is highly localized on the reactive Carbon ($>0.8e$)^[70], while in Tyr radical studied in the present work, spin density on the ortho carbon is small ($<0.3e$). The low spin density on the ortho carbon could also contribute to explain why the C-C coupling step also exhibits a very high barrier (see below). In any case, the high barrier for hydroxylation is consistent with recent findings using mass-spec showing that hydroxylation of cYY was not detected, even for the minority (ca. $<2\%$) side products, different from mcyc.^[16]

Second, it is important to note that in each reaction cycle *Mt*CYP121 requires two electrons to reduce the dioxygen to Cpd I. It is important that each oxygen reduction cycle results in cYY oxidation, and not in unproductive oxygen reduction to water. For this sake, a tight coupling of the co-substrates (cYY and molecular oxygen) is needed. In many heme proteins, a Tyr residue located at hydrogen bonding distance to the dioxygen bound heme, is the main responsible for its stabilization and an overall high oxygen affinity. This high affinity is a key element for assuring the efficiency of subsequent oxidative reactions^[38,71]. Given, thus the above mentioned observations related with the cYY *close*-Tyr position and its interactions with heme dioxygen ligand, we computed the oxygen binding energy to reduced *Mt*CYP121 in the presence and absence of cYY. Our results show that the presence of cYY increases the oxygen binding energy in ca. 7 kcal/mol. This value is similar to that reported in our previous work when comparing wt *Mycobacterium* truncated hemoglobin N (displaying the TyrB10 residue hydrogen bonded to the dioxygen ligand) to the TyrB10 to Ala mutant, which results in an over 100 fold decrease in oxygen affinity^[38,71]. Thus it is clear that the presence of cYY significantly increases *Mt*CYP121 oxygen affinity, thus promoting its reduction only when the enzyme substrate is present, and thus decreasing non

productive consumption of reducing power. The relevance of Tyr-OH for the oxidation reaction is consistent with recent observations showing that when Tyr is replaced by Phe, it can be found in the site close to the heme, but it is oxidized at a very slow rate. Moreover, in the corresponding crystal structure the iron to ligand bound water molecule network is also missing ^[16].

SCHEME 2

Concerning the second Tyr radical formation step, the results presented here clearly favor an internal electron transfer (or charge rearrangement) process, that can be described as Proton Coupled Electron Transfer. Specifically, the presence of a *close*-Tyr radical slightly decreases the *far*-Tyr proton affinity (thus favoring its release to the water environment), while the protein matrix, in part due to the presence of Arg386, favors negative charge localization in the *close*-Tyr. These results in a net driving force for the formation of a *far*-Tyr radical, and a close negatively charged Tyr. Moreover, Cpd II although capable of abstracting an hydrogen from the *close*-Tyr to generate a Tyr radical, does it with a ca. 4 kcal/mol energy cost (although it should be noted, that the barrier and energy cost is smaller than that observed for Cpd II hydrogen abstraction from alkyl substrates of ca. 20 kcal/mol and ca. 13 kcal/mol, respectively ^[4,35]). Strikingly, it can act efficiently as an electron sink, producing the cY^*Y^* bi-radical and regenerating the ferric hydroxo enzyme resting state. The alternative proposal (without IET) requires that cY^*Y leaves the protein and re-binds in the opposite orientation. This proposal would however, not only diminish the enzyme overall rate, but lead to the presence of significant amount of cY^*Y^* in solution, which do to its radical nature could lead to non desired side products, like for example cYY dimers (due to C-C bond formation of Tyr radical from different molecules).

Looking at both Tyr radical generation steps together, it is clear that the rate limiting step should be the formation of the strong oxidant Cpd I. Once Cpd I is present, it abstracts the phenolic hydrogen

leading to the *Close*-Tyr radical, which would then proceed fast and directly to the cYY biradical through the PCET. Thus both Tyr radicals are indeed formed almost in a concerted fashion in one heme reduction cycle.

Once the cY^*Y^* is established, last reaction step concerns Tyr-Tyr coupling or C-C bond formation. Our results show that mechanistic steps of this reaction are same for the reaction in vacuum, in water or inside *MtCYP121*. In all cases, rate limiting step is formation of the C-C bond, leading to the first intermediate IS_A , where the C-C bond is already established, with the two ortho carbons adopting transiently an sp^3 conformation. This first reaction step is further diffculted by the fact that a triplet to singlet spin transition is required, since in the cY^*Y^* the triplet state, with one unpaired electron localized in each aromatic ring is the ground state. Concerning the energy and geometry of the different states, our results are in excellent agreement with previous calculations of Tyr-Tyr coupling reaction in vacuum using B3LYP, where the relative energy of stable intermediates along the reaction were computed, showing that IS_A like state lies ca. 12 kcal/mol below the cY^*Y^* . In that case no barriers were reported [72].

Besides the mechanism, when comparing the energetics of the reaction in the three studied environments it is clear that the protein framework does not promote it. Indeed although the MD simulation data show that reactive C-C carbons can be found close to each other, the energy barrier for establishing the bond is significantly higher inside the protein compared to the reaction in vacuum/water. In other words the protein active site seems shaped for properly accommodating the substrate and not the product, and thus possibly the C-C bond formation step occurs outside the protein in the bulk solution. The fact that *MtCYP121* active site is unable to properly accommodate the product, is consistent with recent results showing that mcyc product does not inhibit the protein possible due to an incapacity of binding (no binding was observed at $80\mu M$) [16]. The occurrence of

the C-C bond formation step in solution has several interesting implications. First, similarly to the cY^*Y , it could be argued, that a cY^*Y^* in solution could give rise to undesired side products.

● However, it is important to note that since **both** Tyr radicals belong to the same molecule, they are linked by the diketopiperazine cycle formed by the respective Tyr backbone groups, the chance of intramolecular bond formation is far more likely than intermolecular formation of cYY dimers, specially if the biradical concentration is small. Also important, Tyr radicals are quite stable to oxidation by molecular oxygen, compared to other residues, such as Trp or Cys ^[73] and thus live long enough in solution to yield the desired product. Another implication of non catalyzed C-C bond formation, is related to the development of selective inhibitors based on substrate and or product analogs. Our results suggest that substrate analogs are better candidates to bind MtCYP121 compared to product analogs, since the product is formed outside the protein matrix.

An important issue of the present mechanistic proposal, is that several predictions are made, that could be confirmed by experimental means. For example addition of an external spin trap to the enzyme in the presence of substrate, could lead to trapping/identification of the cY^*Y^* , but not the cY^*Y . Moreover, mutation of Arg386 to a neutral residue (like Gln or even Leu), could by retarding the IET, to some accumulation of cY^*Y and/or significant changes in the overall rate. The same or even an additive effect, could be achieved by performing mutations that stabilize a negative charge in the *far*-Tyr, as for example Gln385Arg and/or Phe76Arg/Lys.

Last but not least, the results presented contribute to a deeper understanding of the general mechanisms of CYP enzymes, in particularly those that perform C-C coupling reactions. In this context, an interesting comparison can be performed with previous QM/MM studies of CYP StaP which catalyzes the intramolecular C-C bond formation in Chromopyrrolic acid in a key step of Staurosporine biosynthesis ^[74]. In the mentioned work the authors were able to show that since the

reactive carbon is too far from Cpd I ferryl oxygen, the first radical is formed by Cpd I, through a PCET process, in which a proton from the substrate (derived from an NH group adjacent to the reactive carbon) is transferred to Cpd I through a water mediated hydrogen bond array, concomitantly with an ET process. The barrier for this step is 16.3 kcal/mol. The second radical in this case is formed concertedly with C-C bond formation and a PCET to Cpd II, yielding a barrier of 20 kcal/mol. Thus in this case since both reactive carbons are far from the reactive ferryl groups, PCET process are operative, as in the second radical generation step presented here for MtCYP121. Also interesting, is the fact that as in the present case, the step involving C-C bond formation displays the larger barrier.

Thus, it would seem that for CYPs involved in C-C coupling reactions, Cpd I is retained as the main oxidant and starter of the reaction cycle and that Cpd II acts as the second oxidant, allowing C-C bond formation to happen in one enzyme oxidation cycle. However, since in one or both radical generation steps the reactive carbon can be positioned far from the iron (possibly to avoid unproductive hydroxylation products), the mechanism operates through a PCET process possibly enhanced by the protein matrix ^[74].

Conclusions

In the present work we have used MD and QM/MM methods to study MtCYP121 reaction mechanism. Our results show that substrate binding promotes formation of the initial oxy complex. That Cpd I is the responsible for first Tyr* formation and that the second Tyr* is formed subsequently, though PCET reaction, promoted by the presence of key residue Arg386, that leads to the heme ferric resting state and avoids OH radical rebinding. The final C-C coupling reaction possibly occurs in bulk solution, thus yielding the product (mcyc) in one oxygen reduction cycle.

Acknowledgements

Computer power was provided by Centro de Computación de Alto Rendimiento (C.E.C.A.R.) at the FCEN-UBA and by the cluster MCG PME No 2006-01581 at the Universidad Nacional de Córdoba, and to Grants PICT-2010-416, UBACyT 2010-2012, PICTO-2012-0057 and Bunge y Born grant for infectious diseases awarded to MAM. This work was granted access to the HPC resources of CINECA made available within the Distributed European Computing Initiative by the PRACE-2IP, receiving funding from the European Community's Seventh Framework Programme (FP7/2007-2013) under grant agreement RI-283493. VGD, AAP and LAD hold CONICET fellowships. AGT and MAM are staff members of CONICET.

References

- [1] A. Koul, E. Arnoult, N. Lounis, J. Guillemont, K. Andries, *Nature* **2011**, *469*, 483–490.
- [2] S. T. Cole, R. Brosch, J. Parkhill, T. Garnier, C. Churcher, D. Harris, S. V. Gordon, K. Eglmeier, S. Gas, C. E. Barry, et al., *Nature* **1998**, *393*, 537–544.
- [3] K. J. McLean, D. Clift, D. G. Lewis, M. Sabri, P. R. Balding, M. J. Sutcliffe, D. Leys, A. W. Munro, *Trends in Microbiology* **2006**, *14*, 220–228.
- [4] S. Shaik, H. Hirao, D. Kumar, *Nat. Prod. Rep.* **2007**, *24*, 533–552.
- [5] V. F. Kalb, C. W. Woods, T. G. Turi, C. R. Dey, T. R. Sutter, J. C. Loper, *DNA* **1987**, *6*, 529–537.
- [6] K. J. McLean, K. R. Marshall, A. Richmond, I. S. Hunter, K. Fowler, T. Kieser, S. S. Gurcha, G. S. Besra, A. W. Munro, *Microbiology* **2002**, *148*, 2937–2949.

- [7] H. Ouellet, J. B. Johnston, P. R. Ortiz de Montellano, *Archives of Biochemistry and Biophysics* **2010**, *493*, 82–95.
- [8] K. J. McLean, P. Carroll, D. G. Lewis, A. J. Dunford, H. E. Seward, R. Neeli, M. R. Cheesman, L. Marsollier, P. Douglas, W. E. Smith, et al., *J Biol Chem* **2008**, *283*, 33406–33416.
- [9] P. Belin, M. H. L. Du, A. Fielding, O. Lequin, M. Jacquet, J.-B. Charbonnier, A. Lecoq, R. Thai, M. Courçon, C. Masson, et al., *PNAS* **2009**, *106*, 7426–7431.
- [10] D. Leys, C. G. Mowat, K. J. McLean, A. Richmond, S. K. Chapman, M. D. Walkinshaw, A. W. Munro, *J. Biol. Chem.* **2003**, *278*, 5141–5147.
- [11] H. E. Seward, A. Roujeinikova, K. J. McLean, A. W. Munro, D. Leys, *J. Biol. Chem.* **2006**, *281*, 39437–39443.
- [12] S. A. Hudson, K. J. McLean, S. Surade, Y.-Q. Yang, D. Leys, A. Ciulli, A. W. Munro, C. Abell, *Angewandte Chemie International Edition* **2012**, *51*, 9311–9316.
- [13] K. J. McLean, M. R. Cheesman, S. L. Rivers, A. Richmond, D. Leys, S. K. Chapman, G. A. Reid, N. C. Price, S. M. Kelly, J. Clarkson, et al., *Journal of Inorganic Biochemistry* **2002**, *91*, 527–541.
- [14] K. G. Ravichandran, S. S. Boddupalli, C. A. Hasermann, J. A. Peterson, J. Deisenhofer, *Science* **1993**, *261*, 731–736.
- [15] J. K. Yano, L. S. Koo, D. J. Schuller, H. Li, P. R. O. de Montellano, T. L. Poulos, *J. Biol. Chem.* **2000**, *275*, 31086–31092.
- [16] M. Fonvielle, M.-H. L. Du, O. Lequin, A. Lecoq, M. Jacquet, R. Thai, S. Dubois, G. Grach, M. Gondry, P. Belin, *J. Biol. Chem.* **2013**, DOI 10.1074/jbc.M112.443853.

- [17] N. R. Melo, G. P. Moran, A. G. S. Warrillow, E. Dudley, S. N. Smith, D. J. Sullivan, D. C. Lamb, D. E. Kelly, D. C. Coleman, S. L. Kelly, *Antimicrob. Agents Chemother.* **2008**, *52*, 3718–3724.
- [18] M. Makino, H. Sugimoto, Y. Shiro, S. Asamizu, H. Onaka, S. Nagano, *PNAS* **2007**, *104*, 11591–11596.
- [19] B. Zhao, F. P. Guengerich, M. Voehler, M. R. Waterman, *J. Biol. Chem.* **2005**, *280*, 42188–42197.
- [20] N. Grobe, B. Zhang, U. Fisinger, T. M. Kutchan, M. H. Zenk, F. P. Guengerich, *J Biol Chem* **2009**, *284*, 24425–24431.
- [21] A. Gesell, M. Rolf, J. Ziegler, M. L. D. Chávez, F.-C. Huang, T. M. Kutchan, *J. Biol. Chem.* **2009**, *284*, 24432–24442.
- [22] N. Ikezawa, K. Iwasa, F. Sato, “Molecular Cloning and Characterization of CYP80G2, a Cytochrome P450 That Catalyzes an Intramolecular C–C Phenol Coupling of (S)-Reticuline in Magnoflorine Biosynthesis, from Cultured *Coptis japonica* Cells,” can be found under <http://www.jbc.org>, **n.d.**
- [23] A. Warshel, M. Levitt, *Journal of Molecular Biology* **1976**, *103*, 227–249.
- [24] A. J. Mulholland, *Drug Discovery Today* **2005**, *10*, 1393–1402.
- [25] S. C. L. Kamerlin, M. Haranczyk, A. Warshel, *J. Phys. Chem. B* **2009**, *113*, 1253–1272.
- [26] N. V. Plotnikov, S. C. L. Kamerlin, A. Warshel, *J. Phys. Chem. B* **2011**, *115*, 7950–7962.
- [27] N. V. Plotnikov, A. Warshel, *J. Phys. Chem. B* **2012**, *116*, 10342–10356.
- [28] E. Rosta, M. Klähn, A. Warshel, *J. Phys. Chem. B* **2006**, *110*, 2934–2941.
- [29] A. Warshel, *Annual Review of Biophysics and Biomolecular Structure* **2003**, *32*, 425–443.

- [30] K.-B. Cho, H. Hirao, H. Chen, M. A. Carvajal, S. Cohen, E. Derat, W. Thiel, S. Shaik, *J. Phys. Chem. A* **2008**, *112*, 13128–13138.
- [31] H. Hirao, K.-B. Cho, S. Shaik, *J Biol Inorg Chem* **2008**, *13*, 521–530.
- [32] J. I. Mujika, X. Lopez, E. Rezabal, R. Castillo, S. Marti, V. Moliner, J. M. Ugalde, *Journal of Inorganic Biochemistry* **2011**, *105*, 1446–1456.
- [33] W. Lai, S. Shaik, *J. Am. Chem. Soc.* **2011**, *133*, 5444–5452.
- [34] J. C. Schöneboom, F. Neese, W. Thiel, *J. Am. Chem. Soc.* **2005**, *127*, 5840–5853.
- [35] S. Shaik, S. Cohen, Y. Wang, H. Chen, D. Kumar, W. Thiel, *Chem. Rev.* **2010**, *110*, 949–1017.
- [36] A. Crespo, M. A. Martí, D. A. Estrin, A. E. Roitberg, *J. Am. Chem. Soc.* **2005**, *127*, 6940–6941.
- [37] A. Crespo, M. A. Martí, A. E. Roitberg, L. M. Amzel, D. A. Estrin, *J. Am. Chem. Soc.* **2006**, *128*, 12817–12828.
- [38] A. Crespo, M. A. Martí, S. G. Kalko, A. Morreale, M. Orozco, J. L. Gelpi, F. J. Luque, D. A. Estrin, *J. Am. Chem. Soc.* **2005**, *127*, 4433–4444.
- [39] A. Lewis-Ballester, D. Batabyal, T. Egawa, C. Lu, Y. Lin, M. A. Marti, L. Capece, D. A. Estrin, S.-R. Yeh, *PNAS* **2009**, *106*, 17371–17376.
- [40] A. G. Turjanski, G. Hummer, J. S. Gutkind, *J. Am. Chem. Soc.* **2009**, *131*, 6141–6148.
- [41] D. A. Pearlman, D. A. Case, J. W. Caldwell, W. S. Ross, T. E. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel, P. Kollman, *Computer Physics Communications* **1995**, *91*, 1–41.
- [42] V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg, C. Simmerling, *Proteins: Structure, Function, and Bioinformatics* **2006**, *65*, 712–725.

- [43] M. A. Marti, A. Crespo, L. Capece, L. Boechi, D. E. Bikiel, D. A. Scherlis, D. A. Estrin, *Journal of Inorganic Biochemistry* **2006**, *100*, 761–770.
- [44] L. Capece, M. A. Marti, A. Bidon-Chanal, A. Nadra, F. J. Luque, D. A. Estrin, *Proteins: Structure, Function, and Bioinformatics* **2009**, *75*, 885–894.
- [45] A. Bidon-Chanal, M. A. Martí, D. A. Estrín, F. J. Luque, in *SelfOrganization of Molecular Systems* (Eds.: N. Russo, V.Y. Antonchenko, E.S. Kryachko), Springer Netherlands, **2009**, pp. 33–47.
- [46] L. Boechi, M. A. Martí, M. Milani, M. Bolognesi, F. J. Luque, D. A. Estrin, *Proteins: Structure, Function, and Bioinformatics* **2008**, *73*, 372–379.
- [47] L. Capece, D. A. Estrin, M. A. Marti, *Biochemistry* **2008**, *47*, 9416–9427.
- [48] A. D. Nadra, M. A. Martí, A. Pesce, M. Bolognesi, D. A. Estrin, *Proteins: Structure, Function, and Bioinformatics* **2008**, *71*, 695–705.
- [49] L. Capece, L. Boechi, L. L. Perissinotti, P. Arroyo-Mañez, D. E. Bikiel, G. Smulevich, M. A. Marti, D. A. Estrin, *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* **n.d.**, DOI 10.1016/j.bbapap.2013.02.038.
- [50] C. I. Bayly, P. Cieplak, W. Cornell, P. A. Kollman, *J. Phys. Chem.* **1993**, *97*, 10269–10280.
- [51] A. A. Petruk, S. Bartesaghi, M. Trujillo, D. A. Estrin, D. Murgida, B. Kalyanaraman, M. A. Marti, R. Radi, *Archives of Biochemistry and Biophysics* **2012**, *525*, 82–91.
- [52] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *The Journal of Chemical Physics* **1984**, *81*, 3684–3690.
- [53] Andrew Leach, *Pearson - Molecular Modelling: Principles and Applications*, **n.d.**
- [54] D. A. Case, T. E. Cheatham, T. Darden, H. Gohlke, R. Luo, K. M. Merz, A. Onufriev, C. Simmerling, B. Wang, R. J. Woods, *Journal of Computational Chemistry* **2005**, *26*, 1668–1688.

- [55] A. Crespo, D. A. Scherlis, M. A. Martí, P. Ordejón, A. E. Roitberg, D. A. Estrin, *J. Phys. Chem. B* **2003**, *107*, 13728–13736.
- [56] J. P. Perdew, K. Burke, M. Ernzerhof, *Phys. Rev. Lett.* **1996**, *77*, 3865–3868.
- [57] D. E. Bikiel, L. Boechi, L. Capece, A. Crespo, P. M. D. Biase, S. D. Lella, M. C. G. Lebrero, M. A. Martí, A. D. Nadra, L. L. Perissinotti, et al., *Phys. Chem. Chem. Phys.* **2006**, *8*, 5611–5628.
- [58] G. M. Torrie, J. P. Valleau, *Journal of Computational Physics* **1977**, *23*, 187–199.
- [59] J. Torras, G. de M. Seabra, A. E. Roitberg, *J. Chem. Theory Comput.* **2009**, *5*, 37–46.
- [60] A. Laio, M. Parrinello, *PNAS* **2002**, *99*, 12562–12566.
- [61] L. L. Perissinotti, M. A. Marti, F. Doctorovich, F. J. Luque, D. A. Estrin, *Biochemistry* **2008**, *47*, 9793–9802.
- [62] L. Capece, A. Lewis-Ballester, S.-R. Yeh, D. A. Estrin, M. A. Marti, *J. Phys. Chem. B* **2012**, *116*, 1401–1413.
- [63] L. Capece, A. Lewis-Ballester, D. Batabyal, N. D. Russo, S.-R. Yeh, D. A. Estrin, M. A. Marti, *J Biol Inorg Chem* **2010**, *15*, 811–823.
- [64] S. Hammes-Schiffer, A. V. Soudackov, *J. Phys. Chem. B* **2008**, *112*, 14108–14123.
- [65] J. Mongan, D. A. Case, J. A. McCammon, *Journal of Computational Chemistry* **2004**, *25*, 2038–2048.
- [66] T. Simonson, J. Carlsson, D. A. Case, *J. Am. Chem. Soc.* **2004**, *126*, 4167–4180.
- [67] E. J. Wood, *Biochemical Education* **1987**, *15*, 97–97.
- [68] J. Gao, P. Müller, M. Wang, S. Eckhardt, M. Lauz, K. M. Fromm, B. Giese, *Angewandte Chemie International Edition* **2011**, *50*, 1926–1930.
- [69] B. Giese, M. Graber, M. Cordes, *Current Opinion in Chemical Biology* **2008**, *12*, 755–759.

- [70] J. C. Schöneboom, S. Cohen, H. Lin, S. Shaik, W. Thiel, *J. Am. Chem. Soc.* **2004**, *126*, 4017–4034.
- [71] D. N. Ferreiro, L. Boechi, D. A. Estrin, M. A. Martí, *Journal of Inorganic Biochemistry* **2013**, *119*, 75–84.
- [72] I. L. Shamovsky, R. J. Riopelle, G. M. Ross, *J. Phys. Chem. A* **2001**, *105*, 1061–1070.
- [73] E. P. Hunter, M. F. Desrosiers, M. G. Simic, *Free Radic. Biol. Med.* **1989**, *6*, 581–585.
- [74] Y. Wang, H. Chen, M. Makino, Y. Shiro, S. Nagano, S. Asamizu, H. Onaka, S. Shaik, *J. Am. Chem. Soc.* **2009**, *131*, 6748–6762.

Figure 1. Structure of the ferric cYY bound MtCYP121. The cYY substrate (labeling the Close and Far Tyr), the heme group, and residues Cys345 (the proximal ligand), Arg386 and Ser237 are shown as sticks. Active site waters are shown as cyan spheres. Substrate binding pocket is shown as yellow isosurface.

Scheme 1. A: Overall reaction catalyzed by MtCYP121, showing both the structure of the reactive ciclo-di-tyrosine (cYY) and product mycocylosine (mcy). B: Proposed reaction mechanism of MtCYP121. Relevant heme oxidation/coordination states are labeled as numbers: 0: Ferric resting state, 1: Compound I with substrate bound, 2 & 3: Compound II with one Tyr radical located at the close and far-Tyr respectively, 4: Ferric di-radical bound and 5: ferric mycocylosine product (mcy) bound state. Reactions are labeled as (letters): (a) oxygen binding and reduction, (b) O-O bond cleavage and water release (c) first radical formation in the close-Tyr, (d) Tyr Tyr intramolecular electron transfer inside the protein, d') cYY radical release and reenter to MtCYP121

with radical located in the far-Tyr, e) second Tyr radical formation and water release, f) C-C bond formation inside MtCYP121 and mcyc release, f') C-C bond formation in solution.

Figure 2. a) Energy profile for the hydrogen atom abstraction and Tyr* formation using Cpd I as the oxidant. b) Energy profile for the hydrogen atom abstraction and Tyr* formation using Cpd II as the oxidant. c) Energy profile for the Cpd II OH rebound step. In Figure A) and C) Red and Blue lines show the quartet and doublet spin states, respectively. In figure B) Red and Blue lines correspond to triplet and singlet spin states, respectively.

Figure 3. Classical analysis of Tyr* to Tyr PCET using TI. $\Delta\Delta G$ values correspond to the free energies for each process (step) inside the protein with respect to the same process in water. Values for wt CYP121 and R386A mutant are shown in red and blue respectively.

Figure 4. Histograms for all possible four relevant C-C distances obtained from simulation of cY*Y* free in solution (blue) and inside protein (red).

Figure 5. (A) Energy profile as a function of the reaction coordinate for the complete reaction in vacuum. For the first step of the reaction, the calculations were done for both S (blue) and T (red) spin states. For the subsequent steps, only the profiles for S state are shown. The // symbol in the X-axis indicates the change in the reaction coordinate. (B) Energy profile for C-C bond formation in solution (solid line) and inside the protein (dotted lines) for both S and T spin states. (C) Key structures along the reaction of mcyc (mycrocyclosine) formation.

Accepted Article

Scheme 2. Proposed four steps MtCYP121 reaction mechanism.

Table I. Heme-ligand close-Tyr interactions. Reported values were taken from the respective MD and correspond to average distances (\AA) \pm SD. Values in the last row shows FeO1-HOH-TyrO angle.

Parameter ^a	Protein Heme State		
	Oxy	CpI	CpII
FeO1-TyrO _{OH}	3.61 \pm 0.47	5.19 \pm 0.29	5.31 \pm 0.28
FeO1-TyrH _{OH}	2.83 \pm 0.68	4.31 \pm 0.28	4.43 \pm 0.28
FeO1-TyrH _{CA}	2.68 \pm 0.35	4.96 \pm 0.40	5.12 \pm 0.32
FeO2-TyrH _{OH}	2.21 \pm 0.66	-	-
FeO2-TyrH _{CA}	2.86 \pm 0.31	-	-
FeO1-H _{OH} -TyrO _{OH}	169°	158°	166°

^a) FeO1 corresponds to the oxygen atom bound to the iron, FeO₂ corresponds to the oxygen atom in the O₂ molecule which is not bound to the iron, TyrO_{OH} Corresponds to the Tyrosine phenol oxygen atom, TyrH_{OH} Corresponds to the Tyrosine phenol hydrogen and TyrH_{CA} corresponds to the Tyrosine aromatic hydrogen in ortho position.

Table II. Selected distances (*d*), Mulliken (*q*) and spin (*s*) populations relevant for the first close-Tyr* (*cY**Y) formation step in MtCYP121.

Parameter	System					
	Cpd I		TS		Cpd II	
	D	Q	D	Q	D	Q
ΔE spin	1.0	0	n.c	n.c	3.0	0
d Fe-O	1.68	1.69	1.74	1.74	1.83	1.83
d FeO-TyrH _{OH}	1.79	1.65	1.16	1.18	1.00	1.00
d TyrO-TyrH _{OH}	1.00	1.01	1.26	1.25	1.99	1.97
q close-Tyr	-0.012	0.205	-0.048	-0.097	-0.075	-0.070
q OX(FeOX)	-0.337	-0.348	-0.284	-0.241	-0.189	-0.190
q TyrH _{OH}	-0.125	-0.133	-0.154	-0.172	-0.115	-0.115
q Heme-Cys	0.131	0.136	0.329	0.330	0.263	0.262
s close-Tyr	-0.118	0.164	-0.234	0.475	-0.981	0.990
s OX(FeOX)	0.697	0.762	0.656	0.605	0.261	0.270
s TyrH _{OH}	-0.001	-0.003	0.000	-0.006	0.005	-0.001
s Heme-Cys	0.424	2.075	0.581	1.920	1.982	1.744

Table III. Selected distance (*d*), and Mulliken (*q*) and spin (*s*) populations relevant for the second Tyr* formation step in MtCYP121. *L* corresponds to the heme bound ligand, i.e the OH in Cpd II and water molecule in the product.

Parameter	System						
	CpdII		TS		Fe3-H ₂ O		
	S	T	S	T	S	T	Q
ΔE spin	3.5	0	n.c	n.c	1.0	0	
d Fe-O	1.93	1.91	2.00	2.02	2.01	2.02	2.32
d FeO-TyrH _{OH}	1.55	1.62	1.20	1.08	1.08	1.08	1.03
d TyrH _{OH} -TyrO	1.04	1.02	1.24	1.40	1.40	1.40	1.56
q Tyr	-0.041	-0.019	-0.092	-0.146	-0.120	-0.144	-0.102
q L	-0,393	-0,410	n.c	n.c	-0,093	-0,073	-0.003
q Heme-Cys	0.435	0.414	0.299	0.262	0.213	0.229	0.183
s Tyr	-0.143	0.125	0.454	0.508	0.610	0.568	0.764
s L	0,049	0,088	n.c	n.c	-0,013	0,019	0.029
s Heme-Cys	0.095	1.792	-0.435	1.472	-0.599	1.415	3.209

Table IV. Free Energy changes associated with radical migration as determined by Thermodynamic Integration. Free energy values are in kcal/mol. Absolute error is estimated as the SD of the obtained values from two independent simulation sets.

Step	In water	Inside protein		$\Delta\Delta G$ wild type – water	$\Delta\Delta G$ R386A – water
		Wild type	R386A		
1. Far Tyr deprotonation	-65.0 ± 0.2	-47.7 ± 0.8	-46.8 ± 0.6	17.3 ± 1.0	18.2 ± 0.8
2. Far to close Tyr ET	-1.0 ± 0.1	-29.8 ± 1.5	-11.5 ± 0.5	-28.8 ± 1.6	-10.5 ± 0.6
3. Close Tyr protonation	60.9 ± 0.2	58.4 ± 0.4	60.6 ± 0.4	-2.5 ± 0.6	-0.3 ± 0.6
Total $\Delta\Delta G$				-14.0 ± 3.2	7.6 ± 2.0

Table V. Mulliken (q) and spin (s) populations for mono-radical and negatively charged substrate (cYY^{*-}) directly related with the IET process.

Parameter	QM Sub-System		
	[cYY] ^{*-}	CpII [cYY] ^{*-}	
		D	Q
q <i>close</i> -Tyr	-0.778	-0.007	-0.065
q <i>far</i> -Tyr	-0.222	-0.325	-0.299
s <i>close</i> -Tyr	0.430	-0.997	0.909
s <i>far</i> -Tyr	0.570	0.673	0.703
q heme-Cys	n.c	-0.188	-0.146
s heme-Cys	n.c	1.207	1.262
q L	n.c	-0.476	-0.486
s L	n.c	0.114	0.126

Table VI. Energetic, selected distance (d), and Mulliken (q) and spin (s) populations for the C-C bond formation reaction in vacuum (reference), solvent and inside protein.

Parameter	Energetic Parameters					
	Vacuum		Solvent		Protein	
	T	S	T	S	T	S
ΔE_{spin}	0	10	0	19	0	11
$E_{\text{act}}^{\text{cY}^*\text{Y}^* \Rightarrow \text{IS}}$	22	10	23	10	40	25
$E_{\text{act}}^{\text{-Effective}}$	16		16		32	
$\Delta E_{\text{reaction}}$	-7		-15		6	
Structural Parameters, Mulliken and Spin Populations for the complete reaction in vacuum						
	$\text{cY}^*\text{Y}^*(\text{T})$	$\text{TS}_\text{A}(\text{S}/\text{T})$	$\text{IS}_\text{A}(\text{S})$	$\text{TS}_\text{B}(\text{S})$	$\text{IS}_\text{B}(\text{S})$	$\text{mcy}(\text{S})$
$d_{\text{C}\varepsilon\text{-C}\varepsilon}$	4.52	2.59	1.59	1.56	1.55	1.52
$d_{\text{C}\varepsilon\text{-H}\varepsilon}$	1.11	1.11	1.13	1.51	2.42	2.38
$d_{\text{H}\varepsilon\text{-O}\zeta}$ (same)	2.65	2.67	2.69	3.04	3.30	3.85
$d_{\text{H}\varepsilon\text{-O}\zeta}$ (diff)	5.25	2.64	2.49	1.24	1.00	0.99
C ε planar angle	89.64°	65.60	61.04°	58.32°	62.95°	68.42°
q_{Tyr1}	0.029	0.019	0.040	0.198	0.350	0.014
q_{Tyr2}	-0.032	-0.021	-0.038	-0.194	-0.350	-0.010
s_{Tyr1}	1.013	0.000	0.000	0.000	0.000	0.000
s_{Tyr2}	0.988	0.000	0.000	0.000	0.000	0.000

Table VII. Summary of the energetic results (in kcal/mol) for each Step of the MtCYP121 reaction mechanism.

Step	Spin-state ^a	ΔE	ΔE^\ddagger	Method
First radical formation by Cpd I	Q	-	4 ± 0.1	QM-MM
OH Rebound	Q	-	16 ± 0.2	QM-MM
PCET-Classical Contribution ^b	D	-14.0 ± 3.2	n.c ^d	TI
Second radical formation by Cpd II	T	3.5 ± 0.1	3.5 ± 0.1	QM-MM
C-C bond formation in vacuum and solution ^c	T => S	-7 ± 0.1 (-15 ± 0.1)	16 ± 0.1 (16 ± 0.1)	QM and (QM-MM)

^a) Reference Spin state corresponds to the ground state where the reaction occurs. If the reaction involves spin states crossing the two states are mentioned. ^b) In this cases value corresponds to free energy (i.e ΔG) instead of potential energy. ^c) Values in parenthesis correspond to those obtained in solution. ^d) not computed

Accepted

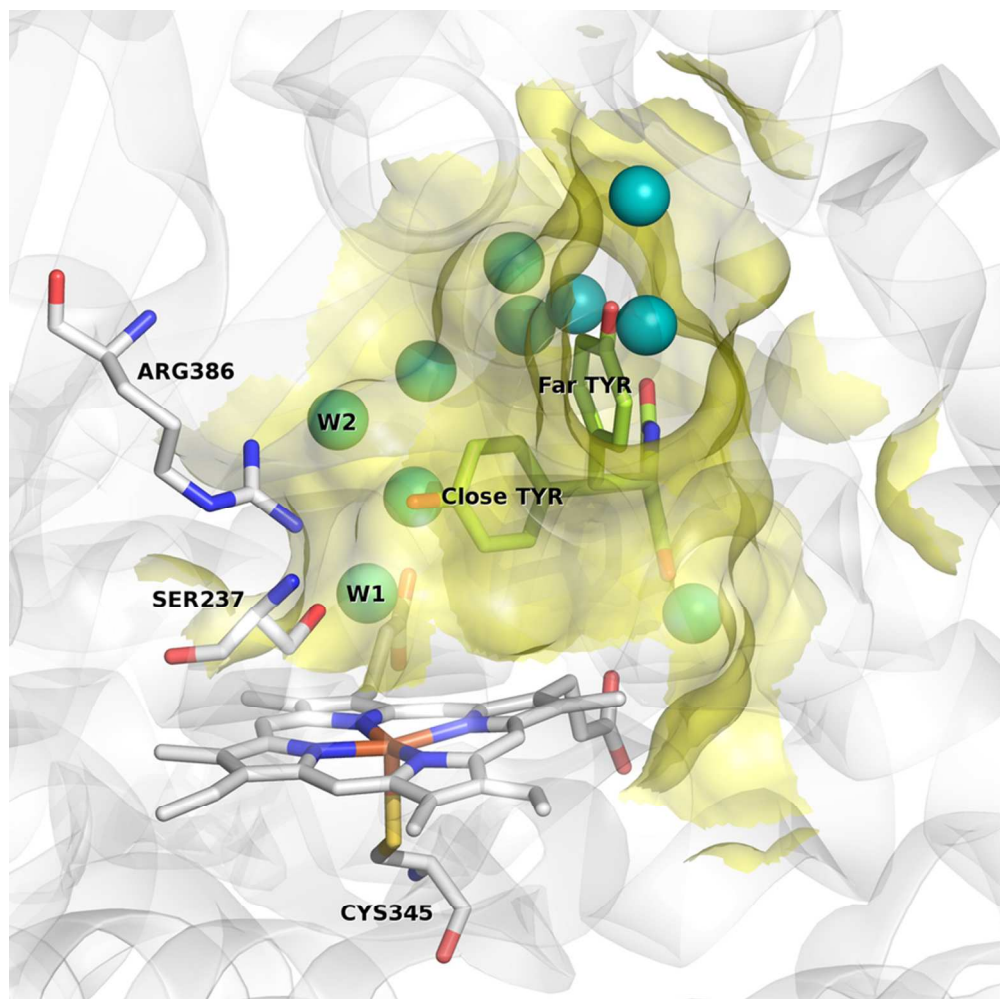


Figure 1. Structure of the ferric cYY bound MtCYP121. The cYY substrate (labeling the Close and Far Tyr), the heme group, and residues Cys345 (the proximal ligand), Arg386 and Ser237 are shown as sticks. Active site waters are shown as cyan spheres. Substrate binding pocket is shown as yellow isosurface.
79x79mm (300 x 300 DPI)

Acc

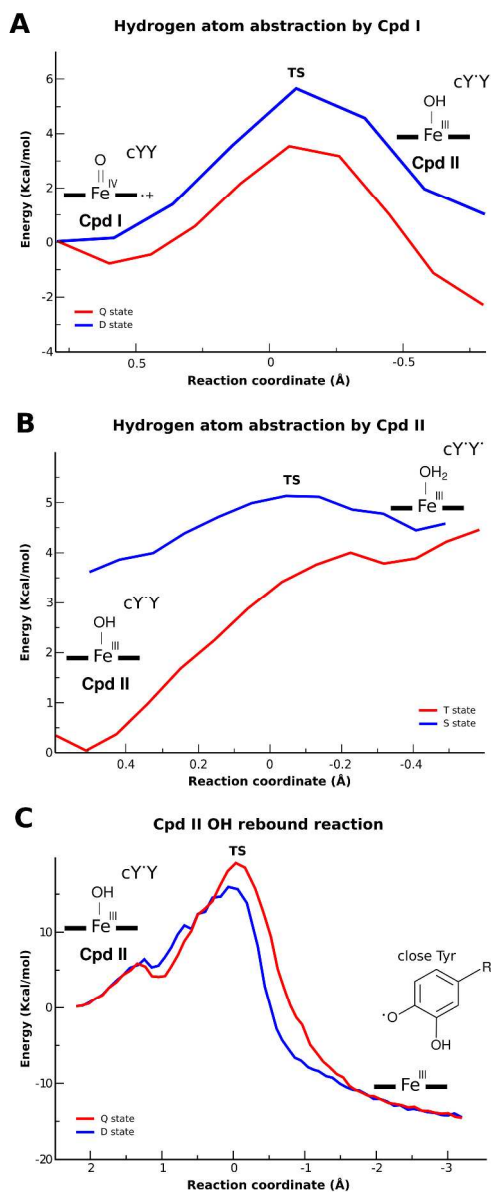


Figure 2. a) Energy profile for the hydrogen atom abstraction and Tyr* formation using Cpd I as the oxidant. b) Energy profile for the hydrogen atom abstraction and Tyr* formation using Cpd II as the oxidant. c) Energy profile for the Cpd II OH rebound step. In Figure A) and C) Red and Blue lines show the quartet and doublet spin states, respectively. In figure B) Red and Blue lines correspond to triplet and singlet spin states, respectively.

326x772mm (300 x 300 DPI)

A

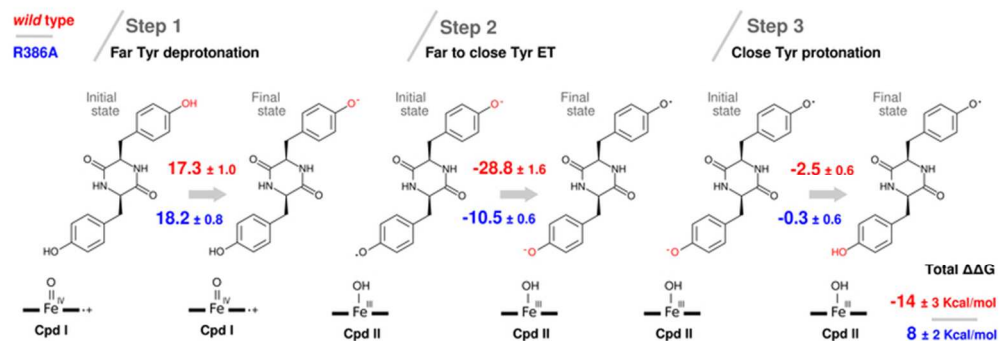


Figure 3. Classical analysis of Tyr* to Tyr PCET using TI. $\Delta\Delta G$ values correspond to the free energies for each process (step) inside the protein with respect to the same process in water. Values for wt CYP121 and R386A mutant are shown in red and blue respectively.
68x24mm (300 x 300 DPI)

Accepted

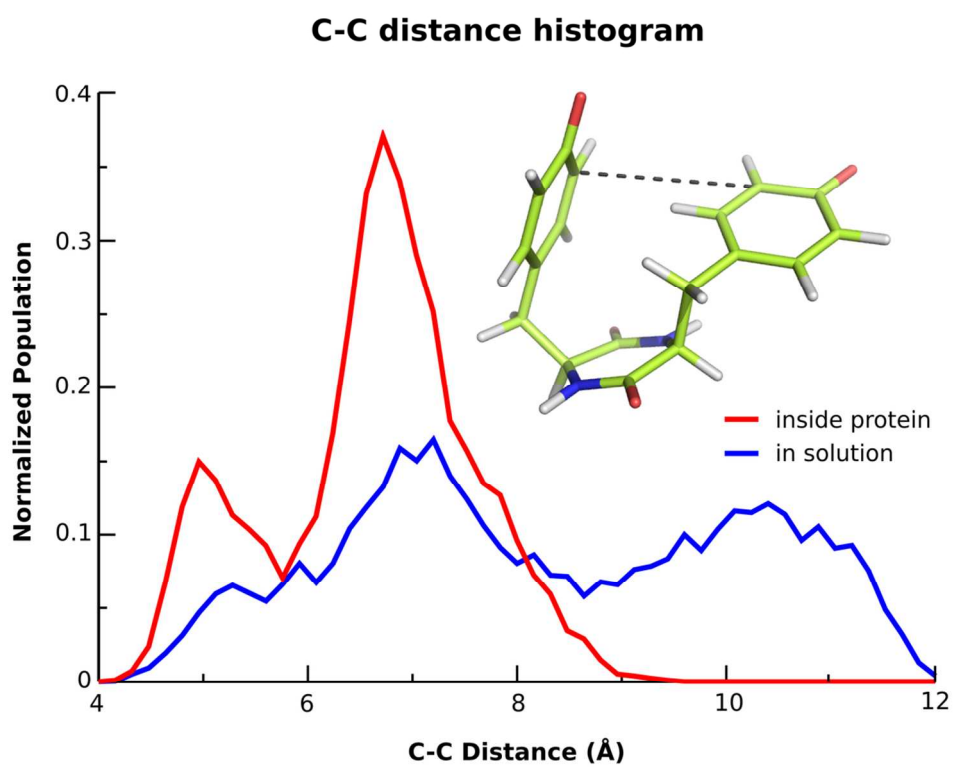


Figure 4. Histograms for all possible four relevant C-C distances obtained from simulation of cY*Y* free in solution (blue) and inside protein (red).
107x84mm (300 x 300 DPI)

Accepted

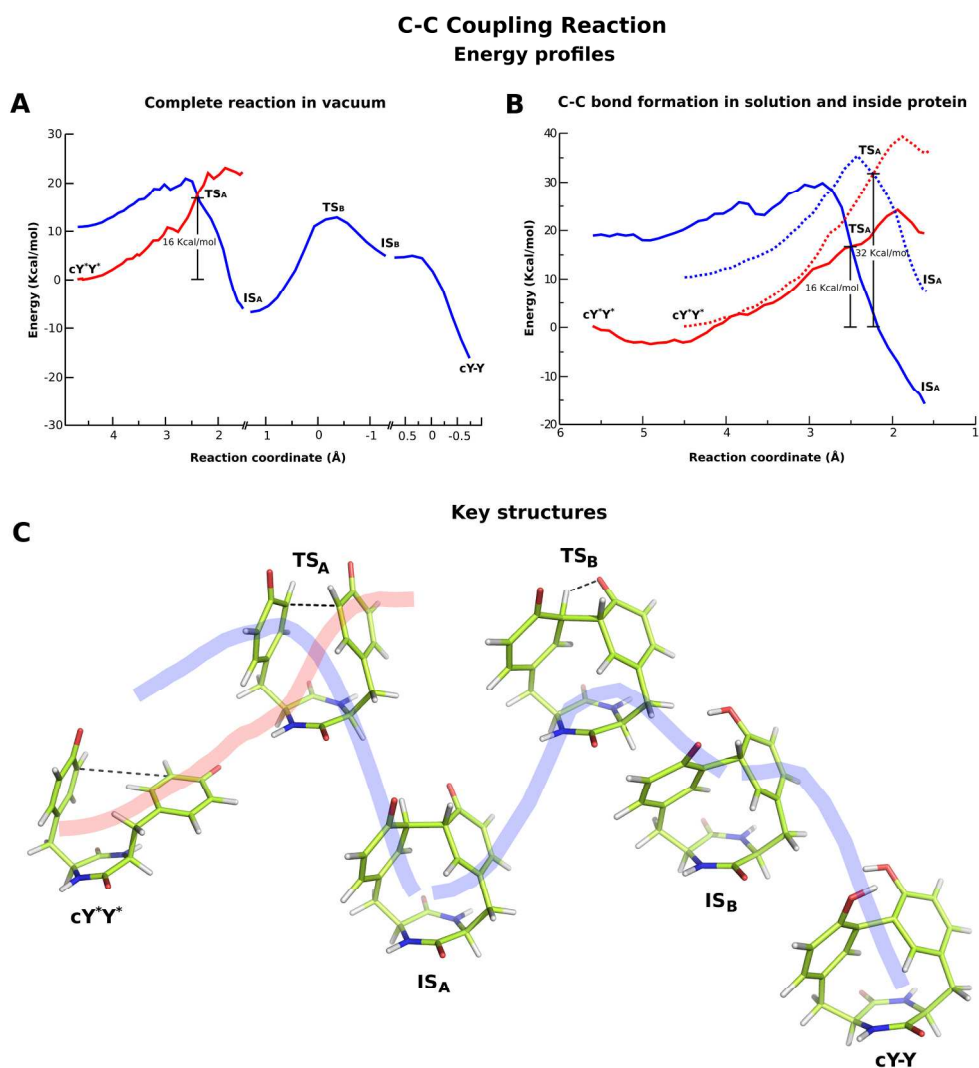
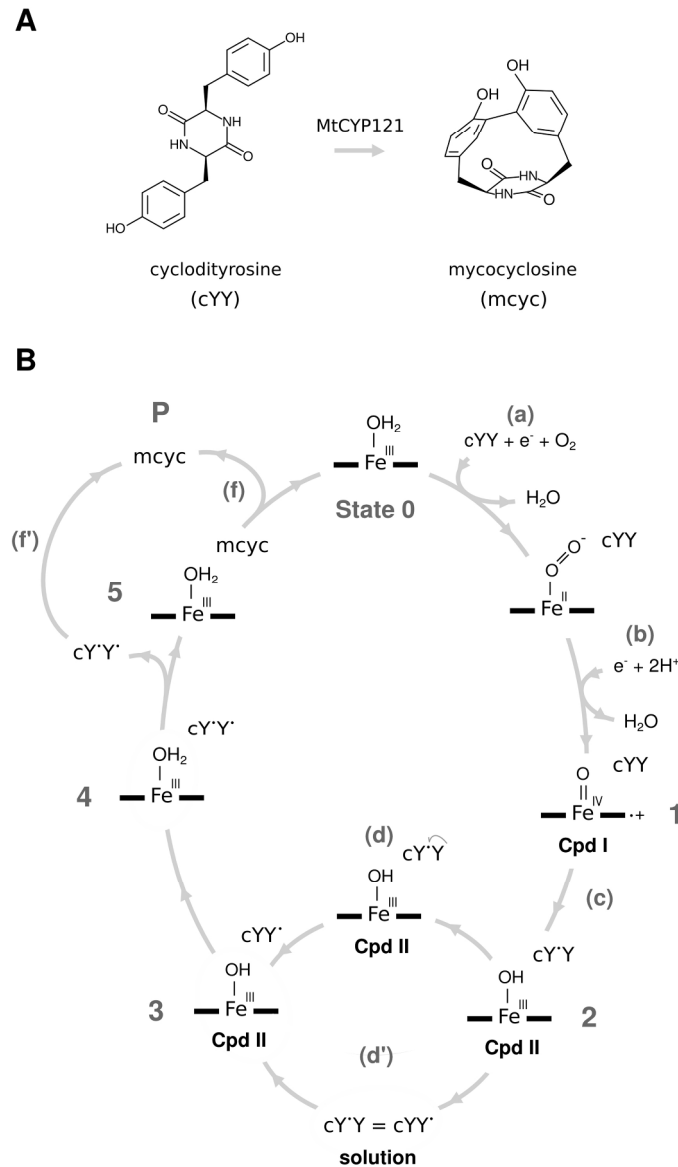


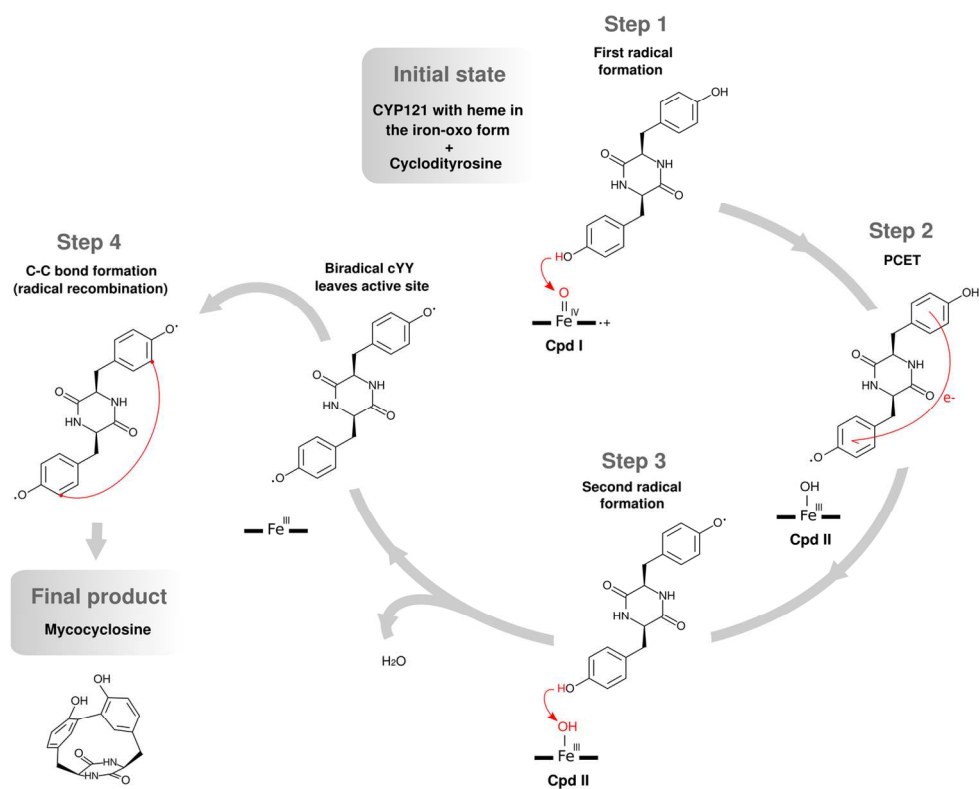
Figure 5. (A) Energy profile as a function of the reaction coordinate for the complete reaction in vacuum. For the first step of the reaction, the calculations were done for both S (blue) and T (red) spin states. For the subsequent steps, only the profiles for S state are shown. The // symbol in the X-axis indicates the change in the reaction coordinate. (B) Energy profile for C-C bond formation in solution (solid line) and inside the protein (dotted lines) for both S and T spin states. (C) Key structures along the reaction of mcyc (mycrocyclosine) formation.
213x233mm (300 x 300 DPI)

A



Scheme 1. A: Overall reaction catalyzed by MtCYP121, showing both the structure of the reactive ciclo-di-tyrosine (cYY) and product mycocyclosine (mcyC). B: Proposed reaction mechanism of MtCYP121. Relevant heme oxidation/coordination states are labeled as numbers: 0: Ferric resting state, 1: Compound I with substrate bound, 2 & 3: Compound II with one Tyr radical located at the close and far-Tyr respectively, 4: Ferric di-radical bound and 5: ferric mycocyclosine product (mcyC) bound state. Reactions are labeled as (letters): (a) oxygen binding and reduction, (b) O-O bond cleavage and water release (c) first radical formation in the close-Tyr, (d) Tyr Tyr intramolecular electron transfer inside the protein, (d') cYY radical release and re-entrance to MtCYP121 with radical located in the far-Tyr, (e) second Tyr radical formation and water release, (f) C-C bond formation inside MtCYP121 and mcyC release, (f') C-C bond formation in solution.

155x263mm (300 x 300 DPI)



Scheme 2. Proposed four steps MtCYP121 reaction mechanism.
141x116mm (300 x 300 DPI)

Accep