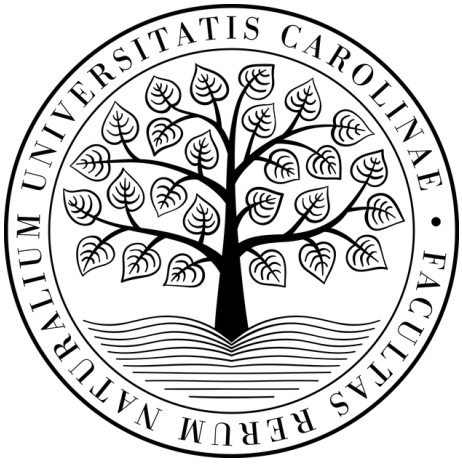
Charles University

Faculty of Science

Study programme:
Modeling of Chemical Properties of Nano- and Biostructures



Agnieszka Stańczak, M.Eng.

Theoretical Investigation of the Enzymatic Reactivity of Coupled Binuclear Copper System(s) "Strongly Correlated" with Experiments

Teoretické studium enzymové reaktivity systémů se dvěma spřaženými ionty mědi "silně korelované" s experimentem

DOCTORAL THESIS

Supervisor:
Prof. Mgr. Lubomír Rulíšek, CSc. DSc.

Co-advisor: RNDr. Martin Srnec Ph.D.

Prague 2024

I declare that I carried out this dissertation thesis independently and that all sources of information are properly cited. Neither this thesis nor its substantial part was used to obtain any other academic degree.

In Prague signature of the author

Abstract

Copper ions are, along with iron and zinc, one of the most commonly occurring metal ions in biological systems. Their electronic structure - especially when it comes to (coupled) polynuclear Cu centers in metalloproteins - makes them one of the most challenging systems in contemporary bioinorganic (and theoretical) chemistry.

On the other hand, the presence of copper ions leads to unique spectroscopic properties. Several spectroscopic "fingerprints" exist which are used to characterize the individual, usually short-lived intermediates in the catalytic cycles of copper metalloenzymes. By correlating experimental and theoretical data, the reaction mechanism of the copper metalloenzymes can be revealed and ultimately fully understood at the atomic and electronic levels. This not only allows us to understand the physicochemical principles underlying fundamental biological processes, but also opens up possibilities to construct artificial (biomimetic) systems that carry the same or even better function as the original biological system.

The aim of this dissertation is to characterize and understand the reaction mechanism of coupled binuclear copper (CBC) (metallo)enzyme(s). The redox chemistry catalyzed by CBC enzymes often employs molecular oxygen as a cofactor. This enables them to activate and subsequently catalyze a cascade of chemical processes leading to substrate hydroxylation. In particular, the reaction mechanism of CBC enzyme tyrosinase (Ty) which catalyzes the hydroxylation of L-tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA) and its subsequent oxidation to (L-DOPA-quinone), was elucidated by means of quantum and molecular mechanical (QM/MM) calculations. The complementary - experimental - part of the project was carried out in the group of Prof. Edward Solomon (Stanford University, U. S. A.) by employing various spectroscopic techniques, such as low-temperature resonance Raman and electron paramagnetic resonance spectroscopy, kinetic measurements, and biochemical experiments. It can be mentioned that throughout the "strongly correlated" experimental and computational efforts on the "strongly correlated" CBC enzyme (Ty), a comprehensive view of the Ty catalytic action has been achieved. Throughout the work on the project, it was very encouraging to find out how theory assisted in planning and designing experiments whereas the experimental data guided us through the manifolds of plausible reaction pathways, described in the literature or suggested by our QM/MM calculations.

The computational part, which comprises the main part of the dissertation, includes state-of-the-art methods of computational chemistry, including multireference *ab initio* calculations and (in our opinion still non-trivial) QM/MM modeling. Successful completion of the PhD project opens new horizons in understanding the fundamental biophysics of bioinorganic systems and helps in the development of new nature-inspired catalytic systems.

Keywords: $DFT, QM/MM, [Cu_2O_2]$, CBC, tyrosinase

Abstrakt

Ionty mědi patří, spolu se železem a zinkem, k nejčastěji se vyskytujícím kovovým iontům v biologických systémech. Jejich elektronová struktura, zejména pokud jde o (spřažená) polynukleární centra v metaloproteinech, z nich činí jedny z nejnáročnějších systémů v současné bioanorganické (a teoretické) chemii.

Na druhou stranu přítomnost iontů mědi vede k jedinečným spektroskopickým vlastnostem. Existuje několik spektroskopických znaků („fingerprints“), které se používají k charakterizaci jednotlivých, obvykle krátce žijících meziproduktů v katalytických cyklech měďn(at)ých metaloenzymů. Pečlivou korelací experimentálních a teoretických dat lze popsat reakční mechanismus těchto metaloenzymů a nakonec jej plně pochopit na atomární či dokonce elektronové úrovni. Získané porozumění nejen umožňuje pochopit fyzikálně-chemické principy, které jsou podstatou základních biologických procesů, ale také otevírá možnosti pro návrhy umělých (biomimetických) systémů, které mají stejnou nebo dokonce lepší funkci jakou měl původní biologický systém.

Cílem disertační práce je charakterizovat a pochopit reakční mechanismus (metalo)enzymů se dvěma spřaženými ionty mědi (tzv. “CBC enzymes”). Oxidačně-redukční chemické reakce katalyzované CBC enzymy často využívající vyžadují molekulu kyslíku jako kofaktor. To jim umožňuje aktivovat a následně katalyzovat celou kaskádu elementárních chemických procesů vedoucích k hydroxylaci daného substrátu. Konkrétně byl pomocí kvantově a molekulově mechanických výpočtů (QM/MM) objasněn reakční mechanismus reprezentativního CBC enzymu, tyrosinázy (Ty), který katalyzuje hydroxylaci L-tyrosinu na L-3,4-dihydroxyfenylalanin (L-DOPA) a jeho následnou oxidaci na L-DOPAchinon. Komplementární – tj. experimentální – část projektu byla prováděna ve skupině Prof. Edwarda Solomona (Stanfordova Universita, Spojené státy americké) za použití rozličných pokročilých spektroskopických technik, jakými jsou například nízkoteplotní rezonanční Ramanova spektroskopie či elektronová paramagnetická rezonance, a dale též kinetických měření a biochemických experimentů. Lze zmínit, že v průběhu celého „silně korelovaného“ experimentálního a výpočetního úsilí na tzv. „silně korelovaném“ CBC enzymu (Ty) bylo dosaženo komplexního pohledu na katalytický mechanismus Ty. V průběhu práce na projektu jsme navíc došli k poměrně povzbudivému zjištění, jak moc může být výpočetní teorie nápomocna při plánování a navrhování experimentů, a naopak, jak nás experimentální data spolehlivě provedou skrz rozsáhlou množinou možných reakčních cest, ať už popsaných v literatuře nebo navržených pomocí našich QM/MM výpočtů.

Výpočetní část, která tvoří hlavní část této disertační práce, zahrnuje použití nejmodernějších metod výpočetní chemie, včetně multireferenčních *ab initio* výpočtů a (podle našeho názoru stále netriviálního) QM/MM modelování. Úspěšné dokončení doktorského projektu pak otevírá nové obzory v pochopení základní biofyzikální podstaty bioanorganických systémů či procesů a pomáhá při vývoji nových, přírodou inspirovaných, katalytických systémů.

Klíčová slova: $DFT, QM/MM, [Cu_2O_2]$, CBC, tyrosináza

The primary goal of computational chemistry, however, is not to provide numbers, but to provide understanding.
$\sim$ Frank Jensen, in his book

Acknowledgements

I extend my deepest appreciation to my PhD supervisor, Prof. Lubomír Rulíšek, for his invaluable guidance, unwavering support, constructive feedback, and continuous patience throughout every stage of this research journey.

I would like to offer my sincere thanks to Dr. Martin Srnec, my co-supervisor, and Prof. Edward I. Solomon for their mentorship. Thank you for your insightful comments and suggestions.
I am also truly grateful to the Rulíšek Group for their unwavering support and belief in me.

I would like to express my gratitude to the participants of this study, my collaborators from both experimental and computational fields, whose willingness to share their experiences added depth and richness to my research. Your contributions are truly appreciated.

List of abbreviations

AF2/3	AlphaFold 2/3
ANO	Atomic natural orbital
AOx	*o*-Aminophenol oxidase
Asn	Asparagine
B	Becke exchange functional
B3	Becke three-parameter exchange functional
B97	Becke (1997) functional
BS	Broken-symmetry
CaOx	Catechol oxidase
CAM	Coulomb-attenuating method
CAS	Complete active space
CASPT2	CAS second-order perturbation theory
CBC	Coupled binuclear copper
CC	Coupled cluster
CCSD(T)	Coupled cluster singles doubles perturbative triples
COSMO(-RS)	Conductor-like screening model (for realistic solvation)
CT	Charge-transfer
Cys	Cysteine
D	Dispersion correction
D3BJ	Dispersion correction with Becke and Johnson damping
DBED	*N,N'*-di-tert-butyl-ethylenediamine
DFT	Density functional theory
DKH	Douglas–Kroll–Hess
DKH2	Second-order Douglas-Kroll-Hess
DMRG	Density-matrix renormalization group
DNA	Deoxyribonucleic acid
DZ	Double-zeta
EAS	Electrophilic aromatic substitution
ECP	Effective core potential
EDG	Electron donating group
EPR	Electron paramagnetic resonance
ET	Electron transfer
EWG	Electron withdrawing group
EXAFS	Extended X-ray absorption fine structure
GGA	Generalized gradient approximation
Glu	Glutamate
Hc	Hemocyanin
HF	Hartree–Fock
His	Histidine
IBCO	Catechol oxidase from sweet potato
KIE	Kinetic isotope effect
KS	Kohn–Sham
L-DOPA	L-3,4-dihydroxyphenylalanine
LH	Local-hybrid

LUMO	Lowest unoccupied molecular orbital
LYP	Lee–Yang–Parr correlation functional
M06	Minnesota (2006) functional
MC-pDFT	Multiconfiguration pair-density functional theory
MD	Molecular dynamics
MM	Molecular mechanics
MR	Multireference methods
MRCI	Multireference configuration interaction
NR	Non-relativistic
NTO	Natural transition orbital
O	Bis-μ-oxo
ONIOM	Our own N-layered integrated molecular orbital and MM
P	$\mu - \eta^2 : \eta^2$-peroxide
P86	Perdew (1986) correlation functional
PBE	Perdew–Burke–Erzenhof functional
PCET	Proton-coupled electron transfer
PCM	Polarizable continuum model
PDB	Protein data bank
PES	Potential energy surface
QM/MM	Quantum mechanics/molecular mechanics
QM	Quantum mechanics
QZ	Quadruple-zeta
RFQ	Rapid-freeze quenching
RLS	Rate-limiting step
RNA	Ribonucleic acid
rR	Resonance Raman
SCF	Self-consistent field
SF-Abs	Stopped-flow absorption spectroscopy
SHE	Standard hydrogen electrode
SKIE	Solvent kinetic isotope effect
T1Cu	Type I copper centers
T2Cu	Type II copper centers
T3Cu	Type III copper centers
TD	Time-dependent
TPSS	Tao–Perdew–Staroverov–Scuseria functional
TPSSh	Hybrid functional (10 % HF) using the TPSS functional
TS	Transition state
Ty	Tyrosinase
Tyr	Tyrosine
TZ	Triple-zeta
U	Unrestricted
UV-Vis	Ultraviolet–Visible
W	Water
WF(T)	Wave function (theory)
X2C	Exact two-component
XAS	X-ray absorption spectroscopy
ZORA	Zeroth-order regular approximation

Contents

1. Introduction

Arieh Warshel in his Nobel lecture from 8^{th} of December 2013 stated:

The ability to model complex molecular systems is crucial for advances in understanding biological systems and in rational progress in molecular medicine, as well as in the rational design of new materials and devices. However, progress in this direction was hindered by the fact that rigorous modeling of complex systems requires enormous computational power. [...] it is possible to decompose the system to parts where the quantum mechanical description is essential (e.g. parts where bonds are being broken), and other parts that can be represented on a simpler level with empirical force fields. This idea [...] led to the development of the combined quantum mechanics/molecular mechanics (QM/MM) model. Here, the coupling between the electrostatic effects of the quantum and classical subsystems has eventually become a key to advances in describing the function of enzymes and other biological molecules.[1]

Enzymes are biomolecules essential to sustaining life because they are involved in almost all metabolic processes that occur in the cell. They perform numerous functions in living organisms, ranging from signal transduction through cell regulation and DNA/RNA/protein biosynthesis to the creation of cooperative metabolic pathways in which several enzymes can work together. They are known to be involved in the catalysis of more than 5,000 types of biochemical reactions.

Enzymes also attract attention due to their possible commercial use, for instance in industrial processes or in synthesis of medicines. Until late 1990's, studying the enzyme's reaction mechanisms using theoretical methods was very limited. It was almost impossible *not* to compromise between the accuracy of a computational treatment and the size and complexity of the (full) enzymatic systems. Typically, the models were limited to the active center and possibly a few additional second-sphere residues.

The development of the QM/MM methodology, which combines, at least in principle, the accuracy of the quantum mechanics with the simplicity and low cost of the molecular mechanics, helped to overcome restrictions related to enzymatic model size. It allowed researchers to study whole proteins, that is to include the outer environment of the active site (protein framework). It also enabled them to investigate and address, for instance, the influence of the rest of the enzyme on catalytic process occurring in the active site. From multiple experimental and theoretical studies, it has been understood that enzymes with seemingly similar active sites may catalyze different reactions.

In this thesis I will show how I have used this powerful computational tool (QM/MM method) in studying enzymatic reaction mechanism of the coupled binuclear copper (CBC) metalloenzyme tyrosinase (Ty).

1.1 Enzymes

Enzymes are gentle, yet so strong. They are one of the most efficient catalysts in fundamental (bio)chemical processes, usually operating under mild ambient conditions. Most of them work efficiently at room temperature and a pH close to neutral.[2] They can accelerate reaction thousand, even million folds.[3] Enzymes can thus change 'impossible' into possible in terms of (bio)chemical transformations. The increase in the reaction rate is achieved by lowering the activation energy of the process; however, enzymes do not change the overall reaction free energy.[4]

Enzymes comprise the protein (*apoenzyme*) and non-protein (*cofactor*) part, which together form catalytically active species called a *holoenzyme* (**Figure 1**).[5] Separately, none of them exhibits catalytic activity. Cofactors can be further divided into two subgroups: (1) coenzymes, which are typically small organic molecules, and (2) metals.[6]

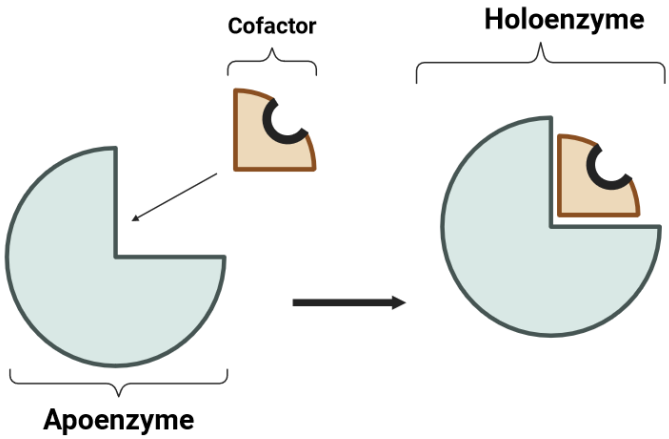


Figure 1: Structure of enzyme. Apoenzyme and non-protein part form together catalytically active form - holoenzyme.

There are also non-protein enzymes, such as ribozymes (**ribo**nucleic acid (RNA) en**zyme**s),[7] and deoxyribozymes (DNAzymes).[8] Those RNA and DNA molecules also exhibit catalytic activity towards some specific reactions. Nevertheless, such catalysts (made of genetic material) constitute a much smaller portion of enzymes.

Enzymes are highly specific. The following can be distinguished: i) substrate specificity, when the enzyme is able to catalyze only a transformation of a specific substrate, ii) spatial specificity, which involves the substrate matching with the appropriate configuration to the spatial arrangement in the active center, iii) chemoselectivity, i.e. the ability to catalyze the transformation of one specific functional group even in the presence of other similar groups, iv) regioselectivity, i.e. the formation of an excess of one of the structural isomers, and v) stereoselectivity, i.e. the ability to catalyze reactions containing or creating only one of the possible enantiomers on a chrial (mostly carbon) atom.[2,5]

The high specificity is ensured by the structure of the active center. The place where the substrate binds and the reaction occurs occupies a small part of the protein's interior. This spatial arrangement usually has the shape of a pocket or a slit. Less often, the catalytic site may be a shallow cleft on the surface of the protein. Active site amino acids are divided into several groups: i) those that take

part in substrate binding and are directly involved in catalysis, ii) those having a supporting role, acting, for example, as proton or electron donors/shuttles, iii) those stabilizing the spatial structure of the active center, for example through hydrogen, electrostatic or stacking interactions.[2,4,5]

Some proteins contain metal ions. Most metals bind by forming a coordinate (sometimes denoted coordinate covalent) bond with the N- , O- , and S-containing side chains of amino acids such as: histidine, tyrosine, cysteine, methionine, glutamic acid, aspartic acid.[2,6]

1.2 Metalloenzymes

Enzymes which require metal ions to perform their catalytic action are called metalloenzymes. Roughly half of all proteins accommodate metal ion(s).[6] Ordering metal ions by decreasing content in organisms is as follows: iron, zinc, copper, molybdenum, cobalt, chromium, vanadium, and nickel.[9] The role of metals is to stabilize the spatial conformation of the protein (structural role) and/or directly participate in enzymatic catalysis (functional role).[2]

Metalloenzymes play an extremely important role in nature. Due to the presence of metal ion(s) inside, they possess the ability to catalyze reactions which are difficult to achieve in organic chemistry. For instance, they activate molecular oxygen, perform redox reactions or catalyze (formally) spin-forbidden reactions.[6]

One may also distinguish between metalloenzymes and metalloproteins. The latter are proteins which contain metal ion(s) but they are not involved in chemical reactions. They serve, for instance as oxygen carriers, play a role in metal-ion storage, and in electron-transfer (ET) processes.[4,6]

In this thesis I have been focusing on systems with copper ions inside their active site.

1.3 Copper-containing proteins

There are several types of copper-containing proteins categorized according to their spectroscopic characteristics:[10] type I copper centers (T1Cu), type II copper centers (T2Cu), type III copper centers (T3Cu), trinuclear (T2+T3) copper centers in multi-copper oxidases, binuclear copper A centers, copper B centers, and a tetranuclear Copper Z center.[11]

1.4 CBC protein family

Specifically, the focus of this thesis is on the binuclear T3Cu centers. Enzymes with such centers are also called proteins with the coupled binuclear copper center. CBC protein family consists of few members sharing a common motif in their active site. It always comprises a pair of copper ions, each of them coordinated by three histidine (His) residues located on antiparallel α-helices (**Figure 2**).[12]

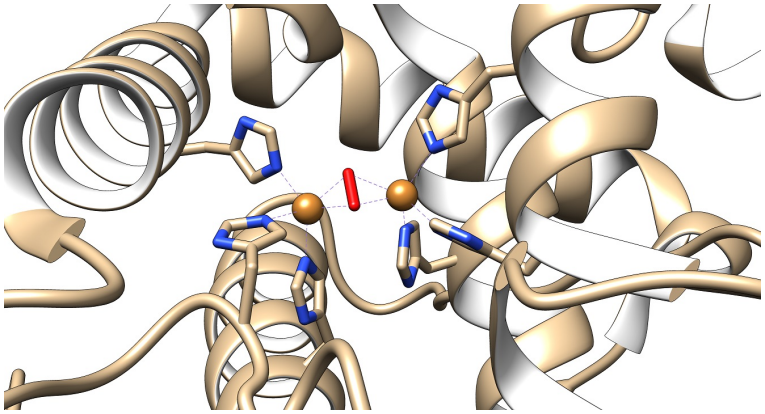


Figure 2: Active center of oxy-tyrosinase (Protein Data Bank (PDB) ID: 1WX2[13]) as an example of the CBC site. Copper ions are shown as brown balls, histidine residues as sticks, dioxygen molecule as red stick, the rest of protein as a cartoon. Visualization was made using the UCSF Chimera program.[14]

Among members of CBC protein family are both catalytically active and catalytically inactive representatives (**Figure 3**). The latter includes hemocyanin (Hc),[15] which is a dioxygen carrier in arthropods and mollusks. The former include catechol oxidase (CaOx) found mostly in plants responsible for oxidation of various catechols to quinones,[16–19] and tyrosinase, which performs monooxygenation of phenols and also oxidation of catechols.[13,20–26] The most recently discovered member of this family, *o*-aminophenol oxidase (AOx) exhibits the same reactivity as tyrosinase, but in addition, it can convert *o*-aminophenols to corresponding *o*-nitrosophenol products.[27,28]

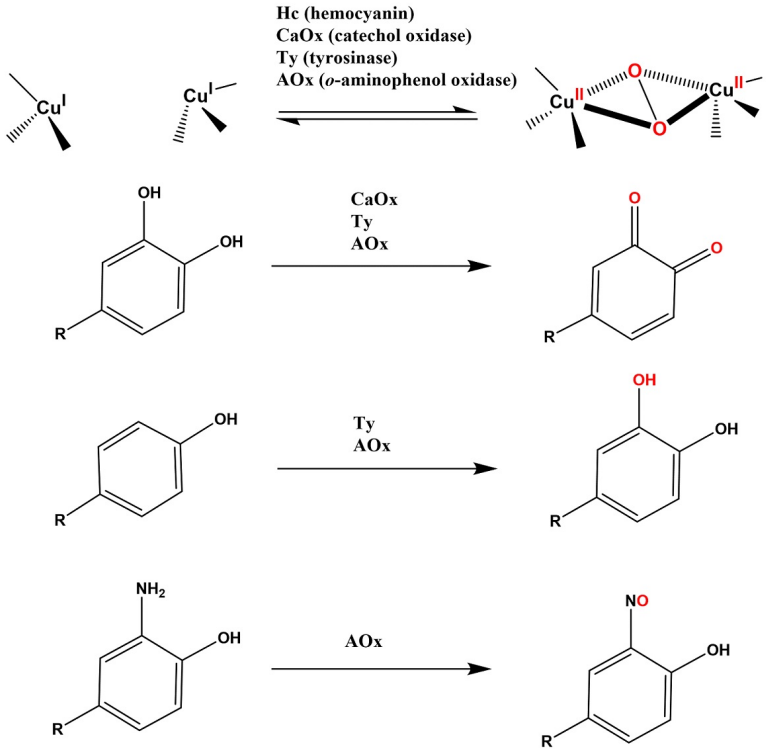


Figure 3: Different reactivities among the CBC protein family members.

All three groups of catalytically active CBC metalloenzymes mentioned above have buried active site with the gatekeeper (usually aromatic amino acid) at the entrance, which must be removed prior to reaction.[29]

Each CBC enzyme can have several forms (**Figure 4**) based on the specific state of the CBC site: *oxy*, *met*, *deoxy*, and also putative *hydroperoxide*.[30] The deoxy form is the fully reduced state that contains Cu(I) ions separated by the distance of approximately 4.5 Å. Deoxy-CBC can reversibly bind O_2 molecule and subsequently generate the oxy state. In oxy state, two isomers may exist, depending on the state of the (originally) O_2 molecule: $\mu - \eta^2 : \eta^2$ (side-on peroxo bridged, seen in most of oxy-CBC), where the two Cu(II)'s are separated by $\sim$ 3.6 Å and the so-called bis-μ-oxo state/isomer (found in some catechol oxidases[30]) with distance between copper(III) ions $\sim$ 3.0 Å. The resting form of a CBC enzyme (met form) then comprises the μ-hydroxo dicopper(II) site (Cu-Cu distance $\sim$ 3.5 Å).

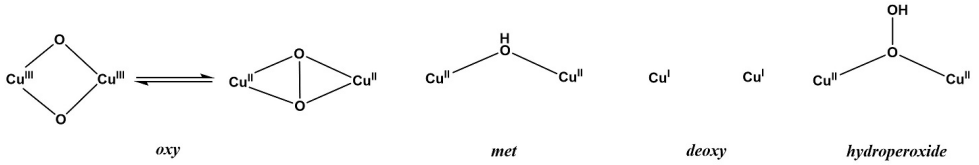


Figure 4: Different forms of CBC active sites.

By sharing the same motif in active site, they exhibit similar spectroscopic properties. Oxy-state is electron paramagnetic resonance (EPR) silent due to antiferromagnetic coupling between two Cu(II) ions. It features two characteristic bands in the Ultraviolet–Visible (UV-Vis) absorption spectrum at $\sim$ 345 nm (ϵ = 19 000 cm^{-1} M^{-1}) and at $\sim$ 600 nm (ϵ = 1 000 cm^{-1} M^{-1}) - corresponding to peroxo O_2^{2-}→Cu(II) charge-transfer (CT) transitions.[21] Characteristic O-O stretching frequency is observed at $\sim$ 740 cm^{-1}, and the Cu-Cu stretch at 280 cm^{-1}.[29,31,32]

1.4.1 Tyrosinase

Tyrosine is a well-known and important representative of the CBC protein family.[13,21] This ubiquitous oxidase is involved in two distinct reactions: (1) conversion of L-tyrosine to L-3,4-dihydroxyphenylalanine (L-DOPA) and (2) subsequent oxidation of L-DOPA to L-DOPAquinone (**Figure 5**). These elementary reactions are the initial and rate-limiting steps (RLS) in melanogenesis (synthesis of melanins).

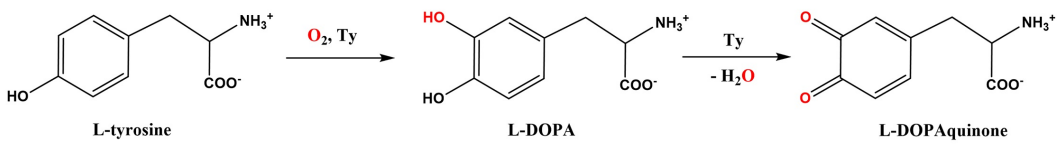


Figure 5: Reactions catalyzed by tyrosinase. Hydroxylation of L-tyrosine to L-DOPA and subsequent oxidation to L-DOPAquinone.

Melanins are omnipresent photoprotective biopolymer pigments with major implications ranging from fruit browning to severe human diseases, such as skin

cancer. Tyrosinase is also involved in Parkinson's disease, because of its ability to oxidize dopamine to form melanin in the brain.[33] Due to the key catalytic role of tyrosinase in melanogenesis, it is an important molecular and therapeutic target.

Until now, the reaction mechanism of tyrosinase has been very poorly understood despite extensive efforts. The three key issues in previously proposed mechanisms were: i) the character of the reactive species (i.e., whether it has a bis-μ-oxo (**O**) or side-on-peroxo (**P**) moiety in the active site), ii) the binding of substrate by either CuA or CuB and also iii) the character of the reaction itself. Some researchers postulated that hydroxylation is best described as electrophilic aromatic substitution (EAS),[20,34–37] whereas others favored a radical mechanism (radical coupling) involving phenoxyl radical(s).[16,38–41]

Part of the presented thesis provides new insight into elusive catalytic intermediate as well as sheds light on the enzymatic process carried out by this CBC protein.

1.4.2 Previous research on CBC systems with particular emphasis on Ty

The CBC proteins have been subject of scientific interest for more than 50 years. Initially, they were mostly studied experimentally, but with the developments in computational methodology and accumulated expertise in theoretical modeling of enzymatic reactions, experiments were complemented by theoretical studies.

In 1999, Siegbahn and co-workers reported the first quantum chemical study of tyrosinase reaction mechanism.[42] They used the so-called QM-cluster model approach, employing a very small model of the active site – ammonias or formimines (instead of imidazoles) as ligands of the $[Cu_2O_2]$ moiety. Owing to the absence of the X-ray crystal structure of Ty (at the time of their study), they utilized the crystal structure of hemocyanin to obtain the cluster model (QM system). Their computational results supported radical character of the reaction mechanism. They postulated that the O-O bond cleavage in the peroxide moiety (**P**) represents the rate-limiting step. The study, even if done on different protein crystal structure and on the small model, provided a lot of initial insight and inspiration into current computational research of CBC proteins.

In the following study, Siegbahn and Wirstam examined whether the bis-μ-oxo isomer (**O**) is a plausible intermediate in the Ty catalytic cycle (again, using Hc crystal structure as the model template).[43] They used larger cluster model (with imidazole rings as copper ions ligands) and found O-O bond cleavage highly endothermic (23.3 kcal/mol). This led to the conclusion that the Cu_2(III,III) bis-μ-oxo state (**O**) is not a likely intermediate of Ty.

Two years later (i.e., in 2003), this finding led Siegbahn to revisit Ty reaction mechanism. He concluded that the first step in the mechanism is the peroxide attack on the phenolate ring, and only then the O_2 cleavage occurs.[44] His model was based on the X-ray structure of catechol oxidase.[45] He studied the whole catalytic Ty cycle, but mostly concentrated on the initial step of the reaction that involves formation and breaking of covalent bonds. Among his conclusions two important observations can be highlighted: i) the dissociation of one histidine

upon substrate binding to Cu, and ii) finding a product that was lower in energy (with respect to the previous study). The product was best described as the Cu(II)Cu(I)-semiquinone radical. However, this structure was in disagreement with experimental data (with an energy which was too low to be consistent with the experimentally observed catalytic turnover).

A year later, in 2004, Siegbahn published a computational study dealing with the full catalytic cycle of catechol oxidase.[46] He based his research on two crystal structures of CaOx (PDB ID: 1BT3 and 1BUG), differing mostly by Cu-Cu distance (2.9 Å *vs.* 4.2 Å). He studied the oxidation of the catechol, release of the quinone and binding of another catechol to form the second quinone product. He argued in favor of the proton-coupled electron transfer (PCET) mechanism and radical character of substrate during the reaction. The highest barrier of 15.2 kcal/mol, related to O-O cleavage, was found to be the RLS. It was in good agreement with experimental data ($\sim$ 13 kcal/mol).

In 2007, another QM study was carried out in the Siegbahn group, focusing on the reaction mechanism of catechol oxidase.[47] The model was constructed employing the available crystal structure of catechol oxidase from sweet potato (IBCO). It contained $[Cu_2O_2]$ core, six imidazole ligands as models of His residues and SCH_3 moiety mimicking the cysteine (Cys) side chain which is covalently (C-S bond) linked to one of the Cu-coordinating His residues (in the CaOx crystal structures). In the first step, bridging peroxide abstracted H-atom from substrate, resulting in radical formation. The next step (O-O bond cleavage) was found as the RLS with activation energy of 12.1 kcal/mol. It was in a reasonable agreement with the experimental one (13 kcal/mol). In the last step, the second H-atom was transferred from the substrate to the $[Cu_2OOH]$ core to form the met state $[Cu_2(OH)_2]$ and resulted in quinone formation (on the substrate side). Again, as in the case of tyrosinase, O-O bond cleavage of peroxide was found as a rate-limiting step of the reaction.

Another impactful study was published by Inoue et al. in 2008.[39] They performed QM/MM calculations to examine conversion of L-tyrosine to L-DOPAquinone carried out by Ty. They modeled full-protein system, which was based on an X-ray structure of oxy-Ty published in 2006,[13] employing the QM/MM approach. However, they used a quite limited model of the QM region. It comprised $[Cu_2O_2]$ core, six histidines and the native substrate. There were no additional residues, hence no additional proton acceptor site was included in the QM part of their QM/MM model. This observation is worth noting because they postulated external proton acceptor of hydrogen/proton from the *ortho*-hydroxyl group of the phenol ring. Due to the restricted model size, they had to use one of the coordinated His residues as the proton acceptor site. This led to chemically unlikely intermediate with CuA coordinated by only one histidine. They also proposed radical character of the reaction (with stable phenoxy radical), and similarly to Siegbahn they found O-O cleavage as RLS. Another computational prediction made by the authors was that oxidation step is much faster than the (first) hydroxylation. Despite these shortcomings, the proposed reaction mechanism provided a lot of mechanistic insight into the CBC research.

A slightly different approach in studying Ty was adopted by Deeth and Diedrich in 2010.[48] They performed molecular dynamics (MD) simulation of oxy-Ty without a caddie protein which is its chaperone responsible for loading

the copper ions into the active site of Ty and also blocking the entrance to its active site. Their goal was to explore the dynamics of the type III copper center. Results of this study proved the rigidity of overall protein structure as well as of the active site with the $[Cu_2O_2]$ moiety. They found that the existence of butterfly angle of the copper-oxygen center is caused by hydrogen bond of tyrosine (Tyr98) from caddie protein. Without the presence of chaperone, the $[Cu_2O_2]$ core is flat.

In 2011, Siegbahn and Borowski once again revisited the Ty reaction mechanism, mostly addressing a few issues in the above study of Inoue et al.[39] They employed both QM and QM/MM methods.[49] One of their findings was that the proton has likely to be removed from the enzyme before catalysis. However, they used small model in their calculations - $[Cu_2O_2]$ and imidazoles as ligands – thus it posed additional difficulties in proving aforementioned statement about proton removal. They also suggested that Inoue et al.[39] QM/MM structure geometry converged to an incorrect reactant state, which was the reason of discrepancies between the QM and QM/MM studies.

Finally, in 2020, Jiang and Lai published their QM/MM study on the monophenolase (cresolase) and catecholase activity of catechol oxidase, with the catechol as a model substrate.[16] Catechol oxidases cannot hydroxylate tyrosine, but they do hydroxylate smaller phenolic substrates. In their proposed reaction mechanism, hydroperoxide species was considered as the reactive moiety. The PCET mechanism was postulated for both studied pathways. Moreover, in both reactions, the O-O bond cleavage was postulated as the rate limiting step. The energy barrier associated with the *o*-quinone formation has been computed at 13.2 kcal/mol, much lower than the barrier for the formation of hydroxylated product (20.0 kcal/mol). Therefore, the catecholase activity of this enzyme should be much higher than cresolase activity (by approximately five orders of magnitude).

Although to a certain extent different from the main goals of the study, it can be also mentioned that the whole field of small molecule CBC catalysts evolved in the last two or three decades. These biomimetic compounds can be used in industrial processes, including oxidation reactions[37] as well as defluorination.[35] In addition, these systems are considered as excellent model systems for benchmarking the computational methods. More details on the small CBC systems will be discussed in Chapter 5, section 5.1.

2. Computational modeling of bioinorganic systems

2.1 Modeling enzyme reactivity

Modeling enzyme reactions presents unique challenges compared to modeling smaller molecules. In addition to selecting the appropriate 'model chemistry' (a combination of electronic structure method, basis set, and solvation model), the crucial component in protein modeling is the initial set-up of the system. Ideally, a high-quality experimental 3D structure (e.g. X-ray or cryo-EM) is available, preferably with a substrate or its analogue bound to the active site. However, even when starting with a known structure (protein, metalloenzyme, protein-ligand complex), preparing the final 'computational model' can be a complex process, with each step potentially introducing challenges and ambiguities. The process typically involves defining the exact position of the substrate (or any other cofactor), assigning protonation states to all protic residues (especially those near the active site and the substrate itself, if necessary), adjusting the protein solvation sphere or solvation box, assigning oxidation and spin states to complex redox active enzymes (considering that crystal structures may represent a mixture of redox states), adding counterions to neutralize the system, addressing missing parts in the X-ray structure, and resolving Cys-Cys links. While most of these tasks are adequately handled by crystallographic and/or molecular-modeling software, care and experience are needed to resolve potential ambiguities and to validate the initial structure. This is especially true if the initial structure is unavailable and is built using a protein-structure predictor, such as AlphaFold 2/3 (AF2/3)[50,51] and/or if ligands or substrates are computationally docked.

A typical workflow for preparing an initial structure involves several rounds of equilibration (**Figure 6**). Our group, being inspired by the work of Ryde and coworkers,[52] has converged to the following protocol: (1) minimize, at the MM level, the positions of all hydrogen atoms *added* to the initial crystal structure corresponding to the standard protonation states of amino acid side chains (typically at pH = 7) and any added non-hydrogen atoms, such as missing links or cofactors. These added atoms are referred to as the "*non-fixed* list". The atoms resolved in the original X-ray structure remain fixed throughout all equilibration steps; (2) run a short (∼1 ns) simulated annealing molecular dynamics on the previously minimized structure, maintaining the same *fixed/non-fixed* sets of atoms at elevated temperature; (3) perform the final MM minimization using the same *fixed/non-fixed* atom sets. Next, a solvation sphere is added with a radius exceeding the protein boundary by at least 2-3 Å, and Na^+ or Cl^- ions are introduced to neutralize the overall charge of the protein. These ions are generally located in the (added) solvent sphere. The equilibration procedure is then repeated: (4) minimize the positions of all atoms, except those in the original *fixed* list (i.e. X-ray resolved); (5) run a 1 ns simulated annealing molecular dynamics simulation and (6) perform the final minimization of the whole system, using the same set of *fixed/free* atoms as in (4) and (5). During step (5),

a centric force is applied to the water molecules in the solvation sphere to prevent their dissociation from the system during higher-temperature (up to 400 K) MD annealing. The final structure is considered to be a 'robust' computational model of the (metallo)enzyme and can be used for further theoretical investigations of the reactivity.

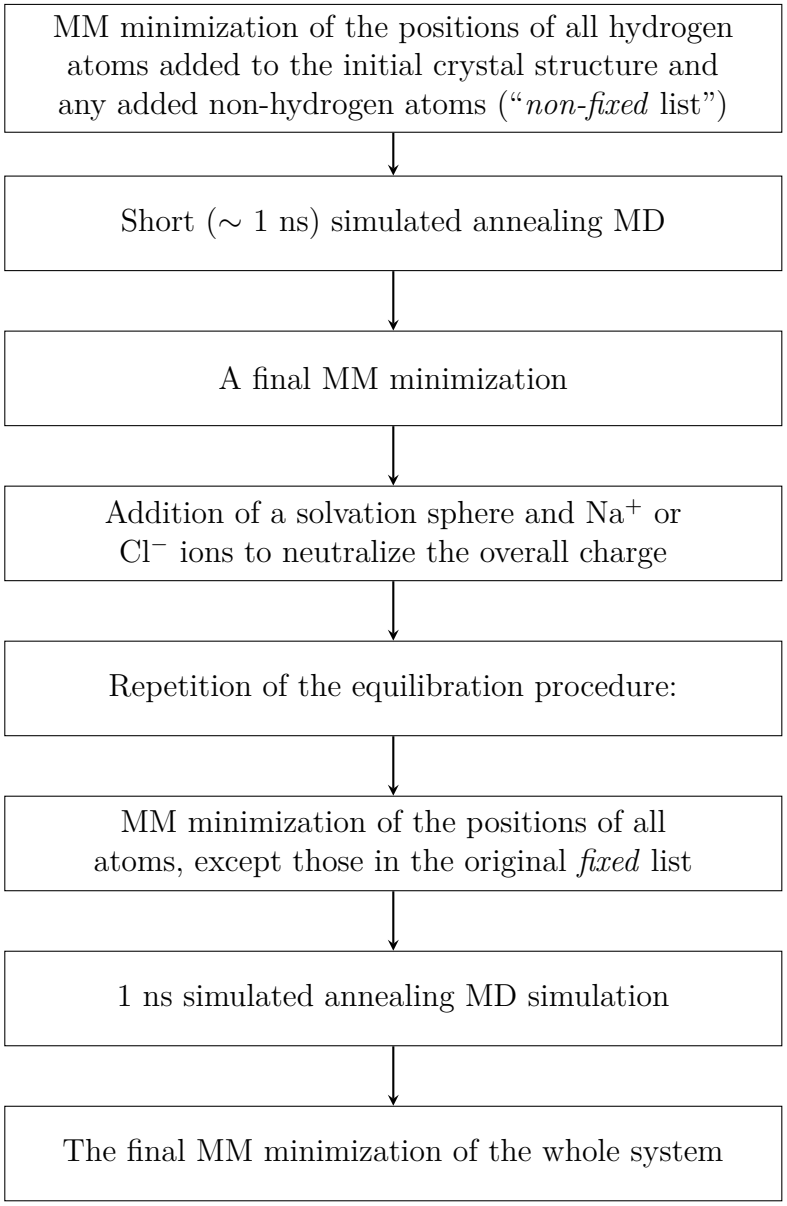


Figure 6: A typical workflow of an initial structure preparation for further QM/MM modeling.

There are two most common (and to some extent complementary) approaches to the theoretical study of enzyme reactivity, each with its own advantages and limitations. The conceptually simpler approach is the QM-cluster method (or model)[53,54] in which the active center (cluster) is represented by a large model (nowadays, ∼150-300 atoms is recommended) of catalytically important residues. The cluster model can be prepared directly from the X-ray structure or from the equilibrated protein model (as presented above). The entire system is described at the QM level, while the rest of the protein is approximated by a polarized continuum ($\epsilon_r = 4$ is often used to represent a 'generic protein'

environment). To avoid artificial and nonrealistic conformational changes, some atoms, usually backbone Cα's of truncated residues in the protein, are kept fixed at their initial positions. A good indicator of the reliability of the cluster-model calculations is the convergence of computed properties with respect to the size of the QM cluster. Unlike QM/MM calculations (discussed below), the cluster-model approach is conceptually straightforward, however, it may suffer from a certain rigidity due to the fixed positions of the outer atoms. This may lead to insufficient flexibility in the active site throughout the enzymatic cycle ('breathing of the active site'). These artifacts sometimes manifest in reaction coordinates with multiple minima and transition states that might be absent in QM/MM calculations (*cf.* QM cluster model study of peptide bond hydrolysis by amino acid peptidase published by Himo and coworkers[55] and QM/MM study of the same reaction catalyzed by the glutamate carboxypeptidase II published by Klusak et al.[56]). Additionally, the absence of long-range electrostatic interactions from the surrounding protein may result in an incomplete description of certain effects in enzyme catalysis, such as the internal electric fields in enzymes.[57] Despite these limitations, Siegbahn recently reviewed the performance of the cluster-model approach for a wide range redox-active metalloenzymes, concluding that it generally performs well when employing density functional theory (DFT) with B3LYP functional as the QM method.[54] However, in our opinion, the cluster model approach is less suitable for processes involving larger conformational changes in the active site or where second sphere interactions of the protein pocket with the substrate influence its interaction with the metal active site. In such cases, QM/MM should be used instead.

The flexibility along the reaction coordinate is inherent to QM/MM calculations. In this approach, the system is divided into the QM region (active site, binding site, reaction center) and the MM region (rest of the protein). Various coupling schemes are then available. In the most common approach, link-atoms are introduced. Link-atom has a dual character and represents original MM atom (typically carbon) in the QM calculation (where is presented as a hydrogen). Alternatively, one can use localized-orbital schemes schemes – placing hybrid orbitals on the boundary, some of which are frozen.[58] We further partition the MM part into two subsystems: $\sim$ 2-6 Å vicinity of the QM system (2-3 Å for the larger QM systems, $\sim$ 6 Å for the smaller ones), defined on a per-residue basis. This "QM neighborhood" (denoted as System 2) typically consists of 600-1000 atoms and is relaxed during QM/MM calculations. The outer parts of the protein (denoted as System 3) are kept fixed. This partitioning (i.e. fixation of the outer part of the protein) prevents divergence from the experimental crystal structure. Fixing the atoms of System 3 also saves computational time as the non-bonding interaction terms within System 3 are constant and do not need to be evaluated. QM/MM modeling can be broadly categorized into two main approaches: *electronic embedding*, where the electrostatics of the MM part are included in the QM/MM calculations (e.g. as the MM point charges in the QM Hamiltonian) and *mechanical embedding*, where they are not. We prefer the former approach, acknowledging that it may lead to larger fluctuations and instabilities in the QM/MM energies due to the long-range character of the electrostatic interactions. Consequently, the computed QM/MM (*el. embedding*) energy values must be interpreted with extreme care.

Ideally, both approaches (QM cluster model and QM/MM) should converge to the same values as the size of quantum system increases. Recently, our group and other researchers have advocated for a combination of the two approaches: employing QM/MM simulations to obtain equilibrium geometries (reactants, products, intermediates) and transition states, and then recalculating single-point QM/PCM(COSMO) energies for the quantum system (without any additional geometry optimization) to obtain the final energy values.[59,60] The combination of the QM/MM modeling with the cluster model approach (combined QM//QM/MM approach) seemed to be superior to the QM-only (cluster model) and QM/MM methods. Geometry optimization at the QM/MM level with inclusion of the whole protein in the model let to maintain its 'real' conformation, especially in case of second-sphere residues (which often are cut and/or partially fix in place in QM-only approach). Subsequent single-point energy calculations at higher level of theory (usually larger basis set) in QM-only method lead to more accurate results/energetics (when correlated with experimental data).

As in any thriving field in molecular sciences, there are still challenges in computer modeling of enzymes. An open issue is the efficient implementation of second derivatives into the QM/MM procedure to obtain *entropic* contributions to the transition-state barriers and reaction energies. Some QM/MM codes, such as the ONIOM method implemented in Gaussian, offer practical solutions to calculate these contributions. However, it is unclear whether more robust and numerically stable results are obtained by considering the full QM/MM system or by using a cluster model approach. In the cluster model approach, imaginary frequencies arising from fixed atoms are eliminated by assigning them infinite masses. Thermochemistry (including part of entropic contributions) is then calculated according to the same recipes as in (standard) modeling of chemical reactivity. Still, further research and benchmarking are needed to establish best practices for accurately computing entropic contributions in QM/MM systems, ensuring reliable predictions of reaction and activation free energies. In our group, we have converged to the following methodology for estimating entropic contributions: we first perform QM/MM optimization of the entire system, then subject the QM/MM-optimized QM region to cluster model re-optimization (with fixed link atoms) and subsequent vibrational frequency calculations (with infinite masses of the 'link' atoms). The latter is then used to compute the entropic and enthalpic contribution of the system of interest using standard thermochemical formulas (i.e. *via* translational, rotational and vibrational partition functions). We have to acknowledge that there are still some important components related to the protein conformational flexibility, leading to multitude of close-in-energy states, missing.

2.2 Quantum chemical methods for strongly electronically correlated systems: the $[Cu_2O_2]$ core

(Bio)molecules featuring the $[Cu_2O_2]$ core pose significant challenges for electronic structure methods. Accurate treatment requires careful consideration of the static and dynamic correlation, basis set convergence, relativistic effects, and possibly also solvent interactions. These contributions can have effects of the same or opposite sign on calculated properties, making it difficult to achieve reliable results and not relying on fortuitous error cancellation. Rigorous validation against experimental data is essential to ensure the robustness of computational approaches for these systems.

The complex electronic structure of the $[Cu_2O_2]$ core has sparked a renewed interest in the application of *ab initio* wave function theory (WFT) methods.[61,62] The equilibrium between the $\mu - \eta^2 : \eta^2$-peroxide dicopper(II) (formally $[Cu_2{}^{II}(O_2)^{2-}]^{2+}$ oxidation state, **P**) and its bis-μ-oxo dicopper(III) (formally $[Cu_2{}^{III}(O^{2-})_2]^{2+}$, **O**) isomer has become a benchmark system in computational bioinorganic chemistry. In the late 2000's, it was among the first examples used to test the density-matrix renormalization group (DMRG) approximations to the multireference (MR) methods (refs.[63,64]). MR methods, such as complete active space self-consistent field (CASSCF), CAS second-order perturbation theory (CASPT2), and multireference configuration interaction (MRCI),[65] due to their ability to recover correlation energy, seemed to be obvious choices for studying these highly correlated systems. Nevertheless, due to their computational cost, their use is limited to a very small model comprising the $[Cu_2O_2]$ core and a few additional first-sphere coordinating atoms.[61,63,66–69] An essential aspect of the MR methods is the choice of the active space, which consists of a subset of electrons and orbitals. In MRCI methods, this selection involves choosing the reference electronic configurations. The selection of the active space is crucial to obtain accurate results, even qualitatively. However, including all valence electrons into the active space is computationally prohibitive, even for smaller systems. The number of active space orbitals and electrons considered in calculations introduces another variable in the complex computational protocol, making the *ab initio* method of choice become very expensive for calculations. To obtain sufficiently converged results with respect to the size of the active space, at least a (24-in-28) active space is recommended for the **P**/**O** equilibrium,[63] or the combination of a smaller space with the Davidson correction.[70]

Recently, there has been a growing interest in the application of the coupled cluster (CC) methods in bioinorganic chemistry.[71] This method, in principle, allows for inclusion of *infinite-order correlation corrections*.[72] While standard CC formulations, such as CCSD(T), are primarily designed for single-reference systems and may not fully capture static correlation, they can still provide accurate results for many multireference (or multi-configurational) systems. In computational chemistry terms, the correct treatment of dynamical correlation, which is well-described by CC methods, might be more important than capturing static correlation, which is better described by MR-SCF methods. Therefore, the number of bioinorganic systems that can be studied using CC calculations is

larger than previously believed.[68] An interesting variant of the CC calculations, supported by many researchers,[73–75] is the use of the spin unrestricted coupled cluster method based on Kohn-Sham orbitals (UCC/UKS) instead of canonical Hartree-Fock (HF) orbitals. Many implementations of advanced post-SCF methods allow for efficient inclusion of relativistic effects, which is very important for heavy elements. As shown in the literature, even for polynuclear sites comprising 3d transition metal ions, (mostly scalar) relativistic effects are non-negligible.

In recent theoretical work on CBC biomimetic systems, scientists evaluated various relativistic approximations, namely effective core potentials (ECPs), Douglas–Kroll–Hess (DKH), and Zeroth-Order Regular Approximation (ZORA).[61,68,76] Additionally, we examined two-component schemes, such as X2C,[77] by employing the X2C-UCCSD(T)/UKS approach to investigate the $\mathbf{P} \leftrightarrow \mathbf{O}$ conversion in the model $[Cu_2O_2(NH_3)_6]^{2+}$ system (**Paper I**, *vide infra*).[78] In many cases, copper systems are well-described by non-relativistic quantum chemistry, and the intricacies in QM treatment are mostly related to electronic structure rather than relativity. However, two factors play a role in CBC's, especially in $\mathbf{P/O}$ equilibrium: a change in the formal oxidation state of the copper ion and the presence of two copper ions, which approximately doubles the (scalar) relativistic effects. This leads to a contribution of up to 8 kcal/mol.[61,78] Witte et al., by examining relativistic effects on the $[Cu_2O_2]$ $\mathbf{P/O}$ equilibrium, showed that the relativistic effects stabilize Cu-O and Cu-N bonds to a greater extent for the **O** isomer, thus shifting the $\mathbf{P/O}$ equilibrium towards **O**.[76] Numerous studies have demonstrated that scalar relativistic effects must be carefully considered when examining CBC, as sometimes treating relativity with ECPs may not be sufficient.[61,63,68]

While WFT methods have made significant contributions to bioinorganic chemistry, it must be acknowledged that DFT methods have become the workhorse of the field due to their efficiency and low computational cost. Despite being considered single-reference methods, DFT approaches often yield satisfactory results for strongly correlated systems. However, one of the principal drawbacks is that DFT methods do not produce pure spin states, as unrestricted Kohn-Sham determinants are not eigenfunctions of the $\hat{S}^2$ operator. This issue is particularly pronounced for lower spin states, which mostly correspond to the antiferromagnetically coupled metal centers in polynuclear metal systems. Spin-coupling mechanisms, originally formulated by Noodleman and coworkers[79,80] in the field of bioinorganic chemistry, have recently been given a rigorous quantum chemical footing by Chalupský and Srnec.[81] Currently, the common practice in treating various oxidation and spin states of open-shell systems involves broken-symmetry (BS) calculations employing Kohn-Sham determinants, acknowledging that they are eigenfunctions of $\hat{M}_S$, rather than $\hat{S}^2$ operator.[82–85] Different projection schemes may be used to estimate the energies of the 'pure' spin states.[79,86–88]

The choice of a functional is a crucial aspect in the approximate Kohn-Sham DFT method. While B3LYP[89–94] (or B3LYP*) remains the prevailing 'model chemistry' in bioinorganic chemistry, as advocated by Siegbahn and others,[53,54,95] there are competitive alternatives. It has been observed that among many 'variables' in an *approximate* DFT functional, the amount of exact exchange

is the primary factor contributing to variations in electron and spin densities, electronic energies, and other quantum mechanical observables. This effect can be evaluated by varying the amount of exchange in one particular functional (e.g., Siegbahn's B3LYP with 20%, 15%, 10% or 0% HF),[95] or performing calculations with two or three different functionals with (inherently) varying levels of HF exchange (e.g. at least one generalized-gradient, GGA or meta-GGA, and one hybrid). Employing few functionals in a particular project may also provide a rough estimate of the systematic error inherent in the calculated values.

Evaluating the performance of DFT functionals for $[Cu_2O_2]$ systems (in particular the **P** $\leftrightarrow$ **O** equilibrium) allowed us, somewhat surprisingly, to conclude that the inexpensive non-relativistic (NR) TPSS[89,90,96,97] functional in combination with double-zeta (DZ) basis set, def2-SVP,[98] performed quite well (**Paper I**, *vide infra*). This is not an unprecedented phenomenon in the literature, the usage of TPSS/TPSSh[89,90,96,97,99] functionals has been previously recommended for $[Cu_2O_2]$ systems.[36,66,67] It must be admitted that the good performance of the NR-TPSS functional with a small basis set is a consequence of the fortuitous cancellation of errors, as the small basis set and relativistic effects compensate each other. Nevertheless, we highly recommend it as the method of choice for the initial stages of QM or QM/MM modeling of CBC enzymes, which may involve scanning tens or hundreds of initial structures to, for example, test the convergence of the QM size (calculation of a $[Cu_2O_2]$ QM system of size $\sim$ 300 atoms takes only a few hours on a few CPU cores at the TPSS/def2-SVP level), and perform frequency calculations, etc.[40,41,78,100] To obtain final energies that are in a good agreement with experimental data, hybrid DFT functionals such as TPSSh or B3LYP, employing triple-zeta (TZ) (or larger) basis set are recommended.[101,102] Our experience showed that the two-component relativistic X2C-TPSSh/def2-TZVP[98] was closest to the experimental activation energies in the initials stages of phenol hydroxylation catalyzed by the CBC enzyme Ty (**Paper IV**, *vide infra*).[41]

In the last two decades, multiconfigurational pair-density functional theory (MC-pDFT)[103,104] has emerged as a promising approach that combines relative cost-efficiency with the inclusion of both static and dynamic correlation. We tested the performance of several 'MC-tailored' functionals within the MC-pDFT framework on $[Cu_2O_2]$ systems and obtained satisfactory results for the **P** $\leftrightarrow$ **O** equilibrium with the trevPBE and ftrevPBE functionals,[105] in combination with an (8-in-6) active space (**Paper I**, *vide infra*).[78]

2.3 Theoretical spectroscopy

CBC enzymes featuring the $[Cu_2O_2]$ center have been probed by many spectroscopic techniques, including UV-Vis absorption, resonance Raman (rR), and X-ray absorption spectroscopy (XAS). Correlation of experimental spectra with theoretical calculations is a direct method to extract information about local geometric and electronic structure (including oxidation states, and the peroxide binding mode).

In our study of $[Cu_2O_2]$ systems, (**Paper I**) we focused on the calibration of the O-O, Cu-Cu, and Cu-O vibrational modes, which are obtained experimentally using rR spectroscopy. By comparing computed and experimental

data on a relatively large and diverse set of inorganic $[Cu_2O_2]$ complexes,[78] we recommend to use pure (i.e., non-hybrid) functionals for frequency calculations (e.g. the TPSS functional, combined with a double-zeta basis set (def2-SVP), which has been used to define hydrogen bond network in the oxy-Ty, *vide infra*, **Paper II**).[100] As an important methodological conclusion, we investigated the influence of the polarized continuum solvent model (e.g. conductor-like screening model, COSMO[106]) on the calculated frequencies. As expected, the computed differences *in vacuo* and in the implicit solvent were marginal. However, using the implicit solvent model for frequency calculations might be safer for charged and highly polarized models, such as $[Cu_2O_2]$.

Over the past thirty years, accumulated evidence has supported the use of time-dependent (TD) DFT as a reliable and relatively user-friendly method for analyzing absorption spectra and electronic transitions. This conclusion is based on comparisons with experimental data, both for biomimetic model complexes[102,107–109] and protein active sites.[40] For instance, relatively small deviations between theory and experiments were observed in the $[Cu_2O_2\{HC(3\text{-}tBuPz)_2(CH_2Py)\}_2[SbF_6]_2]$ spectrum. The deviations were ~ 1500 cm^{-1} for the in-plane transition of the $\pi_{\sigma*}$-orbital of the ${O_2}^{2-}$ unit to the d-orbital of the Cu-atoms (experimental: 360 nm, calculated: 381 nm) and ~ 2100 cm^{-1} for the out-of-plane transition of the $\pi_{\nu*}$-orbital of the ${O_2}^{2-}$ unit to the d-orbital of the Cu-atoms (experimental: 560 nm, calculated: 634 nm).[102] For the CBC catalytic center of Ty, both the intensities and the absorption energies of peaks in the UV-Vis spectra were predicted with reasonable accuracy, supporting the molecular structure assignment of the observed ternary intermediate (*vide infra*, **Paper III**).[40]

Due to the lack of an EPR signal, XAS is the most direct method to probe the LUMO of the $[Cu_2O_2]$ active site. Analysis of XAS data provides valuable insights into the structure and reactivity of CBC systems, notably the formal oxidation states of the metal ion. This is especially useful for transient intermediates that are captured using techniques, such as rapid-freeze quenching (RFQ). XAS may also serve as a reference for benchmarking theoretical methods such as DFT functionals (e.g. estimating covalency of metal-ligand bonds). Qayyum et al. tested several functionals with varying percentages of Hartree-Fock exchange to accurately describe the metal center, experimentally validated by L-edge XAS spectra.[37] They concluded that the B3LYP functional best reproduces the electronic structures of the **P** and **O** $[Cu_2O_2]$ complexes.

3. Motivation

The remarkable role of tyrosinase in the biosynthesis of melanin(s) was the primary reason for taking up this topic during my PhD studies.

Owing to its role in the nature, Ty is of great interest in many fields of science: ii) in medicine, due to its implication in skin cancer and Parkinson's disease, ii) in the cosmetic industry, to prevent albinism and skin discoloration, iii) in agriculture, to prevent browning of vegetables and fruits.

Notably, my aim was to reveal the reaction mechanism of this enzyme. A full understanding of the mechanism substrate(s) binding and its conversion by Ty may help prevent and treat the above-mentioned illnesses, as well as extend the shelf life of agricultural products. The two reactions catalyzed by tyrosinase, monooxygenation and oxidation of the substrate, profoundly differ in their chemical nature. Despite extensive efforts, the reaction mechanism of Ty has so far been very poorly understood. It resulted in many different proposals, ambiguities, and inconsistencies. I believe that understanding this mechanism requires not only high-quality experimental data, but above all, state-of-the art computations and analyses.

4. Aims of the doctoral thesis

My doctoral project was part of a larger project that also involved other persons, which is why I chose to use the term 'we' when describing the outcomes of the presented research.

The main goal of this dissertation was to characterize and understand the reaction mechanisms of copper-containing enzyme tyrosinase, specifically focusing on the monophenol hydroxylation to catechol and the subsequent oxidation to quinone.

Considering the importance of the project, the following four main aims were proposed:

1. To calibrate DFT functional(s) of choice to available experimental data and higher-level (correlated) wave function methods, to determine the proper methodology for further study of the real enzymatic system. In particular, calibration was done for $[Cu_2O_2]$ compound(s), and their $\mu{-}\eta^2:\eta^2$-peroxide and bis-μ-oxo equilibrium.

2. To build the best possible model of oxy-Ty, with a focus on its calibration against the experimental data as an initial step for exploration of enzymatic reaction mechanism.

3. To elucidate the structure of ternary intermediate (Ty/O_2/substrate) and reveal monooxygenation reaction mechanism, also exploring different substituent in the *para*-group of monophenols.

4. To characterize the subsequent oxidation of catechol to quinone that completes the whole tyrosinase catalytic cycle.

Concerning the first task, we chose several biomimetic compounds, well-characterized in the literature, alongside with the 'iconic' artificial model of $[Cu_2O_2(NH_3)_6]^{2+}$ which has been previously studied by many researchers. We performed calculations by employing various (kinds of) DFT functionals with different basis set size, with or without X2C relativistic treatment and benchmarked them with respect to the X2C-CCSD(T) values and experimental data, when available. In addition, we studied the effect of solvation and computed vibrational spectra to compare them with the experimental resonance Raman spectra **(Paper I)**.

Considering the second goal, we performed 500 ns MD simulation and prepared more than sixty plausible QM/MM models to explore hydration level of the active site of oxy-Ty to find its best possible structural description of this state. All calculations were correlated with the rR spectra **(Paper II)**.

The third aim was achieved by exhaustive QM/MM and QM-cluster modeling correlated to kinetic and spectroscopic experiments. By this, we clearly defined and characterized the elusive ternary catalytic intermediate ($Ty{+}O_2$+substrate).

This allowed us to explore the monooxygenation with methyl 4-hydroxybenzoate analog of the native substrate (Tyr). Finally, we studied aforementioned reaction with different *para*-substituent monophenols to show interesting reversal in the RLS of the reaction (**Paper III** and **Paper IV)**.

The fourth objective, oxidation of catechol to quinone, was again obtained by correlating experimental spectroscopic and kinetic data to the QM/MM reaction coordinate. Computationally, this is simply the second part of the overall catalytic cycle. Experimentally, the second part (catechol oxidation) is conveniently studied by addition of monocatecholate to the met-Ty **(manuscript in preparation)**.

In summary, we believe that we have provided the fundamentally new insight into the structure of an elusive catalytic intermediate of Ty, as well as shed light on the enzymatic process carried out by CBC protein(s). Finally, we proposed a reliable methodology for studying such complicated biological system(s) with two (coupled) copper ions inside the active site.

5. Results and discussion

5.1 Theoretical view of the $[Cu_2O_2]$ isomerization equilibrium (Paper I)

The complexity of the electronic structure of $[Cu_2O_2]$ systems led us to a thorough computational investigation of the performance of quantum chemical methods for these 'entangled' or 'strongly correlated' species. As mentioned above, the 'production-level' QM and QM/MM calculations on enzymatic systems require the usage of DFT methods. The cost to be paid for their efficiency and (somewhat ironically) for the availability of hundreds of DFT functionals is a need for their calibration against experimental data, or more conveniently against higher-level WFT calculations. Thus, prior to modeling the reaction mechanism(s) of CBC enzymes, we wanted to estimate the anticipated errors in the computed data by carrying out such calibration. Last but not least, we also wanted to select the presumably most accurate DFT functional for problem at hand. Emphasis was laid on $\mu - \eta^2 : \eta^2$ (side-on peroxo) and bis-μ-oxo (**P** *vs.* **O**) isomerism (equilibrium). In addition, we also explored vibrational frequencies, geometrical parameters (interatomic distances) and thermochemical properties of the known $[Cu_2O_2]$ complexes (to compare with available rRaman, crystallographic data, and experimental enthalpy values).

For our study, we chose several experimentally characterized biomimetic compounds with $[Cu_2O_2]$ core (molecules **1-10** shown in **Figure 7**) as well as 'iconic' artificial $[Cu_2O_2(NH_3)_6]^{2+}$ model (structure **11** in **Figure 7**).

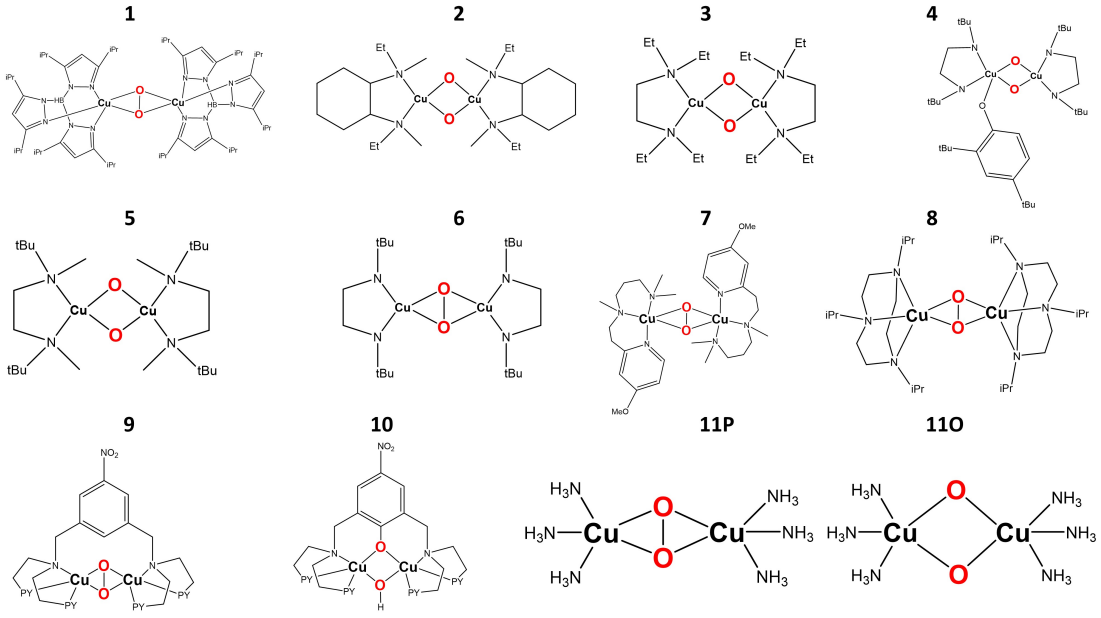


Figure 7: Biomimetic compounds **1-10** and model system **11** (both **P** and **O** forms of **11** are shown) studied in this work. Figure adapted from **Paper I** with some modifications.

The following computational methods were selected for benchmarking: 11 DFT functionals, MC-pDFT, MRCI, CCSD(T), and CASSCF. In addition,

we investigated the influence of basis set size, relativistic and solvation effects. To evaluate the performance of DFT functionals, we tested the following types: i) GGA functionals: BP86, B97D, and PBE, ii) meta-GGA, represented by TPSS, iii) hybrid: TPSSh, B3LYP, PBE0 and M06 (with 10 %, 20 %, 25 %, and 27 % of the exact exchange admixture, respectively), iv) range-separated hybrid functionals CAM-B3LYP and ωB97XD, v) and the local hybrid: Lh14t-calPBE.

We started our survey by exploring the small artificial model, compound **11**. Namely, we evaluated the **O** to **P** reaction coordinate as a relaxed scan (restrained geometry optimization) of the O-O bond distance with a step size of 0.05 Å around the minima and the transition state, and with 0.1 Å for the rest of the reaction coordinate, employing the X2C-TPSS/x2c-TZVPall level of theory. Then, for each optimized point, single-point energy calculations were carried out employing different theoretical methods: UCCSD(T), MRCI, MC-pDFT, and DFT, with and without considering relativistic effects. The obtained results are shown in **Figure 8**.

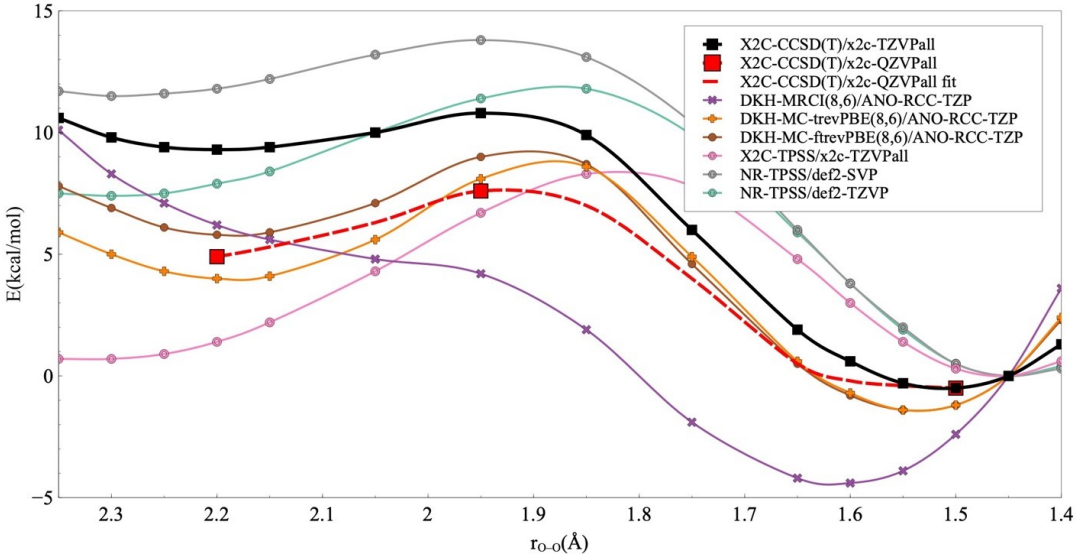


Figure 8: Comparison of selected computational levels for the relaxed scan defining the **11O** $\leftrightarrow$ **11P** reaction coordinate. Figure adapted from **Paper I**.

As can be seen, all tested quantum chemical methods, except MRCI with the 8-in-6 active space, predicted two stable minima (**O** and **P** isomers) and a transition state connecting them. It has been previously shown that a small active space size may not yield sufficiently accurate results, hence at least (24-in-28) active space size is recommended for the $[Cu_2O_2]$ system(s).[63] We chose UCCSD(T)/UKS as our 'benchmark', our 'gold standard' for further calibration. Interestingly, it can be seen (**Figure 9**) that several MC-pDFT functionals (e.g., MC-trevPBE and MC-revPBE) yielded results that are in good agreement with the X2C-UCCSD(T)/x2c-QZVPall potential energy curve. The X2C-UCCSD(T)/UKS with the quadruple-zeta (def2-QZVP) basis set, yielded consistent results comparable to previous, more elaborate MRCI approaches.[110] However, the standard ('canonical') X2C-UCCSD(T)/UHF scheme did not perform as well, presumably due to the spin-contaminated UHF reference wave function. On the contrary, the D_1 diagnostic in the

X2C-UCCSD(T)/UKS approach, which loosely correlates with the multireference character of the system, was acceptably low, around 0.1. This is the value that we typically find also for closed-shell complexes of e.g. Zn^{2+}. These findings suggest that X2C-UCCSD(T)/UKS might be a viable and *conceptually* simpler alternative to multireference wave function methods, provided that certain care is taken in the selection of the reference wave function (UKS, in this case). It is worth noting that similar approaches have been used previously to study iron systems by Radoń and Pierloot.[111–114] The X2C-UCCSD(T)/x2c-QZVPall results served as a reference for further calibration involving a set of various DFT functionals (**Figure 9**).

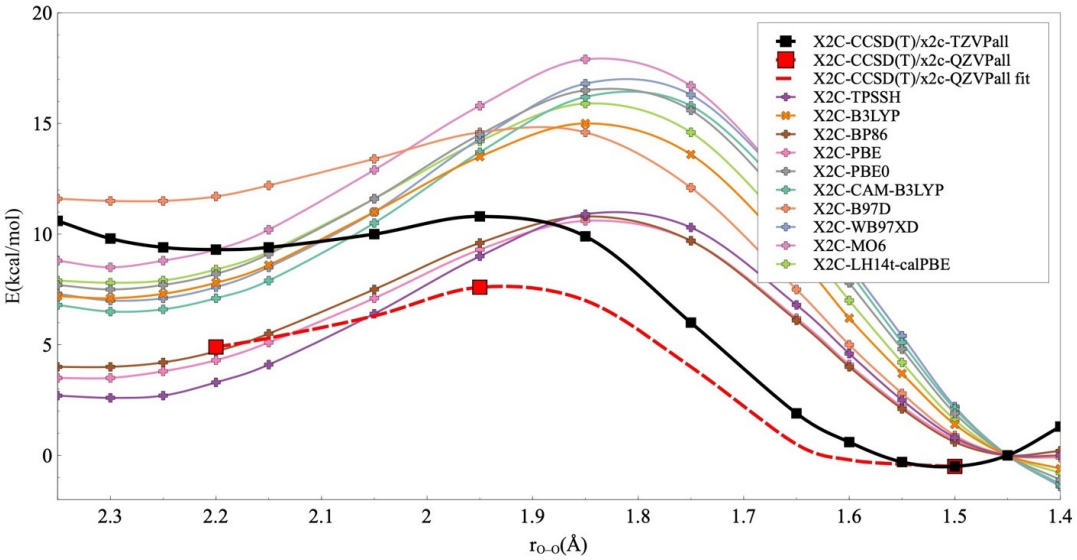


Figure 9: Comparison of various DFT functionals for the new **11O** ↔ **11P** isomerization coordinate. Figure adapted from **Paper I**.

The overall performance of various DFT functionals was rather satisfactory, much better than shown previously,[66] which was an encouraging result. We may only hypothesize whether the problem with previous benchmarking leading to discrepancies up to 30 kcal/mol might have been influenced by 'Cartesian interpolation' between **P** and **O**, or some other issues in DFT calculations. However, our findings enable us to study larger models (up to hundreds of atoms) using DFT calculations for which *ab initio* approaches are not feasible. Generally, DFT functionals were prone to overestimating the activation energy barrier by ∼ 5-10 kcal/mol (compared to the UCCSD(T)/QZ results), nevertheless, about half of the tested functionals accurately predicted the thermodynamics between these two isomers (∼ 6–8 kcal/mol). The closest energy barrier values compared to UCCSD(T)/QZ benchmark were obtained using PBE and BP86 (surprisingly), and with TPSSh functionals. The usage of the latter has been previously recommended for $[Cu_2O_2]$ systems.[66,67]

In the next step we focused on synthetic models. We investigated the accuracy and robustness of various DFT functionals for $[Cu_2O_2]$ systems by employing X2C approximation with x2c-TZVPall basis set. For the MC-pDFT, we used the ANO-RCC-TZP/DZP(H) basis set and the DKH2 scalar-relativistic approximation.

Last but not least, we also included in our research a practical and very cheap non-relativistic meta-GGA TPSS functional with def2-SVP basis set for comparison. It has been shown in the previous study to yield quite good performance.[67]

First, we focused on interatomic distances. The results obtained with pure functionals were, somewhat surprisingly, in better agreement with experiment than those calculated with the hybrid functionals. Moreover, the 'cheapest' NR-TPSS/def2-SVP gave very good agreement with the experimental data obtained by X-ray diffraction and extended X-Ray absorption fine structure (EXAFS) methods. Next, we examined the thermodynamics of the **P**/**O** equilibrium for the complexes **1-9**. For most complexes, experiments have only identified the more stable isomer of the studied $[Cu_2O_2]$ system. To evaluate (semi-)quantitatively the performance of the DFT functionals, we introduced a simple binary *scoring function*. It increments its value by 1, if the particular functional predicts correctly the more stable isomer, otherwise there is no change in the value of the scoring function. Thus, the result m/n says that m minima (**P** or **O**) out of the total number n synthetic complexes were accurately predicted by the specified method. The obtained scores are shown in **Table 1**.

Table 1: The 'score' results of various DFT functionals for the **O** $\leftrightarrow$ **P** isomerization of systems **1-9**.

Method of choice	**Score**
X2C-PBE/x2c-TZVPall	0.67
X2C-BP86/x2c-TZVPall	0.67
X2C-B97D/x2c-TZVPall	0.67
X2C-TPSS/x2c-TZVPall	0.56
X2C-TPSSh/x2c-TZVPall	**0.78**
X2C-B3LYP/x2c-TZVPall	0.56
X2C-M06/x2c-TZVPall	0.67
X2C-LH14t-calPBE/x2c-TZVPall	**0.88**
X2C-CAM-B3LYP/x2c-TZVPall	0.67
DKH-MC-ftrevPBE/ANO-RCC-TZP & DZP(H)	0.67
NR-TPSS/def2-SVP	**0.81**

In general, the performance of the selected DFT methods in predicting the more stable isomer is very good. TPSSh and LH14tcalPBE deserve to be highlighted with a 80-90 % success rate. NR-TPSS/def2-SVP gave a score of 0.81, and can be recommended as an excellent price/performance 'solution' in practical and larger-scale studies of $[Cu_2O_2]$ systems. Nevertheless, it must be emphasized that such good results are rather an effect of cancellation of errors, as the small basis set and relativistic effects counterbalance each other (since the small basis underestimates the **O** isomer, whereas the relativistic effect stabilizes it to a greater extent).

The next step in our research focused on vibrational frequency calculations (normal mode analysis). We wanted to compare the calculated frequency values (obtained with different functionals) with the experimental data obtained from rR spectroscopy, in particular for three $[Cu_2O_2]$ vibrational modes: O-O(A_g), Cu-Cu(A_g), and Cu_2O_2(B_{2u}) (shown in **Figure 10**), as these are the fingerprints of CBC systems.

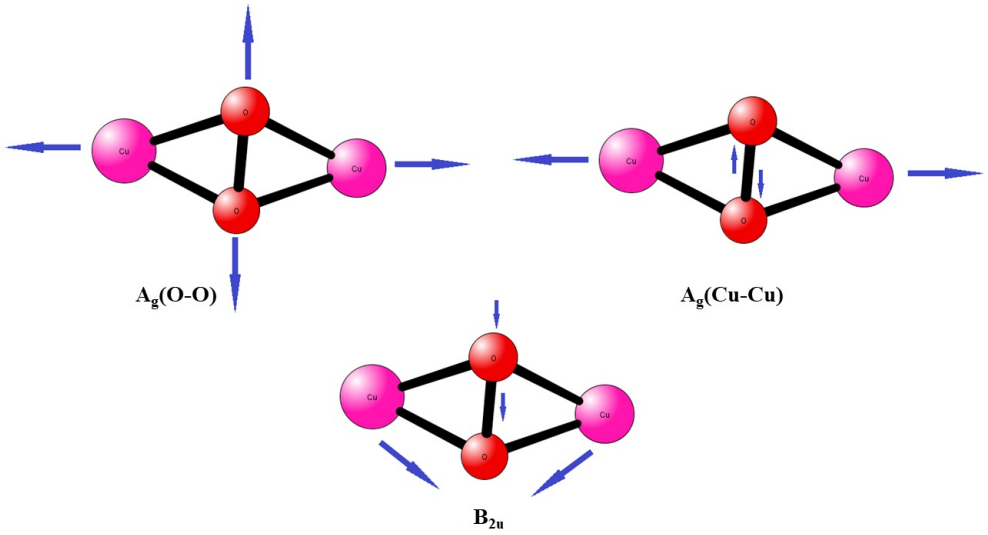


Figure 10: O-O(A_g), Cu-Cu(A_g), and Cu_2O_2(B_{2u}) vibrational modes of $[Cu_2O_2]$ core in D_{2h} symmetry. Figure adapted from **Paper I**.

As can be seen in **Table 2**, the non-hybrid functionals performed better than the hybrids. Particularly, NR-TPSS/def2-SVP can be highlighted. It gave the best agreement with the experimental values. TPSS with relativistic effects included and in combination with the larger basis set (which should be in principle a way superior treatment) performed only slightly worse. The good agreement with the experimental values obtained at the NR-TPSS/def2-SVP level of theory is noteworthy, because in the case of frequency calculations, the size of basis set (DZ *vs.* TZ) significantly affects the calculation time.

Table 2: Comparison between experimental and theoretical vibrational frequencies. All calculations were done employing the X2C relativistic approximation, x2c-TZVPall basis set and the COSMO solvation model. All values in cm^{-1}. Is. - isomer; Vib - vibration; Sys. - system (**1-10**, *cf.* **Figure 7**). Note that isomer **H** is the hydroxylated product for which we examined C-O vibrational mode.

Is.	Vib.	Sys.	PBE	BP86	B97D	TPSS	TPSSh	B3LYP	Exp.
O	$^{16}O_2$	2	619	624	612	635	653	650	610
	$^{18}O_2$	2	590	595	583	606	622	619	587
	$^{16}O_2$	4	599	604	587	620	639	633	590
	$^{18}O_2$	4	570	576	560	592	610	604	570
	$^{16}O_2$	5	611	617	605	627	645	643	600
	$^{18}O_2$	5	581	587	576	597	614	612	624
	$Cu_2{}^{16}O_2$	7	560	563	556	601	623	599	585
	$Cu_2{}^{18}O_2$	7	535	538	531	574	596	572	561
	$^{16}O_2$	8	588	595	588	607	625	624	592
	$^{16}O_2$	9	582	587	583	599	621	622	600
P	$^{16}O_2$	1	843	830	853	811	870	872	741
	$^{18}O_2$	1	796	783	804	766	821	823	719
	Cu-Cu	5	289	286	306	285	306	340	294
	$^{16}O_2$	6	754	729	788	716	797	853	721
	$^{18}O_2$	6	712	689	744	678	753	806	681
	Cu-Cu	7	270	273	277	274	300	291	287
	$^{16}O_2$	8	818	799	832	786	818	832	713
	$^{16}O_2$	9	929	918	942	892	904	902	747
	$^{18}O_2$	9	876	865	889	841	852	851	707
H	$C^{16}O_2$	10	1417	1409	1454	1399	1443	1500	1320

Is.	Vib.	Sys.	M06	LH14t-calPBE	CAM-B3LYP	NR-TPSS	Exp.
O	$^{16}O_2$	2	654	659	681	614	610
	$^{18}O_2$	2	623	628	649	586	587
	$^{16}O_2$	4	649	639	669	592	590
	$^{18}O_2$	4	618	609	638	565	570
	$^{16}O_2$	5	647	649	673	607	600
	$^{18}O_2$	5	616	618	640	579	624
	$Cu_2{}^{16}O_2$	7	599	608	635	548	585
	$Cu_2{}^{18}O_2$	7	573	581	607	524	561
	$^{16}O_2$	8	629	5631	658	597	592
	$^{16}O_2$	9	638	625	654	584	600
P	$^{16}O_2$	1	937	-	-	839	741
	$^{18}O_2$	1	884	-	-	792	719
	Cu-Cu	5	331	314	334	302	294
	$^{16}O_2$	6	882	908	868	766	721
	$^{18}O_2$	6	832	857	819	723	681
	Cu-Cu	7	301	301	318	286	287
	$^{16}O_2$	8	891	892	885	806	713
	$^{16}O_2$	9	969	959	952	944	747
	$^{18}O_2$	9	913	904	897	891	707
H	$C^{16}O_2$	10	1546	-	1527	1385	1320

To summarize, we attempted to provide guidelines for studying $[Cu_2O_2]$ systems, including CBC active sites. We used the UCCSD(T)/X2C-QZVPall results as the reference point for further calibration of cheaper methods, because we consider CCSD(T) a reliable, and conceptually simpler choice compared to multireference WF methods. To correctly describe the electronic structure and properties of CBC systems, a relativistic approximation and a basis set of at least TZ quality must be used. To obtain good interatomic distances of the $[Cu_2O_2]$ core, pure functionals are recommended. The same applies to vibrational frequencies. Pure functionals (and TPSSh) also work well for energetics, especially when it comes to **P**/**O** equilibrium. Eventually, the cheap approach represented by the NR-TPSS/def2-SVP is an alternative method for studying larger $[Cu_2O_2]$ systems. It gives fairly good results in terms of geometric quantities, energetics, and vibrational frequencies. Therefore, this level of theory is recommended for larger-scale production studies of CBC systems.

5.2 Computational and spectroscopic definition of oxy-Ty and of the elusive ternary intermediate. Suggested *ortho*-hydroxylation and oxidation reaction coordinates

Most of the theoretical results presented in the thesis were obtained using QM/MM calculations. Specifically, the ComQum software of Ryde and coworkers was used[115,116] which (typically) combines Turbomole[117] for the QM part and AMBER[118] for the MM part. For QM region description DFT method was used by employing TPSS[89,90,96,97] meta-GGA functional with combination of double-zeta def2-SVP[98] basis set, including the dispersion correction with Becke and Johnson damping (D3BJ).[119] In the MM part the Amber ff14SB force field was used.[120] In ComQum, the subtractive scheme was applied.[121] The electrostatic embedding scheme was used on the QM and MM border.

QM/MM model was built based on the crystal structure of oxy-tyrosinase from *Streptomyces castaneoglobisporous* obtained at the resolution of 1.8 Å (PDB code: 1WX2[13]).

Detailed information about the research methodology used in this scientific project (including experimental part) can be found in attached articles.

The catalytic cycle of Ty is sketched in **Figure 11**. To avoid confusion and ambiguities, we decided to present the version that has been the result of our computational and spectroscopic studies (presented later in this thesis).

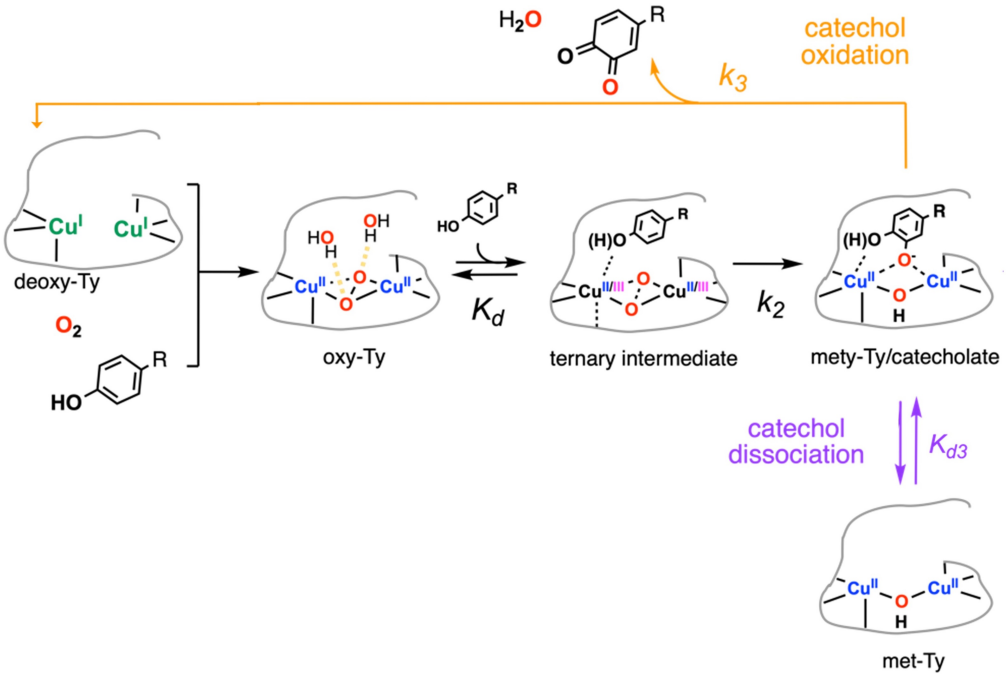


Figure 11: Tyrosinase catalytic cycle. Figure adapted from **Paper IV** with some modifications.

5.2.1 Evidence for hydrogen bonding interactions to the $\mu - \eta^2 : \eta^2$-peroxide of oxy-Ty (Paper II)

The second goal of my doctoral study was to construct the best model of oxy-Ty, which would then serve as a reference for exploring its enzymatic reaction.

For oxy-Ty, the crystal structure exists, and it is also a well-defined state for spectroscopic investigations. In our case, the experimental guidance was provided by high-quality rRaman spectra obtained both in H_2O and D_2O solutions. The obtained spectra featured the characteristic $[Cu_2O_2]$ peaks with interesting changes upon isotopic substitution. In deuterated water, Cu-O(B_{3u}) vibrational mode increased by 4 ± 2 cm $^{-1}$ (**Figure 12**), while remaining modes of $[Cu_2O_2]$ core (O-O and Cu-Cu stretches) were unperturbed.

The likely factor lying behind this phenomenon was the interaction of the peroxide moiety with water molecules. The active center of Ty is solvent-accessible, *unlike* in hemocyanin, for which a similar shift has not been observed. To confirm the hypothesis that the direct interactions between water molecules and the $[Cu_2O_2]$ core are responsible for Cu-O stretch upshift in D_2O, we performed extensive computational modeling.

First, we performed 500 ns MD simulation. Inspection of the trajectory clearly showed that almost all water molecules are highly dynamic, both within the

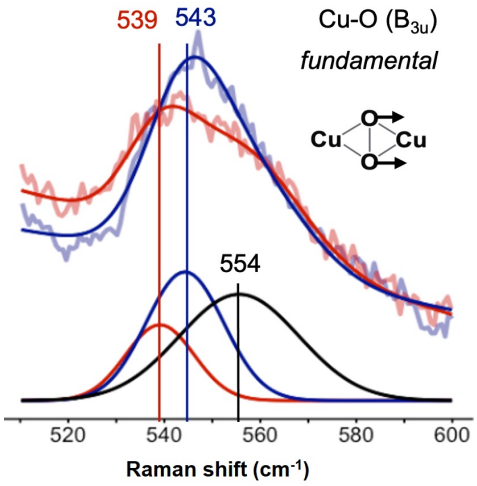


Figure 12: Resonance Raman (351 nm laser excitation) spectral region for the Cu-O(B_{3u}) fundamental stretch of oxy-Ty in H_2O (red) and D_2O (blue); the black Gaussian indicates a non-isotope sensitive peak present in both solvents. Figure adapted from **Paper II** with some modifications.

protein and also on its surface (see MD trajectory from **Paper II** SI[1]). However, in the immediate vicinity of the dicopper active site there are 3-4 water molecules, mostly hydrogen-bonded to peroxide, which are much less dynamic and remain there throughout the whole simulation (**Figure 13**).

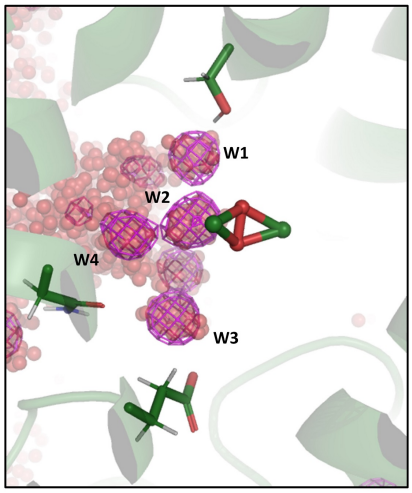


Figure 13: Visualization of water molecules from a MD simulation showing the conserved and distinct high-density water occupancy positions (indicated by mesh) for W1-W4, *vs.* the continuous and non-discreet distribution of water molecules in the bulk solvent outside the protein pocket (seen in the background). Figure adapted from **Paper II** with some modifications.

While admitting that this may be influenced to some extent by the (fixed) point charges at the $[Cu_2O_2]$ site, we consider this finding as strong qualitative evidence for a well-defined hydration network around the active site. One of the three water molecules was previously found in the crystal structure (W3), however additional two (W1 and W2) were not observed before. They form altogether with peroxide and second-sphere residues H-bonding network (**Figure 14**).

[1] `https://www.rsc.org/suppdata/d2/cc/d2cc00750a/d2cc00750a3.mpg`

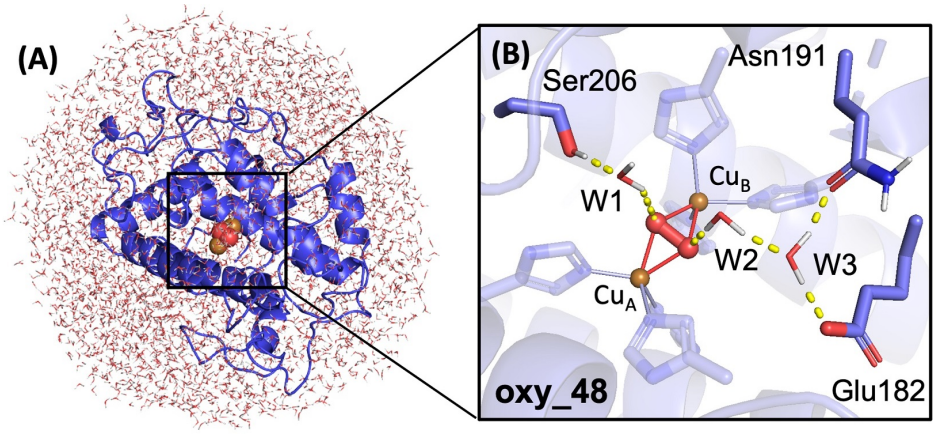


Figure 14: (A) The QM/MM-optimized structure for the experimentally-calibrated oxy-Ty active site (model oxy_48, blue), and (B) its active site. Figure adapted from **Paper II** with some modifications.

In the next step, we generated and optimized 64 QM/MM models of oxy-Ty differing in the number (2-8) and positions of water molecules in the substrate pocket of the enzyme's active site. Our task was to assess the hydration level of the catalytic centre of a monophenol-free enzyme (oxy-Ty state). A systematic evaluation of the thermodynamic stability of oxy-Ty at different levels of hydration, by releasing water molecule from the active site one by one showed a decrease in energy until the substrate pocket was filled by four water molecules, indicating that the catalytic center should be occupied by approximately four water molecules prior to the monophenol binding (**Figure 15** and **Table 3**). Again, we admit that the computed data should be considered qualitatively. Despite the use of the well-established QM(DFT-D3)/COSMO-RS computational protocol, previously tested for calculations of pK_a's, redox potentials, stability constants, and ligand binding, the absolute binding free energies of a water molecule (or the ligand in general) to the complex protein structure is a huge challenge for computational chemistry.[122–126] Nevertheless, our most thermodynamically stable structure with four water molecules (oxy_48) exhibits a network of hydrogen bonds between aforementioned W1 and W2 water molecules and peroxide (**Figure 14**), which is consistent with the MD simulation data.

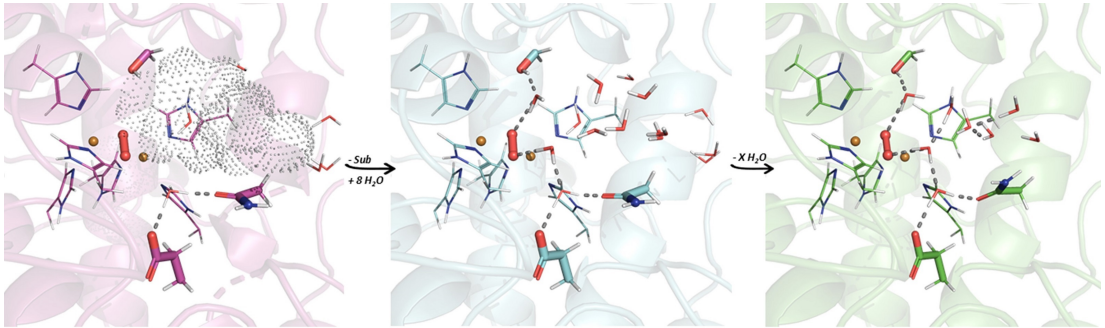


Figure 15: Schematic representation of examining hydration level of oxy-Ty catalytic centre shown by QM/MM-optimized structures of oxy-Ty models with substrate shown as volume (left), eight (middle) and two (right) water molecules in the place of the substrate in the active-site. Figure adapted from **Paper II** with some modifications.

Table 3: The evaluation of hydration level of oxy-Ty catalytic centre - free-energies (kcal/mol) of dehydration indicate the hydration level in the active site of oxy-Ty. $\Delta G_{dehydration}$ calculated as a $G_{products}$ (i.e. right hand side of the equation [oxy-Ty.$(H_2O)_n]^{c+} \rightarrow$ [oxy-Ty.$(H_2O)_{n-1}]^{c+}$ + H_2O) - $G_{reactants}$ (i.e. l.h.s. of the aforementioned equation). The free energy value was calculated as a sum of the gas-phase molecular energy (BP86-D3BJ/def2-TZVPD), the solvation energy (COSMO-RS method), the zero-point vibrational energy, and the entropic terms obtained from the rigid-rotor/harmonic oscillator approximation.

Dehydration	**H_2O**	**1-octanol**	**diethyl ehter**
oxyTy_8_WAT $\rightarrow$ oxyTy_7_WAT + H_2O	-2.5	-3.2	1.4
oxyTy_7_WAT $\rightarrow$ oxyTy_6_WAT + H_2O	-18.2	-24.4	-34.1
oxyTy_6_WAT $\rightarrow$ oxyTy_5_WAT + H_2O	-2.7	-2.4	-5.1
oxyTy_5_WAT $\rightarrow$ oxyTy_4_WAT + H_2O	-24.2	-26.7	-28.0
oxyTy_4_WAT $\rightarrow$ oxyTy_3_WAT + H_2O	-2.1	0.3	-0.6
oxyTy_3_WAT $\rightarrow$ oxyTy_2_WAT + H_2O	0.0	0.2	-1.1

To examine contribution of these interactions to the experimentally measured upshift of Cu-O stretch in deuterated water, we performed vibrational frequency calculations for the QM region of all 64 QM/MM-optimized models. By performing isotopic substitution and analyzing the computed Cu-O mode shift ($H_2O \rightarrow D_2O$), we observed a change in this stretch in D_2O for structures where water molecules interact with peroxide. It is worth noting that 52 out of 61 structures, with at least one H-bonded water to the $\mu - \eta^2 : \eta^2$-peroxide, exhibited the $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) upshift, and 18 out of 19 structures with two observed H-bonding interactions showed such an upshift. In particular, for the most thermodynamically stable model with four water molecules (oxy_48), the calculated shift was 10 cm^{-1}, which is quite consistent with the experimental value of 4 $\pm$ 2 cm $^{-1}$. Reference structures without water molecules near the peroxide exhibit no change in the Cu-O(B_{3u}) mode upon isotopic substitution (**Figure 16**).

The molecular mechanism behind this phenomenon originates from coupling of 'rR-active' vibrational modes of the [Cu_2O_2] core with the twisting and rocking modes of water molecules. For the H_2O, the rocking modes were calculated to be at 484 and 655 cm^{-1}, and twisting at 537 and 710 cm^{-1}, for W1 and W2, respectively. They are above 425 cm^{-1}, which was the calculated value for the Cu-O(B_{3u}) mode for oxy-Ty model *without* W1 and W2 present in the vicinity of peroxide (i.e. unperturbed [Cu_2O_2] site). For D_2O, twisting and rocking modes of W1 are lowered by approximately 150 cm^{-1} (calculated values are 346 and 394 cm^{-1} , for rocking and twisting modes, respectively) and are below the Cu-O(B_{3u}) mode. As a result, in H_2O the energy of the Cu-O(B_{3u}) mode is lowered because of its mixing with the higher-energy twisting mode of water, whereas in D_2O the energy of the Cu-O(B_{3u}) mode is increased due to its mixing with lowered-energy rocking mode. Therefore, the total effect of the decrease in Cu-O(B_{3u}) frequency in H_2O and the increase in D_2O through this mode-mixing mechanism provides the experimentally measured and reproduced by computation the $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) upshift.

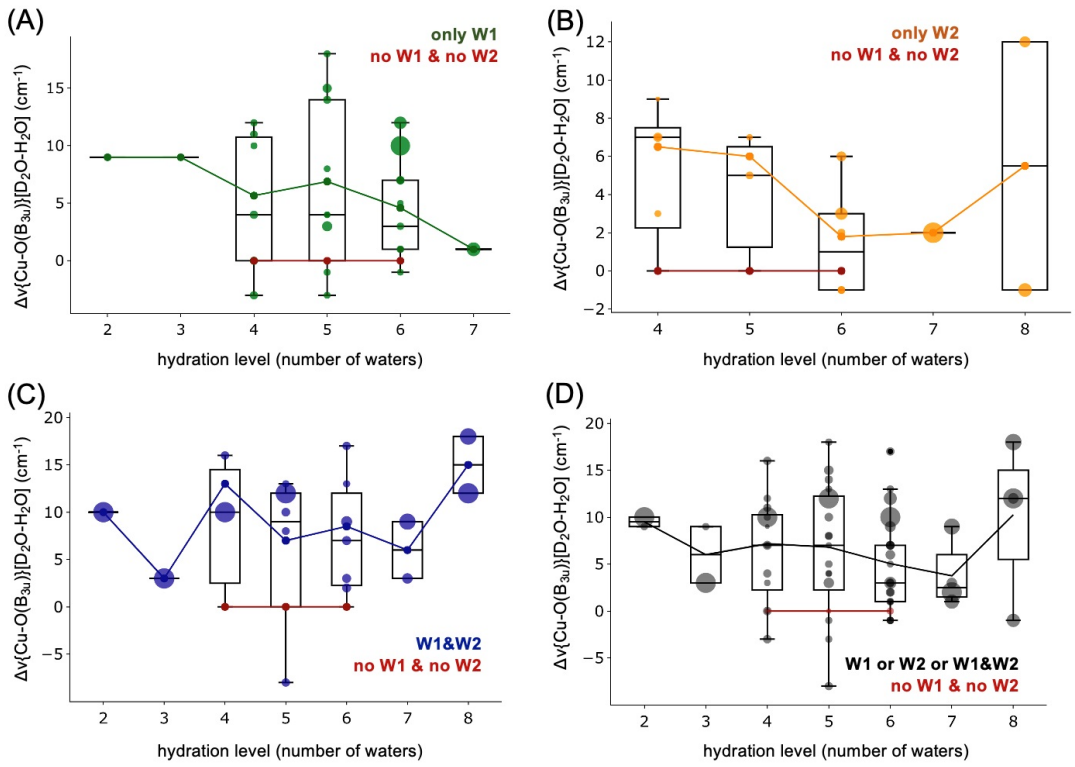


Figure 16: Boxplot distributions of the $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) frequency shifts in oxy-Ty at different active site hydration levels for structures with the $\mu - \eta^2 : \eta^2$-peroxide H-bonding with water(s) at the positions (A) only W1, (B) only W2, (C) W1 and W2 (i.e. 1 H-bond to each of the two peroxide O-atoms), (D) any position (i.e. at least one H-bond to either peroxide atom(s)). The red points/lines serve as references indicating structures that had no water molecules at the W1 and W2 positions. Figure adapted from **Paper II** with some modifications.

In summary, our results showed that the most realistic model of oxy-tyrosinase has four water molecules occupying the substrate pocket in the catalytic center, and two of them form hydrogen bonds with the $\mu - \eta^2 : \eta^2$-peroxide. These four waters are anticipated to have long residential times (slow exchange with solvent and with the rest of protein interior).

5.2.2 Elucidation of the Ty/O_2/monophenol ternary intermediate that dictates the monooxygenation mechanism (Paper III)

The third aim of my PhD project was to characterize the initial Ty/O_2/substrate structure (sometimes denoted as Michaelis complex, henceforth referred to as the ternary intermediate) of the monooxygenation reaction and reveal the mechanism of monophenol hydroxylation carried out by Ty. With the experimentally calibrated oxy-Ty model (*cf.* **Chapter 5.2.1**), serving as the reference for further exploration of the reaction coordinate, we therefore proceeded to study the first half of the Ty enzymatic reaction.

Ternary intermediate

Experimentally, the elusive ternary intermediate was captured and characterized by the stopped-flow absorption spectroscopy (SF-Abs) by mixing the deoxy-Ty with substrate in O_2-saturated borate buffer. Employing methyl 4-hydroxybenzoate (monophenol with electron withdrawing *para*-group (EWG)) as the substrate and borate system, our experimental colleagues were able to uncouple hydroxylation from the subsequent oxidation. The obtained ultraviolet-to-visible absorption spectrum for the ternary intermediate (**Figure 17**) showed two bands (in the region of 345 and 650 nm) similar to the transitions exhibited by oxy-Ty (peroxide $\rightarrow$ Cu(II) CT). This result demonstrated that active site consisted of the $\mu - \eta^2 : \eta^2$-peroxide moiety.

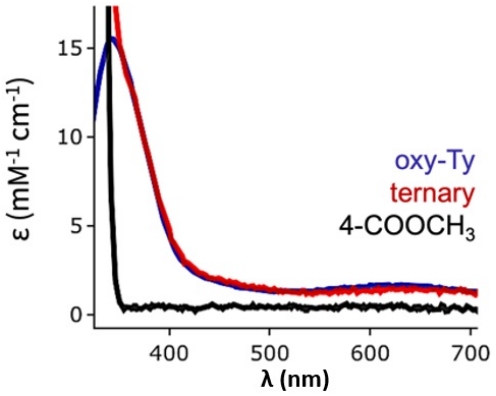


Figure 17: UV-Vis absorption spectra for the ternary intermediate, oxy-Ty, and 4-$COOCH_3$. Figure adapted from **Paper III** with some modifications.

Spectroscopic definition of the ternary intermediate was also obtained by resonance Raman experiments, which confirmed the $\mu - \eta^2 : \eta^2$-peroxide (by presence of characteristic peak of O-O stretch at 742 cm^{-1}). Comparison between rR spectra of oxy-Ty and ternary intermediate in case of O-O and Cu-O vibrational modes showed only minor perturbations caused by substrate binding. O-O(A_g) increased by 3 cm^{-1} and Cu-O(B_{3u}) overtone by 24 cm^{-1}, which would equal to 12 cm^{-1} in (a nontotally symmetric) fundamental mode (**Figure 18**).

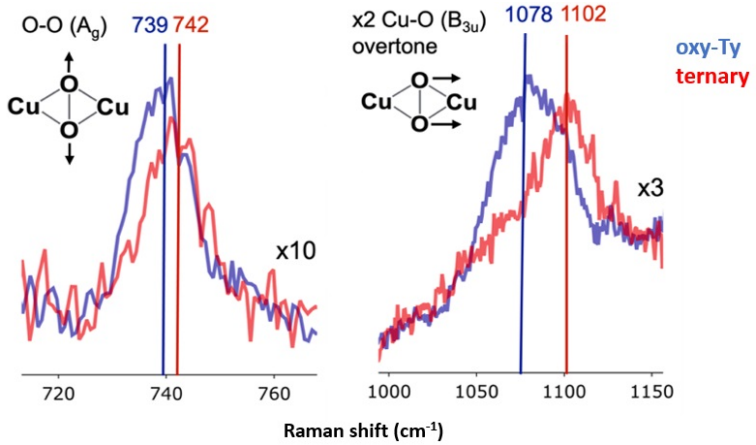


Figure 18: rR spectra of oxy-Ty and the ternary intermediate (Ty/O_2/4-$COOCH_3$). Figure adapted from **Paper III** with some modifications.

The next thing that needed to be evaluated was the protonation state of the substrate during the formation of the intermediate. Based on values of k_2 and K_d obtained from pH dependence kinetic experiments (former is independent, while latter decreases with decreasing pH) it was shown that substrate binds to the catalytic center in its fully protonated form or as a monophenolate with proton accepted by a base present in the active site. This has been also in accordance with the above rR spectrum: similarity between the oxy-Ty and the ternary intermediate suggested only small perturbation of the active site by the substrate binding and was incompatible with phenolate binding of the **O** isomer in the active site.

To obtain further structural insight into the ternary intermediate, QM/MM calculations were employed. From multitude ($\sim$ 30) of various structural alternatives investigated, five plausible models were further investigated to more details, all of them corresponding to the $\mu{-}\eta^2 : \eta^2$-peroxide (**P**) isomer (structures **1P-5P** in **Figure 19**). We evaluated the possibility of substrate coordination to one of the copper ions (CuA) and various positions of phenolic proton (substrate, axial His or Glu182/water(W3)/Asn191 cluster nearby).

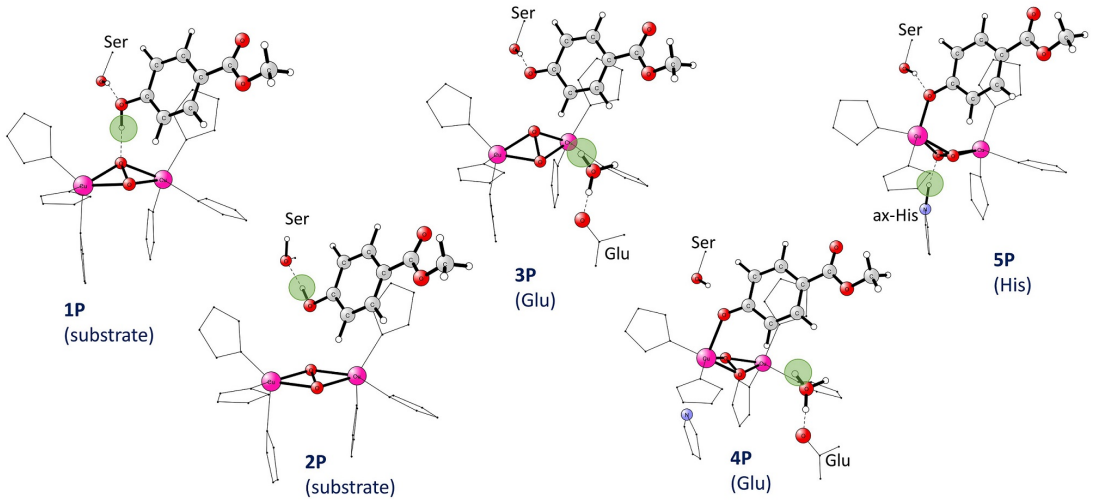


Figure 19: Possible models of ternary intermediate with consideration of different phenolic proton acceptor sites (shown as green circles).

For each structure, we calculated vibrational frequencies and TD-DFT spectra to compare the obtained results with experimental data. The huge differences in the computed frequencies (with respect to the experimental values) allowed us to exclude structures **2P-5P** from further analysis (**Figure 20**). They exhibited a large increase or decrease in one of the two stretches (**2P** +69 cm^{-1} and **4P** +81 cm^{-1} for Cu-O(B_{3u}), wheres **3P** +24 cm^{-1} and **5P** -50 cm^{-1} for O-O(A_g)). The intermediate **1P** has values closest to experiment (+18 cm^{-1} and -3 cm^{-1} for Cu-O(B_{3u}) and O-O(A_g); the measured changes in the rR spectra were +12 $\pm$ 3 cm^{-1} and +3 $\pm$ 2 cm^{-1}, respectively). Similarly, TD-DFT spectra (**Figure 21**) showed a shift of intense absorption band ($\sim$ 345 nm) towards higher energy for **2P**, **4P**, and **5P** models, whereas **3P** intermediate exhibited this peak at $\sim$ 400 nm, while **1P** species the most closely reproduced oxy-Ty spectrum (as seen in experimental optical spectra; **Figure 17**).

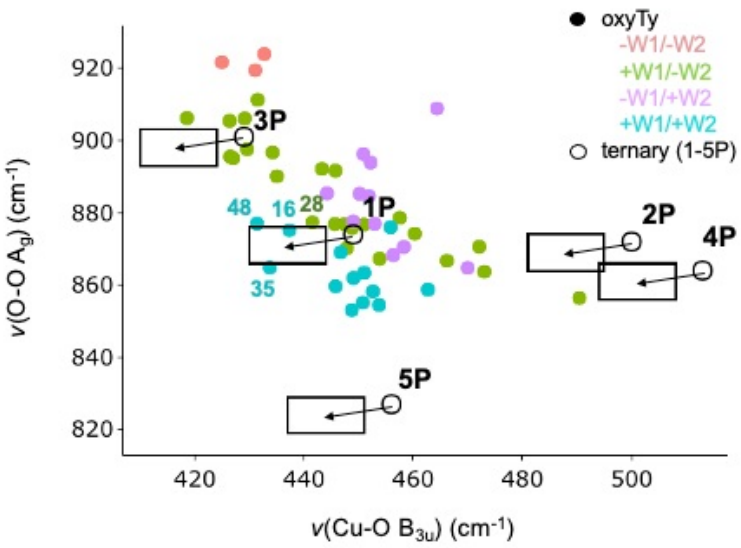


Figure 20: Correlation of the O-O(A_g) and Cu-O(B_{3u}) vibrational modes changes upon methyl 4-hydroxybenzoate binding to oxy-Ty between experiment (rR) and calculations (QM). Filled circles represent the 64 QM/MM-optimized oxy-Ty structures (**Paper II**), and colored based on the different water H-bonding interactions to the peroxide. The black circles represent the possible ternary intermediates (**1P-5P**) and their corresponding black arrows indicate the O-O(A_g) and Cu-O(B_{3u}) modes of oxy-Ty that are required to reproduce the rR data (the black boxes indicate the cumulative associated errors; presumed as ± 3 cm^{-1}). Note, that oxy_48 was the most thermodynamically-stable and experimentally-validated structure for the oxy-Ty active site (**Paper II**), and together with **1P** ternary model, it can reproduce experimental rR perturbations. Figure adapted from **Paper III** with some modifications.

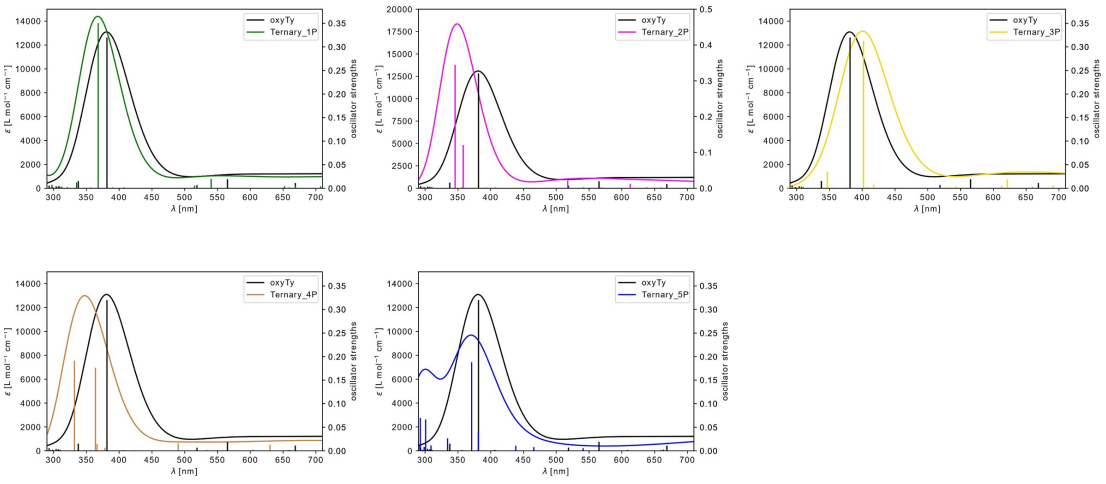


Figure 21: TD-DFT spectra of the possible ternary intermediates (**1P-5P**) *vs.* oxy-Ty (in black). Figure adapted from **Paper III** with some modifications.

Taken together, the correlation of experimental and computational data revealed that model **1P**, in which substrate binds fully protonated and forms a hydrogen bond with the peroxide is the catalytically relevant ternary intermediate (**Figure 22**). Importantly, it was also the structure with the lowest QM/MM and QM/PCM energy among all the candidates for the ternary intermediate (**2P**, **3P**, **4P**, and **5P** were higher in energy by +5.1, +21.4, +22.5, and +28.1 kcal/mol,

respectively), and the calculated frequency changes for the key vibrational modes of $[Cu_2O_2]$ were closest to experimental values. Also, TD-DFT **1P** spectrum (top left in **Figure 21**) thoroughly reproduces oxy-Ty absorption band (as a reference we used previously calibrated oxy-Ty structure (**Paper II**)).

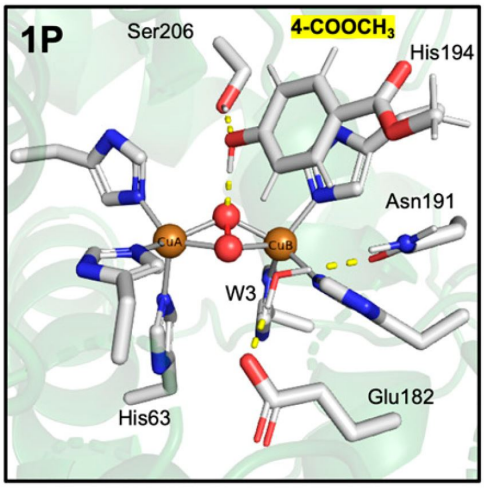


Figure 22: The active site of the QM/MM-optimized structure of ternary intermediate **1P** ($Ty/O_2/4\text{-}COOCH_3$). Figure adapted from **Paper III**.

Having well-defined and characterized ternary intermediate (**1P**) we were able to start exploring monooxygenation reaction.

Monooxygenation reaction

The kinetic data for the monooxygenation reaction were obtained in the single-turnover experiment monitored by SF-Abs. Thanks to the borate system trapping it was possible to uncouple monophenol hydroxylation from the subsequent catechol oxidation for substrates with the EWG *para*-group. Catechol product (after dissociation from met-Ty) was trapped by condensation with borate complex.

Enthalpy barrier of the RLS in monooxygenation reaction, based on Eyring plot, was calculated as $\Delta H^{\ddagger} = +10.3 \pm 0.7$ kcal/mol (from experimental rate constant). Experimentally measured solvent kinetic isotope effect (SKIE) on k_2 with value of $k_{H/D} = 1.5 \pm 0.1$ indicated no involvement of phenolic proton in the rate-limiting step of the reaction. Moreover, no KIE was observed for the dideuterated (in *ortho*-position) monophenol ($k_{H/D} = 1.05 \pm 0.05$).

To correlate calculations with the experimental data, we used the same substrate, methyl 4-hydroxybenzoate. In the first step, we ruled out the direct *ortho*-hydroxylation of protonated monophenol (in both isomers: bis-μ-oxo, **O**, and side-on peroxo, **P**) due to the unrealistically high activation barriers (>+25 kcal/mol, **Figure 23**), which was an indication that deprotonation of the substrate must occur first.

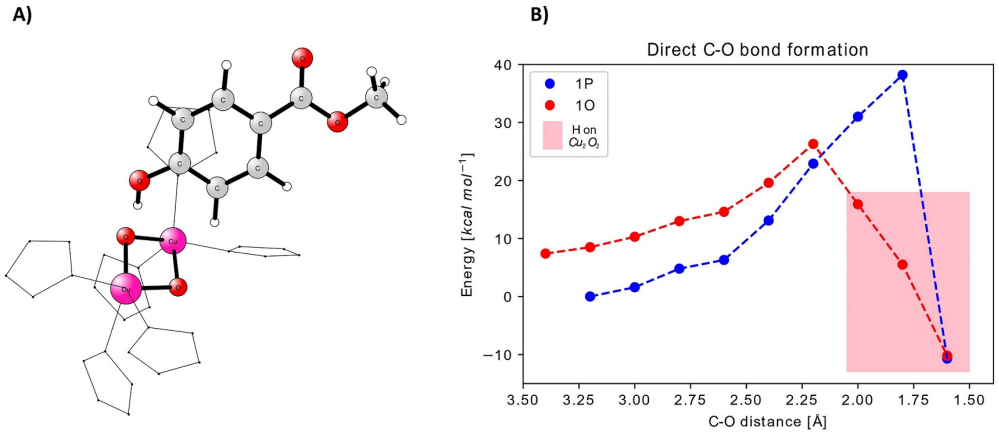


Figure 23: A) Intermediate **1O**, and B) 1D scans for **1P** and **1O** (blue and red dots, respectively) to C-O bond formation; the shaded regions indicate the structures after the spontaneous phenolic proton transfer to the $[Cu_2O_2]$ site during optimization. Figure adapted from **Paper III** with some modifications.

In addition to the previously evaluated possible proton acceptor sites (axial His and Glu182/water(W3)/Asn191 cluster in the models **3P-5P**), we investigated the $[Cu_2O_2]$ core at different stages of the monooxygenation reaction (i.e. before and after the O-O peroxide bond cleavage, **P** *vs.* **O**), as a potential H-acceptor site. In fact, it turned out to be the eventual proton-acceptor site, especially in its bis-μ-oxo isomer (intermediate **2O**, *vide infra*) due to high thermodynamic stability (*cf.* previous section, **Figure 24** and **Table 4**).

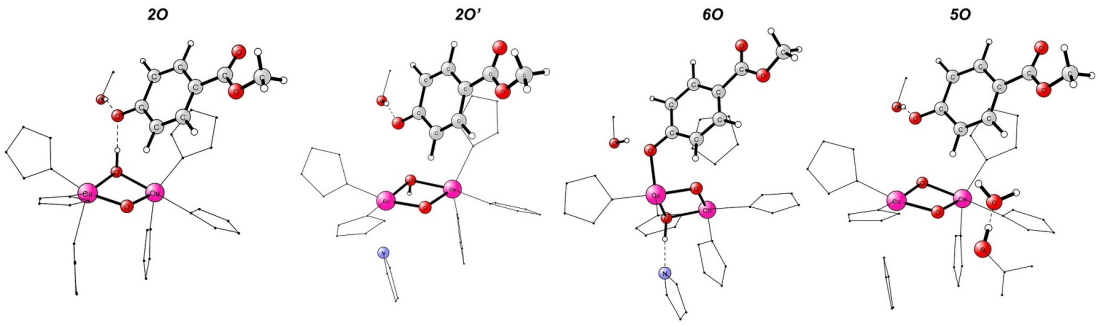


Figure 24: The structures for different proton acceptor sites (in addition to possible ternary intermediates **1P-5P**) at different stages in the monooxygenation reaction. Figure adapted from **Paper III** with some modifications.

Table 4: Characterisation and evaluation of the thermodynamic stability of the structures for various proton acceptor sites presented in **Figure 24**. The **1P** structure is included as the reference.

Structure	O-O cleaved?	Substrate coordinated?	Proton position	Axial His dissociated?	Energy (kcal/mol)
1P	no	no	4-$COOCH_3$	no	0.0
2O	yes	no	Cu_2O_2	no	1.5
2O'	yes	no	Cu_2O_2	yes	7.0
6O	yes	yes	Cu_2O_2	yes	0.1
5O	yes	no	Glu/Asn/W3	yes	20.4

To investigate the aforementioned **2O** intermediate and find the relevant transition state, we performed the 2D reaction coordinate scan along the peroxide O-O bond elongation as well as for the $H_{phenolic}$-O_{oxo} distance (O-O bond cleavage and concomitant phenolic proton transfer to the $[Cu_2O_2]$ moiety) and identified a low energy (+7.0 kcal/mol) transition state (**TS1**) leading to this thermodynamically stable (+1.5 kcal/mol) μ-oxo-μ-hydroxo intermediate (**2O**). Starting from the **2O** structure we performed another 2D scan corresponding to the concomitant formation of the C_{ortho}-O_{oxo} and CuA-$O_{phenolate}$ bonds and found a transition state (**TS2**) with the barrier of +14.1 kcal/mol to produce stable (-18.7 kcal/mol) intermediate (**3O**), in which the *ortho*-hydroxylated substrate is bridging the copper-hydroxo core. Mulliken population analysis performed for the key intermediates and transition states showed transfer of electron in the deprotonation step (to form **2O** intermediate) indicating the radical character of this step in the hydroxylation reaction (**Table 5**).

Table 5: Mulliken spin densities for selected atoms for key reaction intermediates and transition states. Table adapted from **Paper III** with some modifications.

	CuA	*CuB*	*O (H-bonded)*	*O (not H-bonded)*	*Σ Cu2O2*	*Σ SUB*
1P	-0.37	0.38	0.01	-0.01	0.01	0.00
1O	-0.02	-0.04	0.08	0.02	0.04	0.00
2O	-0.09	-0.07	-0.08	-0.41	-0.65	0.63
3O	0.56	-0.55	-0.03	0.02	0.00	0.01
TS1	0.03	0.05	-0.06	-0.05	-0.03	0.01
TS2	0.40	-0.25	0.04	0.00	0.20	-0.22

To confirm the proposed reaction mechanism (shown in **Figure 25**), we preformed subsequent frequency calculations in order to obtain thermochemistry correction to energy as well as calculate KIE. Computed enthalpy of **TS2** was in excellent agreement with experimental energy barrier ($\Delta H^{\ddagger}_{calc}$ = +11.4 kcal/mol *vs.* $\Delta H^{\ddagger}_{exp}$ = +10.3 ± 0.7 kcal/mol). Similarly, the calculated kinetic isotope effects of the **TS2** reproduced experimental values. For this second step, the C_{ortho}-O_{oxo} and CuA-$O_{phenolate}$ bonds formation, we calculated solvent KIE = 1.3, while after the substitution in the *ortho*-position (*ortho*-C-H/D) KIE = 1.0 (experimentally obtained were 1.5 ± 0.1 and 1.05 ± 0.05, respectively.)

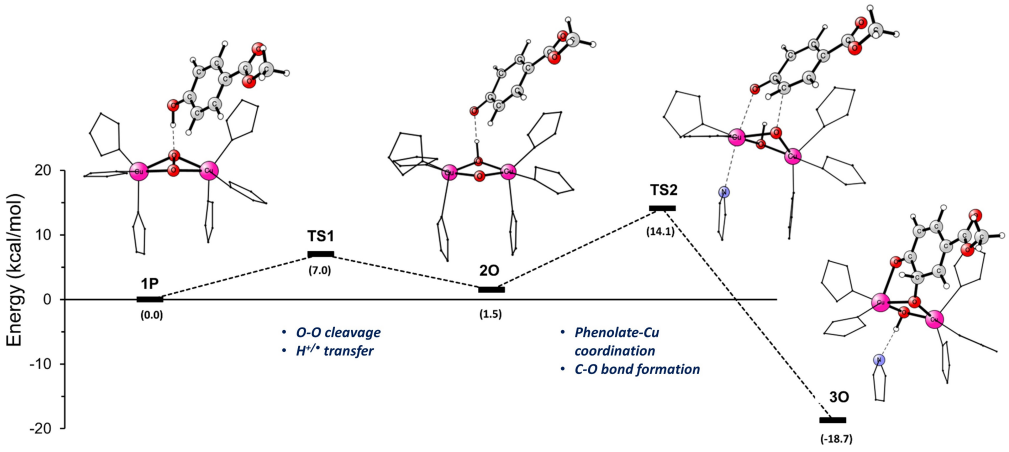


Figure 25: Experimentally supported QM/MM mechanism of the key steps in the monooxygenation reaction of Ty. Figure adapted from **Paper III** with some modifications.

Combined experimental and theoretical efforts have therefore provided fundamental insights into the key steps of the hydroxylation reaction mechanism carried out by Ty, where 1) O-O bond cleavage of peroxide followed by substrate deprotonation leads to μ-oxo-μ-hydroxo intermediate (**2O**), and 2) in subsequent rate-limiting step occurs concerted coordination of monophenolate to CuA with *ortho*-C hydroxylation by the non-protonated O_{oxo}.

5.2.3 Dependence of the substrate *para*-group on the hydroxylation reaction (Paper IV)

After revealing the monooxygenation reaction mechanism of Ty with methyl 4-hydroxybenzoate (monophenol with EWG *para*-group) we decided to validate our proposal by employing different substrates. Our computational investigation was driven by experimental findings, namely k_2 values, dependent on the EWG/EDG character of the *para*-group. Pre-steady-state/single-turnover kinetic data indicated that this reaction rate was up to $\sim$ 3 orders of magnitude faster than steady-state rate of the overall catalytic cycle (that includes the catechol oxidation). This outcome shows that former studies[23,127] (based on the steady-state rate Hammett plots) presumably reflected other RLS, therefore do not fit to the monooxygenation mechanism.

In this part of the project, kinetic data (monooxygenation rate constant, k_2) were obtained for series of monophenols differing by their *para*-groups from the most EWG (4-$COOCH_3$) to the most electron donating group (EDG) (4-OCH_3). A Hammett plot (**Figure 26**) was then constructed to show the dependence of the monooxygenation rate constant (in particular, $\log(k_2)$) on the ability of the substrates' *para*-groups to donate electrons (presented by their σ^+ values).

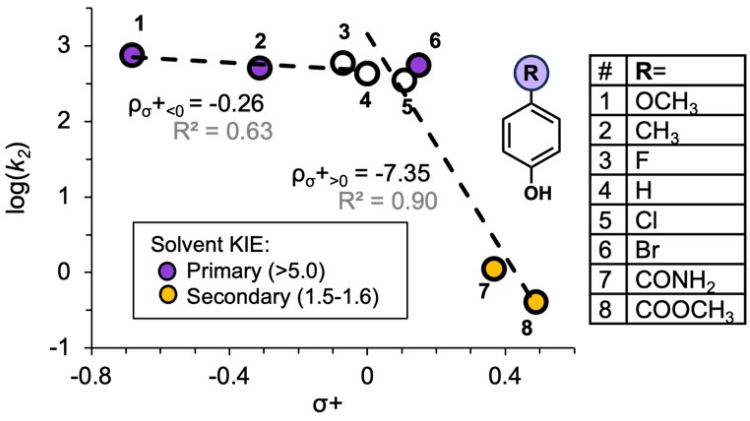


Figure 26: Hammett plot for the tyrosinase monooxygenation reaction. Figure adapted from **Paper IV** with some modifications.

It can be seen that for monophenols with EWGs there is a linear relationship between the monooxygenation rate constant (k_2) and the electron-donating ability (σ^+) of the monophenol *para*-groups ($\log_{10}(k_2)$ values increase with gradual accumulation of electron density in aromatic ring), whereas the rate of monooxygenation reaction with substrates with EDGs depends only marginally on the σ^+ value. Furthermore, this set of monophenols exhibited differences in

their solvent KIEs. In the contrast to the substrates with EWGs,[40] there is a large primary SKIE (>5.0) measured for monophenols with EDGs.

Summarizing the experimental data, the observed biphasic substrate *para*-group dependence on the rate constant and SKIE of the hydroxylation reaction implies a remarkable alteration in the rate-limiting TS correlating with the ability of donating an electron density to the ring by the substrate substituent in position 4 (with respect to -OH).

To investigate and understand these phenomena (the difference between two TS's and change in the RLS) we performed exhaustive QM/MM modeling. Based on our experience and results from previous study (**Paper III**), we used **1P** intermediate with methyl 4-hydroxybenzoate as a template and modified substrate *para*-group into: $-CONH_2$, -Br, $-CH_3$, and $-OCH_3$ (i.e. employing *para*-groups with increasing EDG). We optimized ternary intermediate and performed 2D potential energy surfaces (PES) scans from **1P** to **2O**, and subsequently from **2O** to **3O**, to reproduce (for each monophenol) previously revealed monooxygenation reaction (i.e., to obtain minima and identify transition states along the pathway).

The computed energetics of tyrosinase hydroxylation reaction with various substrates showed the following trend: with the donation of electron density to the aromatic ring, the **2O**, **TS2**, and **3O** structures are more stabilized, whereas **TS1** remains unchanged (**Figure 27**).

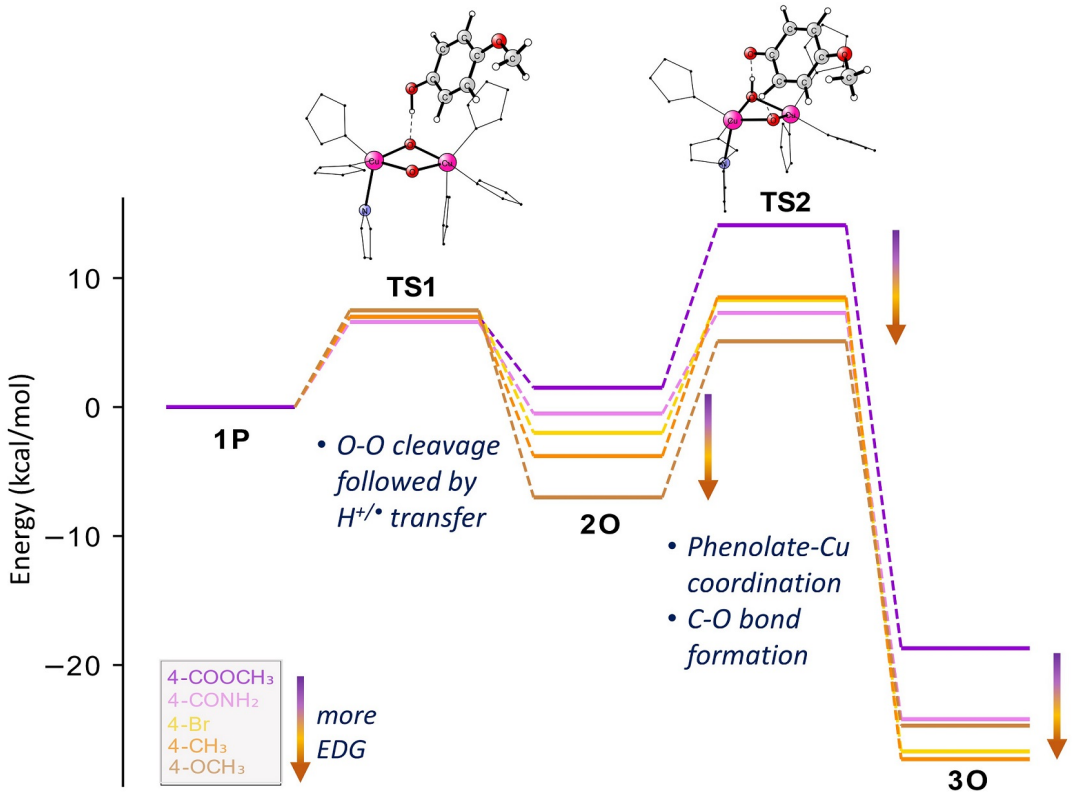


Figure 27: Monooxygenation reaction coordinate for monophenols with different *para*-groups.

Therefore, for EDG's monophenols the **TS2** becomes lower-in-energy transition state, consequently the **TS1** appears to be the RLS. This theoretical result corresponds in part to the experimental data shown in the Hammett plot

(**Figure 26**), since **TS1** (peroxide O-O bond cleavage) does not depend on the electron density accumulation in the aromatic ring. A weak dependence of the k_2 values on the substrate para-group is correct, *however*, the calculations did not provide large SKIE for any of the TS's (range of 1.1-1.3 and 1.3-1.4 for **TS1** and **TS2**, respectively).

A large SKIE indicates that H-transfer is involved in the RLS. In a previous study (**Paper III**) we showed that O-O bond cleavage of peroxide (**TS1**) is followed by monophenol deprotonation to form **2O** intermediate. A closer analysis of **TS1** allowed us to notice that as the electron density on the ring increases (EDGs substrates), the cleavage of the O-O bond in peroxide and phenolic H-transfer to the $[Cu_2O_2]$ site occur in more concerted way (than in the EWGs' case). In fact, for the 4-OCH_3 monophenol we were not able to obtain an intermediate with only the O-O bond broken (structure **1O**), as a result of the spontaneous transfer of the phenolic H-atom from the substrate to the bis-μ-oxo $[Cu_2O_2]$ core, already forming **2O** intermediate. To gain more insight into above-mentioned concertedness and to identify transition state which exhibits large SKIE, we explored TS's from 2D PESs reflecting at the same time O-O bond cleavage and H-transfer (labeled **TS1c**; **Figure 28**).

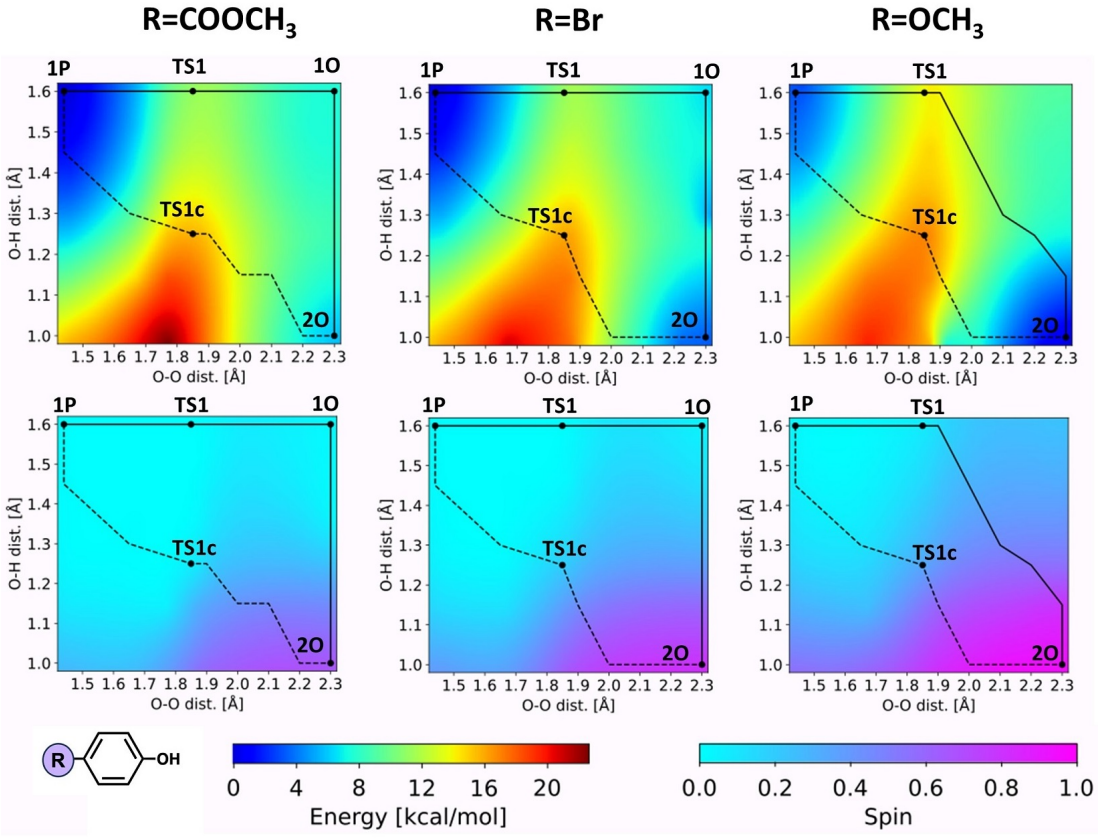


Figure 28: The 2D potential energy surfaces for the **1P** → **2O** reaction coordinate for 4-$COOCH_3$ (left; the most EWG), 4-Br (middle), and 4-OCH_3 (right; the most EDG). The O-O axis indicates the O-O peroxide bond cleavage and O-H axis indicates the phenolic proton transfer to oxygen of bis-μ-oxo $[Cu_2O_2]$ core. Solid black lines show stepwise process, while dashed lines indicate the concerted mechanism *via* **TS1c**. Top plots show changes in the energy, while bottom ones indicate electron spin density on the substrate along the **1P** → **2O** reaction coordinate. Note that electron transfer is much faster (and occurs together with proton transfer) for EDG substrates.

For the **TS1c** of each substrate, the calculated SKIE values were consistent with experiments (varying from 4.3 to 5.2, **Table 6**). This is also worth noting that at the QM/MM level of theory (QM(TPSS-D3BJ/def2-SVP)/MM(ff14SB)), **TS1c** was only slightly higher in energy than **TS1** (+1.0; +3.5; +3.8; +1.6; +0.9 kcal/mol for -$COOCH_3$, -$CONH_2$, -Br, -CH_3, and -OCH_3, respectively).

Table 6: Experimental and calculated solvent KIE values for **TS1**, **TS1c** and **TS2**.

Monophenol	**TS1**	**TS1c**	**TS2**	**Experimental SKIE**
4-OCH_3	1.2	5.2	1.4	5.8 ± 0.4
4-CH_3	1.3	5.0	1.4	5.6 ± 0.5
4-Br	1.2	4.3	1.4	7.1 ± 0.3
4-$CONH_2$	1.2	4.5	1.4	1.6 ± 0.1
4-$COOCH_3$	1.1	4.7	1.3	1.5 ± 0.1

The outcome of this part of the project provided clear proof for the transfer of phenolic H to the $\mu - \eta^2 : \eta^2$-peroxide and showed that the RLS of tyrosinase hydroxylation reaction may be altered by the monophenol *para*-group. The significance of experimental validation of calculated predictions (and *vice versa*) in the study of (metallo)enzymatic reaction is also highlighted.

It needs to be emphasized that the recently revealed mechanism of oxy-Ty monooxygenation (with phenolic H-transfer to bis-μ-oxo [Cu_2O_2] core in the initial step), later verified by studying the reaction with a broad set of substrates is substantially distinct from past proposals that suggested coordination of the monophenolate substrate with a Cu(II) or Cu(III) site.

5.2.4 The catechol oxidation mechanism in the monophenolase catalytic cycle of Ty (manuscript in preparation)

The fully described and experimentally verified mechanism of the tyrosinase hydroxylation reaction was an excellent basis for moving towards the final goal of the PhD project: characterization of the oxidation of catechol to quinone which is the last missing step to elucidate the entire tyrosinase catalytic cycle.

Completing monooxygenation cycle of tyrosinase to provide a connection between oxy-Ty/monophenol and met-Ty/catecholate reactions

The first part of the last section of this study, concerning the relationship between the hydroxylation and oxidation reactions, was of a purely computational nature. The reason was that in order to examine and experimentally characterize the oxidation reaction and its related intermediates, the starting point of the experiment was the met-Ty/catecholate complex which is a presumable experimental (re-)connection to the oxidation part of the Ty catalytic cycle.

Our calculations started with the monooxygenated sp^3 intermediate (**3O**, with an energy of -24.1 kcal/mol relative to the ternary intermediate (**1P**); which is the single-point QM/COSMO energy on top of the QM/MM-optimized geometries employing the x2c-TPSSh-D3BJ/TZVPall/$\epsilon_r = 8$ level of theory, *cf.* **Paper III & IV**). We first evaluated two possible reaction coordinates for the $sp^3 \rightarrow sp^2$ conversion to form the catechol(ate) intermediate. This step involves the transfer of the H^+ from the C(sp^3) center to one of the oxygens of the monooxygenated substrate (O1 corresponds to the hydroxyl *para* to the methyl ester and O2 at the *meta* position, *cf.* **Figure 29**).

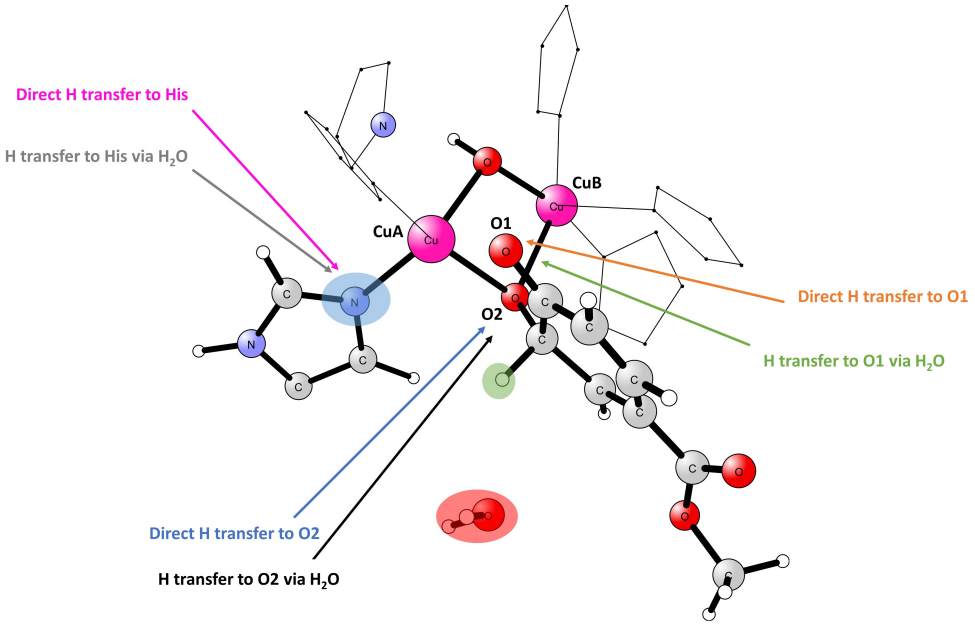


Figure 29: Sp^3 intermediate (**3O**, **Paper III**) with plausible proton transfer pathways to obtain catecholate substrate. *ortho*-H, N from histidine, and mediated-water molecule are highlighted by green, blue, and red circle, respectively.

First, we have shown that H^+(sp^3) transfer to the O2 on the same sp^3 carbon (H···O2 distance = 2.05 Å) proceeds (somewhat expectedly, owing to the confined transition state 4-membered ring structure) *via* a high TS (**TS3_O2**; +28.1 kcal/mol). It results in the formation of **4a** (-51.1 kcal/mol, **Figure 30**, with the substrate coordinated to CuA). Second, direct H^+(sp^3) transfer to O1 (H···O1 distance = 2.80 Å) proceeds *via* a slightly higher barrier (**TS3_O1**; +35.2 kcal/mol, **Figure 30**) and results in the formation of **4b** (-53.0 kcal/mol, **Figure 30**, followed by CuB coordination). Both **4a** and **4b** represent equivalent coordination modes of the monocatecholate substrate to the μ-OH dicopper(II) site of met-Ty (i.e., a $\mu - \eta^1 : \eta^2$ coordination) but each with different hydroxyl(ate) groups serving as the bridging catecholate ligand.

To investigate whether the solvent molecule(s) can mediate a lower-barrier proton transfer pathway, we evaluated the transfer of H^+(sp^3) to a nearby water molecule that is present in the protein pocket. This is computed to proceed *via* a lower TS (+8.8 kcal/mol, **TS3a** in **Figure 30**) and resulted in hydronium ion near a catecholate/met-Ty intermediate (-16.1 kcal/mol, structure **3O'** in **Figure 30**) where subsequent H^+ transfer to O2 proceed barrierlessly *via* **TS3b_O2** (-6.1 kcal/mol, note that it is lower in energy than **3O'**) to form **4a**, or *via*

TS3b_O1 (+15.8 kcal/mol) to form **4b**. It can be concluded that the solvent-mediated formation of catechol is energetically feasible and it leads to the intermediate **4a**.

In addition, we evaluated one of the nearby histidine residues (His54, coordinated to CuA) as a mediator (proton shuttle) in the proton transfer step (H···N distance = 3.41 Å), previously proposed by Inoue et al.[39] Corresponding TS's of proton transfer to His54 (we evaluated both direct H-transfer: **TS3_H54**: +25.4 kcal/mol and with aforementioned water molecule assistance: **TS3b_H54**: +15.3 kcal/mol) were higher in energy than **TS3a** in water-mediated pathway (*cf.* **Figure 30**) to create **3O'**. Thus, we excluded this His-mediated proton transfer pathway as well as the aforementioned direct H-transfers to either of oxygens.

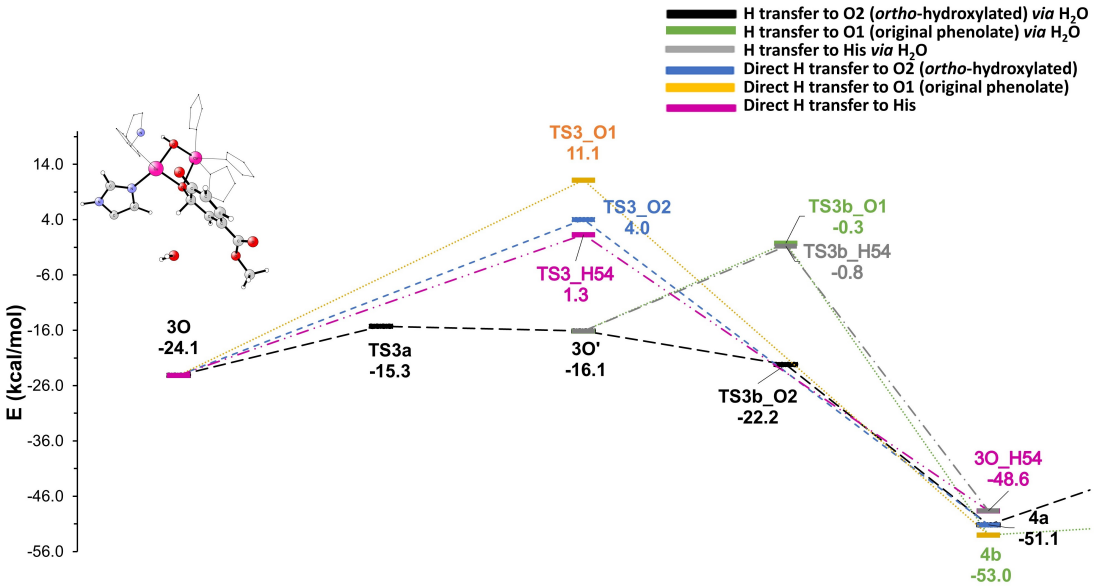


Figure 30: The calculated reaction coordinates for the $sp^3 \rightarrow sp^2$ conversion started from **3O**. Energies are calculated as single points upon the QM/MM-optimized geometries at the x2c-TPSSh-D3BJ/TZVPall level of theory in a homogeneous dielectric continuum with $\epsilon_r = 8$.

Notably, the plausible reaction coordinates converged to two structures for the catecholate-bound met-Ty intermediate that are of similar energy (**4a**: -51.1 kcal/mol *vs.* **4b**: -53.0 kcal/mol) and have similar electronic structure (i.e., two antiferromagnetically coupled Cu(II) ions bound to the monocatecholate). Moreover, the two coordination modes are also very similar (the difference is only the coordination *via* the O1 and O2, in **4a** and **4b**, respectively. In the intermediate **4a**, O1 coordinates to CuA, while in **4b**, O2 to CuB, *cf.* **Figure 31**).

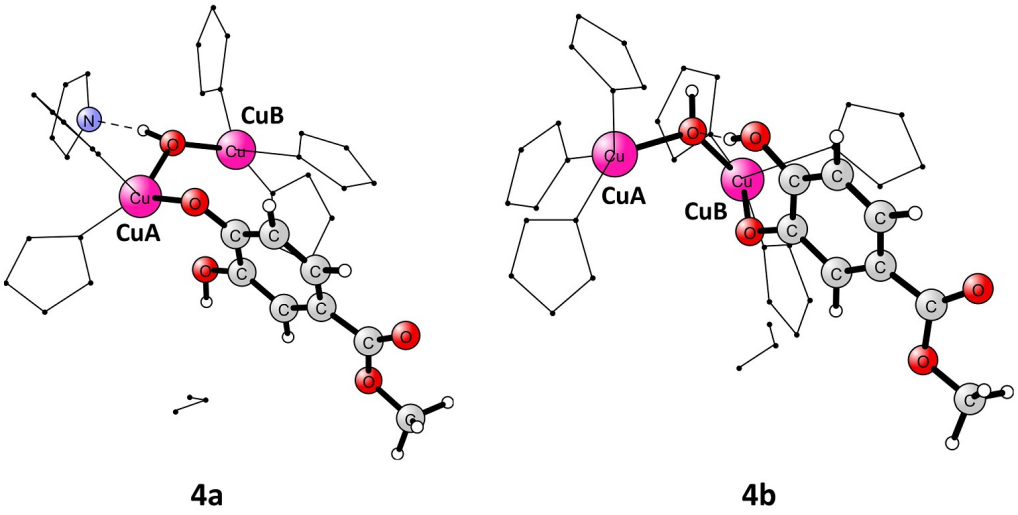


Figure 31: Comparison of possible met-Ty/catecholate intermediates: **4a** and **4b**.

Despite being close in energy, these two met-Ty/catecholate intermediates (**4a** and **4b**) exhibit distinct features in TD-DFT spectra (**Figure 32**). Peak corresponding to catecholate → Cu(II) charge transfer for species **4a** is found in lower-energy region (718 nm), whereas it is in higher-energy region (around 545 nm) for the structure **4b**. This difference is due to the dissociation of the axial histidine upon substrate biding to CuA (in **4a**) which resulted in four-ligand coordinated intermediate (*cf.* **Figure 31**), while this phenomenon is not observed for **4b** (CuB is penta-coordinated).

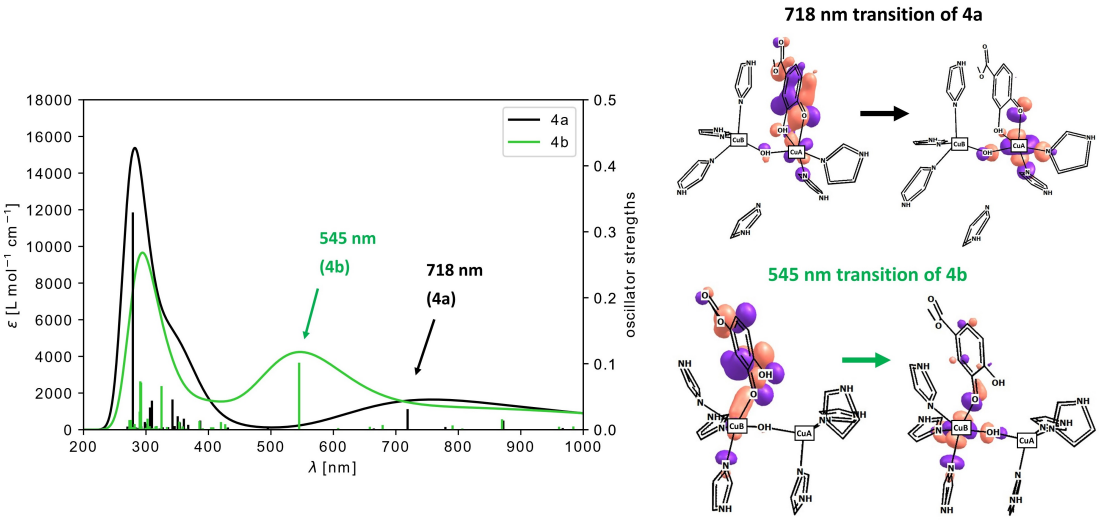


Figure 32: TD-DFT spectra of **4a** (black) and **4b** (green) and their corresponding natural transition orbitals (NTOs) for ligand (catecholate) → Cu(II) CT.

Correlation of these spectra to experimentally measured UV-Vis (*vide infra*) enabled us to exclude intermediate **4a** as met-Ty/catecholate complex and continue evaluating reaction coordinate from **4b**.

Still, we need to keep in mind that the water-assisted pathway from sp^3 intermediate (**3O**) to met-Ty/catecholate led directly to intermediate **4a**. Thus, we also evaluated the possibility of the interconversion between **4a** and **4b**

(to avoid high-energy **TS3b_O1**, +15.8 kcal/mol) by performing 1D reaction coordinate PES (with scanned distance of original *ortho*-H in structure **4a** to O1) from species **4a** to **4b**. The associated transition state was computed to have an energy barrier of +8.4 kcal/mol (**TS4a-b** in **Figure 33**), which is lower than the above mentioned **3O'** → **4b** reaction coordinate (**TS3b_O1**: +15.8 kcal/mol).

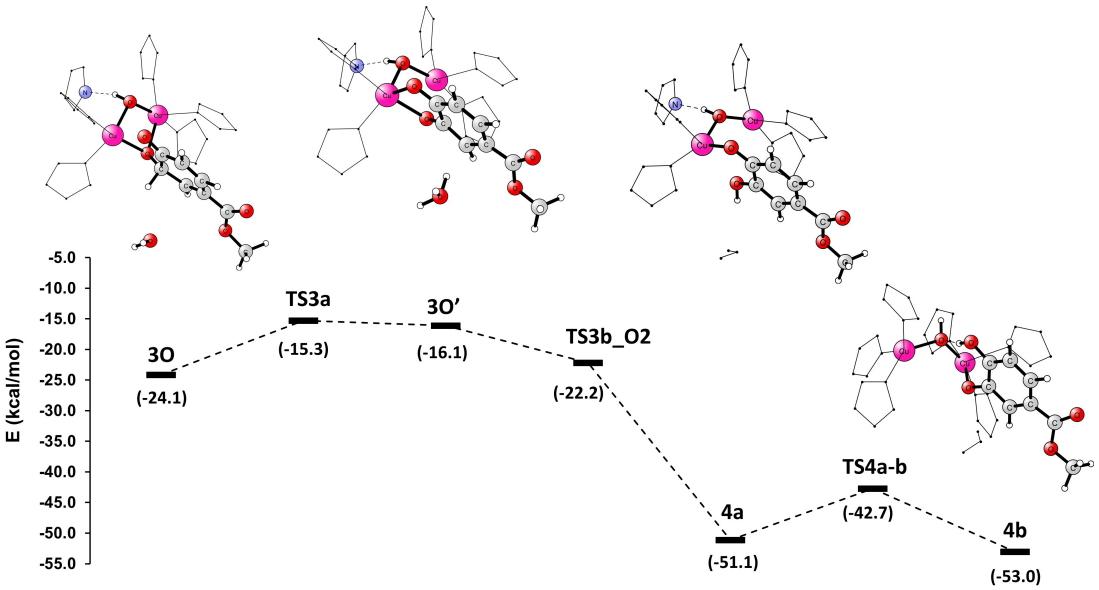


Figure 33: Postulated mechanism for the $sp^3 \rightarrow sp^2$ conversion of **3O** to met-Ty/catecholate intermediate that completes the Ty monooxygenation cycle. Energies are calculated as single points on top of the QM/MM-optimized geometries at the x2c-TPSSh-D3BJ/TZVPall level of theory in a homogeneous dielectric continuum with $\epsilon_r = 8$. Structures of **3O**, **3O'**, **4a**, and **4b** intermediates are depicted in the figure.

Catechol(ate) oxidation by met-Ty: preliminary experimental data obtained for 4-$COOCH_3$ *para*-substituted substrate

To obtain experimental insight into the catechol oxidation reaction (and to reveal the structure of putative intermediates) in the monophenolase catalytic cycle of Ty, the experimental efforts carried out in the laboratory of our collaborator, Prof. Solomon (Stanford University) focused on the reaction of met-Ty with methyl 3,4-dihydroxybenzoate (further denoted 3,4-$COOCH_3$) under anaerobic conditions at pH 9.0. We hypothesized that under these specific conditions, the reaction of met-Ty with 3,4-$COOCH_3$ would allow us to probe the catechol oxidation reaction of the monophenolase cycle. One factor supporting this hypothesis was that the free 3,4-$COOCH_3$ substrate ($pK_a = 8.2$) exists primarily in its monoanionic state under the basic reaction conditions (pH=9.0) and thus, its binding to met-Ty could form the short-lived met-Ty/catecholate intermediate (note that structures **4a** and **4b** discussed above are the met-Ty and monocatecholate complexes).

Preliminary experimental data (SF-Abs results) show the transient formation of a reaction intermediate with distinct absorption features, including an intense 400 nm band. The dependence of the formation/decay of the intense 400 nm

feature on the 3,4-$COOCH_3$ concentration allowed us to obtain the rate constants for its formation/decay. This also allowed us to probe the reaction in D_2O solvent. Notably, no solvent kinetic isotope effect for the formation/decay of the met-Ty/catecholate intermediate has been observed, suggesting that the exchangeable proton of its hydroxyl group does not participate before or at the rate-limiting transition states of the formation/decay steps. Experimental rate constant measured for the decay of the intermediate (leading to deoxy-Ty + quinone) was $k_2 = 0.4\ s^{-1}$, which corresponds to $\Delta G^{\ddagger} \sim +17$ kcal/mol. Performing the reaction at pH 7.0 (phosphate buffer) did not show accumulation of the 400 nm intermediate, suggesting that met-Ty indeed catalyzes the conversion of monoionic substrate.

We then proceeded to cryogenically trap this intermediate and spectroscopically characterized it by EPR, rR, and XAS (work done at Stanford and SLAC). The intermediate is EPR silent, which excludes the possibility that the 400 nm feature corresponds to a free semiquinone byproduct from partial oxidation of the 3,4-$COOCH_3$ substrate. It is also consistent with the XAS data which suggests this 400 nm intermediate to be met-Ty (both ions in the (+2) oxidation state) with catecholate substrate. We were not able to obtain the rR spectrum of this intermediate, due to the high background fluorescence from met-Ty.

All experiments will be repeated (work in progress), prior to publication of this joint computational and experimental material. Nevertheless, they have already provided a clue for computer modeling of the oxidation reaction.

Exploring catechol(ate) oxidation by met-Ty

We used structure **4b** (henceforth referred to as **4O**) as a starting and reference point (its energy set to 0.0 kcal/mol) for the catechol oxidation. We postulate that **4b** is the experimentally observed monocatecholate/met-Ty catalytic intermediate.

First, we evaluated the proton transfer pathway from its hydroxyl group (O1) to the bridging hydroxide (H···O distance = 1.52 Å). It resulted in the formation of semiquinone radical and Cu(I)Cu(II) oxidation state of the dicopper active center (structure **5** in **Figure 5.2.4**, +4.6 kcal/mol). The **TS4** associated with this proton transfer has been calculated at +8.8 kcal/mol (*via* 1D scan of the O···H coordinate). Based on the calculated spin densities along the reaction coordinate, we postulate that the studied proton transfer is accompanied by the electron transfer (i.e., it is PCET process) from the catechol substrate to the CuA site, resulting in the (+1) oxidation state of CuA (*cf.* Mulliken population analysis in **Table 7**). This is consistent with the calculated redox potential(s) for catechol substrate(s), where it is shown that for monocatecholate it is easier to transfer proton first (or simultaneously with electron, PCET) than electron (1.0 eV reduction potential of monoion *vs.* 0.3 eV calculated for diion, *cf.* **Figure 34**).

Table 7: Mulliken spin densities for selected atoms in for intermediates **4O**, **5**, and **TS4** calculated at the QM(B3LYP-D3BJ/def2-SVP)/MM(ff14SB) level of theory.

	4O	TS4	5
CuA	0.60	0.28	0.04
CuB	-0.56	-0.58	0.57
O (met)	0.05	-0.03	0.06
O2	-0.12	0.05	-0.13
O1	0.00	0.09	-0.17
Σ SUB	-0.15	0.43	-0.81

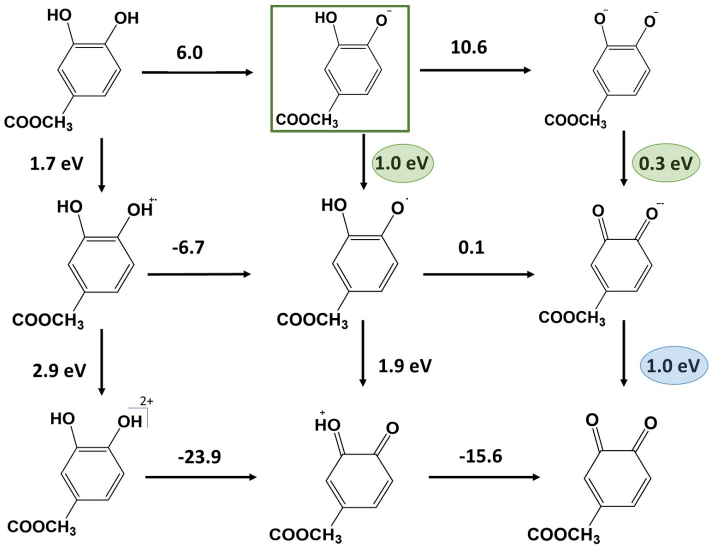


Figure 34: Thermochemistry of the methyl 3,4-dihydroxybenzoate/methyl 3,4-dioxo-1,5-cyclohexadiene-1-carboxylate system in water, with pK_a values above horizontal arrows, redox potential (in eV *vs.* standard hydrogen electrode (SHE)) besides vertical arrows. The monocatecholate substrate of the studied reaction is shown in the green rectangle, the values of the reduction potential of the 1^{st} electron transfer (discussed above) are marked with green ellipses, and of the 2^{nd} ET with a blue ellipse.

In the next step, starting from structure **5**, we performed 2D scan for releasing the newly formed water molecule and (semi)quinone from the active site, to obtain deoxy-Ty and quinone as final species, as well as the corresponding **TS5** related to 2^{nd} electron transfer. As the reaction coordinate we chose CuB – O2 distance and O_{met} – O_{W3}, where O_{met} is oxygen from newly formed water molecule (initial hydroxide of met-Ty) and O_{W3} stands for oxygen from the conserved and crystallographically-observed water molecule that is H-bonded to the Asn191 and Glu182 residues (*cf.* **Paper II**). This O_{W3} atom was fixed in place during the PES calculations and served as an anchor point to attract newly formed water molecule to detach it from the active site with the concomitant reduction of the copper site to deoxy-Ty state.

The obtained 2D-PES (**Figure 35**) exhibited a transition state (**TS5**) with energy barrier of +19.6 kcal/mol; quite comparable to the experimental value of $\sim$ 17 kcal/mol. It is associated with the 2^{nd} electron transfer to CuB, resulting in the formation of quinone and the reduction of the copper site (structure **6** in **Figure 35**, +18.4 kcal/mol). Monitoring both the substrate C-O bond distances and the electron spin density on the substrate and copper ions (**Figure 36**),

led us to the conclusion that water molecule loss is synchronized with product formation and release.

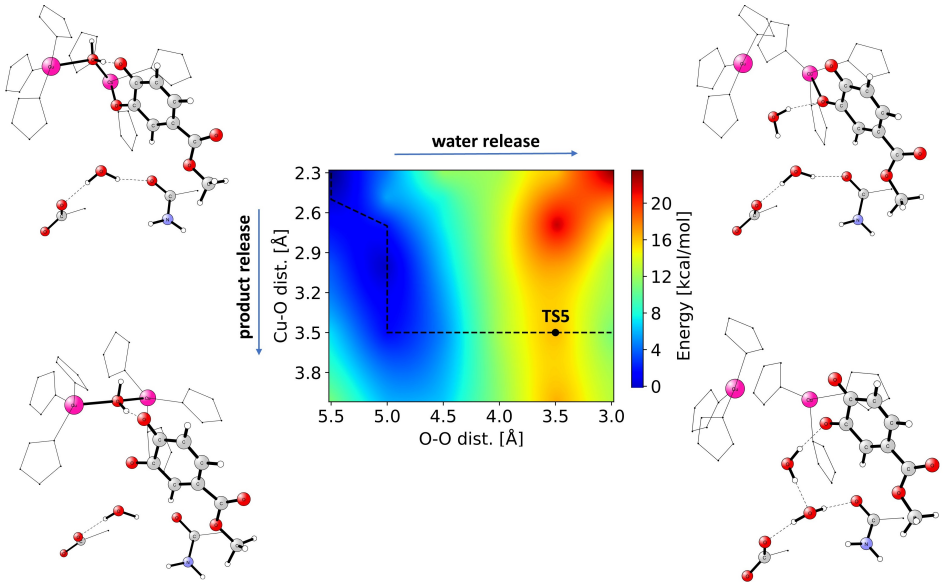


Figure 35: The 2D PES for water and quinone release for methyl 4-hydroxybenzoate (4-$COOCH_3$) performed at the QM(B3LYP-D3BJ/def2-SVP)/MM level of theory. The lowest-energy **5** → **6** path *via* **TS5** is shown as a dashed black line. Presented structures are from each corner of the 2D scan (top left: **5**; top right: released water, semiquinone bound to CuB; bottom left: semiquinone released, water bridged; bottom right: **6** – water, quinone, and deoxy-Ty.

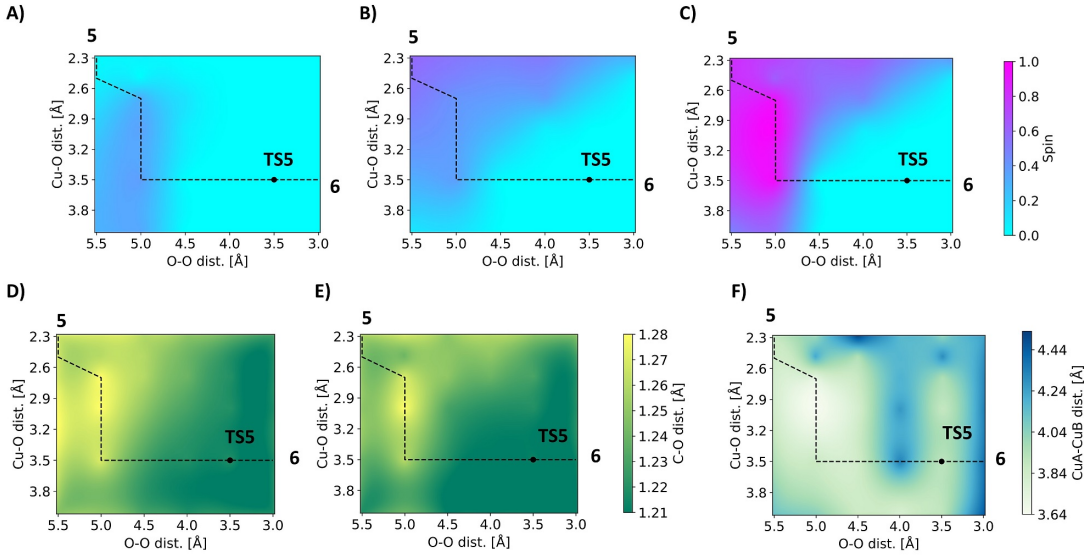


Figure 36: The electron spin density on CuA (A), CuB (B), and substrate/product (C), C-O (D, E), and CuA-CuB (F) bond distances extracted from the 2D PES for water and quinone release (**5** → **6**) for methyl 4-hydroxybenzoate (4-$COOCH_3$) performed at the QM(B3LYP-D3BJ/def2-SVP)/MM level of theory (**Figure 35**). Note that plots A), B), and C) have common scale, as well as D) and E). Mulliken spin density obtained after single-point energy calculations upon their QM/MM-optimized geometries at the x2c-B3LYP/SVPall level of theory in a homogeneous dielectric continuum with $\epsilon_r = 8$. Note that 2D PES was calculated by employing B3LYP functional, as it was shown that hybrids functionals with ∼ 20 % of HF exchange better captured copper ion electronic description than pure, that is crucial to monitor electron transfer.[128]

The formation of quinone is computed to be an endergonic process, however, releasing the product (with energy of -29.6 kcal/mol calculated by employing COSMO-RS protocol, + ΔG correction obtained from the rigid-rotor/harmonic oscillator approximation) from the active site and replacing it by four water molecules is (due to entropic effects) thermodynamically favorable, therefore the entire reaction is exergonic. Oxidation reaction mechanism with electronic energy and key intermediates is shown in **Figure 5.2.4**.

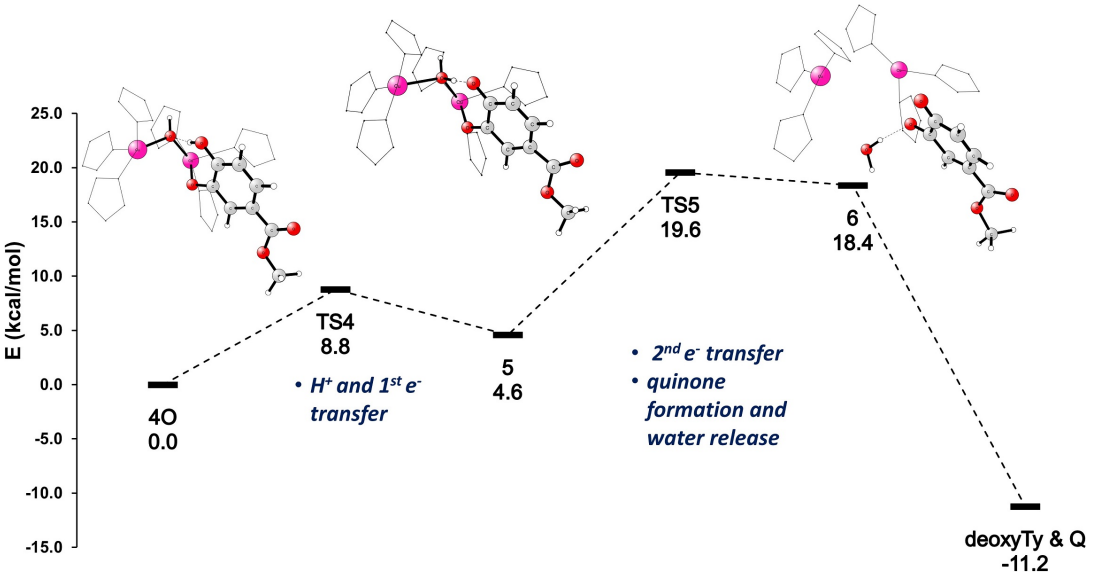


Figure 37: Reaction coordinate and structures of intermediates **4O**, **5**, and product **6** for catecholate → quinone reaction coordinate. Energies are calculated as single points on top of the QM/MM-optimized geometries at the x2c-M06-D3/TZVPall level of theory in a homogeneous dielectric continuum with $\epsilon_r = 8$. Note that final energies are reported with usage of M06 hybrid functional, as it resulted in the energy barrier (**TS5**) closest to experimentally obtained (among B3LYP, TPSSh, and LH14t-calPBE). The energy value of **DeoxyTy & Q** is the electronic energy value of **6** + aforementioned free energy of the product release.

To summarize our QM/MM-supported mechanism: after proton transfer (**4O** → **5**) from monocatecholate to μ-OH of the met-Ty dicopper site, CuA is reduced and likely to lose coordination with the newly formed water. The loss of water molecule (which increases the copper reduction potential due to the accommodation of the tetrahedral coordination geometry by CuB ion, i.e. more in favor of Cu(I) state) seems to be synchronous with the quinone formation, by triggering electron transfer. This is computed to be the RLS of catechol oxidation reaction. This is consistent with calculated reduction potential for deprotonated substrate (removal of the second electron is much higher in energy; 1.0 *vs.* 0.3 eV, in **Figure 34**), where it is shown that the second electron transfer is higher-in-energy step. Experimental data support this finding by lack of SKIE that would indicate PCET as RLS.

Summary of the Ty catalytic cycle

In the previous study (**Paper III**), we trapped and spectroscopically characterized the elusive catalytic ternary intermediate (Ty/O_2/monophenol) in the monooxygenation reaction of Ty. By correlation to QM/MM calculations we revealed the geometric and electronic structure of this intermediate, as well as demonstrated hydroxylated mechanism of Ty, distinct from previous proposals, and further validated monooxygenation by employing various monophenol substrate (**Paper IV**) confirming substrate deprotonation in initial steps of the reaction. In this study, we continued exploring tyrosinase catalytic cycle starting from sp^3 intermediate (**3O**). **3O** is one proton-transfer step (from the sp^3-carbon to the phenolic oxygen of the ensuing catechol ring) before the met-Ty/catecholate complex which is experimentally accessible by mixing met-Ty with catechol(ate). We demonstrated that catecholate formation in Ty likely occurs *via* the deprotonation of the sp^3 intermediate by external base (water molecule). Direct H-transfer to any of the two phenolic oxygens was highly unfavorable. Our finding is in agreement with previous proposals for this step, where exogenous base in form of methyl amine or acetate was probed in model system study, and the authors suggested involvement of solvent in the enzyme.[129] Matoba et al. indicated that it is a water molecule, but different from the conserved and crystallographically-observed water that is H-bonded to Asn191 and Glu182.[130] In contrary, Inoue et al., in their QM/MM study of the Ty reaction mechanism proposed His54 as the proton acceptor.[39] According to our computed data, this may have been an artifact of their small QM region. In summary, we were able to obtain met-Ty/catecholate-bound intermediate through the low energy pathway, which together with our experimental findings, is slightly different from previous proposals.

Our catalytic intermediate **4b**, with its characteristic absorption band at 400 nm, no observed solvent KIE of its formation/decay, and in the Cu(II)Cu(II)-catecholate state (XAS preliminary data), is distinct from previously reported both model and enzymatic complexes. For example, Opt'Holt et al. reported mechanism for the reaction of *N,N'*-di-tert-butyl-ethylenediamine (note that DBED is the bidentate ligand, while in enzymes copper ions are coordinated by three histidines) with 2,4-di-tert-buylphenolate *via* three intermediates.[129] One of them was proposed to be an asymmetrically bonded catecholate-dicopper(II) species, however, no characteristic feature around 400 nm was reported. On the other hand, Matoba et al. examined Ty in complex with the caddie protein (different from our fully catalytic caddie-free enzyme), and speculated about observed 400 nm species to be Cu(II)-semiquinone.[20,22] Our study, however, disproves this hypothesis.

With Cu(II)Cu(II)-catecholate (**4b**) intermediate as a starting point of the reaction, we provide new mechanistic insight into catechol oxidation reaction catalyzed by Ty. It consists of two steps: i) PCET to form Cu(I)Cu(II)-semiquinone complex, and ii) ET, being the RLS of the oxidation reaction as it involves the concerted loss of the newly formed water molecule with the formation and release of the quinone product. Our calculations showed that is much easier for substrate to transfer the first electron with proton, whereas the second electron transfer is the RLS due to higher electron potential (for 4-$COOCH_3$), which is confirmed experimentally by the lack of SKIE (that suggests that PCET is *not*

the RLS of the oxidation reaction).

An interesting consideration in the above reaction was copper ion to which substrate was coordinated. In our mechanism, catecholate dissociates from CuA to bind to CuB, and oxidation reaction proceeds. Some previous studies have argued for CuA site to be the one of substrate binding (for both mono- and diphenols),[131,132] whereas Inoue et al.[39] performed their whole QM/MM study (hydroxylation and oxidation reactions) *via* CuB. On the other site, past proposals suggested that those two reactions proceed with substrate coordination to distinct copper ions (i.e. monooxygenation *via* CuA, and oxidation *via* CuB),[133,134] which is confirmed in the present study. It could be one of plausible hypotheses to investigate further, as this can be the reason why CaOx cannot hydroxylate monophenols. They lack access to CuA (due to rigidity of CuA site caused by the Cys-His link presence) and steric hindrance (bulk gate residue), which blocks the catalytic site. Additional investigations are under way to determine the factor responsible for lack of monooxygenation activity among catechol oxidases.

Importantly, the results of this study not only reveal the catechol oxidation reaction carried by Ty, but they also address some important mechanistic considerations in the monophenolase cycle of Ty, which is the connection between phenol monooxygenation and catechol oxidation (with water molecule assistance as an exogenous base in sp^3-intermediate deprotonation step).

Overall, by identifying the catalytic 400 nm (met-Ty/catecholate) intermediate and defining its geometric and electronic structure, this study elucidates the key steps in the oxidation reaction of Ty. This mechanism completes the entire catalytic cycle of tyrosinase and is of a broad fundamental interest in catalysis of CBC-enzymes (and related biomimetics). It serves as the key reference for correlating the diverse reactivities among other members of the CBC protein family, including CaOxs, which can oxidize *o*-catechols to *o*-quinones but lack monooxygenase reactivity, and the recently discovered *o*-aminophenol oxidases, which are able to monooxygenate both monophenols to catechols and *o*-aminophenols to *o*-nitrosophenols.

6. Conclusions

In my PhD thesis, I showed how the complementarity and "strong correlation" of computational and experimental results can be used to understand the complex enzymatic cycle of the coupled binuclear copper enzyme, tyrosinase.

Starting from a comprehensive benchmarking study of small $[Cu_2O_2]$ models, we proposed guidelines for studying $[Cu_2O_2]$ systems, including CBC active sites. We showed that the UCCSD(T)/UKS level of theory can be considered as a reliable and conceptually simple choice compared to multireference WF methods and may serve as a reference for $[Cu_2O_2]$ systems. At the same time, we showed that the very cheap NR-TPSS/def2-SVP method might be an interesting alternative for larger $[Cu_2O_2]$ systems and/or for larger-scale production studies of CBC systems (**Paper I**).

Next, we explored the Ty catalytic cycle. To this aim, we structurally defined the oxy-Ty structure (**Paper II**). Our results showed that the most realistic model of oxy-Ty has approximately four water molecules occupying the substrate pocket of the catalytic center. Two of them form hydrogen bonds with the $\mu - \eta^2 : \eta^2$-peroxide. We proved that this water hydrogen-bonding network is the source of the (experimentally observed) Cu-O vibrational mode shift ($H_2O \rightarrow D_2O$).

In the following step, we characterized the elusive ternary intermediate **1P**, in which the substrate (phenol) binds in a fully protonated form and forms a hydrogen bond with the peroxide. This ternary intermediate (Michaelis complex) is the starting point for the hydroxylation reaction (**Paper III**). We provided fundamental insight into the key steps of the hydroxylation reaction mechanism catalyzed by Ty, where i) O-O bond cleavage of the peroxide followed by deprotonation of the substrate leads to the μ-oxo-μ-hydroxo intermediate (**2O**), and ii) the subsequent rate-limiting step corresponds to the concomitant coordination of monophenolate to CuA and *ortho*-C hydroxylation by the non-protonated bridging oxygen (with a non-negligible oxyl radical character).

We validated the proposed mechanism by employing a set of different substrates (**Paper IV**). We provided clear evidence for the transfer of phenolic H to the $\mu - \eta^2 : \eta^2$-peroxide and showed that the RLS of the Ty hydroxylation reaction may be altered by the monophenol *para*-group (EWG *vs.* EDG).

In the last part of the project, we attempted to provide new insight into the catechol oxidation reaction in the monophenolase catalytic cycle of Ty (**manuscript in preparation**). We demonstrated coupling between monooxygenation and oxidation reactions through a water-mediated proton transfer. We identified and characterized the catalytic catecholate/met-Ty intermediate and elucidated the oxidation reaction mechanism. We showed that it likely consists of two steps: i) PCET, which results in the formation of bridging water and the Cu(I)Cu(II)-semiquinone complex, and ii) ET which can be the RLS, where the release of water is accompanied by the semiquinone oxidation and release of the quinone product. We showed that Ty catalyzes the hydroxylation of monophenols *via* CuA, and the oxidation through CuB (which may be considered a factor responsible for the lack of monophenolase activity in CaOx's).

It is important to emphasize the importance of experimental validation of calculated predictions (and *vice versa*). The interplay between experiment and theory (sometimes led by one side, sometimes by the other) enabled us to learn and understand the catalytic cycle of Ty. It should also be highlighted that our efforts provided new information in the field of CBC enzymes and their biomimetics, distinct from previous proposals.

7. Impact of this work and outlook

CBC proteins, although they share the same structural motif in the active site, exhibit different functions. Hemocyanin is not catalytically active. Its role is to transport oxygen molecules *via* reversible binding of O_2. Catechol oxidase possesses only catecholase activity, while *o*-aminophenol oxidase can, in addition to monooxygenation and subsequent oxidation of phenols, catalyze the oxidation of *o*-aminophenols.

Unraveling the mechanism of the reactions carried out by tyrosinase is the first step towards understanding the differences between CBC proteins. Why do very similar biomolecules exhibit various reactivities? Why do they perform different reactions? Deep insight into the enzymatic processes carried out by one member of this protein family may help in understanding the others.

This study and the presented mechanistic observations may also be of interest for further development of small catalysts. For instance, inclusion of truncated CBC sites into various ligand scaffolds or into solids (e.g., zeolites), can be applied in industrial oxidation and defluorination processes.

Last but not least, this study may serve as an importance guidance in ongoing research and efforts aimed at finding effective tyrosinase inhibitors.

References

(1) Warshel, A. *Angewandte Chemie International Edition* **2014**, *53*, 10020–10031.

(2) Copeland, R. A. In *Enzymes*; John Wiley & Sons, Ltd: 2000.

(3) Engel, P., *Enzymes: A Very Short Introduction*; Oxford University Press: 2020.

(4) Murray, R. K.; Granner, D. K.; Mayes, P. A.; Rodwell, V. W., *Harper's Illustrated Biochemistry (Harper's Biochemistry)*, 27th ed.; McGraw-Hill Medical: 2006.

(5) Stryer, L. In *Biochemia*; Wydawnictwo Naukowe PWN: 1997.

(6) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chemical Reviews* **1996**, *96*, 2239–2314.

(7) Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147–157.

(8) Köhler, T.; Patsis, P. A.; Hahn, D.; Ruland, A.; Naas, C.; Müller, M.; Thiele, J. *ACS Omega* **2020**, *5*, 7059–7064.

(9) In Bertini, I.; Gray, H. B.; Lippard, S.; Valentine, J. *Bioinorganic Chemistry*; University Science Books: 1994.

(10) Holm, R. H.; Kennepohl, P.; Solomon, E. I. *Chemical Reviews* **1996**, *96*, 2239–2314.

(11) Koval, I. A.; Gamez, P.; Belle, C.; Selmeczi, K.; Reedijk, J. *Chemical Society Reviews* **2006**, *35*, 814–840.

(12) Kaintz, C.; Mauracher, S. G.; Rompel, A. In *Metal-Containing Enzymes*, Christov, C. Z., Ed.; Advances in Protein Chemistry and Structural Biology, Vol. 97; Academic Press: 2014, pp 1–35.

(13) Matoba, Y.; Kumagai, T.; Yamamoto, A.; Yoshitsu, H.; Sugiyama, M. *Journal of Biological Chemistry* **2006**, *281*, 8981–8990.

(14) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *Journal of Computational Chemistry* **2004**, *25*, 1605–1612.

(15) Kato, S.; Matsui, T.; Gatsogiannis, C.; Tanaka, Y. *Biophysical Reviews* **2018**, *10*, 191–202.

(16) Jiang, H.; Lai, W. *Organic & Biomolecular Chemistry* **2020**, *18*, 5192–5202.

(17) Solem, E.; Tuczek, F.; Decker, H. *Angewandte Chemie International Edition* **2016**, *55*, 2884–2888.

(18) Kampatsikas, I.; Rompel, A. *ChemBioChem* **2021**, *22*, 1161–1175.

(19) Prexler, S. M.; Frassek, M.; Moerschbacher, B. M.; Dirks-Hofmeister, M. E. *Angewandte Chemie International Edition* **2019**, *58*, 8757–8761.

(20) Matoba, Y.; Kihara, S.; Bando, N.; Yoshitsu, H.; Sakaguchi, M.; Kayama, K.; Yanagisawa, S.; Ogura, T.; Sugiyama, M. *PLOS Biology* **2019**, *16*, 1–22.

(21) Lerch, K.; Ettlinger, L. *European Journal of Biochemistry* **1972**, *31*, 427–437.

(22) Matoba, Y.; Kihara, S.; Muraki, Y.; Bando, N.; Yoshitsu, H.; Kuroda, T.; Sakaguchi, M.; Kayama, K.; Tai, H.; Hirota, S.; Ogura, T.; Sugiyama, M. *Biochemistry* **2017**, *56*, 5593–5603.

(23) Yamazaki, S.-i.; Itoh, S. *Journal of the American Chemical Society* **2003**, *125*, 13034–13035.

(24) Itoh, S.; Fukuzumi, S. *Accounts of Chemical Research* **2007**, *40*, 592–600.

(25) Jolley, R.; Evans, L.; Mason, H. *Biochemical and Biophysical Research Communications* **1972**, *46*, 878–884.

(26) Himmelwright, R. S.; Eickman, N. C.; LuBien, C. D.; Solomon, E. I.; Lerch, K. *Journal of the American Chemical Society* **1980**, *102*, 7339–7344.

(27) Ginsbach, J. W.; Kieber-Emmons, M. T.; Nomoto, R.; Noguchi, A.; Ohnishi, Y.; Solomon, E. I. *Proceedings of the National Academy of Sciences* **2012**, *109*, 10793–10797.

(28) Noguchi, A.; Kitamura, T.; Onaka, H.; Horinouchi, S.; Ohnishi, Y. *Nature Chemical Biology* **2010**, *6*, 641–643.

(29) Decker, H.; Schweikardt, T.; Nillius, D.; Salzbrunn, U.; Jaenicke, E.; Tuczek, F. *Gene* **2007**, *398*, 183–191.

(30) Penttinen, L.; Rutanen, C.; Saloheimo, M.; Kruus, K.; Rouvinen, J.; Hakulinen, N. *PLOS ONE* **2018**, *13*, 1–15.

(31) Henson, M. J.; Mahadevan, V.; Stack, T. D. P.; Solomon, E. I. *Inorganic Chemistry* **2001**, *40*, 5068–5069.

(32) Eickman, N. C.; Solomon, E. I.; Larrabee, J. A.; Spiro, T. G.; Lerch, K. *Journal of the American Chemical Society* **1978**, *100*, 6529–6531.

(33) Asanuma, M.; Miyazaki, I.; Ogawa, N. *Neurotoxicity Research* **2003**, *5*, 165–176.

(34) Güell, M.; Luis, J. M.; Solà, M.; Siegbahn, P. E. M. *JBIC Journal of Biological Inorganic Chemistry* **2009**, *14*, 229–242.

(35) Besalú-Sala, P.; Magallón, C.; Costas, M.; Company, A.; Luis, J. M. *Inorganic Chemistry* **2020**, *59*, 17018–17027.

(36) Kieber-Emmons, M. T.; Ginsbach, J. W.; Wick, P. K.; Lucas, H. R.; Helton, M. E.; Lucchese, B.; Suzuki, M.; Zuberbühler, A. D.; Karlin, K. D.; Solomon, E. I. *Angewandte Chemie International Edition* **2014**, *53*, 4935–4939.

(37) Qayyum, M. F.; Sarangi, R.; Fujisawa, K.; Stack, T. D. P.; Karlin, K. D.; Hodgson, K. O.; Hedman, B.; Solomon, E. I. *Journal of the American Chemical Society* **2013**, *135*, 17417–17431.

(38) Herres-Pawlis, S.; Verma, P.; Haase, R.; Kang, P.; Lyons, C. T.; Wasinger, E. C.; Flörke, U.; Henkel, G.; Stack, T. D. P. *Journal of the American Chemical Society* **2009**, *131*, 1154–1169.

(39) Inoue, T.; Shiota, Y.; Yoshizawa, K. *Journal of the American Chemical Society* **2008**, *130*, 16890–16897.

(40) Kipouros, I.; Stańczak, A.; Ginsbach, J. W.; Andrikopoulos, P. C.; Rulíšek, L.; Solomon, E. I. *Proceedings of the National Academy of Sciences* **2022**, *119*, e2205619119.

(41) Kipouros, I.; Stańczak, A.; Dunietz, E. M.; Ginsbach, J. W.; Srnec, M.; Rulíšek, L.; Solomon, E. I. *Journal of the American Chemical Society* **2023**, *145*, 22866–22870.

(42) Lind, T.; Siegbahn, P. E. M.; Crabtree, R. H. *The Journal of Physical Chemistry B* **1999**, *103*, 1193–1202.

(43) Siegbahn, P. E. M.; Wirstam, M. *Journal of the American Chemical Society* **2001**, *123*, 11819–11820.

(44) Siegbahn, P. E. M. *JBIC Journal of Biological Inorganic Chemistry* **2003**, *8*, 567–576.

(45) Klabunde, T.; Eicken, C.; Sacchettini, J. C.; Krebs, B. *Nature Structural Biology* **1998**, *5*, 1084–1090.

(46) Siegbahn, P. E. M. *JBIC Journal of Biological Inorganic Chemistry* **2004**, *9*, 577–590.

(47) Güell, M.; Siegbahn, P. E. M. *JBIC Journal of Biological Inorganic Chemistry* **2007**, *12*, 1251–1264.

(48) Deeth, R. J.; Diedrich, C. *JBIC Journal of Biological Inorganic Chemistry* **2010**, *15*, 117–129.

(49) Siegbahn, P. E. M.; Borowski, T. *Faraday Discussions* **2011**, *148*, 109–117.

(50) Jumper, J. et al. *Nature* **2021**, *596*, 583–589.

(51) Abramson, J. et al. *Nature* **2024**, *630*, 493–500.

(52) Bergmann, J.; Oksanen, E.; Ryde, U. *Current Opinion in Structural Biology* **2022**, *72*, 18–26.

(53) Himo, F. *Journal of the American Chemical Society* **2017**, *139*, 6780–6786.

(54) Siegbahn, P. E. M. *RSC Advances* **2021**, *11*, 3495–3508.

(55) Chen, S.-L.; Marino, T.; Fang, W.-H.; Russo, N.; Himo, F. *The Journal of Physical Chemistry B* **2008**, *112*, 2494–2500.

(56) Klusák, V.; Bařinka, C.; Plechanovová, A.; Mlčochová, P.; Konvalinka, J.; Rulíšek, L.; Lubkowski, J. *Biochemistry* **2009**, *48*, 4126–4138.

(57) Chaturvedi, S. S.; Bím, D.; Christov, C. Z.; Alexandrova, A. N. *Chemical Science* **2023**, *14*, 10997–11011.

(58) Senn, H. M.; Thiel, W. *Angewandte Chemie International Edition* **2009**, *48*, 1198–1229.

(59) Bím, D.; Navrátil, M.; Gutten, O.; Konvalinka, J.; Kutil, Z.; Culka, M.; Navrátil, V.; Alexandrova, A. N.; Bařinka, C.; Rulíšek, L. *The Journal of Physical Chemistry B* **2022**, *126*, 132–143.

(60) Hu, L.; Söderhjelm, P.; Ryde, U. *Journal of Chemical Theory and Computation* **2011**, *7*, 761–777.

(61) Neese, F.; Liakos, D. G.; Ye, S. *JBIC Journal of Biological Inorganic Chemistry* **2011**, *16*, 821–829.

(62) Ghosh, A. *JBIC Journal of Biological Inorganic Chemistry* **2011**, *16*, 819–820.

(63) Kurashige, Y.; Chalupský, J.; Lan, T. N.; Yanai, T. *The Journal of Chemical Physics* **2014**, *141*, 174111.

(64) Marti, K. H.; Ondík, I. M.; Moritz, G.; Reiher, M. *The Journal of Chemical Physics* **2008**, *128*, 014104.

(65) Shamasundar, K. R.; Knizia, G.; Werner, H.-J. *The Journal of Chemical Physics* **2011**, *135*, 054101.

(66) Cramer, C. J.; Włoch, M.; Piecuch, P.; Puzzarini, C.; Gagliardi, L. *The Journal of Physical Chemistry A* **2006**, *110*, 1991–2004.

(67) Lewin, J. L.; Heppner, D. E.; Cramer, C. J. *JBIC Journal of Biological Inorganic Chemistry* **2007**, *12*, 1221–1234.

(68) Liakos, D. G.; Neese, F. *Journal of Chemical Theory and Computation* **2011**, *7*, 1511–1523.

(69) Phung, Q. M.; Wouters, S.; Pierloot, K. *Journal of Chemical Theory and Computation* **2016**, *12*, 4352–4361.

(70) Langhoff, S. R.; Davidson, E. R. *International Journal of Quantum Chemistry* **1974**, *8*, 61–72.

(71) Harvey, J. N. *JBIC Journal of Biological Inorganic Chemistry* **2011**, *16*, 831–839.

(72) Bartlett, R. J. *The Journal of Physical Chemistry* **1989**, *93*, 1697–1708.

(73) Harvey, J. N.; Aschi, M. *Faraday Discussions* **2003**, *124*, 129–143.

(74) Petit, A. S.; Pennifold, R. C. R.; Harvey, J. N. *Inorganic Chemistry* **2014**, *53*, 6473–6481.

(75) Phung, Q. M.; Martín-Fernández, C.; Harvey, J. N.; Feldt, M. *Journal of Chemical Theory and Computation* **2019**, *15*, 4297–4304.

(76) Witte, M.; Herres-Pawlis, S. *Physical Chemistry Chemical Physics* **2017**, *19*, 26880–26889.

(77) Peng, D.; Reiher, M. *Theoretical Chemistry Accounts* **2012**, *131*, 1081.

(78) Stańczak, A.; Chalupský, J.; Rulíšek, L.; Straka, M. *ChemPhysChem* **2022**, *23*, e202200076.

(79) Noodleman, L. *The Journal of Chemical Physics* **1981**, *74*, 5737–5743.

(80) Noodleman, L.; Davidson, E. R. *Chemical Physics* **1986**, *109*, 131–143.

(81) Chalupský, J.; Srnec, M.; Yanai, T. *The Journal of Physical Chemistry Letters* **2021**, *12*, 1268–1274.

(82) Dai, D.; Whangbo, M.-H. *The Journal of Chemical Physics* **2003**, *118*, 29–39.

(83) Pantazis, D. A. *Journal of Chemical Theory and Computation* **2019**, *15*, 938–948.

(84) Massolle, A.; Neugebauer, J. *Faraday Discussions* **2020**, *224*, 201–226.

(85) Brandenburg, J. G. et al. *Faraday Discussions* **2020**, *224*, 309–332.

(86) Nishino, M.; Yamanaka, S.; Yoshioka, Y.; Yamaguchi, K. *The Journal of Physical Chemistry A* **1997**, *101*, 705–712.

(87) Ruiz, E.; Cano, J.; Alvarez, S.; Alemany, P. *Journal of Computational Chemistry* **1999**, *20*, 1391–1400.

(88) Singh, G.; Gamboa, S.; Orio, M.; Pantazis, D. A.; Roemelt, M. *Theoretical Chemistry Accounts* **2021**, *140*, 139.

(89) Dirac, P. A. M.; Fowler, R. H. *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **1929**, *123*, 714–733.

(90) Slater, J. C. *Physical Review* **1951**, *81*, 385–390.

(91) Becke, A. D. *Physical Review A* **1988**, *38*, 3098–3100.

(92) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.

(93) Vosko, S. H.; Wilk, L.; Nusair, M. *Canadian Journal of Physics* **1980**, *58*, 1200–1211.

(94) Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785–789.

(95) Siegbahn, P. E. M.; Blomberg, M. R. A. *Frontiers in Chemistry* **2018**, *6*, DOI: 10.3389/fchem.2018.00644.

(96) Perdew, J. P.; Wang, Y. *Physical Review B* **1992**, *45*, 13244–13249.

(97) Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Physical Review Letters* **2003**, *91*, 146401.

(98) Weigend, F.; Ahlrichs, R. *Physical Chemistry Chemical Physics* **2005**, *7*, 3297–3305.

(99) Staroverov, V. N.; Scuseria, G. E.; Tao, J.; Perdew, J. P. *The Journal of Chemical Physics* **2003**, *119*, 12129–12137.

(100) Kipouros, I.; Stańczak, A.; Culka, M.; Andris, E.; Machonkin, T. R.; Rulíšek, L.; Solomon, E. I. *Chemical Communications* **2022**, *58*, 3913–3916.

(101) Paul, M. et al. *Chemistry – A European Journal* **2020**, *26*, 7556–7562.

(102) Dalhoff, R.; Schmidt, R.; Steeb, L.; Rabatinova, K.; Witte, M.; Teeuwen, S.; Benjamaâ, S.; Hüppe, H.; Hoffmann, A.; Herres-Pawlis, S. *Faraday Discussions* **2023**, *244*, 134–153.

(103) Gagliardi, L.; Truhlar, D. G.; Li Manni, G.; Carlson, R. K.; Hoyer, C. E.; Bao, J. L. *Accounts of Chemical Research* **2017**, *50*, 66–73.

(104) Li Manni, G.; Carlson, R. K.; Luo, S.; Ma, D.; Olsen, J.; Truhlar, D. G.; Gagliardi, L. *Journal of Chemical Theory and Computation* **2014**, *10*, 3669–3680.

(105) Carlson, R. K.; Truhlar, D. G.; Gagliardi, L. *Journal of Chemical Theory and Computation* **2015**, *11*, 4077–4085.

(106) Klamt, A.; Schüürmann, G. *Journal of the Chemical Society, Perkin Transactions 2* **1993**, 799–805.

(107) Rohrmüller, M.; Herres-Pawlis, S.; Witte, M.; Schmidt, W. G. *Journal of Computational Chemistry* **2013**, *34*, 1035–1045.

(108) Herres-Pawlis, S.; Haase, R.; Verma, P.; Hoffmann, A.; Kang, P.; Stack, T. D. P. *European Journal of Inorganic Chemistry* **2015**, *2015*, 5426–5436.

(109) Hoffmann, A.; Herres-Pawlis, S. *Physical Chemistry Chemical Physics* **2016**, *18*, 6430–6440.

(110) Rode, M. F.; Werner, H.-J. *Theoretical Chemistry Accounts* **2005**, *114*, 309–317.

(111) Phung, Q. M.; Feldt, M.; Harvey, J. N.; Pierloot, K. *Journal of Chemical Theory and Computation* **2018**, *14*, 2446–2455.

(112) Radoń, M. *Physical Chemistry Chemical Physics* **2019**, *21*, 4854–4870.

(113) Radoń, M. In *Computational Chemistry*, van Eldik, R., Puchta, R., Eds.; Advances in Inorganic Chemistry, Vol. 73; Academic Press: 2019, pp 221–264.

(114) Drabik, G.; Szklarzewicz, J.; Radoń, M. *Physical Chemistry Chemical Physics* **2021**, *23*, 151–172.

(115) Ryde, U. *Journal of Computer-Aided Molecular Design* **1996**, *10*, 153–164.

(116) Ryde, U.; Olsson, M. H. M. *International Journal of Quantum Chemistry* **2001**, *81*, 335–347.

(117) TURBOMOLE V7.4.1 2017, a development of University of Karlsruhe and Forschungszentrum Karlsruhe GmbH, 1989-2007, TURBOMOLE GmbH, since 2007; available from `http://www.turbomole.com`.

(118) Case, D.A., and Babin, V., and Berryman, J.T., and Betz, R.M., and Cai, Q., and Cerutti, D.S., and Cheatham, T.E. III, and Darden, T. A., and Duke, R.E., and Gohlke, H., and Goetz, A.W., and Gusarov, S., and Homeyer, N., and Janowski, P., and Kaus, J., and Kolossváry, I., and Kovalenko, A., and Lee, T.S., and LeGrand, S., and Luchko, T., and Luo, R., and Madej, B., and Merz, K.M., and Paesani, F., and Roe, D.R., and Roitberg, A., and Sagui, C., and Salomon-Ferrer, R., and Seabra, G., and Simmerling, C.L., and Smith, W., and Swails, J., and Walker, R.C., and Wang, J., and Wolf, R.M., and Wu, X., and Kollman, P.A (2014), AMBER 14, University of California, San Francisco.

(119) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *The Journal of Chemical Physics* **2010**, *132*, 154104.

(120) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. *Journal of Chemical Theory and Computation* **2015**, *11*, 3696–3713.

(121) Cao, L.; Ryde, U. *Frontiers in Chemistry* **2018**, *6*, DOI: 10.3389/fchem.2018.00089.

(122) Rulíšek, L. *The Journal of Physical Chemistry C* **2013**, *117*, 16871–16877.

(123) Gutten, O.; Rulíšek, L. *Inorganic Chemistry* **2013**, *52*, 10347–10355.

(124) Smola, M.; Gutten, O.; Dejmek, M.; Kožíšek, M.; Evangelidis, T.; Tehrani, Z. A.; Novotná, B.; Nencka, R.; Birkuš, G.; Rulíšek, L.; Boura, E. *Angewandte Chemie International Edition* **2021**, *60*, 10172–10178.

(125) Mohammed, T. A.; Meier, C. M.; Kalvoda, T.; Kalt, M.; Rulíšek, L.; Shoshan, M. S. *Angewandte Chemie International Edition* **2021**, *60*, 12381–12385.

(126) Štěpánová, S.; Andris, E.; Gutten, O.; Buděšínský, M.; Dejmek, M.; Břehová, P.; Rulíšek, L.; Kašička, V. *ELECTROPHORESIS* **2024**, *45*, 687–705.

(127) Muñoz-Muñoz, J. L.; Berna, J.; del Mar García-Molina, M.; Garcia-Molina, F.; Garcia-Ruiz, P. A.; Varon, R.; Rodriguez-Lopez, J. N.; Garcia-Canovas, F. *Biochemical and Biophysical Research Communications* **2012**, *424*, 228–233.

(128) Solomon, E. I.; Heppner, D. E.; Johnston, E. M.; Ginsbach, J. W.; Cirera, J.; Qayyum, M.; Kieber-Emmons, M. T.; Kjaergaard, C. H.; Hadt, R. G.; Tian, L. *Chemical Reviews* **2014**, *114*, 3659–3853.

(129) Op't Holt, B. T.; Vance, M. A.; Mirica, L. M.; Heppner, D. E.; Stack, T. D. P.; Solomon, E. I. *Journal of the American Chemical Society* **2009**, *131*, 6421–6438.

(130) Matoba, Y.; Oda, K.; Muraki, Y.; Masuda, T. *International Journal of Biological Macromolecules* **2021**, *183*, 1861–1870.

(131) Goldfeder, M.; Kanteev, M.; Isaschar-Ovdat, S.; Adir, N.; Fishman, A. *Nature Communications* **2014**, *5*, 4505.

(132) Fujieda, N.; Umakoshi, K.; Ochi, Y.; Nishikawa, Y.; Yanagisawa, S.; Kubo, M.; Kurisu, G.; Itoh, S. *Angewandte Chemie International Edition* **2020**, *59*, 13385–13390.

(133) Ramsden, C. A.; Riley, P. A. *Bioorganic & Medicinal Chemistry* **2014**, *22*, 2388–2395.

(134) Olivares, C.; García-Borrón, J. C.; Solano, F. *Biochemistry* **2002**, *41*, 679–686.