Electronic Supplementary Material (ESI) for Chemical Communications.

Supporting Information

Evidence for H-bonding interactions to the μ-η²:η²-peroxide of oxy-tyrosinase that activate its coupled binuclear copper active site

Ioannis Kipouros[a,†], Agnieszka Stańczak[b,c,†], Martin Culka[c], Erik Andris[c], Timothy R. Machonkin[d], Lubomír Rulíšek[b,*], Edward I. Solomon[a,e,*]

[a] *Department of Chemistry, Stanford University, Stanford, California 94305, United States*

[b] *Institute of Organic Chemistry and Biochemistry of the Czech Academy of Sciences, Flemingovo náměstí 2, 166 10, Praha 6, Czech Republic*

[c] *Faculty of Science, Charles University, Albertov 2038/6, 128 00 Praha 2, Czech Republic*

[d] *Department of Chemistry, Whitman College, Walla Walla, WA 99362, United States.*

[e] *Stanford Synchrotron Radiation Lightsource, SLAC National Accelerator Laboratory, Stanford University, Menlo Park, California 94025, United States*

Table of Contents:

1. Materials & Methods

1.1. Experimental Methods

1.1.1. Materials

Chemicals and buffers were obtained from commercial sources (Sigma) and used without further purification unless noted. Water (H_2O) was purified to a resistivity >17 MΩ cm-1 with a Barnstead Nanopure deionizer. Deuterated water (D_2O; 99.5% purity) was obtained from Cambridge Isotope Laboratories.

1.1.2. Cloning

The copper chaperone (MelC1) and tyrosinase (MelC2) gene sequences from Streptomyces glaucescens were cloned into the pETDuet-1 vector (Novagen) to generate the pMAC-1 vector. MelC1, was cloned using the restriction sites NcoI (forward; primer: 5' GTCCATGGCCGCCGGTCACCACCC 3') and HindIII (reverse; primer: 5' GCG AAGCTTCAGTTGGCGGGGAACGCC 3'). MelC2 was cloned using the restriction sites NdeI (forward; primer: 5' GCTTTA CATATG ACCGTCCGGAAG AACCAG 3') and XhoI (reverse; primer: 5' CTACTCGAG TCAGTCGGTGTCGAAGGTG 3'). All plasmids were isolated and purified using the GeneJET Plasmid MiniPrep and were either stored at -20°C or used to transform Rosetta™ (DE3)pLysS cells (EMD Millipore), which were then stored as cell stocks (50% glycerol) at -80°C for the next 6 months.

1.1.3. Heterologous Protein Expression and Purification

Rosetta™ (DE3)pLysS cells (EMD Millipore) transformed with pMAC-1 were transferred from a Luria-Bertani (LB) plate in the presence of chloramphenicol (34 µg/mL) and ampicillin (50 µg/mL) to 15 mL of liquid culture and grown overnight. These cultures were used to inoculate 1.0 L of sterile LB auto-induction media (10 g tryptone, 5 g yeast extract, 10 g NaCl, 34 µg/mL chloramphenicol, 50 µg/mL ampicillin, 1 mM $CuCl_2$, 1 mM $MgSO_4$, 1.32 g $(NH_4)_2SO_4$, 2.72 g KH_2PO_4, 2.84 g Na_2HPO_4, 0.5 g glucose, 2.0 g α-lactose monohydrate, 5 g glycerol, pH 7.25). Four 1.0 L cultures were grown for 36 hours at 25 °C, after which the cells were harvested via centrifugation (3,500 rpm for 15 min). The cell pellet was washed once with TRIS-HCl buffered saline (10 mM TRIS-HCl, 150 mM NaCl, 100 µM $CuCl_2$, pH 8.5). Cells were lysed via sonication after re-suspending the pellet in approximately 300 mL of TRIS-HCl buffered saline containing 20 µM phenylmethanesulfonyl fluoride. Nucleotides were digested with 10 µl of Benzonase® Nuclease (EMD Millipore, 25 U/µL) for 30 min at 4 °C before removal of cell debris via centrifugation. PEG 3350 was added (9% w/w) and mixed for 4 hours at 4 °C before solids were removed via centrifugation (15,000 rpm for 20 min). The soluble cell extract was dialyzed overnight in Buffer A (10 mM TRIS-HCl, 10% ethylene glycol, 100 µM $CuCl_2$, pH 8.5). This solution was loaded on a Q-Sepharose fast-flow resin column and eluded with Buffer B (1.0 M NaCl, 10 mM TRIS-HCl, 10% ethylene glycol, 100 µM $CuCl_2$, pH 8.5) gradient. This crude product was dialyzed overnight in Buffer A and frozen (yields of approximately 20 mg of Ty per L of culture).

1.1.4. Preparation of deoxy-Ty

Crude Ty was thawed and further purified with a HiTrap Q HP column (GE Healthcare/Cytvia) by elution with Buffer B. Purified Ty was immediately transferred to a Schlenk flask and stirred overnight at 4 °C and under a constant flow of water saturated nitrogen gas. Under a nitrogen atmosphere, resting Ty was transferred to either a Vivacell 70 concentrator body (PES, 5kDa MWCO) sealed in an airtight centrifuge tube under a positive pressure of nitrogen or to several Amicon ® Ultra centrifugal filters (15mL, 3KDa) sealed with parafilm. This solution was concentrated at 2,000g's to a volume of less than 2.0 mL and mixed hourly while concentrating. Total copper concentrations in the sample were measured by the biquinoline method.[1] The concentrated tyrosinase sample was transferred inside the glove-box and placed on the

cold plate (4 °C) and a fivefold excess of sodium hydrosulfite (with respect to total copper concentration) in Buffer C (0.1M CHES buffer in H_2O or D_2O, at pH/D 9.0) was added and allowed to react for 10 minutes, under slow stirring, to form deoxy-Ty. Buffer exchange of the deoxy-Ty solution with Buffer C was performed in the glove-box either by several re-concentration cycles (0.5mL Amicon® Ultra 3KDa MXCO centrifugal filters; 10,000g ,17 min/per cycle) each cycle followed by Buffer C addition, repeated until complete removal (>98%) of excess hydrosulfite and free copper, to ultimately generate a solution of deoxy-Ty in Buffer C. Concentration of deoxy-Ty in the final solution was measured by the 345nm absorbance of the oxy-Ty (ε_{345nm}=16,000 $M^{-1}cm^{-1}$, oxy-Ty generated by x20 dilution of deoxyTy to an O_2 saturated Buffer C at 4°C). The purity of the enzyme sample was assessed at various stages of the purification protocol by gel filtration electrophoresis (Mini-PROTEAN TGX precast gels) following the instructions provided by the manufacturer.

1.1.5. Preparation of oxy-Ty

Anaerobic solutions of deoxy-Ty ([deoxy-Ty]=1.5mM, 4°C, in Buffer C; 0.1M CHES in H_2O or D_2O, pH/D 9.0, 0.2mL) were mixed 1:1 (v/v) with O_2-saturated solutions of the same buffer system ([O_2]=2.0mM, 4°C) in a NMR tube and were frozen immediately (~3-5 sec) in liquid nitrogen to avoid oxyTy self-decay.[2] The samples were stored in liquid nitrogen dewars at 77K before used to collect the resonance Raman spectra.

1.1.6. Resonance Raman spectroscopy

Resonance Raman spectra were obtained using an Andor Newton CCD detector on a Spex 1877 CP triple monochromator with 1200, 1800, and 2400 grooves/mm holographic spectrograph gratings. Laser excitation at 351nm was provided by an Ar_+ ion (Innova Sabre 25/7) laser with 20 mW of incident power at the sample in a ~135° backscattering configuration. Resonance Raman spectra were collected for 20 minutes on samples cooled to 77 K in a liquid nitrogen finger Dewar (Wilmad) and rotated (~10 rotations/min) to avoid sample photodegradation. Baseline fluorescence was subtracted from all samples with SpectraGryph® (https://www.effemm2.de/spectragryph/) and relative intensities were normalized to the solvent/buffer vibrations. The corrected rR spectra of oxyTy were obtained by subtracting the rR spectra of buffer solutions. The vibrational frequencies of the key Cu_2O_2 modes in the corrected rR spectra of oxyTy, were determined by the mean values of gaussian fits in R-studio, and their standard deviations were estimated by independent replicates (N=3).

1.2. Computational Methods

1.2.1. Protein Setup

All reported calculations are based on the fully equilibrated all-atom QM/MM model (Figure S7) built from the crystal structure of oxy-tyrosinase from *Streptomyces castaneoglobisporous* obtained at the resolution of 1.8 Å (Protein Data Bank accession code 1WX2). Superposition of this model to the Zn-substituted tyrosinase crystal structure with its native substrate, L-tyrosine (L-Tyr), bound and fully-occupying its active site (PDB:4P6R) was used to define the space for the active site cavity. Since the ratio of molecular volumes of L-Tyr to H_2O is ~8:1, we generated the fully-hydrated form of our oxy-Ty model with 8 waters occupying its active site.

1.2.2. Equilibration

The protein was equilibrated using the following procedure:

(a) minimizing the positions of all hydrogen atoms in the initial structure,

(b) adding a solvation sphere with a radius of 35Å (3761 water molecules in total), and two sodium atoms to neutralize the system
(c) running a 1 ns simulated annealing molecular dynamics followed by the final minimization of the whole system (with all non-hydrogens atoms kept at their crystallographic positions throughout).

1.2.3. Charges / Protonation of Residues

Tyrosinase was then neutralized by addition of two Na^+ ions. We assumed the standard protonation states at pH 7 for all amino acids. For the histidine residues, the protonation status was assigned based on a careful inspection of the hydrogen bond network around the residue and the solvent accessibility; i.e., histidines 37, 53, 62, 179, 189, 193, and 215 were assumed to be protonated on the N_δ atom; histidines 214, 229, and 264 on the N_ε atom.

1.2.4. Molecular dynamics simulations of water dynamics

The same oxy-Ty model employed in the QM/MM calculations was used as a starting point for molecular dynamics simulations. The force field parameters were identical as for the QM/MM calculations. Due to inaccurate bonding parameters of Cu_2O_2 core, these copper and oxygen atoms, together with the first shell atoms coordinating the coppers (histidine nitrogens) were kept fixed during the simulation. The protein (including the sphere of water molecules used in the QM/MM calculations) was solvated in 76 x 77 x 76 Å box of water molecules using Amber's tleap tool. All MD simulations were performed using NAMD software.[3] The MD simulations used periodic boundary conditions to emulate bulk solvent. NPT scheme with Langevin temperature and pressure control was applied. Following equilibration protocol was used: (1) 1000 steps of conjugate gradient minimization with 25 $kcal.mol^{-1}Å^{-2}$ restraints on heavy atoms of the protein, (2) heating from 0 to 298 K followed by 10 ps of MD with above restraints, (3) 1000 steps of minimization without restraints, (4) heating from 0 to 298 K followed by 100 ps of unrestrained equilibration. We used 1 fs MD integration step for the above-described equilibration phase and 2 fs for the production phase, which lasted 0.5 µs. The trajectory was analyzed in vmd;[4] the date for mesh representation of water density was generated using VolMap plugin of vmd. Since the behavior of the solvent remained equilibrated throughout the whole 0.5 us simulation, only first 10 ns of the simulation were used for the water density analysis for the sake of computational efficiency. The final snapshots were visualized using PyMol.[5]

1.2.5. ComQum Program

For the QM/MM calculations, we used ComQum program[6,7] that combines Turbomole[8] and AMBER software[9] to couple quantum and molecular mechanics. The full system (protein and solvent) is divided into three subsystems. The System 1 (typically consisting of the protein active site) is treated at the quantum-mechanical level, System 2 includes all atoms of the amino acids and solvent molecules within a specified radius of any atom in the System 1 and is optimized by classical force-field methods, and System 3 comprises all of the remaining atoms, which are fixed at the original positions. Both System 2 and System 3 are therefore treated at the MM level, the only difference is that System 3 (outer part of the protein) is kept fixed. Because of that, there is no need to compute pair interactions (vdW and electrostatics) within the System 3 and these are constant. It reduces computational efforts in the MM part.

1.2.6. Subtractive Scheme

In ComQum, the subtractive scheme[10] is utilized, which can be expressed as:

$$E_{QM/MM} = E_{QM\text{-}pchg} + E_{MM123} - E_{MM1} \quad (Eq.1)$$

Therefore, to obtain the QM/MM energy, three separate calculations are performed on the model system: one QM calculation for the QM region that includes point charges from the MM region in one-electron part of the QM Hamiltonian (System 1; $E_{QM\text{-}pchg}$) and two MM calculations, (i) for the entire model system (Systems 1, 2, and 3; E_{MM123}) and (ii) for the QM region (E_{MM1}). The subtractive scheme ensures that no interactions are double counted and the terms from the junctions between System 1 and System 2 shall cancel out exactly. Note that ComQum thus uses a similar approach to the ONIOM method.

The System 1 is optimized using the sum of the quantum-chemical forces within System 1 and molecular-mechanical forces from System 2 onto System 1. Within each QM geometry-optimization cycle, a full molecular mechanics optimization is performed for the geometry of System 2.

1.2.7. QM/MM Boundary

The boundary between the QM and MM system is treated using the hydrogen link-atom approach; *i.e.*, the QM system is truncated by hydrogen link atoms (HL) replacing the corresponding carbon link atoms (CL) in the real system. The HL atom is covalently bound to a single Q_1 atom in the QM region and is placed along the Q_1-CL bond, as shown in Figure S1.

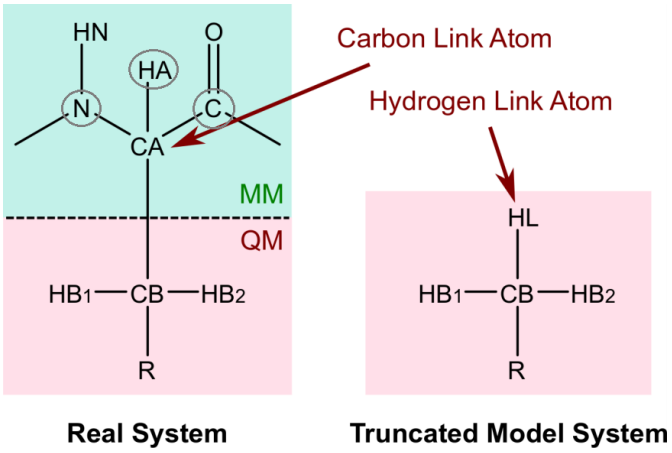


Figure S1. The junction between System 1 (QM region) and System 2 (MM region). The CA (Carbon Link Atom)/CL is replaced by HL (Hydrogen Link Atom) for the quantum-chemical calculation. Charges of atoms in the grey circles are zeroed.

1.2.8. Electrostatic Embedding

We utilize the electrostatic embedding QM/MM approach. In this approach, the QM energy is obtained from the QM calculation of System 1 embedded in the set of the MM point charges (i.e., the partial charges of the atoms of the Systems 2 and 3). At the same time, it excludes the self-energy of the point charges to avoid double-counting in the total E(QM/MM). Thereby, System 1 is polarized by Systems 2 and 3, but not vice versa. Standard charges for Systems 2 and 3 are taken from the AMBER force field,

whereas the charges fitted to the quantum chemical electrostatic potential (ESP charges) using the Merz–Kollman method are employed for atoms in System 1. To avoid double counting of the electrostatic interactions, charges of the QM system are set to zero for the MM energy term evaluations. In addition, the partial charges of the atoms bound to the link atom in the full system are zeroed in the quantum-chemical computations (c.f. Figure S1), and the remaining charges on the truncated amino acids are uniformly scaled to keep the fragment neutral. This avoids the overpolarization of the quantum system and ensures the conservation of the total charge of the enzyme, while allowing the charge transfer within the QM system.

1.2.9. Subsystems Definition

The System 1 (QM region) of oxy-Ty protein was selected to include 203-185 atoms, depending on the number of water molecules inside its active site cavity. To avoid artifacts in the electrostatic interactions, such as overpolarization of System 1, we have divided the System 1 and System 2 of the real system by only creating junctions between non-polar carbon atoms, whenever possible. Comprehensive description of the boundary between QM and MM regions for all of the residues (H37, I41, G52, H53, R54, H62, R177, E182, W183, H189, N190, H193, S205, F211, H214, H215) is included in the diagram in Figure S2. Additionally, the System 1 comprised two Cu ions, two oxygen atoms, and 14 water molecules. For obtaining information of oxy-Ty at different hydration levels, we added different number of waters (2-8) to fill the space within the protein pocket previously occupied by the substrate in our starting crystal structure (*vide supra*, SI section 1.2.1). This varying number of added waters (2-8) defines the active-site hydration level for our 64 QM/MM-optimized oxy-Ty structures (with additional water molecules of the solvent or at the active site boundary also present in System 1). The System 2 (MM region that is allowed to relax during the geometry optimization) comprised of additional 24 residues and water molecules within a radius of 2.0 Å from any atom of the QM region. All of the 24 residues in the System 2 were included as a whole.

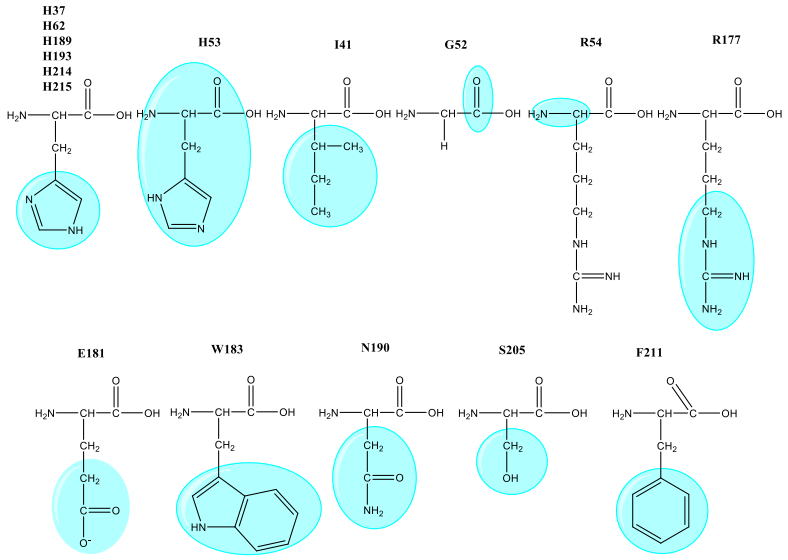


Figure S2: The boundary between QM and MM regions. Highlighted parts of the amino acids were included in System 1, while the remaining parts were treated by classical force field in System 2.

1.2.10. QM Calculations

All quantum chemical calculations, for the QM region in QM/MM and QM-only (cluster model), were performed by density functional theory (DFT) calcualtions using the Turbomole 6.6 (QM/MM) and Turbomole 7.4 (cluster model) program. Geometry optimizations were carried out using the TPSS functional with the def2-SVP basis set, including the empirical zero-damping dispersion correction (D3). DFT calculations were expedited by the RI-J approximation. The broken symmetry M_S = 0 spin state with the two antiferromagnetically-coupled copper ions was considered throughout. The MM calculations were performed employing the Amber ff14SB force field.[11] The sample oxy_1.pdb file with the full protein structure, including point charges used on all atoms, is provided as a separate supplementary material.[8]

1.2.11. Solvation (COSMO-RS) Calculations

By employing the QM parts of the QM/MM optimized structures (cluster model), in the same way as in the frequency calculations, we carried out a series of single point calculations to obtain accurate values of solvation energies which allowed us to (semi)quantitatively address partial dehydrations. Again, all DFT calculations were carried out using the Turbomole 7.4 program.[8] Geometries of the QM region were taken from the respective QMMM optimized structures. In the case of multiple QM/MM structures with the same number of water molecules in their active sites, only the global minimum was considered. Single point calculations employed the dispersion-corrected BP86 functional [BP86-D3(BJ)][12-15] and the def2-TZVPD[16] basis set (the default and recommended COSMO-RS protocol).

The free energy value corresponding to a particular structure/molecule can be then conveniently expressed as:

$$G_S = E_{el} + \Delta G_{solv} + E_{ZPVE} - RT\ln(q_{trans}q_{rot}q_{vib}) + pV \quad \text{(Eq. 2)},$$

where E_{el} is the electrostatic potential energy of the molecule *in vacuo* (gas-phase molecular energy), calculated at the BP86-D3(BJ)/def2-TZVPD level, ΔG_{solv} is the solvation energy calculated by employing the COSMO-RS method (*vide infra*), E_{ZPVE} is the zero-point vibrational energy whereas $RT\ln(q_{trans}q_{rot}q_{vib})$ are the entropic terms obtained from the rigid-rotor/harmonic oscillator (RRHO) approximation in which a free rotor model was applied for low-lying vibrational modes under 100 cm^{-1} with a smoothing function applied (sometimes denoted as quasi-RRHO, or RRFRHO approximation).

The ΔG_{solv} was obtained using Klamt's conductor-like screening model for realistic solvation method (COSMO-RS).[17] COSMO-RS calculations were carried out using *cosmotherm19* software with the parameter file "BP_TZVPD_FINE_C30_1901.ctd" and the recommended protocol: BP86-D3(BJ)/def2-TZVPD single point calculations *in vacuo* (on top of the QM/MM geometries) and in ideal conductor ($\varepsilon_r = \infty$), followed by the COSMO-RS (*cosmotherm19*)[18] calculations in the target solvent (water, 1-octanol and hexane – to mimic varying polarity of the protein site). Finally, a correction of $(1.9 \cdot \Delta n)$ kcal/mol (corresponding to the difference between the concentration of the ideal gas at 298 K and 1 atm and its 1 M concentration; Δn is the change in number of moles in the reaction) has been applied in order that the computed values refer to 1 M standard state.

1.2.12. Definition of an Incremental Hydration Energy.

The following definition of the incremental dehydration (Gibbs free) energy has been used:

$$\Delta G_{dehydration} = G_S \text{ (products, r.h.s of Eq. 4)} - G_S \text{ (reactants, l.h.s of Eq. 4)} \quad \text{(Eq. 3)}$$

The "reaction" used is the following:

$$[\text{oxy-Ty}_{\text{active-site}}.(H_2O)_n]^{c+} \rightarrow [\text{oxy-Ty}_{\text{active-site}}.(H_2O)_{n-1}]^{c+} + H_2O, \qquad \text{(Eq. 4)}$$

In our recent work, we have shown that the approach worked well for predicting the optimal hydration numbers of several divalent ions.[19] From our unpublished data, the current COSMO-RS protocol slightly favors the right-hand side of Equation 3, by approximately 3 kcal/mol. This empirical correction likely originates in the greater conformational entropy on the left hand-side of the Equation 3, which is not considered in the current protocol. Table S1 shows that each dehydration step from oxy_1 (8 waters) to oxy_48 (4 waters) is associated with a significant gain in thermodynamical stability, with any additional dehydrations (from oxy_48) not being favorable. The empirical correction of 3 kcal/mol mentioned above further supports this conclusion. Thus, oxy_48, with an active-site hydration level of 4 waters, seems to be the most computationally favorable model for oxy-Ty.

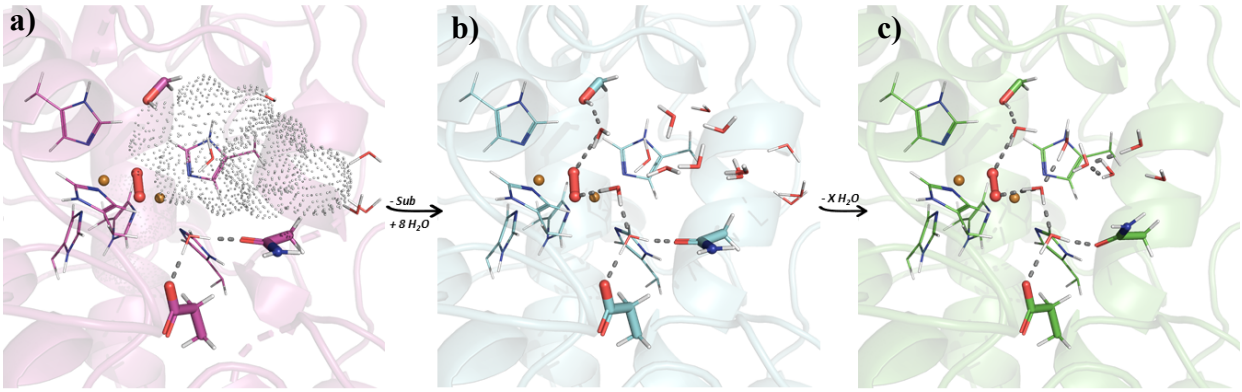


Figure S3. QM/MM optimized structures of oxy-Ty models with substrate shown as volume (a), eight (model oxy_1) (b) and two (model oxy_63) (c) water molecules in the place of the substrate in the active-site.

Table S1: Thermodynamic stability of oxy-Ty at various hydration levels. Free-energies (kcal/mol) of dehydration indicate the hydration level in the active site of oxy-Ty. Dehydration from eight water molecules to two.

Dehydration		H_2O	1-octanol	Diethyl ether
oxyTy_8_wat → oxyTy_7_wat + H_2O	**oxy_1 → oxy_5 + H_2O**	-2.5	-3.2	1.4
oxyTy_7_wat → oxyTy_6_wat + H_2O	**oxy_5 → oxy_9 + H_2O**	-18.2	-24.4	-34.1
oxyTy_6_wat → oxyTy_5_wat + H_2O	**oxy_9 → oxy_31 + H_2O**	-2.7	-2.4	-5.1
oxyTy_5_wat → oxyTy_4_wat + H_2O	**oxy_31 → oxy_48 + H_2O**	-24.2	-26.7	-28.0
oxyTy_4_wat → oxyTy_3_wat + H_2O	**oxy_48 → oxy_61 + H_2O**	-2.1	0.3	-0.6
oxyTy_3_wat → oxyTy_2_wat + H_2O	**oxy_61 → oxy_63 + H_2O**	0.0	0.2	-1.1

1.2.13. Frequency calculations

Vibrational frequency calculations (after QMMM) were performed for the QM region only. The junction atoms are considered to be frozen in place by a steep harmonic potential, using the TPSS functional with the def2-SVP basis set, including the empirical zero-damping dispersion correction (D3), in an implicit solvent with the dielectric constant of $\varepsilon_r = 8$ - employing the COSMO solvation model as implemented in Turbomole 7.4 program.[20] DFT calculations were expedited by the RI-J approximation.

1.2.14. Vibrational analysis.

Since the vibrational modes of the QM region were highly coupled (delocalized over the bigger parts of the QM region), we performed the vibrational analyses (of the normal modes computed as described in 1.2.13) with almost all atoms frozen, except for the $Cu(II)_2O_2$ core, water molecules and histidine ligands bound to both copper atoms. The freezing was performed by zeroing all elements of the vibrational Hessian containing the frozen atoms prior to its diagonalization and projected the $Cu(II)_2O_2$ modes from the coupled normal modes, as previously described.[21]

1.2.15. Frequency of the Cu-O (B_{3u}) mode

To correlate our calculation to the experimentally observed changes of the Cu-O(B_{3u}) resonance Raman mode upon solvent deuteration, the calculated frequency of the Cu-O (B_{3u}) mode was estimated as the weighted average of the frequencies of all vibrational modes with significant B_{3u} character. The selected frequencies were weighed by the magnitude of their largest Cu-O bond length change calculated by the XYZviewer program[22] (magnitude threshold set to 0.05).

1.2.16. Frequencies of the water twisting and rocking modes of W1 and W2

The energies of the unmixed twisting and rocking modes for waters in positions W1 and/or W2 were estimated by freezing all atoms in the QM region except those for the W1 and/or W2 in the QM/MM-optimized active site structure, and then performing QM-frequency analysis (as described in SI section 1.2.14).

2. Additional Supporting Information Tables and Figures.

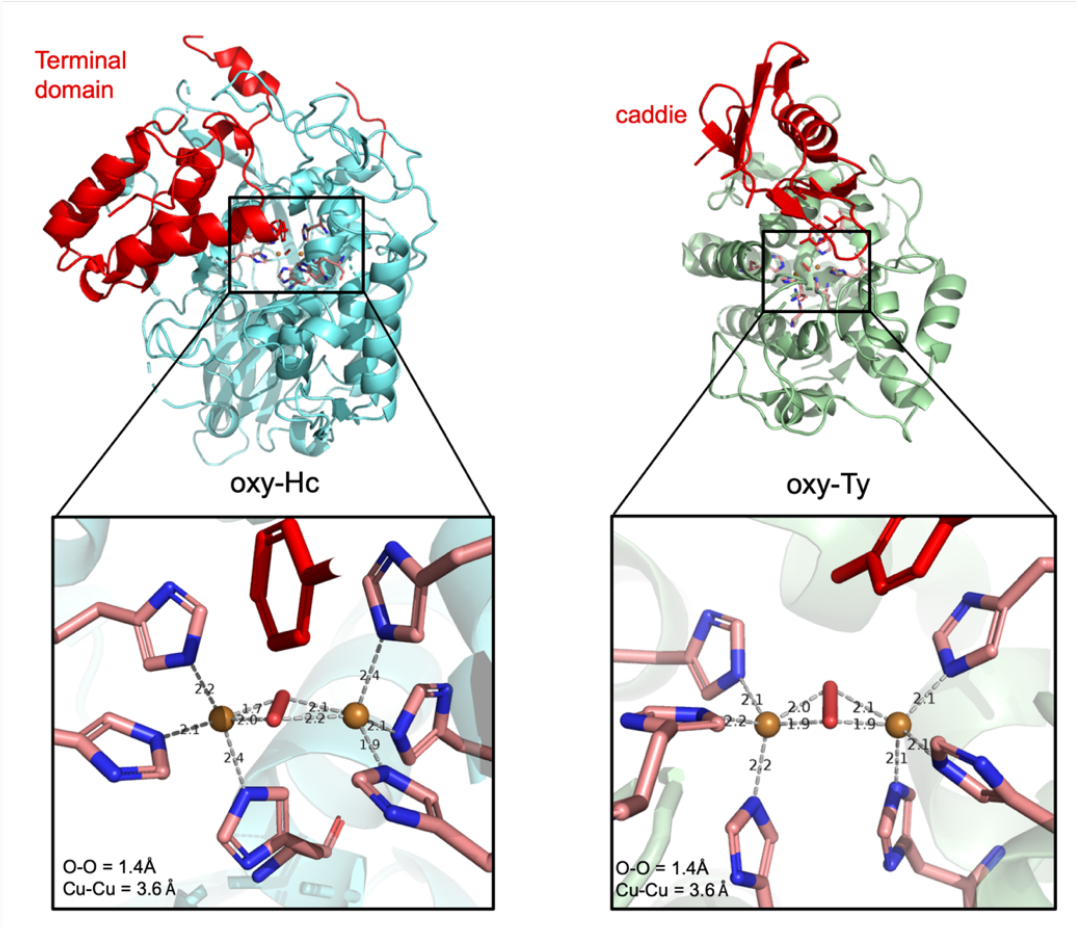


Figure S4. The active site of oxy-hemocyanin, and oxy-tyrosinase in complex with its caddie protein. The crystal structures of oxy-hemocyanin (*left*: oxy-Hc, PDB:1OXY) and oxy-tyrosinase with its caddie protein (*right*: oxy-Ty, PDB:1WX2), with inserts showing their respective active-site structures. The terminal domain of oxy-Hc and the caddie protein of oxy-Ty, both of which restrict access of phenolic substrates to the active site, are shown in red.

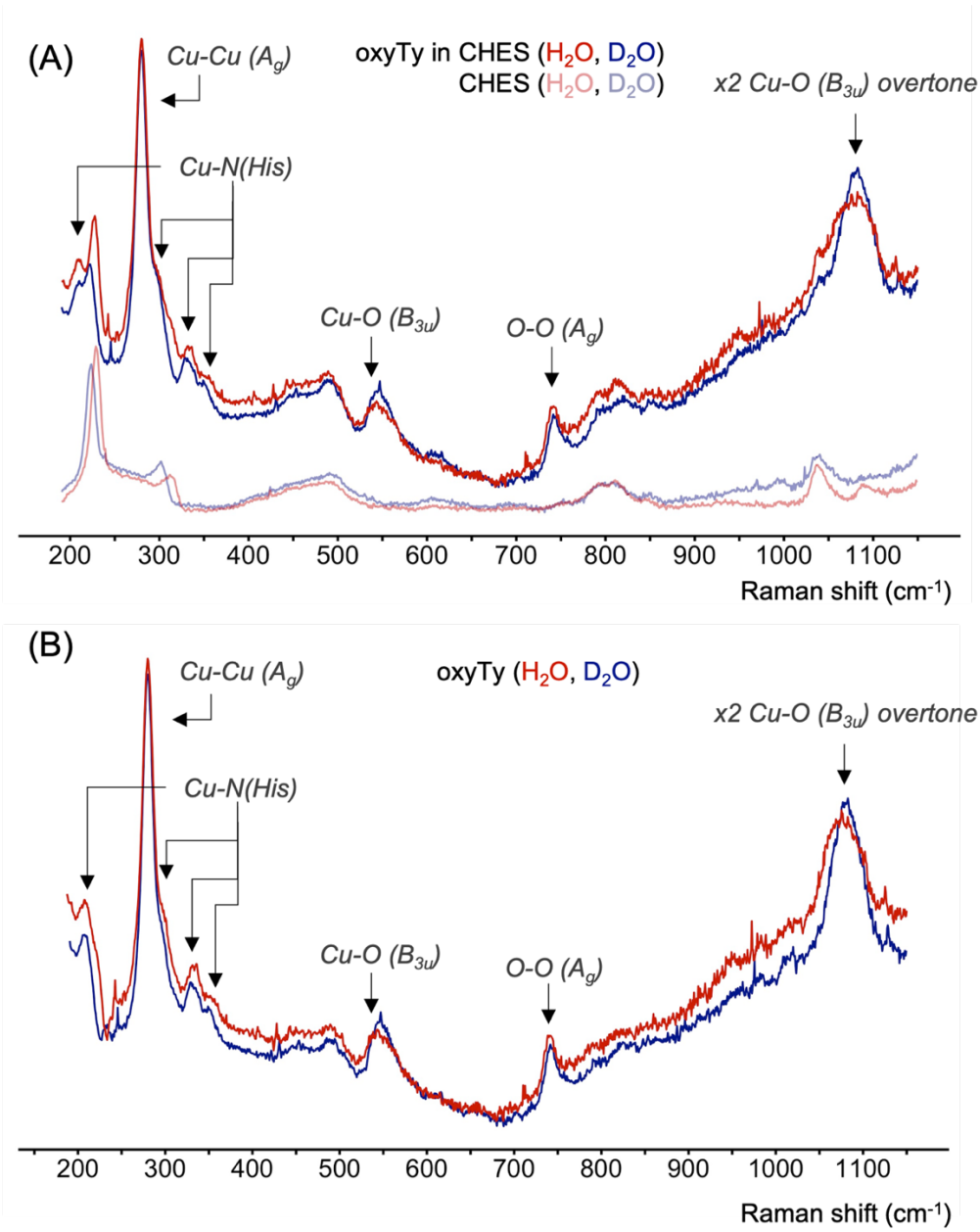
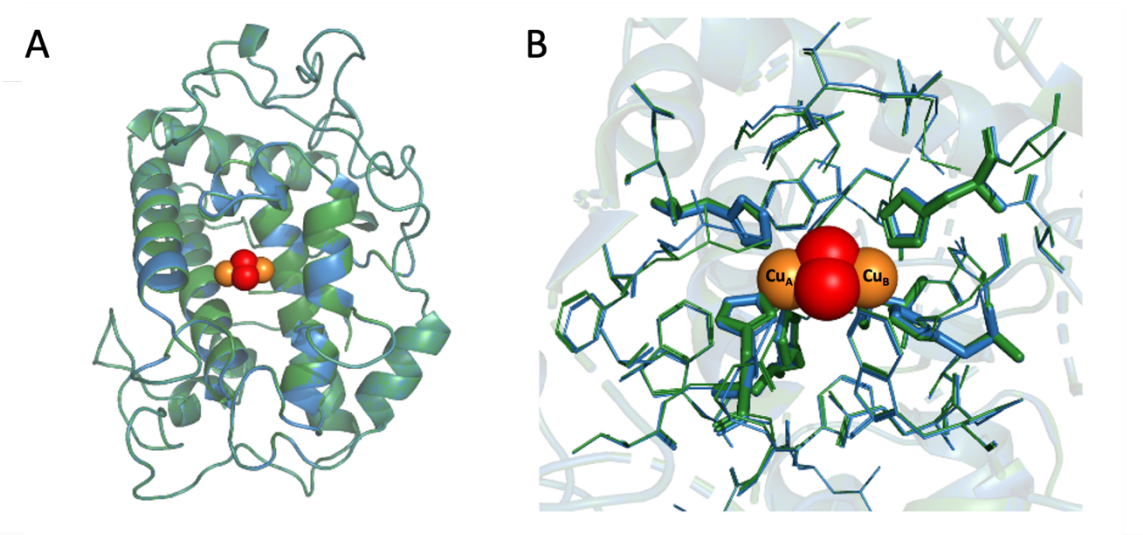


Figure S5. Resonance Raman spectra of oxy-Ty in H_2O vs D_2O, (A) with the contributions from CHES buffer, and (B) after spectral subtraction of buffer contributions. The rR spectra of these oxy-Ty samples ([deoxy-Ty]=0.75mM, [O_2]=1.2mM, 0.1M CHES, pH/D=9.0) were collected at 351nm laser excitation, 20mW, 77K and the background fluorescence contribution were removed with a linear baseline. For detailed analysis on the spectral features corresponding to the main Cu_2O_2 modes (indicated with arrows), see Figure 1 and the main text. Note also: (i) due to its weak intensity (non-totally symmetric, formally forbidden rR vibration) and the background contributions of the nearby higher energy peak (not originating from oxy-Ty or the buffer) it is challenging to estimate any differences of the peak intensity for the fundamental Cu-O(B_{3u}) mode for oxy-Ty in H_2O vs D_2O, and (ii) the baseline-corrected integrated peak area for the Cu-O(B_{3u}) overtone mode is the same between oxy-Ty in H_2O vs D_2O, however, in the former case, we observe some peak broadening that likely reflects a higher degree of conformational flexibility for the active site of oxy-Ty in H_2O than in D_2O.



C

STR_CAST.: Tyrosinase from Streptomyces Castaneoglobisporous (QM/MM model)
STR_GLAU.: Tyrosinase from Streptomyces Glaucescens (Experiments)

Score:487 bits(1254), Expect:0.0
Method: Compositional matrix adjust.

Identities active site: 35/35 (**100%**)
Identities overall: 233/272(**86%**)
Positives:248/272(**91%**)

```
STR_CAST.      1    MTVRKNQATLTADEKRRFVAAVLELKRSGRYDEFVRTHNEFIMSDTDSGERTGHRSPSFL  60
                    MTVRKNQATLTADEKRRFVAAVLELKRSGRYDEFV THN FI+ DTD+GERTGHRSPSFL
STR_GLAU.      1    MTVRKNQATLTADEKRRFVAAVLELKRSGRYDEFVTTHNAFIIGDTDAGERTGHRSPSFL  60

STR_CAST.      61   PWHRRFLLDFEQALQSVDSSVTLPYWDWSADRTVRASLWAPDFLGGTGRSTDGRVMDGPF  120
                    PWHRR+LL+FE+ALQSVD+SV LPYWDWSADRT RASLWAPDFLGGTGRS DGRVMDGPF
STR_GLAU.      61   PWHRRYLLEFERALQSVDASVALPYWDWSADRTARASLWAPDFLGGTGRSLDGRVMDGPF  120

STR_CAST.      121  AAFTGNWPINVRVDSRTYLRRSLGGSVAELPTRAEVESVLAISAYDLPPYNSASEGFRNH  180
                    AA  GNWPINVRVD R YLRRSLG +V ELPTRAEVESVL ++ YD  P+NSAS+GFRNH
STR_GLAU.      121  AASAGNWPINVRVDGRAYLRRSLGTAVRELPTRAEVESVLGMATYDTAPWNSASDGFRNH  180

STR_CAST.      181  LEGWRGVNLHNRVHVWVGGQMATGVSPNDPVFWLHHAYVDKLWAEWQRRHPDSAYVPTGG  240
                    LEGWRGVNLHNRVHVWVGGQMATG+SPNDPVFWLHHAYVDKLWAEWQRRHP S Y+P  G
STR_GLAU.      181  LEGWRGVNLHNRVHVWVGGQMATGMSPNDPVFWLHHAYVDKLWAEWQRRHPGSGYLPAAG  240

STR_CAST.      241  TPDVVDLNETMKPWNTVRPADLLDHTAYYTFD  272
                    TPDVVDLN+ MKPWN   PADLLDHTA+YTFD
STR_GLAU.      241  TPDVVDLNDRMKPWNDTSPADLLDHTAHYTFD  272
```

Figure S6. Comparison of two tyrosinase structures from *Streptomyces castaneoglobisporous* and *Streptomyces glaucescens*. (A) Overlay of the 1.8 Å crystal structure for oxy-Ty from *S. castaneoglobisporous* (green) used for generating the QM/MM model with the Ty homology model from *S. glaucescens* (blue) generated with iTasser. (B) Superposition of the two active sites with the Cu_2O_2 core shown as spheres, the 6 first-sphere coordination histidines as sticks, and the 29 second-sphere coordination residues (defined as residues within 8Å from Cu_2O_2) as lines. (C) Sequence alignment of the two tyrosinases with BLAST, with active site residues (all residues <8Å from Cu_2O_2 core) shown in red, and the 6 first-sphere coordination histidines also highlighted in yellow.

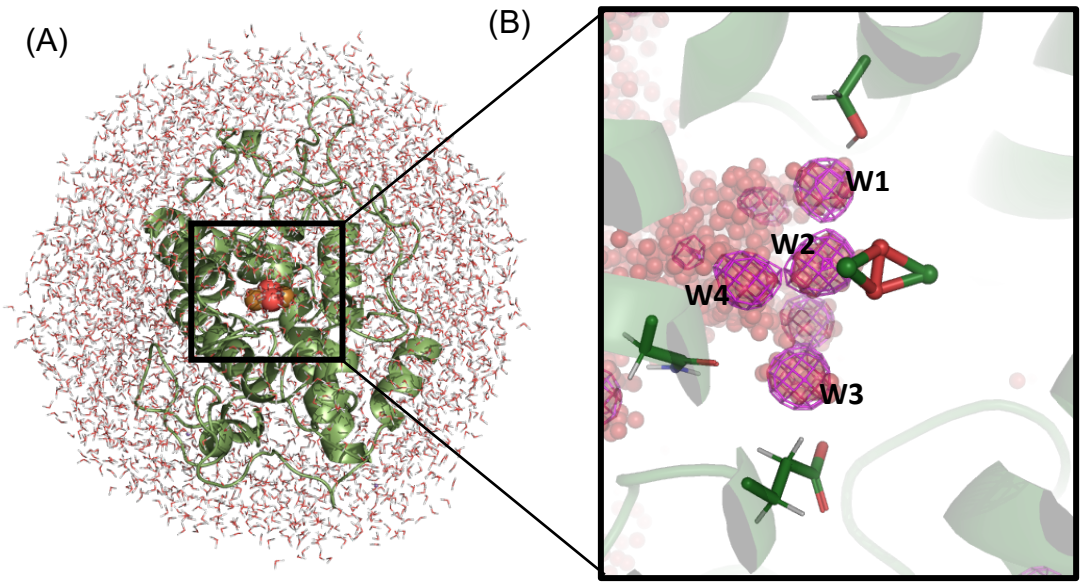


Figure S7. Model of oxy-tyrosinase employed for MD and QM/MM calculations, highlighting the water occupancy in the active site cavity. (A) The $Cu(II)_2O_2$ core is shown as spheres (Cu=brown, O=red), the protein backbone as green cartoon, and the water molecules as sticks. (B) The frames of water molecules from a 10 ns MD simulation segment (see SI MD video) are collapsed to visualize the conserved and distinct high-density water occupancy positions (indicated by mesh) for W1-3 in the vicinity of the Cu_2O_2 core (and the W4 position for an additional water molecule near the entrance to the protein pocket), versus the continuous and non-discreet distribution of water molecules in the bulk solvent outside the protein pocket (seen in the background). Note the view from inside the oxy-Ty protein (front and bottom left) and towards the active site protein pocket and bulk solvent (back and top right). Water molecules are shown as red spheres, the Cu centers as green spheres, and selected second-sphere residues as sticks. Also note that sporadic water occupancy in the active sites of CBC enzymes has been observed in previous crystallographic studies,[23,24] but these waters do not occupy the W1 and W2 positions and would not H-bond to the μ-η^2:η^2-peroxide.

Table S2. **Summary of QM/MM-optimized structures for the oxy-Ty active site at different hydration levels.** The relative energies at each hydration level (in kcal/mol, calculated by QM cluster single-point TPSS-def2-SVP upon QM/MM optimization), the number of water molecules in the active site (i.e. hydration level), the H-bonding interactions of the μ-η^2:η^2-peroxide with waters (W1, W2, W1+W2, or "-" indicating no water present in either position), the calculated frequencies changes upon solvent deuteration for the key Cu_2O_2 vibrational modes, and the method (manual construction or MD simulations) for generating each structure are also indicated. For the cartesian coordinates of all structures see Supporting file SI_xyz.zip.

Model ID	Rel. Energy (kcal/mol)	No. waters	H-bonding	Δν(D_2O-H_2O) (cm^{-1})			Method
				Cu-O (B_{3u})	Cu-Cu (A_g)	O-O (A_g)	
oxy_1	0.0	8	W1+W2	12	0	0	Manual
oxy_2	0.7	8	W1+W2	18	0	2	Manual
oxy_3	2.0	8	W2	-1	-2	2	Manual
oxy_4	4.4	8	W2	12	0	-1	Manual
oxy_5	0.0	7	W2	2	-1	-1	Manual
oxy_6	0.8	7	W1+W2	9	-1	-1	Manual
oxy_7	1.5	7	W1	1	0	1	Manual
oxy_8	4.9	7	W1+W2	3	0	-1	Manual
oxy_9	0.0	6	W1	10	0	-3	MD
oxy_10	2.2	6	W1	12	0	-1	MD
oxy_11	2.9	6	W2	3	-1	-2	Manual
oxy_12	4.2	6	W1+W2	9	-1	-5	Manual
oxy_13	6.3	6	W2	6	0	-5	MD
oxy_14	8.1	6	W1+W2	7	0	-2	MD
oxy_15	8.2	6	W1+W2	3	0	-1	MD
oxy_16	8.3	6	W1+W2	2	-1	1	MD
oxy_17	9.1	6	W1	7	0	-2	MD
oxy_18	10.4	6	W1+W2	17	-2	-1	Manual
oxy_19	10.4	6	W1	7	-1	0	Manual
oxy_20	14.2	6	W2	2	0	1	Manual
oxy_21	14.3	6	W1	3	0	0	Manual
oxy_22	14.4	6	W1	1	0	-1	MD
oxy_23	15.2	6	W2	-1	-1	-1	MD
oxy_24	15.6	6	W2	-1	0	1	Manual
oxy_25	16.2	6	W1+W2	13	-1	-1	Manual
oxy_26	18.2	6	-	0	0	1	Manual
oxy_27	19.4	6	W1	-1	0	0	Manual
oxy_28	19.7	6	W1	5	0	2	Manual
oxy_29	22.4	6	W1	1	0	1	MD
oxy_30	23.6	6	W1	1	0	-1	Manual
oxy_31	0.0	5	W1+W2	12	-1	-2	MD
oxy_32	4.9	5	W1	3	1	0	MD
oxy_33	6.7	5	W1	15	0	-3	MD
oxy_34	9.3	5	W1+W2	10	-1	-1	Manual
oxy_35	10.6	5	W1+W2	-8	-1	-2	Manual
oxy_36	11.6	5	W1	14	0	-2	MD
oxy_37	12.5	5	W1+W2	8	-1	-2	Manual
oxy_38	15.2	5	W2	5	-3	-1	MD
oxy_39	16.8	5	W1	18	0	0	MD

oxy_40	17.1	5	W1	8	0	-4	Manual
oxy_41	17.2	5	W1	-1	1	1	Manual
oxy_42	17.6	5	W1+ W2	13	-1	-1	Manual
oxy_43	17.7	5	W1	-3	0	-1	MD
oxy_44	18.9	5	W1	4	0	-2	MD
oxy_45	24.4	5	W2	7	-1	-2	MD
oxy_46	25.9	5	W1	4	0	0	MD
oxy_47	31.9	5	-	0	0	2	Manual
oxy_48	0.0	4	W1+W2	10	-1	1	Manual
oxy_49	8.2	4	W2	7	-1	-1	MD
oxy_50	9.5	4	W1	0	0	2	MD
oxy_51	9.7	4	W1	-3	0	2	MD
oxy_52	10.5	4	W1+W2	16	-1	-2	MD
oxy_53	10.9	4	W1	4	0	2	MD
oxy_54	12.2	4	W1	11	0	-2	MD
oxy_55	15.9	4	W1	12	0	-2	MD
oxy_56	17.5	4	W1	10	1	1	MD
oxy_57	20.9	4	W2	7	-1	1	MD
oxy_58	21.5	4	W2	3	0	-4	MD
oxy_59	26.3	4	-	0	0	1	Manual
oxy_60	32.8	4	W2	9	0	-4	Manual
oxy_61	0.0	3	W1+W2	3	-1	-2	Manual
oxy_62	17.0	3	W1	9	0	-3	Manual
oxy_63	0.0	2	W1+W2	10	-1	-2	Manual
oxy_64	16.1	2	W1	9	-2	-2	Manual

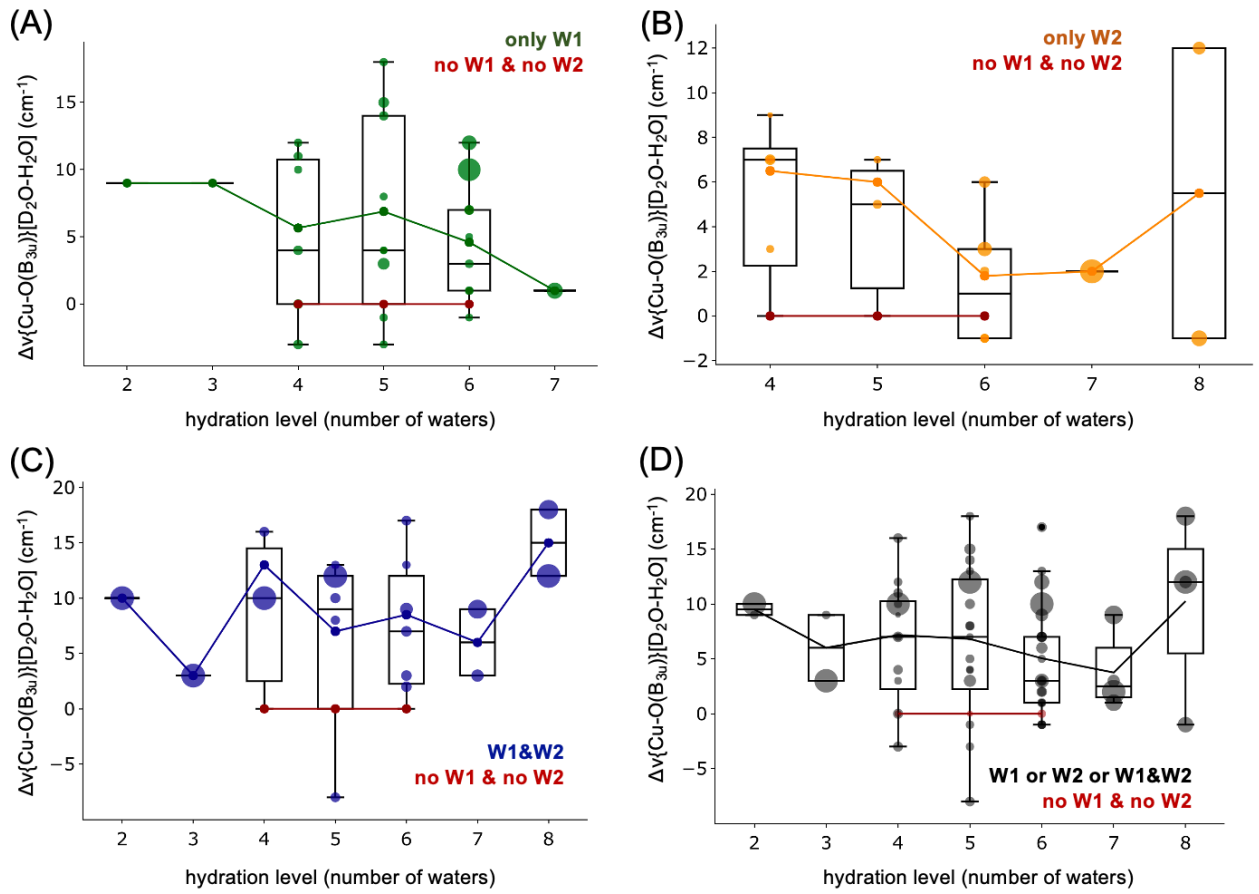


Figure S8. Summary statistics for $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) frequency shifts in oxy-Ty at different hydration levels (from data shown in Table S2). Boxplot distributions of the $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) frequency shifts at each hydration level for structures with the μ-η^2:η^2-peroxide 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)). In all cases the structures with neither W1 nor W2 positions occupied by waters are included as references (red points/lines). The size of each point is scaled based on the thermodynamic stability of each corresponding structure (larger size = lower relative energy). Note that out of the 61 structures with at least one H-bonded water to the μ-η^2:η^2-peroxide, 52 exhibit a $H_2O \rightarrow D_2O$ Cu-O(B_{3u}) upshift (D), and 18 (out of 19) structures with two H-bonding interactions exhibit such an upshift (C). Also note that across all cases (A-D) the most thermodynamically stable structures (larger points) consistently correlate with the expected Cu-O(B_{3u}) upshifts.

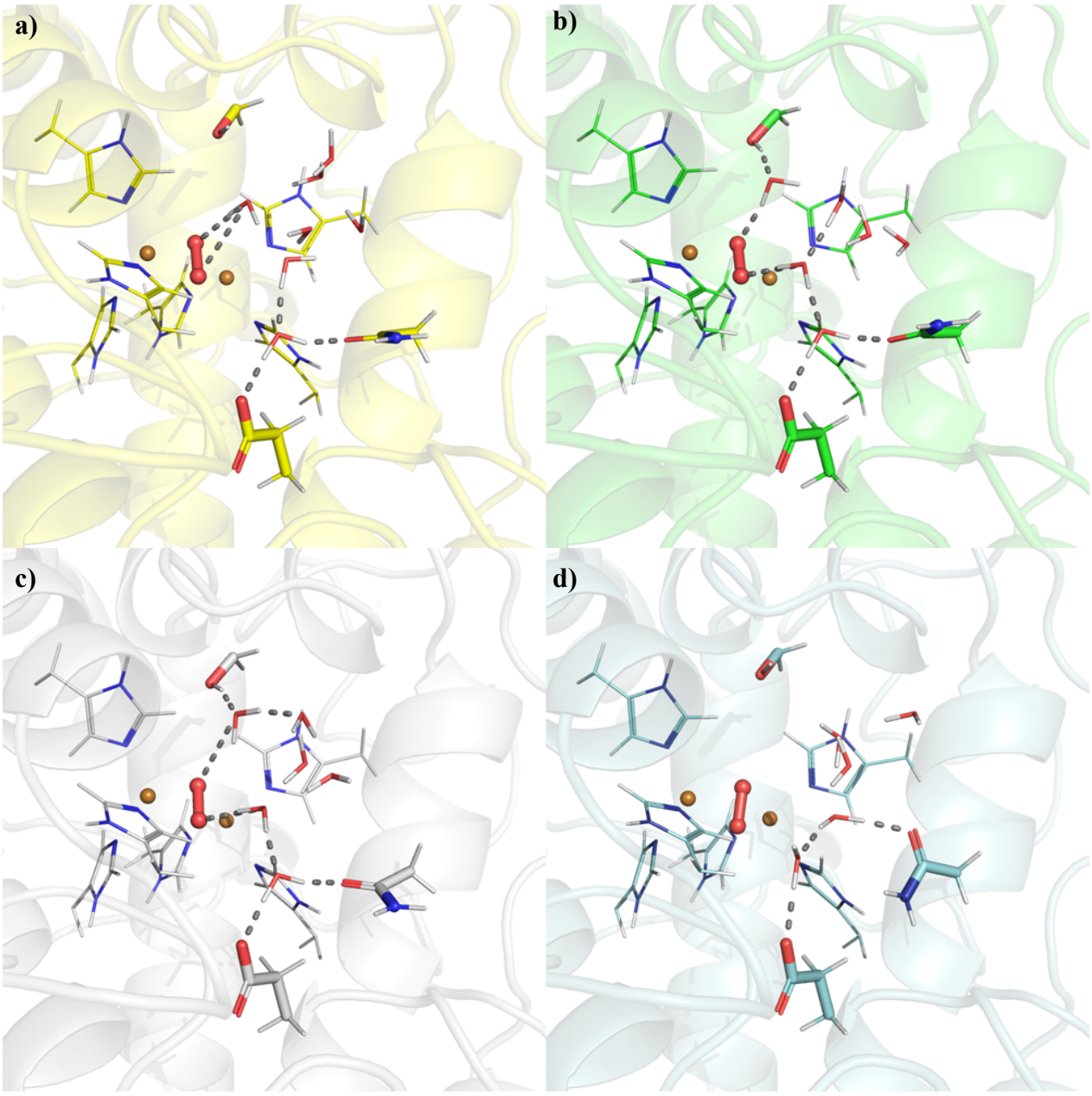


Figure S9. QM/MM optimized structures of most thermodynamically favorable oxy-Ty models of the different hydration level (W6-W4), and an example without water molecules H-bond to peroxide: a) oxy_9, b) oxy_31 c) oxy_48 , d) oxy_59. Copper atoms shown as brown spheres. Only conserved water molecule and those within radius of 6.5 Å from Cu_2O_2 core are shown.

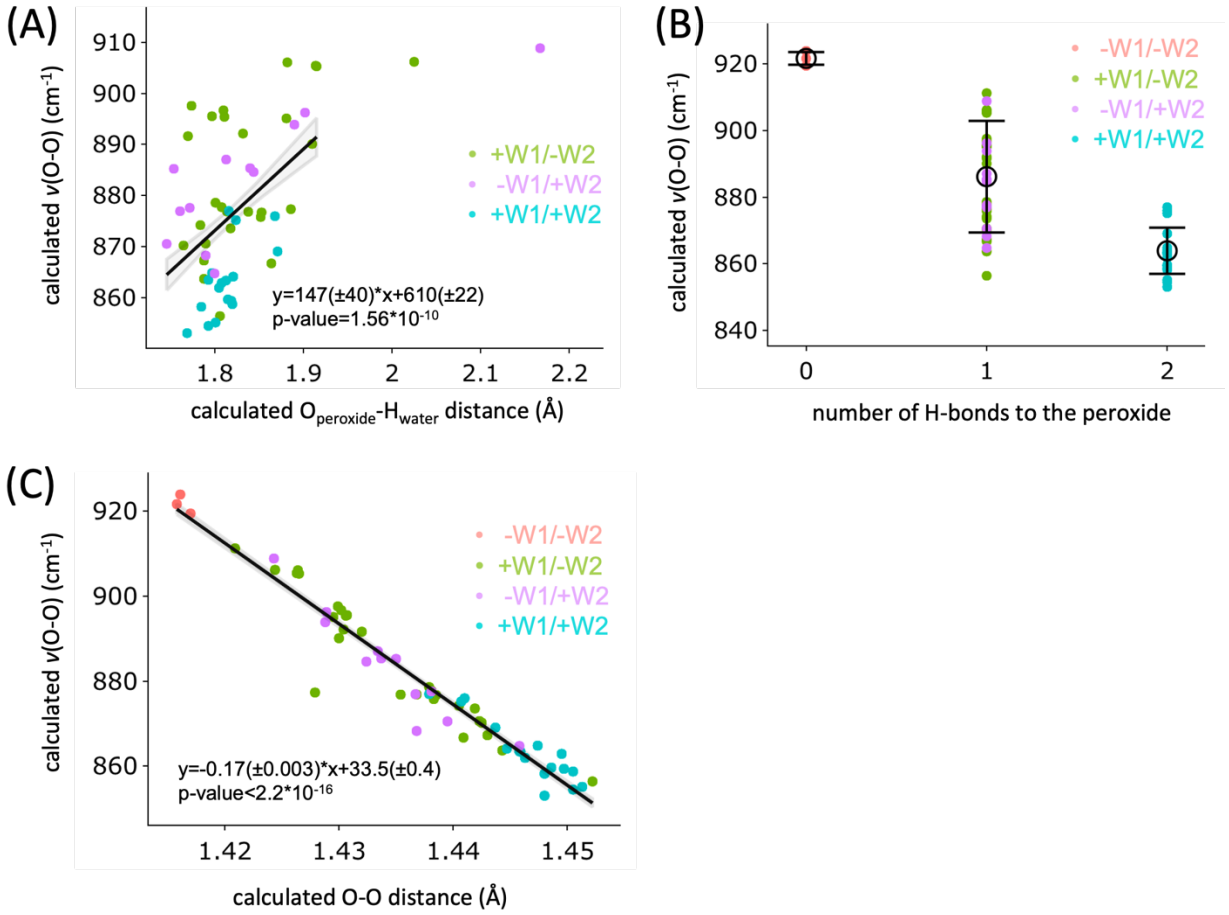


Figure S10. H-bonding activates the μ-η^2:η^2-peroxide O-O bond. Each point corresponds to one of our QM/MM-optimized structures for oxy-Ty listed in Table S2 and is colored based on the number of H-bonded waters (orange=no H-bonded water, green=only W1, purple=onlyW2, blue=both W1 and W2): (A) dependence of the calculated O-O(A_g) frequency to the strength of the H-bond (defined as the distance of the H-atom of water to the O-atom of the peroxide), with the linear regression fit (black line, with standard error as grey area) for moderate/strong H-bonds ($O_{peroxide}$-H_{water} distance <2Å; for the structures with both W1 and W2, the distance for the strongest H-bond is considered). (B) dependence of the calculated O-O(A_g) frequency to the number of H-bonding interactions to the μ-η^2:η^2-peroxide, including the mean values (black circles) and standard deviations for each group. (C) Correlation of the calculated μ-η^2:η^2-peroxide O-O(A_g) frequency to the calculated O-O bond length, with the linear regression fit (black line, with standard error as grey area).

References

(1) Felsenfeld, G. The Determination of Cuprous Ion in Copper Proteins. Archives of Biochemistry and Biophysics **1960**, 87 (2), 247–251. https://doi.org/10.1016/0003-9861(60)90168-5.

(2) Tepper, A. W. J. W.; Bubacco, L.; Canters, G. W. Stopped-Flow Fluorescence Studies of Inhibitor Binding to Tyrosinase from Streptomyces Antibioticus. Journal of Biological Chemistry **2004**, 279 (14), 13425–13434. https://doi.org/10.1074/jbc.M309367200.

(3) Phillips, J.C.; Braun, R.; Wang, W.; Gumbart, J.; Tajkhorshid, E.; Villa, E.; Chipot, C.; Skeel, R.D.; Kalé, L.; Schulten, K., Scalable molecular dynamics with NAMD. Journal of Computational Chemistry **2005**, 26, 1781-1802. https://doi.org/10.1002/jcc.20289

(4) Humphrey, W.; Dalke, A.; Schulten, K., VMD - Visual Molecular Dynamics. Journal of Molecular Graphics **1996**, 14, 33-38. https://doi.org/10.1016/0263-7855(96)00018-5

(5) DeLano, W. L. Pymol: An open-source molecular graphics tool. CCP4 Newsletter On Protein Crystallography **2002**, 40, 82-92.

(6) Ryde, U. The coordination of the catalytic zinc ion in alcohol dehydrogenase studied by combined quantum-chemical and molecular mechanics calculations. Journal of Computer-Aided Molecular Design **1996**, 10, 153 164. https://doi.org/10.1007/BF00402823

(7) Ryde, U.; Olsson, M. H. M. Structure, strain, and reorganization energy of blue copper models in the protein. International Journal of Quantum Chemistry **2001**, 81, 335 347. https://doi.org/10.1002/1097-461X(2001)81:5<335::AID-QUA1003>3.0.CO;2-Q

(8) TURBOMOLE V7.4 2019, University of Karlsruhe and Forschungszentrum Karlsruhe GmbH: **1989-2007**. https://doi.org/10.1063/5.0004635

(9) Case, D.A.; Babin, V.; Berryman, J.T.; Betz, R.M.; Cai, Q.; Cerutti, D.S.; Cheatham, T.E., III; Darden, T.A.; Duke, R.E.; Gohlke, H.; et al. Amber 14; University of San Francisco: San Francisco, CA, USA, **2014**.

(10) Cao, L.; Ryde, U. On the Difference Between Additive and Subtractive QM/MM Calculations. Frontiers in Chemistry **2018**, 6, 89. https://doi.org/10.3389/fchem.2018.00089

(11) Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C. ff14SB: Improving the Accuracy of Protein Side Chain and Backbone Parameters from ff99SB. Journal of Chemical Theory and Computation **2015** 11, 8, 3696-3713.

(12) Becke, A. D., Density-functional exchange-energy approximation with correct asymptotic behavior. Physical Review A **1988**, 38 (6), 3098-3100. https://doi.org/10.1103/PhysRevA.38.3098

(13) Vosko, S. H.; Wilk, L.k; Nusair, M. Accurate spin-dependent electron liquid correlation energies for local spin density calculations: A critical analysis Canadian Journal of Physics **1980**, 58, 1200-11. DOI: 10.1139/p80-159

(14) Perdew, J. P., Density-functional approximation for the correlation energy of the inhomogeneous electron gas. Physical Review B **1986**, 33 (12), 8822-8824. https://doi.org/10.1103/PhysRevB.33.8822

(15) Grimme, S. Supramolecular Binding Thermodynamics by Dispersion-Corrected Density Functional Theory. Chemistry—A European Journa., **2012**, 18, 9955-9964. https://doi.org/10.1002/chem.201200497

(16) Weigend, F.; Ahlrichs, R. Balanced Basis Sets of Split Valence, Triple Zeta Valence and Quadruple Zeta Valence Quality for H to Rn: Design and Assessment of Accuracy. Physical Chemistry Chemical Physics **2005**, 7, 3297-3305. https://doi.org/10.1039/b508541a

(17) Klamt, A. Conductor-like Screening Model for Real Solvents: A New Approach to the Quantitative Calculation of Solvation Phenomena The Journal of Physical Chemistry A **1995**, 99, 2224-2235. https://doi.org/10.1021/j100007a062
(18) Klamt, A.; Koch, L.; Terzi, S. COSMOtherm 19.0.0 (Revision 5239), 19.0.0 (Revision 5239); COSMOlogic GmbH & Co. KG: 1999-2018.
(19) Mohammed, T. A.; Meier, C. M.; Kalvoda, T.; Kalt, M. A.; Rulíšek, L.; Shoshan, M. S.: Potent Cyclic Tetrapeptides for Lead Detoxification. Angew. Chem. Int. Ed. **2021**, 60, 12381-12385. https://doi.org/10.1002/anie.202103217
(20) Klamt, A.; Schüürmann, G. COSMO: a new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. Journal of the Chemical Society, Perkin Transactions 2 **1993**, 2, 799-805. https://doi.org/10.1039/P29930000799]]
(21) Head, J.D. (1997), Computation of vibrational frequencies for adsorbates on surfaces. Int. J. Quantum Chem., 65: 827-838. https://doi.org/10.1002/(SICI)1097-461X(1997)65:5<827::AID-QUA47>3.0.CO;2-U
(22) S. de Marothy, XYZViewer version 0.975, **2010**.
(23) Matoba, Y.; Kihara, S.; Bando, N.; Yoshitsu, H.; Sakaguchi, M.; Kayama, K.; Yanagisawa, S.; Ogura, T.; Sugiyama, M. Catalytic Mechanism of the Tyrosinase Reaction toward the Tyr98 Residue in the Caddie Protein. PLoS Biol **2018**, 16 (12), e3000077. https://doi.org/10.1371/journal.pbio.3000077.
(24) Sendovski, M.; Kanteev, M.; Ben-Yosef, V. S.; Adir, N.; Fishman, A. First Structures of an Active Bacterial Tyrosinase Reveal Copper Plasticity. Journal of Molecular Biology **2011**, 405 (1), 227–237. https://doi.org/10.1016/j.jmb.2010.10.048.