Supporting Information

Towards Solvent Regulated Self-Activation of N-Terminal Disulfide Bond Oxidoreductase-D

Aparna G Nair[a‡], D. Sravanakumar Perumalla[a,b‡], and Padmesh Anjukandi[*a*]

[a] *Department of Chemistry, Indian Institute of Technology, Palakkad-678557, Kerala, India. E-mail: padmesh@iitpkd.ac.in*

[b] *'Present Address': Department of Inorganic and Physical Chemistry, Indian Institute of Science, Bengaluru-560012, India.*

[*] padmesh@iitpkd.ac.in

CONTENTS

S1. METHODS

A. Molecular Mechanics- Molecular Dynamics Simulations

The native structure for the 124 residue $nDsbD_{Ox}$ (PDB ID: 1L6P) is collected from Protein Data Bank [1], and missing atoms for Arg_8 and Lys_{73} are added using PYMOL [2]. Classical Molecular Dynamics (MD) simulations are performed using GROMACS-2016.3 [3, 4]. for understanding the behaviour of the catalytic disulfide bond and its surrounding residues. Considering that 1L6P is a globular protein of size $\approx$ 6nm, the protein-water system is enclosed in a cubic box of length 8nm. Protein is then solvated with TIP3P modelled [5] 16336 water molecules and neutralized for residual charges using 10 Na^+ and 5 Cl^- ions. Following the protocol below, the protein-water system is simulated for 400ns with 2fs time step and Amber99sb-ildn force field [6]. Periodic boundary condition is applied along the three axes and all bonds are constrained using LINCS algorithm [7]. 1.0 nm cut-off is defined for both short-range Coulomb and van der Waals interactions. Long-range electrostatic interactions are accounted using the Particle-mesh Ewald method [8, 9]. Trajectory coordinates are recorded for every 200fs, later visualized and analyzed using VMD 1.9.3 software [10]. MD simulations are also conducted on native $nDsbD_{Ox}$ with (i) de-protonated Tyr_{40} and protonated Asp_{68} ($dTyr_{40}$) (ii) de-protonated Tyr_{42} and protonated Asp_{68} ($dTyr_{42}$) in a similar fashion. As $TyrO^-$ is not a natural state for amino acid, partial charges and topology parameters are derived from tleap tool in Ambertools18 [11]. 400ns MD simulations is also performed with $cDsbD_{Ox}$ (PDB ID: 2FWE)[12] using the same method described as above and with the following changes; protein is enclosed in a cubic box of size 10nm, solvated with 32145 TIP3P water molecules and neutralized with 10 Na^+ and 4 Cl^- ions.

The protocol for all MD simulations conducted here are as follows: (i) Energy minimization of the neutralized protein-water system with Steepest Descent algorithm (ii) Equilibration in NVT ensemble at 300K for 1 ns, followed by 1 ns equilibration in NPT ensemble at 1 bar pressure (iii) Final production run with V-rescale thermostat [13] and Parrinello-Rahman barostat [14] for 400ns.

B. QM/MM Molecular Dynamics

System conformation favourable for the reactions is identified from the MD trajectory and is taken as the input for QM/MM MD simulations [15] implemented in CP2K 6.1 code [16]. The inputs are converted into CP2K readable format with tleap package of AmberTools18 [11], keeping atom positions and simulation box size intact. Topology files obtained so were corrected for solvent-protein Lennard-Jones interaction potential terms and later energetically minimized. The side chain of Tyr_{40}, Tyr_{42}, Asp_{68}, Gln_{101}, Cys_{103}, Cys_{109}, and the water molecules from the first solvation shell (4Å cut-off) around these residues are included in the QM region (fig. S6). These atoms are then placed inside a cubic box of length 2.6nm with a reflective wall along all the sides. QM energies are calculated based on Density Functional Theory (DFT) with Gaussian Plane Wave method (GPW) [17] (BLYP-D3 functional [18–21], DZVP basis set [22] and Goedecker-Teter-Hutter (GTH) pseudo-potential [23, 24] with 300Ry cut-off). For energy calculations on larger systems such as proteins, BLYP-D3/DZVP is on par with CCSD results [25]. Also, BLYP density functional has been earlier adopted in exploring the disulfide chemistry [26–28]. MM region for the present calculations are modelled by Amberff14SB force-field[29] and TIP3P water model. QM/MM boundary is separated with IMOMM [30] link atom approach [31, 32] with hydrogen as the linking atom. Electrostatic interaction between QM and MM region is defined using Gaussian Expansion of Electrostatic Potential (GEEP) method [33, 34]. QM/MM system is then equilibrated for 5ps with 0.5 fs time step at 300K using Nosé–Hoover thermostat chain [35, 36], in NVT ensemble.

C. Metadynamics

Free Energy Surface (FES) for the reactions are explored with QM/MM MD metadynamics (MTD) simulations [37]. As we are looking at a rare event, MTD simulations are run in parallel with eight walkers [38]. This helps sample all possible conformation in the Collective Variable (CV) space for the reaction under study. History dependent Gaussian potential of height 0.25kcal/mol and width 0.15 is deposited with the CV every 300 steps (150fs). Here, CV is defined as the difference in coordination number ($\Delta CN(O_a, O_d, H)$) for acceptor (O_a)

and donor (O_d) oxygen atom, with respect to the hydrogen atom which is being transported (eq. (S2)).

$$CV = \Delta CN(O_a, O_d, H) = CN(O_a, H) - CN(O_d, H) \tag{S1}$$

$CN(O_d, H)$ and $CN(O_a, H)$ are calculated using eq. (S2).

$$CN(O, H) = \frac{1 - \left(\frac{d_{OH}}{d_0}\right)^6}{1 - \left(\frac{d_{OH}}{d_0}\right)^{12}} \tag{S2}$$

where 'd_{OH}' is the distance between oxygen and hydrogen atom; 'd_0' is the equilibrium bond length between the respective atoms, taken here as 0.98 Å. If the proton is located near the acceptor, then the CV ($\Delta CN(O_a, O_d, H)$) will have a positive value otherwise negative. Thus, a value close to 0.45 of any CV is considered the product state along that CV.

D. DFT Calculations

Energetics for static gas-phase nucleophile mediated disulfide cleavage reaction was computed using BLYP-D3/6-31G* [18–21, 39] in ORCA quantum chemistry program suite [40]. $Tyr_{42}O^-$, $Cys_{103/109}$ residues as defined in the QM region only are incorporated for the study. The inputs for the calculation is collected from the $Tyr_{42}O^-$-$Cys_{103}C_\alpha H$ distance scan performed using the QM/MM implementation in Gromacs 2016.3-ORCA combined program. Valency for carbon is satisfied using hydrogen atoms. The reactants and products are optimized within extreme tight SCF convergence criteria. The transition state was confirmed from the imaginary frequency.

S2. NUCLEOPHILE GENERATING RESIDUES NEAR THE ACTIVE SITE

A. Proton Transfer From $Tyr_{40/42/71}OH$ To $Asp_{68}O^-$

The active site in $nDsbD_{O}x$ is surrounded by amino acid residues such as Tyr_{40}, Tyr_{42}, Asp_{68} and Tyr_{71}, whose side chain can act as potential nucleophile generators (fig. S1). Here, ($TyrO^-$) nucleophile can be generated through proton abstraction by Asp_{68} from any of the above tyrosine residues.

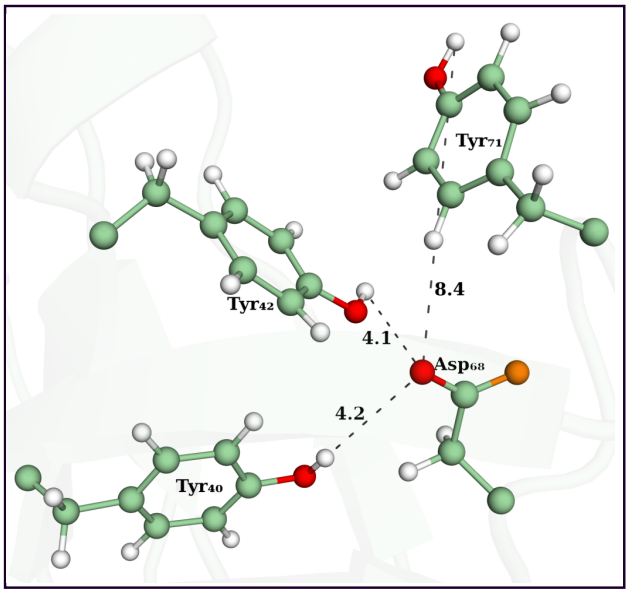


FIG. S1: Possible nucleophile generating residues near the disulfide bond of $nDsbD_{Ox}$. For identification, one of the Asp_{68} oxygen is coloured orange.

B. Distance between $Asp_{68}O^-$ and $Tyr_{40/42}OH$

Possibilities of proton abstraction by Asp_{68} from $Tyr_{40/42}$ to form $TyrO^-$ nucleophile is analyzed here. For this, the evolution of distance with time and corresponding normalized distributions are plotted (figs. S2 and S3). Based on the simulations, distance analysis between the residues are separated into two sections (i) 0 - 150ns and (ii) 150-400ns. The majority of the time up to 150ns, as both $Tyr_{40/42}$ residues are far away from Asp_{68}, there is no significant hydrogen bond interaction. Thus in the distribution plot (fig. S3a) major peak

is due to 4-8Å distance. In rare cases, Asp_{68} comes closer only to Tyr_{40} (t=60ns, fig. S2a). After 150ns, it can be seen that both oxygen's of Asp_{68} forms a hydrogen bond to Tyr_{42}; the trend being retained for ≈25ns and repeated over time. Nevertheless, at ≈200ns, Tyr_{40} and Tyr_{42} together share hydrogen bond interactions with Asp_{68}. The same can be interpreted from the distribution as well. The dominant peak at ≈2Å comes from $Tyr_{42}OH$ - $Asp_{68}O^-$, whereas $Tyr_{40}-Asp_{68}$ distance contributes only one-third of the area (fig. S3b). Thus, it can be concluded that near the active site of $nDsbD_{Ox}$, strong hydrogen bond interaction exists between the side chains of $Tyr_{42}-Asp_{68}$ residues and proton abstraction from Tyr_{42} by Asp_{68}, can generate $TyrO^-$ nucleophile. Tyr_{40} also has a minor contribution towards the formation of $TyrO^-$, but the possibilities are minimal. Hence, our MD simulation results for the Phe_{70} cap fluctuations, together with the proton transfer direct towards the exploration of possibilities and energetics involved in forming nucleophile ($TyrO^-$), which may further lead to disulfide scission.

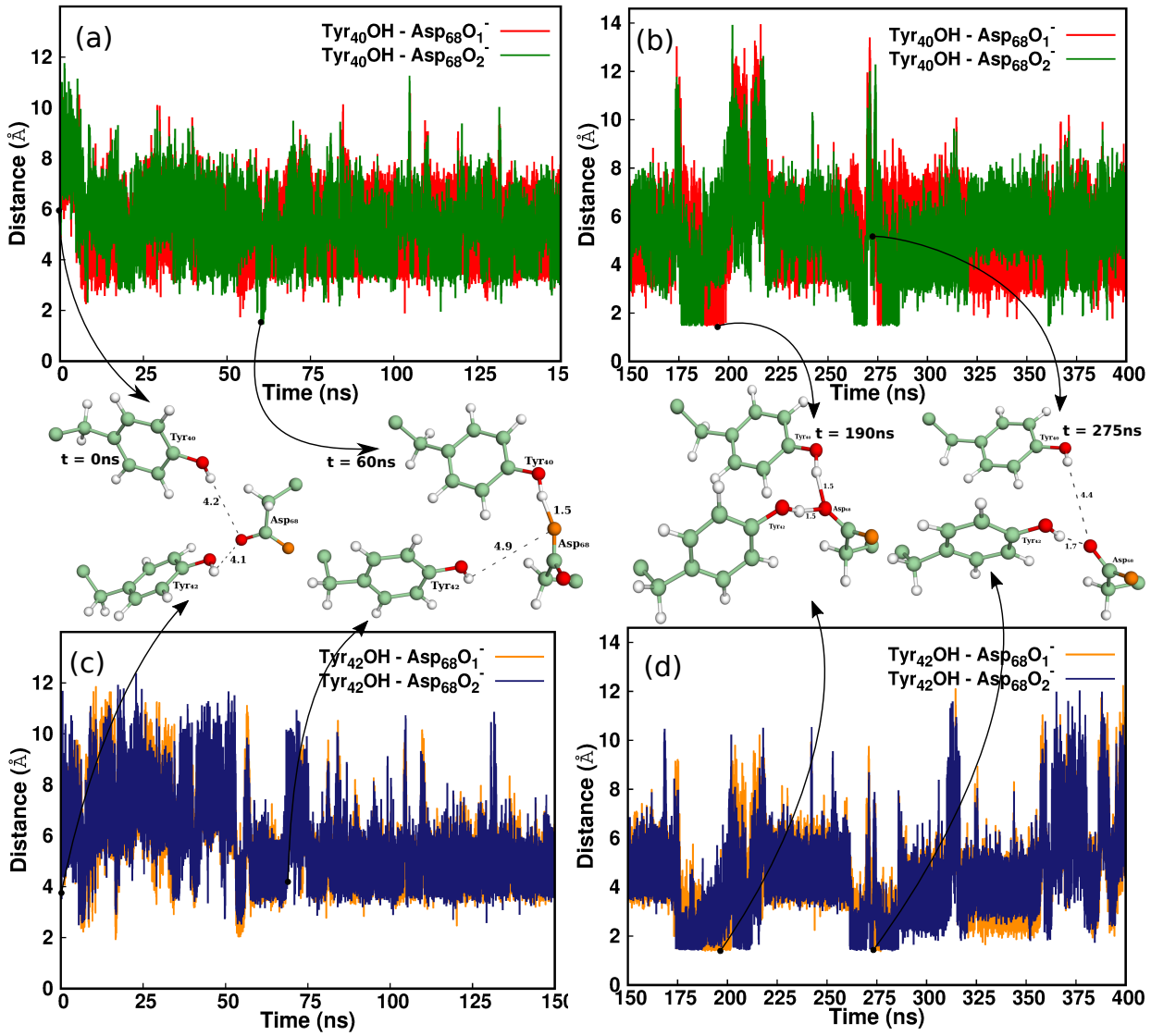


FIG. S2: Time evolution for the distance between $Tyr_{40/42}$ and Asp_{68} separated into multiple sections based on time. $Tyr_{40}-Asp_{68}$ distance for time (a) 0-150ns, (b) 150-400ns and $Tyr_{42}-Asp_{68}$ distance from (c) 0-150ns (d) 150-400ns are given. Snapshots corresponding to selected simulation time are also shown in insight. One of the Asp_{68} oxygen is coloured orange for identification. Strong hydrogen bond interaction can be seen between Tyr_{42} and Asp_{68} residues.

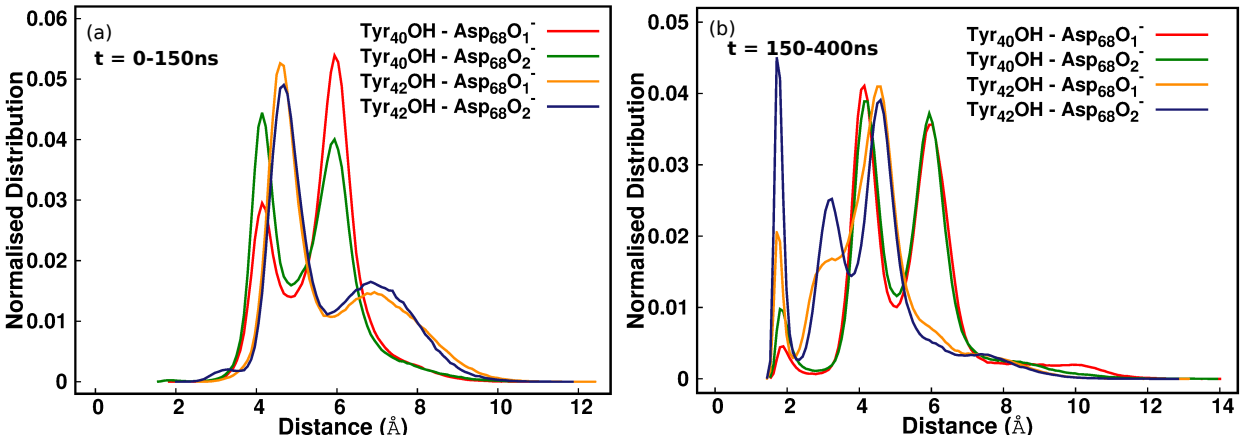


FIG. S3: Normalized distribution for the distance between $Tyr_{40/42}$ and Asp_{68} separated into two-time sections; (a) 0-150ns and (b) 150-400ns.

C. Distance between $Asp_{68}O^-$ and $Tyr_{71}OH$

Tyr_{71} can transfer its proton to Asp_{68} and generate $Tyr_{71}O^-$ nucleophile. So, the distance between the residues is plotted to determine whether Tyr_{71} can be a potential nucleophile generating residue (fig. S4). As seen from the plot, these residues approach down to ≈ 4Åonly. The formation of $TyrO^-$ by the proton transfer between Tyr_{71} and Asp_{68} is very rare, and therefore, for further investigations, Tyr_{71} is excluded from the QM region.

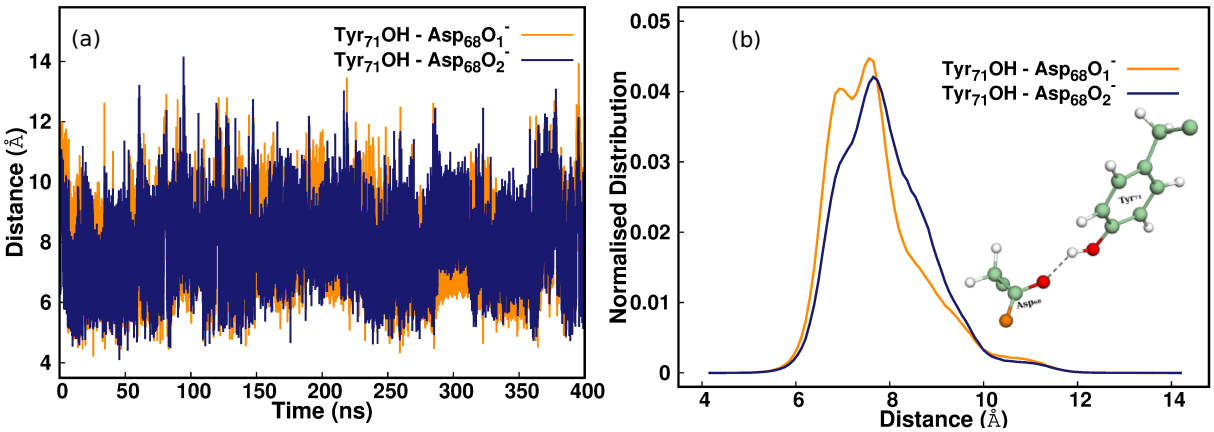


FIG. S4: (a) Time evolution for the distance between $Tyr_{71}OH$ and $Asp_{68}O_{1/2}{}^-$ and (b) corresponding normalized distribution. As can be seen, $Asp_{68}O^-$ is far away from $Tyr_{71}OH$ for proton abstraction, and so $Tyr_{71}O^-$ nucleophile is not a possibility.

D. $Asp_{68}O^-$ as a nucleophile

Chances of direct attack by Asp_{68} as a nucleophile in breaking the disulfide bond is also examined (fig. S5). As the closest distance between Asp_{68} and $Cys_{103}S$ is ≈ 4Å, chances of Asp_{68} to perform a direct attack on Cys_{103} and break the disulfide bond can be ruled out.

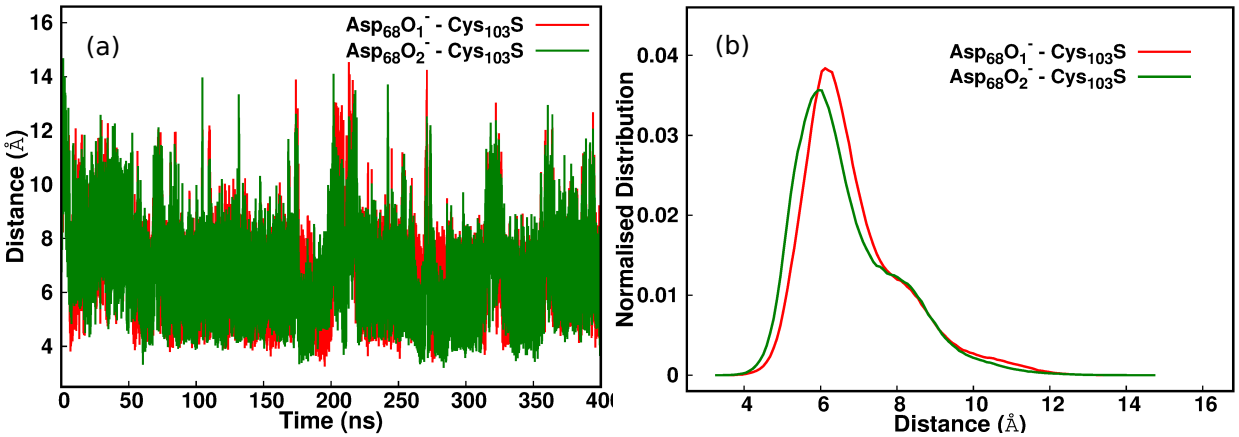


FIG. S5: (a) Time evolution for the distance between $Cys_{103}S$ and Asp_{68} (b) Corresponding normalized distribution. Possibilities for Asp_{68} to attack the disulfide bond is minimal here.

E. Defining QM region

Residues, as shown in fig. S6 in native $nDsbD_{Ox}$ and ten water molecules from the first solvation shell (4Å cut-off) are incorporated in the QM region. The boundary between the QM and MM region is separated using hydrogen atoms, and at most care has been given to place hydrogen between C−C single bond. Same residues are included when defining the QM region in $dTyr_{42}$, together with thirteen water molecules from the first solvation shell.

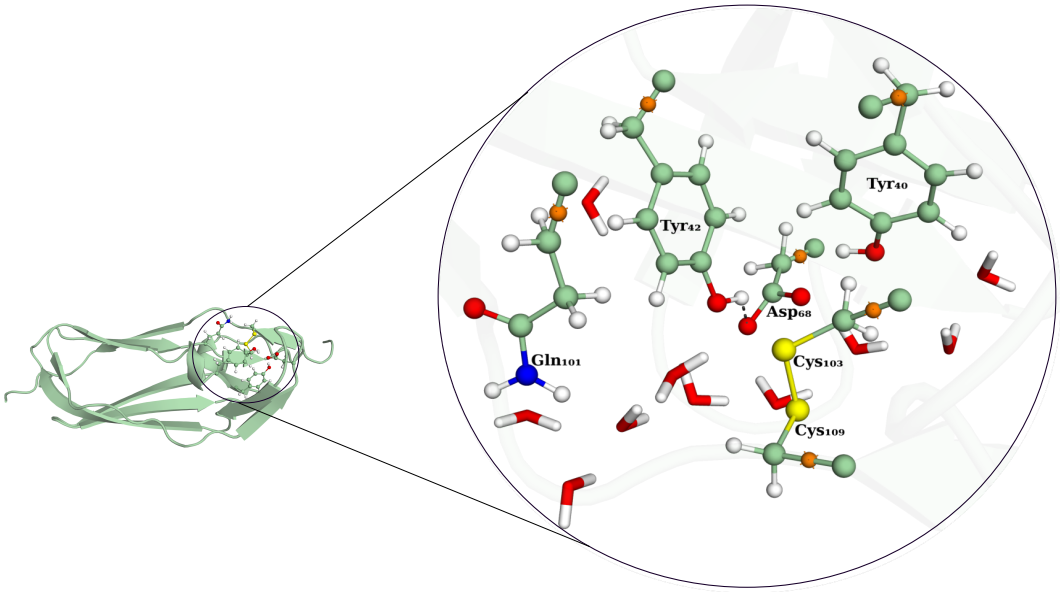


FIG. S6: Native structure of $nDsbD_{Ox}$ (left) (PDB ID: 1L6P). Enlarged on the right are the residues along with ten water molecules included in the QM region. Link atoms separating the boundary between QM and MM are shown as orange balls.

S3. PRESENCE OF CAP-LOOP NEAR THE ACTIVE SITE

A. Opening of Phe_{70} Cap

The active site (including Cys_{103}- Cys_{109} disulfide bond) is enclosed in a cap loop made up of Asp_{68}- Glu_{69} - Phe_{70} - Tyr_{71} - Gly_{72} - Lys_{73} residues [41–43]. Among the loop residues, Phe_{70} forms a cap to Cys_{109} and protects the disulfide from nucleophilic and solvent attack [43]. Conformational changes in the cap loop region have a key role in the flexibility of

nDsbD$_{Ox}$ and its function [44]. The change from a closed to open conformation for Phe$_{70}$ cap can be characterized by the distance from the centre of mass of Phe$_{70}$ ring to the sulphur of Cys$_{103}$ (d$_{Phe_{70}-Cys_{103}}$) and Cys$_{109}$S (d$_{Phe_{70}-Cys_{109}}$). If d$_{Phe_{70}-Cys_{109}}$ is less than 5Å, the cap is considered closed conformation or else open. [44] X-ray structure for nDsbD$_{Ox}$ holds Phe$_{70}$ cap in a closed conformation (distance=3.5Å). Nevertheless, our simulations show a fluctuating behaviour for the cap, in which during the majority of simulation time cap remains in the open state. It is visible from (i) time evolution plot for (d$_{Phe_{70}-Cys_{103/109}}$) distance (fig. S7a) (ii) normalized distribution plot for the same (fig. S7b). Snapshots at different time frames demonstrating Phe$_{70}$ cap opening are also included here fig. S8.

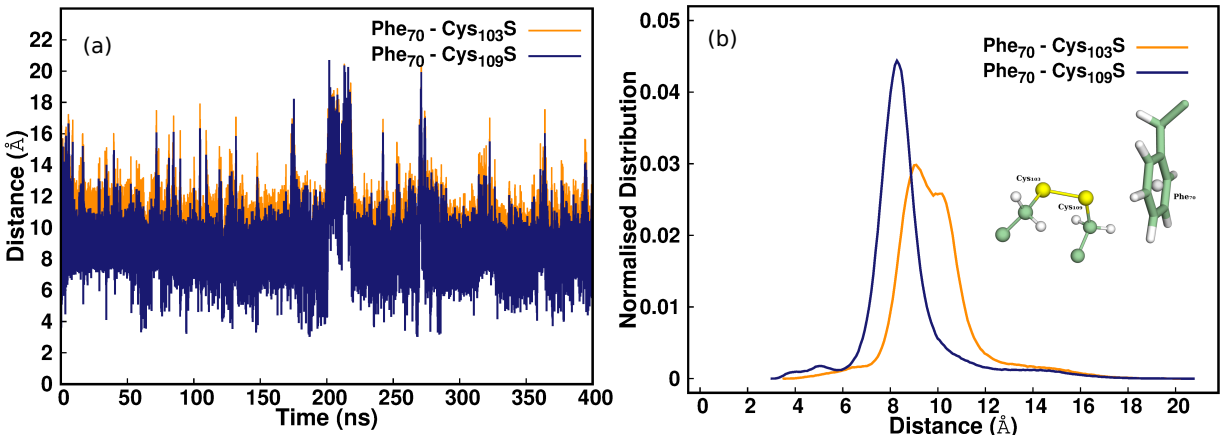


FIG. S7: Fluctuation of Phe$_{70}$ cap with respect to Cys$_{103/109}$S distance. (a)Time evolution for the distance (b) Corresponding normalized distribution plot. It is clear that Phe$_{70}$ cap moves from a closed to open conformation during the simulation.

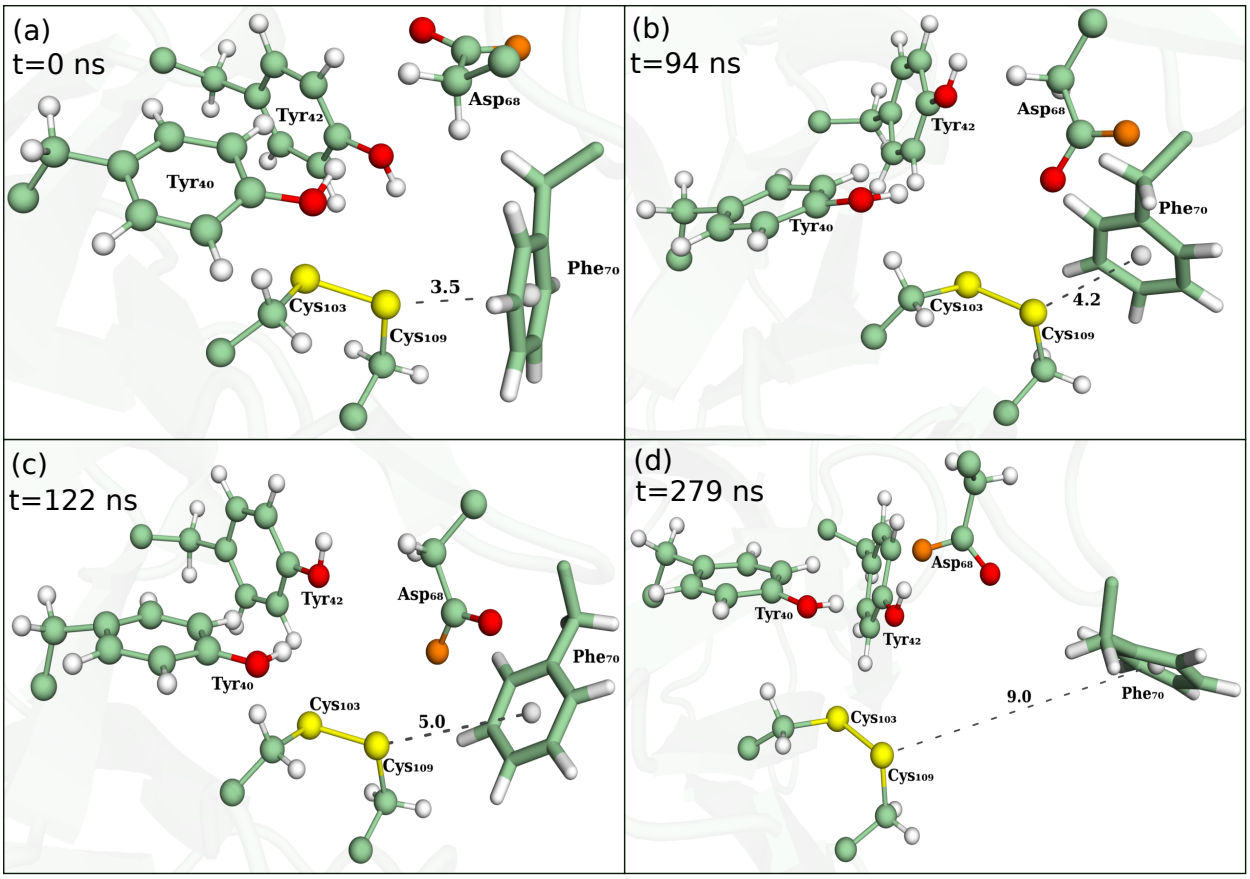


FIG. S8: Snapshots at different time frames from MD simulation showing Phe_{70} Cap opening.

S4. $TYR_{40/42}O^-$ AS NUCLEOPHILE

MD simulations indicate the possibility of proton abstraction by Asp_{68} from $Tyr_{40/42}$ residues to form $TyrO^-$ nucleophile. The fate of these nucleophiles after their formation is analyzed. For this, MD simulations are conducted on $dTyr_{40}$ and $dTyr_{42}$. Following are the observations from the MD data.

A. Opening of Phe_{70} Cap - Distance Analysis

Based on the distance cut-off for Phe_{70} - Cys_{109} in section S3, the cap's opening is evaluated in $dTyr_{40}$, $dTyr_{42}$ systems fig. S9. It is clear that Phe_{70} cap is flexible for $dTyr_{42}$(blue), similar to native $nDsbD_{Ox}$ (green), but remains closed for $dTyr_{40}$ (red).

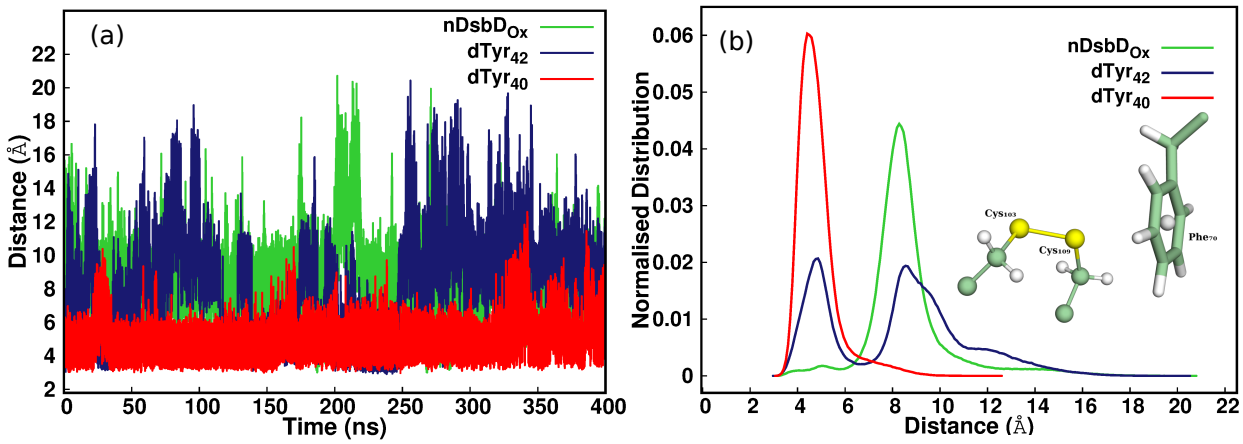


FIG. S9: (a) Variation in Phe_{70} cap - Cys_{109} distance with time for $nDsbD_{Ox}$(green), $dTyr_{42}$ (blue) and $dTyr_{40}$(red) systems (b) Corresponding normalized distribution. When the distance is less than 5Å, which indicates a closed Phe_{70} cap conformation.

B. Opening of Phe_{70} Cap - Torsional Angle Analysis

To measure the presence of local frustration near the active site, χ_1 torsional angle (N−CA−CB−SG) for $Cys_{103/109}$ are measured (fig. S10). χ_1 in gauche (≈60°) suggests that $Cys_{103/109}$ of both $dTyr_{40}$ and $dTyr_{42}$ maintain local frustration like the disulfide of

$nDsbD_{Ox}$.

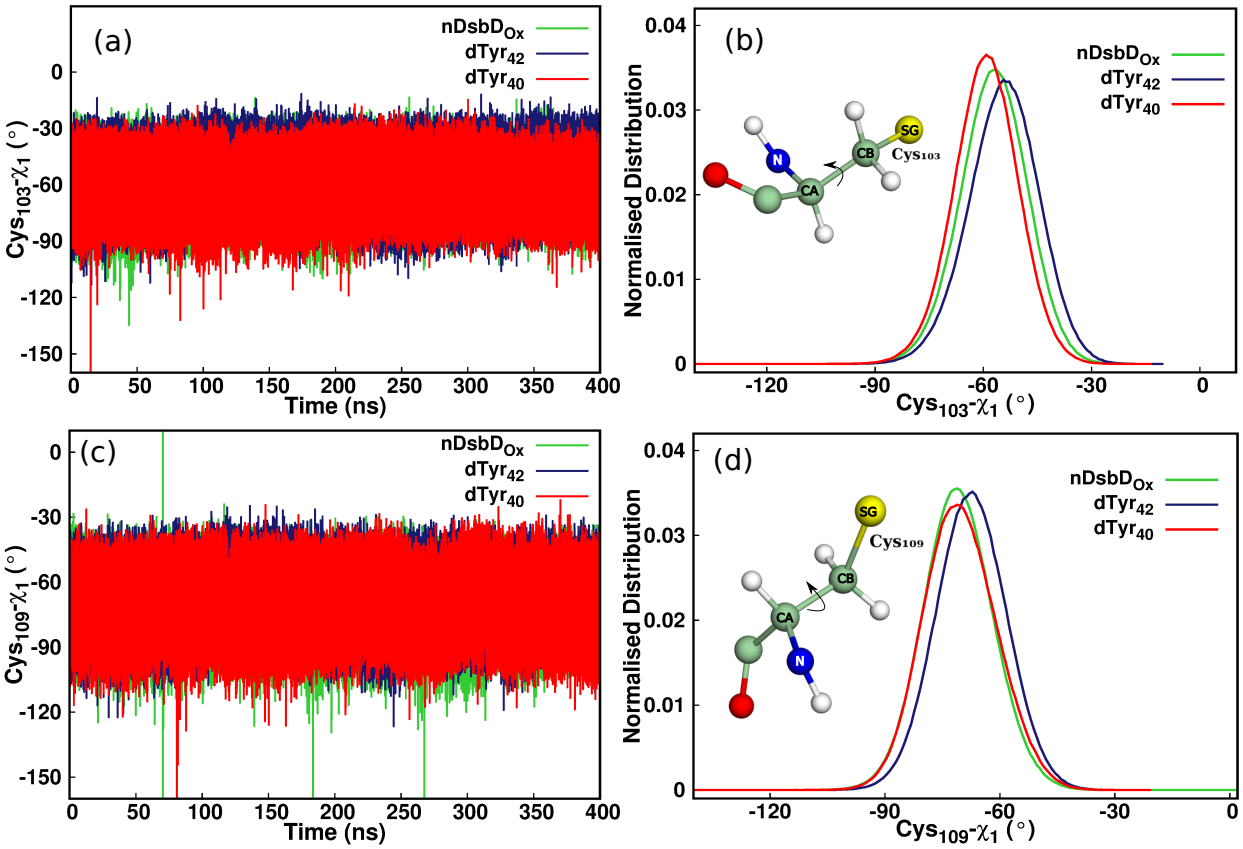


FIG. S10: (a) χ_1 torsional angle and (b) corresponding normalized distribution for Cys_{103} (c) χ_1 torsional angle and (d) corresponding normalized distribution for Cys_{109} in $nDsbD_{Ox}$, $dTyr_{40}$ and $dTyr_{42}$. Values at ≈60° and ≈180° refer to gauche, trans conformations respectively.

χ_1 torsional angle (N−CA−CB−CG) in Phe_{70} is measured (fig. S11). Dominant peak at ≈180° (trans) χ_1 for $nDsbD_{Ox}$ shows a completely open cap, while the presence of two equal peaks at ≈60° (gauche), ≈180° (trans) for $dTyr_{40}$, $dTyr_{42}$ indicate flexible cap.

C. Stability of $Tyr_{40/42}O^-$ nucleophile generated

Now that the formation of $Tyr_{40/42}O^-$ are possible, their fate as a nucleophile is explored. In the case of $Tyr_{40}O^-$, it always stays near the proximity of $Asp_{68}OH$, increasing the chances

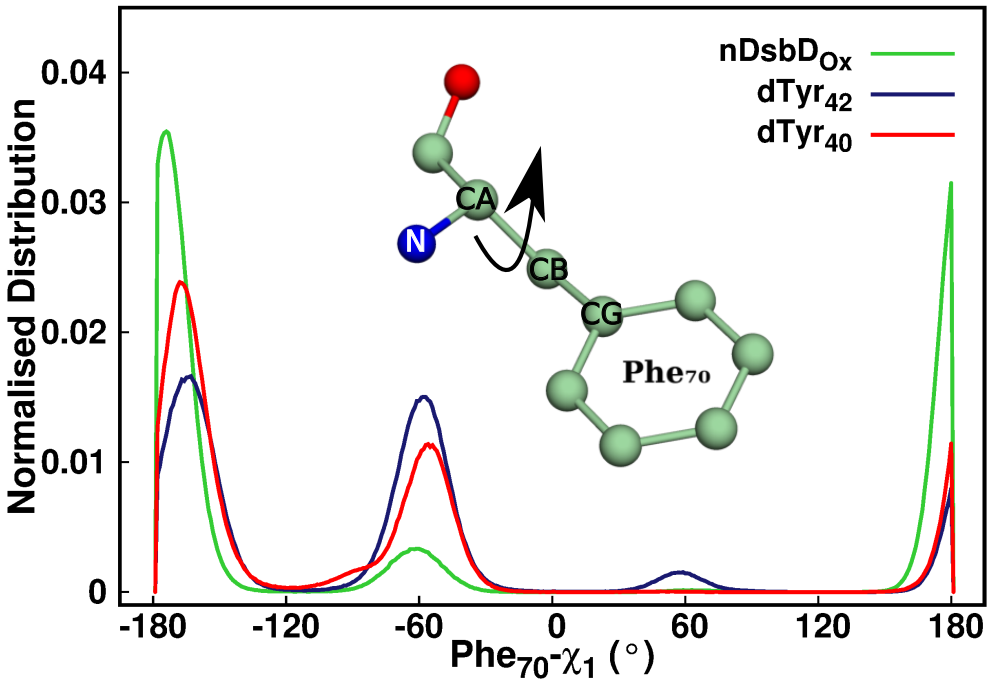


FIG. S11: χ_1 torsional angle for Phe_{70} in $nDsbD_{Ox}$, $dTyr_{40}$ and $dTyr_{42}$. Values at ≈60° and ≈180° refers to gauche, trans conformations, respectively.

of reverse proton transfer and destabilizing the nucleophile (fig. S12a). At the same time, $Tyr_{42}O^-$ drifts away from $Asp_{68}OH$. As a result, proton shuttling between the two residues is hindered, causing so formed $Tyr_{42}O^-$ to stabilize through other means. It is seen that $Tyr_{42}O^-$ approaches $Tyr_{40}OH$ for hydrogen bond interaction, but as seen from fig. S12b, happens very few times during the simulation. Thus, $Tyr_{40}O^-$ is minimal, whereas the drifted $Tyr_{42}O^-$ can be stabilized either by hydrogen bonding interactions with $Tyr_{40}OH$ or through the solvent medium.

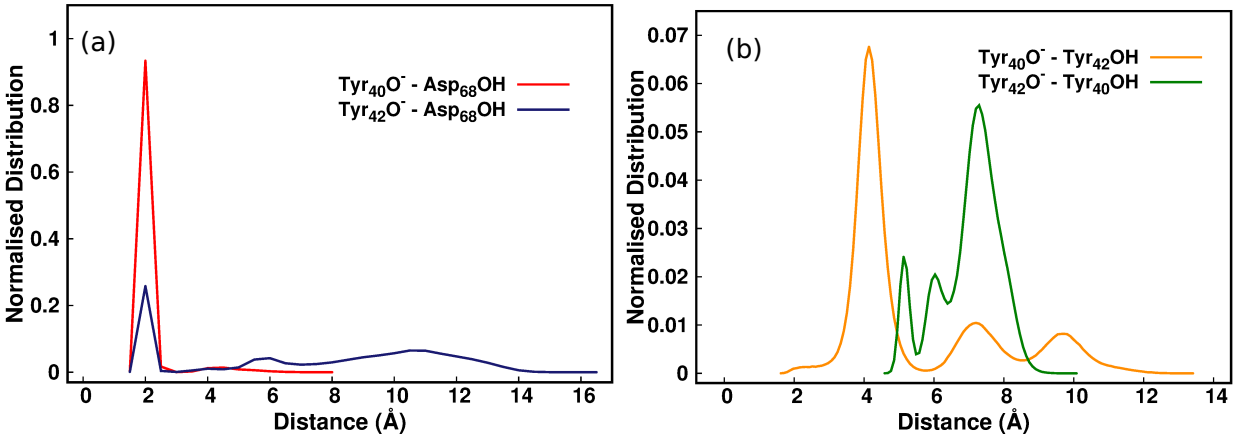


FIG. S12: (a) Distance between $Tyr_{40}O^{-}$- $Asp_{68}OH$ and $Tyr_{42}O^{-}$ - $Asp_{68}OH$. A strong hydrogen bond interaction between $Tyr_{40}O^{-}$- $Asp_{68}OH$ points towards reverse proton transfer. $Tyr_{42}O^{-}$ comes closer to $Asp_{68}OH$, but flies away later. (b) Distance between $Tyr_{40}O^{-}$- Tyr_{42} and $Tyr_{42}O^{-}$ - Tyr_{40}.

D. Solvation around $Tyr_{40/42}O^-$ residues

Since the opening of the Phe_{70} cap increases the solvation around the active site and nearby residues,[41, 42] chances of $Tyr_{42}O^-$ through water is studied. For this, no. of water molecules (cut-off distance 4Å) for the respective residues in are compared with that of the native protein (fig. S13). Tyr_{42} is comparatively more solvated, increasing the chances of $Tyr_{42}O^-$ stabilization through the surrounding water medium.

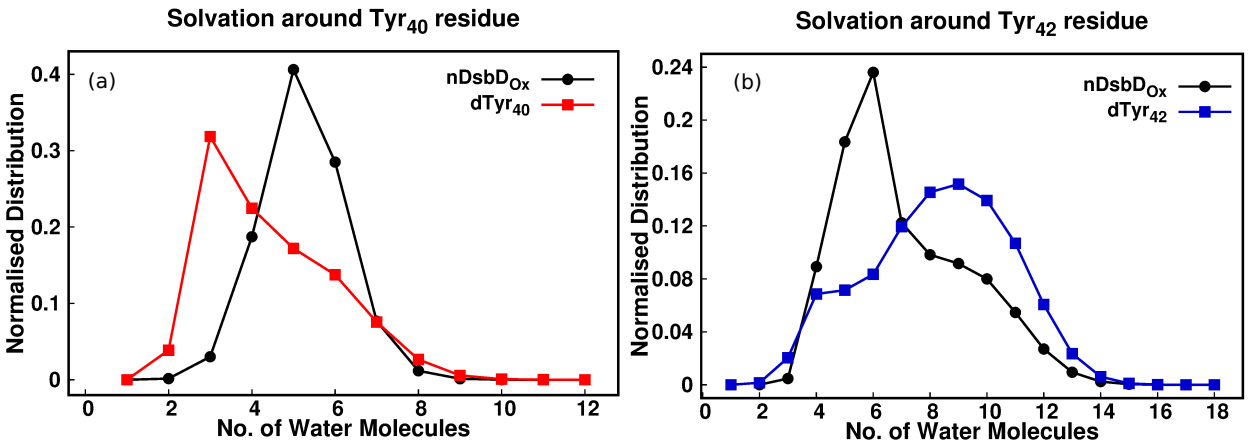


FIG. S13: No. of water molecules around (a) Tyr_{40} in $nDsbD_{Ox}$ and $dTyr_{40}$ (b) Tyr_{42} in $nDsbD_{Ox}$ and $dTyr_{42}$. Here water molecules from the first solvation shell (4Å distance) are only counted. It is well clear that Tyr_{42} is comparatively well solvated.

E. Solvation around $Cys_{103/109}$ residues

No. of water molecules around $Cys_{103/109}$ from first solvation shell (4Å distance) is counted fig. S14. It is well clear that Cys_{109} is more solvated than Cys_{103} as the later residue is buried inside the cap loop. Solvation around the disulfide is least for $dTyr_{40}$ compared to $dTyr_{42}$ and $nDsbD_{Ox}$, and this is expected as the Phe_{70} cap is in a closed state. The disulfide is maximally solvated in native $nDsbD_{Ox}$, which shows the highest cap opening events rate. The cap flexibility affects the solvation around the disulfide bond.

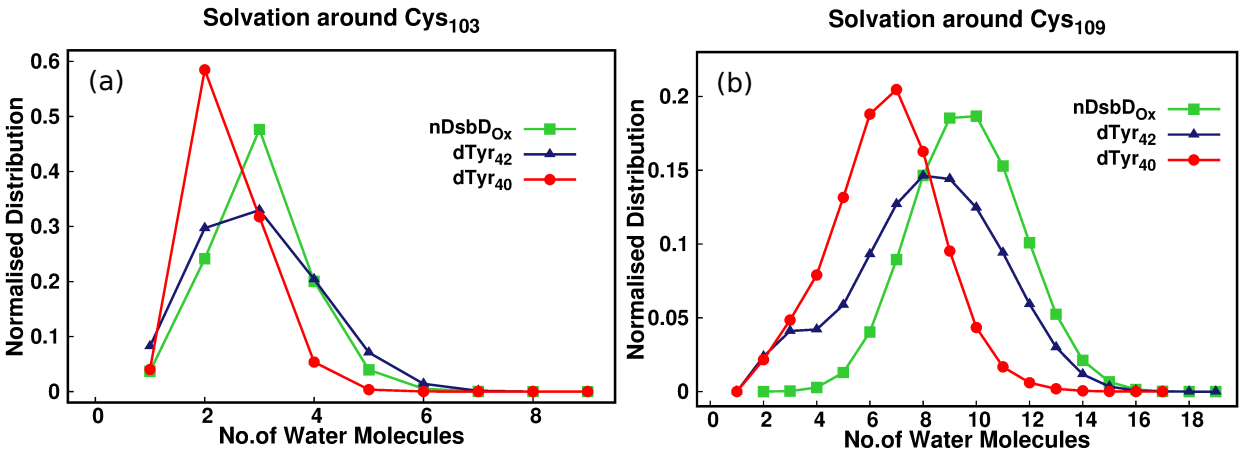


FIG. S14: No. of water molecules around (a) Cys_{103} (b) Cys_{109} in $nDsbD_{Ox}$, $dTyr_{40}$ and $dTyr_{42}$. Here water molecules within 4Å distance is only counted. It is well clear that Cys_{109} is highly solvated than Cys_{103}.

S5. SOLVENT ACCESSIBLE SURFACE AREA (SASA)

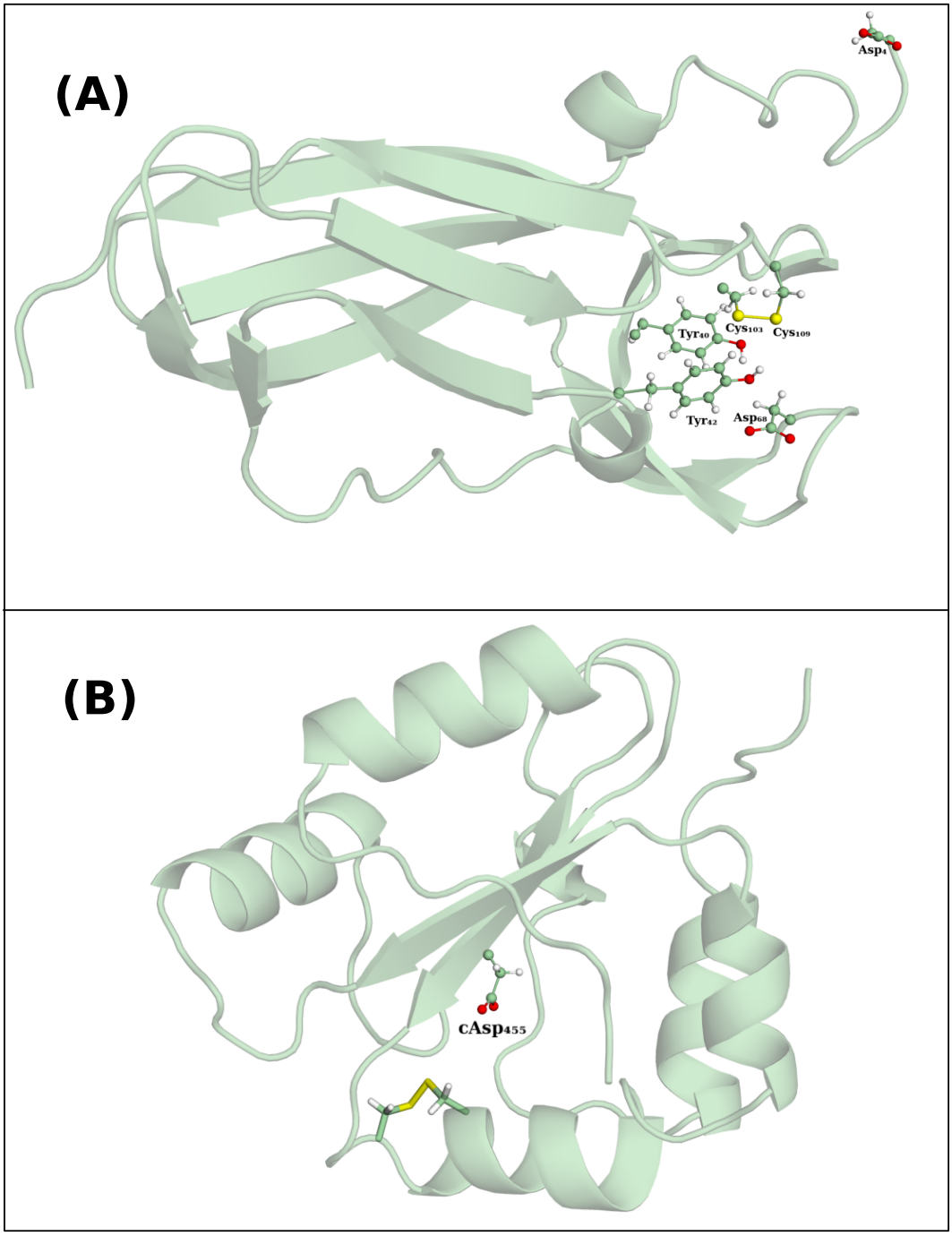


FIG. S15: Residues for which pKa values calculated for (A) $Asp_{4/68}$, $Tyr_{40/42}$ and $Cys_{103/109}$ in $nDsbD_{Ox}$ (Oxidised N-terminal DsbD) and (B) $cAsp_{455}$ in $cDsbD_{Ox}$ (Oxidised C-terminal DsbD)

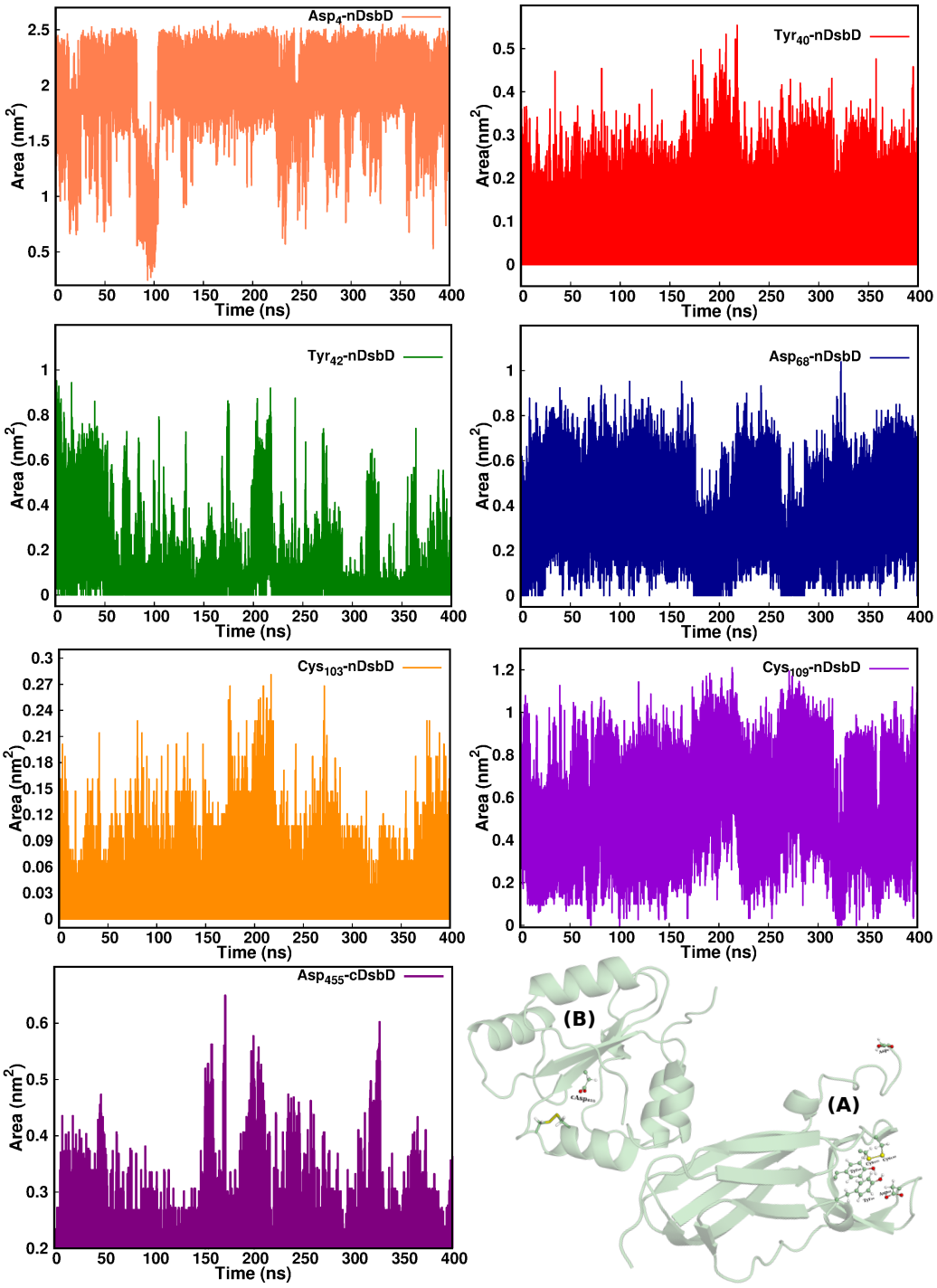


FIG. S16: Evolution of Solvent accessible surface area (SASA) with time calculated for residues $Asp_{4/68}$, $Tyr_{40/42}$ and $Cys_{103/109}$ in $nDsbD_{Ox}$(A) and $cAsp_{455}$ in $cDsbD_{Ox}$ (B). Comparison with deeply buried $cAsp_{455}$ clearly shows that $Tyr_{40/42}$, Cys_{103} are deeply buried while Asp_{68} is partially buried in $nDsbD_{Ox}$.

SASA for Residues						
Protein	**Residue**	**SASA Calculated (Å^2)**				**Nature of the residue from Exp.**[45]
		CIB-server[46]	**PDBePISA**[47]	**STRIDE**[48]	**GETAREA**[49]	
1RGG[50]	Asp1	146.63	141.53	161.70	145.70	Exposed
	Asp17	82.47	83.35	93.10	81.82	Exposed
	Asp79	12.55	12.83	15.80	13.07	Buried
	Asp84	64.26	60.67	62.90	63.63	Exposed
6LYZ[51]	Asp18	50.44	52.19	61.70	50.65	Exposed
	Asp48	85.74	83.12	81.30	83.82	Exposed
1XQ8[52]	Asp2	88.47	NA	107.30	88.50	Exposed
	Asp98	132.50	NA	153.80	132.93	Exposed
1UBQ[53]	Asp32	127.98	129.22	147.00	127.58	Exposed
	Asp39	82.71	81.38	85.30	82.30	Exposed
	Asp58	84.52	85.30	91.30	83.55	Exposed
1QKP[54, 55]	Asp96	0.00	0.00	0.40	0.00	Buried
4KQ8	Asp309	8.08	14.61	9.70	10.70	Buried
1XOA[56]	Asp26	0.00	NA	0.00	0.00	Buried
2FWE*2FWE*[57]	Asp455	4.70	7.40	4.00	5.55	Buried
1L6P	Asp4	161.63	164.83	179.10	161.05	NA
	Tyr40	3.16	4.33	4.60	3.30	NA
	Tyr42	3.71	5.70	3.80	3.38	NA
	Asp68	14.69	14.97	7.40	15.51	NA
	Cys103	0.11	0.34	0.20	0.12	NA
	Cys109	31.89	32.31	31.00	30.84	NA

TABLE S1: Comparison of SASA for buried and exposed residues in different proteins and $nDsbD_{Ox}$, calculated using available web-servers. Nature of the residues as obatined from the experiments are also included.[45] It can be seen that buried residues are having lower SASA values and hence their pKa is expected to be higher than the normal range.[58, 59] For $nDsbD_{Ox}$, there are no available experimental data, but SASA predicts an abnormally higher pKa for Asp_{68}, $Tyr_{40/42}$ and $Cys_{103/109}$ residues.

S6. SCHEME FOR DISULFIDE SCISSION BY NUCLEOPHILE

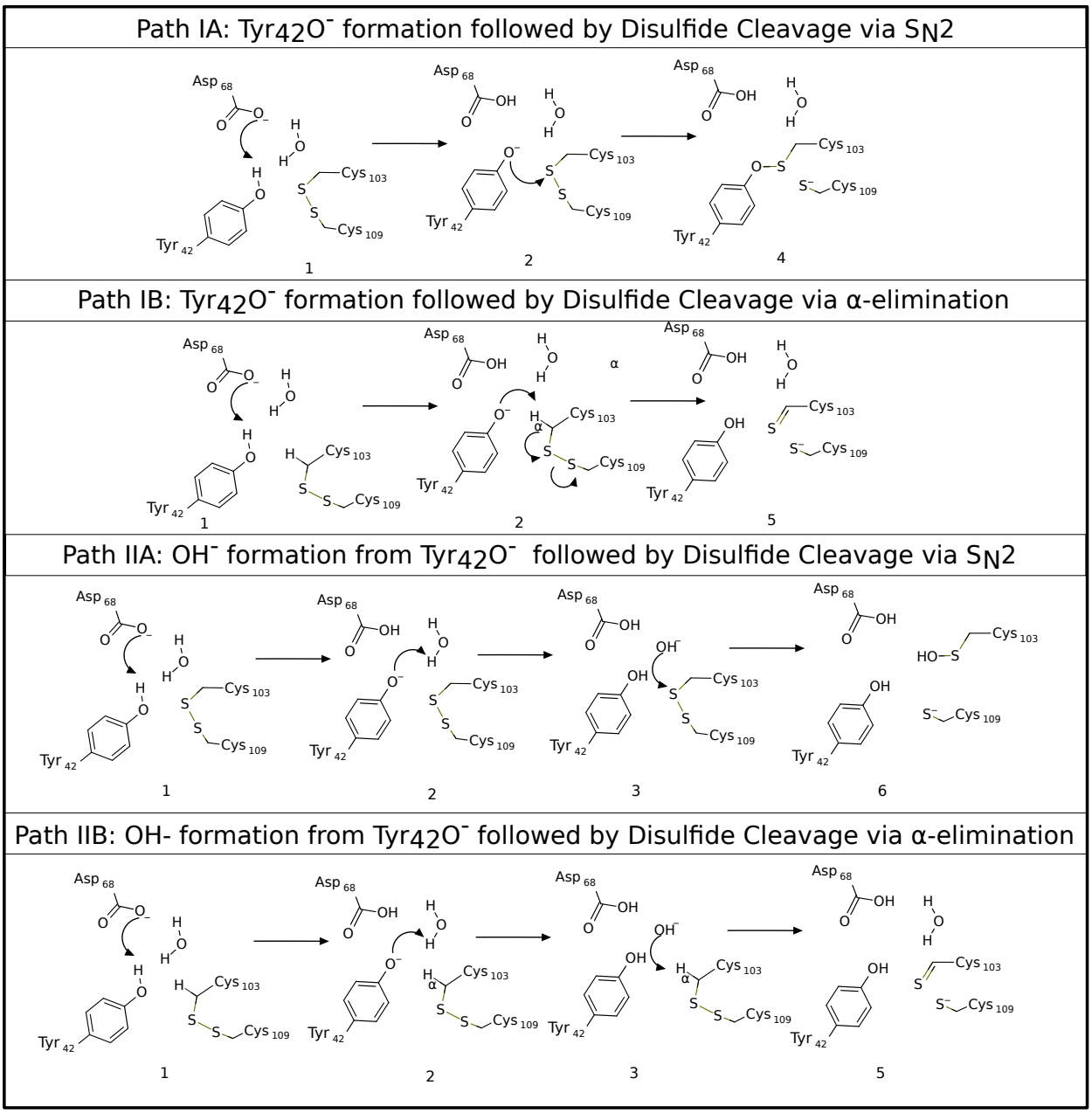


FIG. S17: Scheme for the disulfide cleavage by the nucleophile. Path IA: $Tyr_{42}O^-$ formation followed by $S_N{}^2$ mechanism. Path IB: $Tyr_{42}O^-$ formation followed by α-elimination.Path IIA: OH^- formation followed by $S_N{}^2$ mechanism. Path IIB: OH^- formation followed by α-elimination.

S7. PATH I: DIRECT PROTON TRANSFER

The nucleophile $Tyr_{42}O^-$ can be generated by the direct proton transfer from Tyr_{42} to Asp_{68} (fig. S18A). As both oxygen atoms of Asp_{68} can capture a proton from Tyr_{42}, these possibilities are included in the CV definition (fig. S18B). CV values expected for reactant and products are shown in fig. S19.

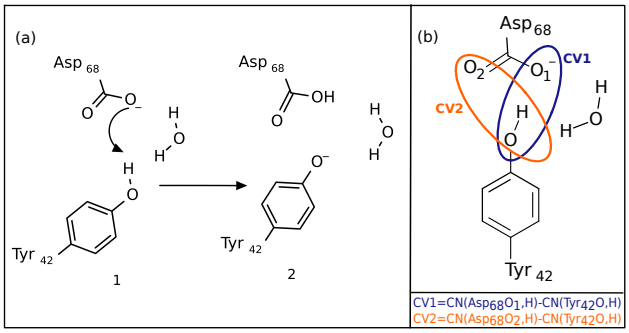


FIG. S18: (a) Reaction scheme for Path I: direct proton transfer (b) Corresponding definition for CV. Encircled in blue and orange are those atoms included in CV1 and CV2 respectively.

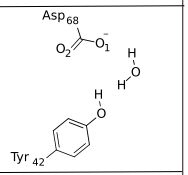
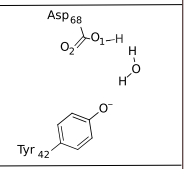
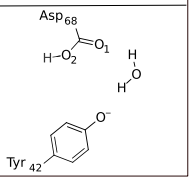
CV1	CV2	Structure
-0.5	-0.5	
+0.5	-0.5	
-0.5	+0.5	

FIG. S19: CV values and the corresponding states expected during QM/MM MTD on Path I.

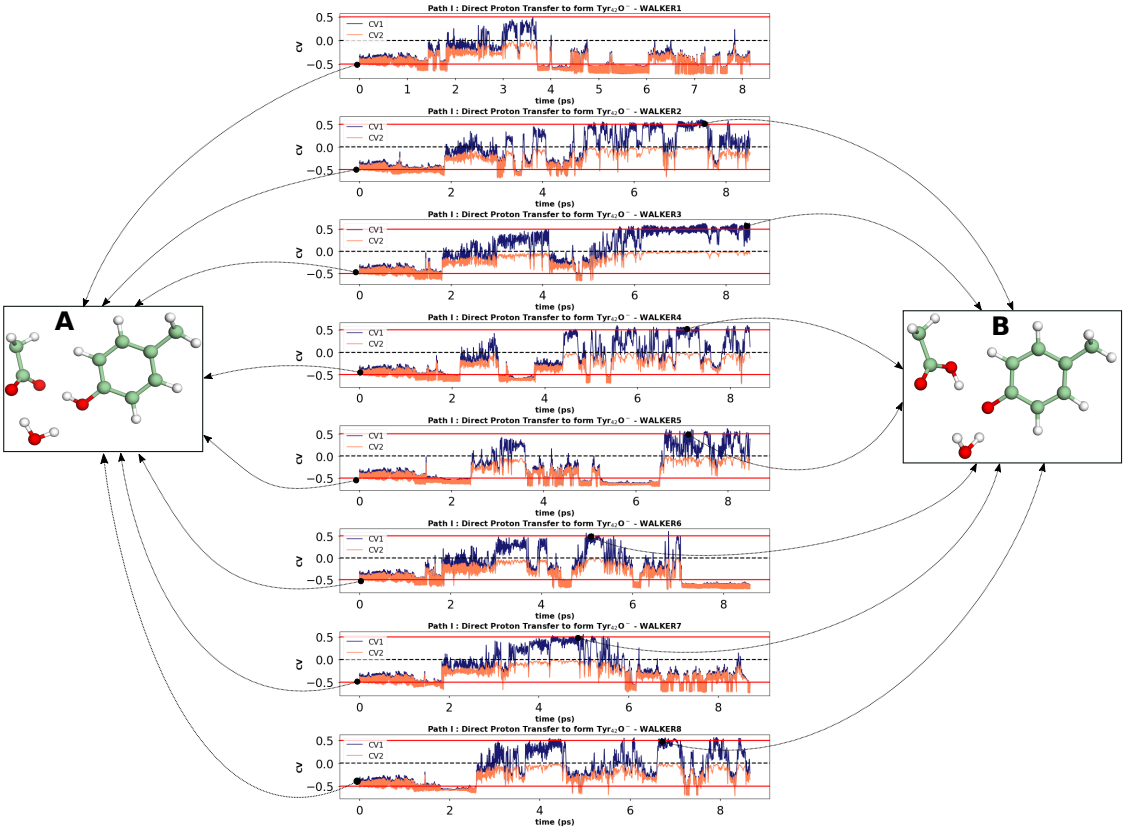


FIG. S20: Evolution of CV with time for Path I. CV1 and CV2 are shown as blue and orange lines, respectively. Images given in the inset are for the reactant (A) and the product (B) state. Formation of $Tyr_{42}O^{-}$ can be visualized along CV1.

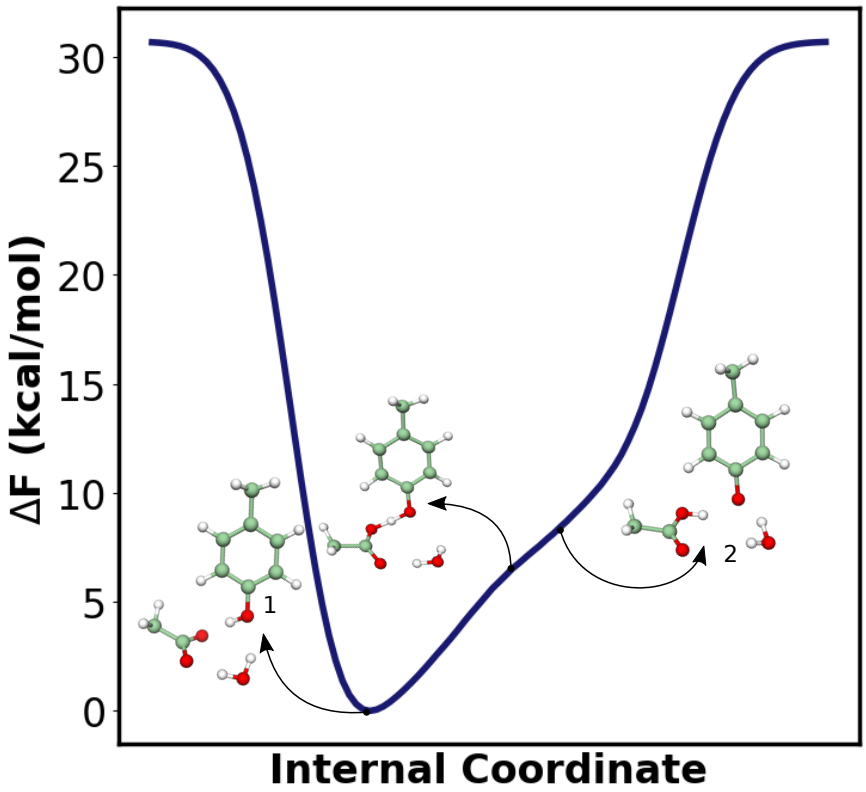


FIG. S21: Minimum Energy Path (MEP) connecting reactant and product for Path I. There is only a single minimum, and that belongs to the reactant(1). The shoulder peak represents the product (2). Here, free energy (from reactant minima to the shoulder) $\Delta F \approx 7$kcal/mol.

S8. PATH II: DIRECT PROTON TRANSFER FOLLOWED BY OH^- FORMATION

$Tyr_{42}O^-$ is solvated, and so there is a possibility that the nucleophile formed in Path I can abstract a proton from nearby water to form OH^-, the second nucleophile (fig. S22a), named as Path II. CV defined according to the scheme is given (fig. S22b). Expected CV values and corresponding structures are shown in fig. S23.

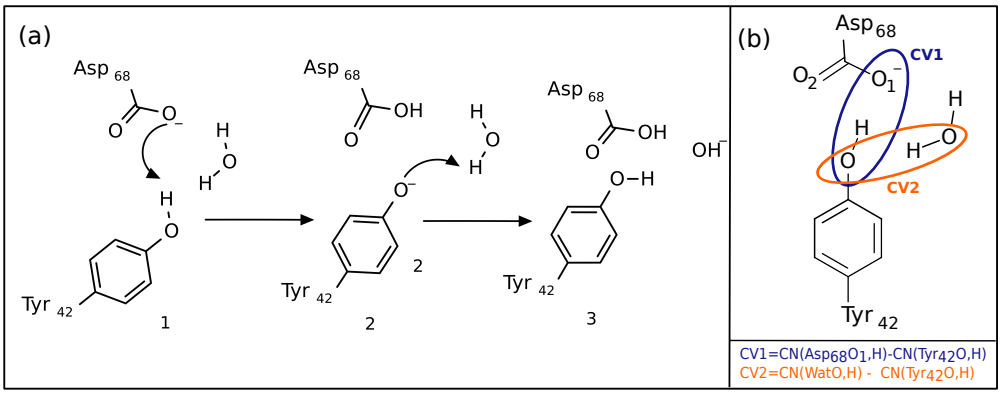


FIG. S22: (a) Reaction scheme for Path II: $Tyr_{42}O^-$ from Path I abstracts a proton from neighbouring water to form OH^-. (b) CV definition for Path II. Encircled in blue and orange are those atoms defined in CV1 and CV2, respectively.

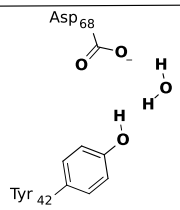
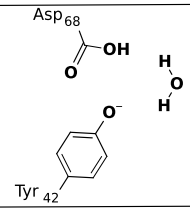
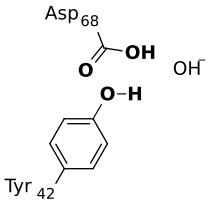
CV1	CV2	Structure
-0.5	-0.5	
+0.5	-0.5	
+0.5	+0.5	

FIG. S23: Values expected for CV in the reactant and product states for Path II.

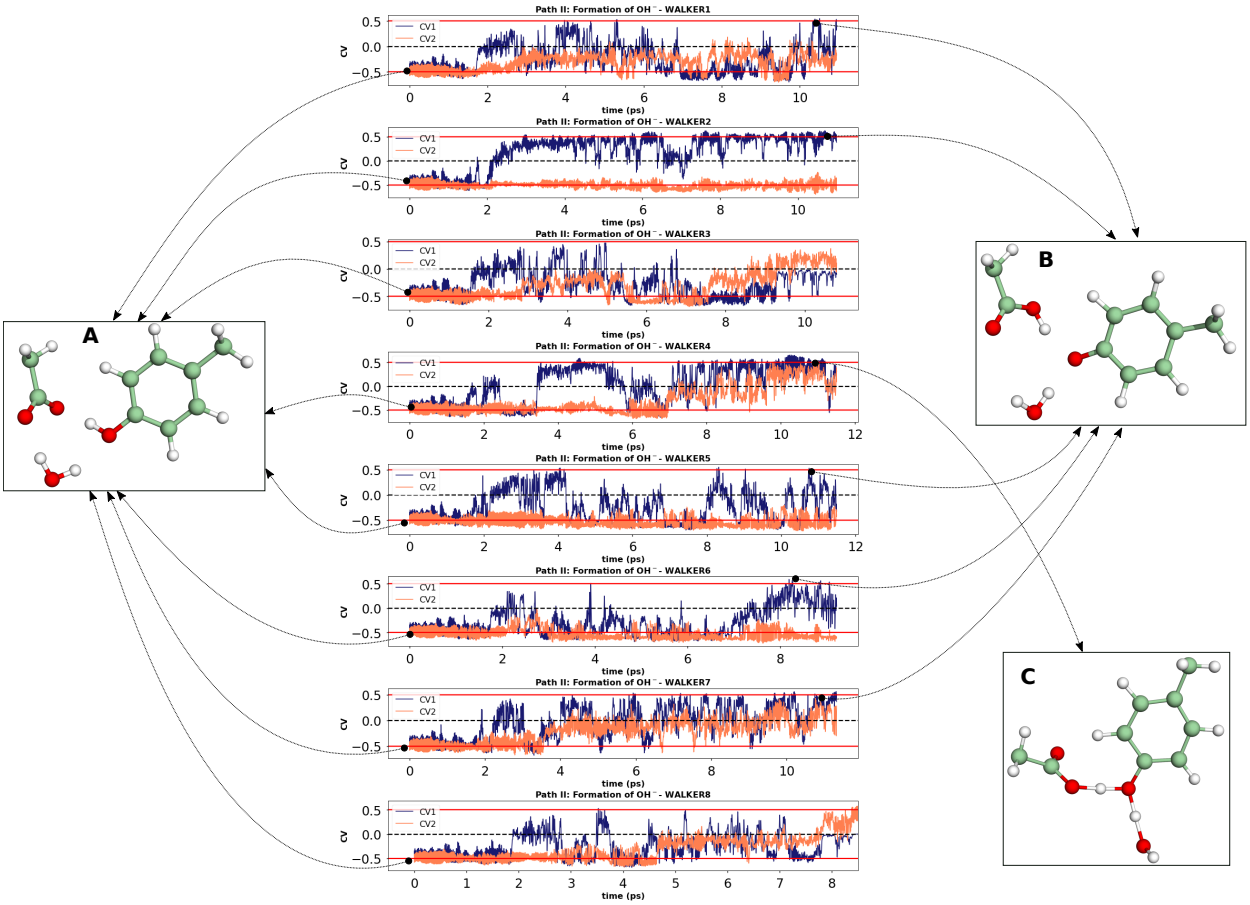


FIG. S24: Time evolution of CV for Path II. CV1 and CV2 are represented by blue and orange lines, respectively. Images in the insight represent reactant (A) and product states for a given CV value. Clearly, five of the walkers show $Tyr_{42}O^-$ (B) formation only, while walker 4 shows the OH^- formation (C), but as seen, it has only transient existence.

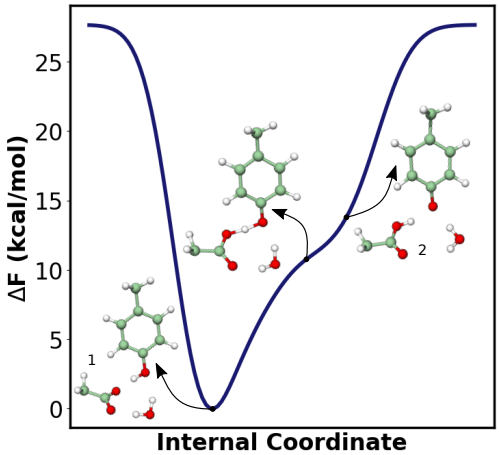


FIG. S25: Minimum Energy Path (MEP) for Path II resembles Path I. Formation of OH^- is not identified in MEP. Here, $\Delta F \approx 10$kcal/mol.

S9. $TYR_{42}O^-$ ATTACK ON DISULFIDE

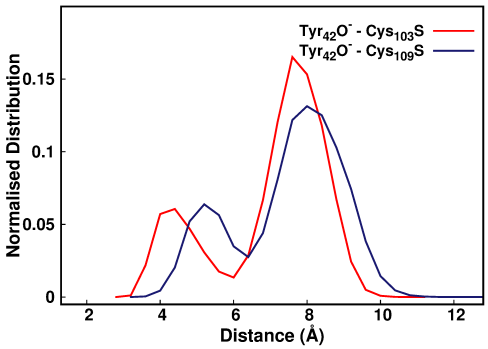


FIG. S26: Normalized distribution for the distance between $Tyr_{40}O^-$-$Cys_{103}S$ and $Tyr_{42}O^-$ - $Cys_{109}S$.

S10. STATIC CALCULATION ON DISULFIDE CLEAVAGE

A. Static Calculation on Path IA: S_N2 by $TyrO^-$

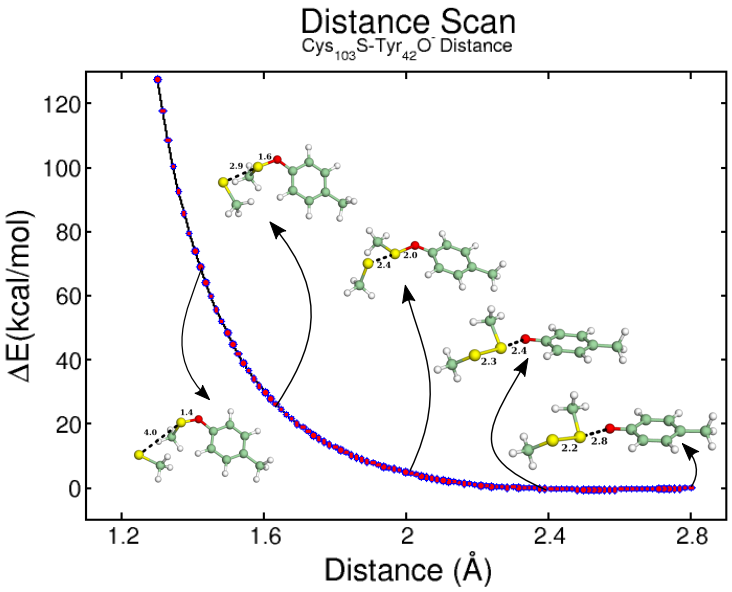


FIG. S27: QM Calculation for Path IA:$Tyr_{42}O^-$ formation followed by S_N2 involved disulfide cleavage

31

C 32.253232416261 87 44.936748477056 08 39.303413881281 2580 C 24.586594554886 88 50.524408655496 6 42.482299916492 52 C 30.857404207092 73 43.726915803768 41 39.859428779511 15
H 31.598557838009 44 44.193915446003 37 38.908347739741 56 H 23.959929560656 43 49.612875061328 56 42.578993702798 65 H 30.146935293833 96 43.197108027749 25 39.201210972886 09
H 33.062439466757 29 45.186710512796 87 38.593041622358 6 H 24.493206940009 63 51.073889046882 27 43.436790798497 27 H 31.878163676122 21 43.473984345984 99 39.521726549932 26
C 31.485541534238 79 46.176448684402 91 39.725580159752 28 C 26.039231077250 95 50.216877935420 70 42.163521660332 46 C 30.618168534551 53 45.227323643143 52 39.843405077948 83
C 30.167200281892 81 46.095626846817 01 40.233575109474 96 C 26.465404682422 39 49.950747501311 85 40.841219938616 70 C 29.359678579452 96 45.766866933363 43 39.501952200928 12
H 29.678295916835 42 45.112316334882 13 40.289429163623 59 H 25.729661638202 79 49.996825480270 7 40.025725202771 10 H 28.547274897504 03 45.089302500526 70 39.213508528093 42
C 29.469080765412 37 47.227121518691 26 40.664875538857 03 C 27.795972275921 88 49.638344213892 92 40.543785809588 38 C 29.124699720742 54 47.148195425753 90 39.519285578407 73
H 28.448445895386 82 47.141131668742 18 41.051880212683 07 H 28.110453934820 11 49.448358137210 27 39.511731020200 8 H 28.149278856874 94 47.559372903416 86 39.246120030901 23
C 30.047847040269 54 48.549879479074 03 40.621923746223 87 C 28.804465769082 39 49.563034283590 76 41.564693281564 89 C 30.153828756113 80 48.050260506203 51 39.886617678078 48
O 29.362494515419 30 49.581280291778 21 41.043876021954 30 O 30.056359998653 46 49.289489999655 67 41.262690000587 41 O 29.856047881906 18 49.375308000464 97 39.889693190635 38
C 31.386053178216 38 48.613365132524 78 40.089322070763 93 C 28.356047441037 44 49.823905776215 54 42.906199972763 52 C 31.425210117384 55 47.521614631367 70 40.228580785434 76
H 31.858554917691 72 49.595123637574 33 40.012886298243 49 H 29.084822765384 71 49.764427038274 89 43.710373200874 82 H 32.227271515775 73 48.202557080870 96 40.523081283009 27
C 32.063911453746 42 47.463871772890 88 39.664554571576 23 C 27.020358564713 61 50.146797257016 57 43.177754526164 28 C 31.636987941865 75 46.135600699116 80 40.201955220118 62
H 33.082969095812 13 47.566146936666 42 39.267044875791 46 H 26.723145633303 12 50.350439690239 44 44.215837084939 96 H 32.626617738039 69 45.748494061015 22 40.471852424468 38
C 31.814571437822 07 50.375256204372 81 43.346085132728 30 C 31.092867581483 97 52.133608556450 06 42.264521712581 68 C 33.787399793910 86 50.754054646184 97 43.002736123170 28
H 32.713393688196 42 51.004904918444 32 43.241964109447 34 H 31.699453018436 34 52.998839358460 39 42.610374853231 01 H 33.418654717141 58 51.470358603761 10 43.771029637748 28
C 30.548848625737 71 51.168594560275 65 43.000126212797 02 C 31.670190000397 17 50.750899999988 04 42.607650000016 95 C 32.699583612421 49 50.335478785815 46 42.041466009380 54
H 30.451138642944 67 52.061383502667 6 43.637632834611 69 H 32.721467857718 74 50.689918060028 06 42.273207779281 60 H 31.922051760715 73 49.681512313496 21 42.465938432508 97
H 29.653242680437 3 50.541567966741 90 43.101867717855 67 H 30.872660002292 36 49.801749999992 77 42.042979997919 53 H 30.676674082883 85 49.937928330669 64 40.103898883772 81
S 30.458954491606 12 51.738336140965 25 41.223480147708 98 S 31.749919868053 03 50.652680709867 18 44.505028073140 09 S 32.392161119806 90 51.132809365177 62 40.601695369352 30
C 34.014519494891 45 52.286489658628 31 40.959768925518 47 C 35.134638871107 11 51.437544815303 34 43.911514178071 46 C 33.037780035443 82 54.660379216187 27 42.330502994949 71
H 33.484718873713 14 51.325406650858 81 41.041942940520 44 H 34.256423650240 27 52.022402526239 08 43.597217674100 95 H 33.220485137577 07 53.681300970215 74 42.801932166234 76
C 33.063790657950 68 53.367410092593 32 40.427529805600 30 C 34.801658696640 59 50.629777395294 55 45.172365323339 92 C 32.635302294610 73 54.474354936670 51 40.855583927465 48
H 32.703454967354 18 53.095402803003 09 39.421404847713 7 H 34.544703412574 82 51.299783837280 94 46.009226692507 02 H 33.458699869420 77 53.981042977418 53 40.307987411333 83
H 33.580312025732 48 54.340470034875 33 40.339661422123 85 H 35.656183633868 76 50.002070624983 82 45.487655766385 56 H 32.478172362540 78 55.463612247968 50 40.385076848882 50
S 31.577092851042 32 53.702635478270 81 41.500099115296 21 S 33.393444496621 85 49.414553561926 65 44.951566188442 41 S 31.093451106582 90 53.467668536112 59 40.622724145232 30
H 31.768849432962 69 50.015995530911 55 44.392824045621 11 H 30.994101627070 14 52.197852338063 79 41.164803300971 90 H 34.197274976003 69 49.884336755472 01 43.555093982800 08
H 34.876280596581 44 52.149166714505 82 40.278408171632 90 H 35.968726377653 82 52.138888435449 54 44.105235402349 00 H 33.960188248821 20 55.271520554662 50 42.428684841094 55
H 32.733174474888 23 44.409836514718 53 40.155674081203 51 H 24.113749293927 54 51.144930293324 1 41.696894911088 28 H 30.741740547167 0 43.293663212726 82 40.872869462353 39
H 31.912989562328 67 49.503043235006 41 42.682773796836 55 H 30.073474372605 20 52.253863265197 7 42.671201687896 14 H 34.616129210834 31 51.252976117271 26 42.470657872777 2
H 34.397412999248 64 52.555023249654 18 41.957946250584 63 H 35.418549614482 83 50.772913874641 18 43.079413011684 45 H 32.229819628588 73 55.156559867740 16 42.895735356091 97

S11. STABILIZATION OF $TYR_{42}O^-$

As it can be seen that a water wire facilitates proton transfer between far away $Tyr_{42}O^-$ and Asp_{68} QM/MM metadynamics simulations were conducted. $Tyr_{42}O^-$ abstracts a proton from nearby water to form OH^-, later OH^- propagates through water network to take a proton from $Asp_{68}OH$ to finally form Asp_{68} as the product (fig. S28a). Definition of CV based on the reaction scheme is given in (fig. S28b).

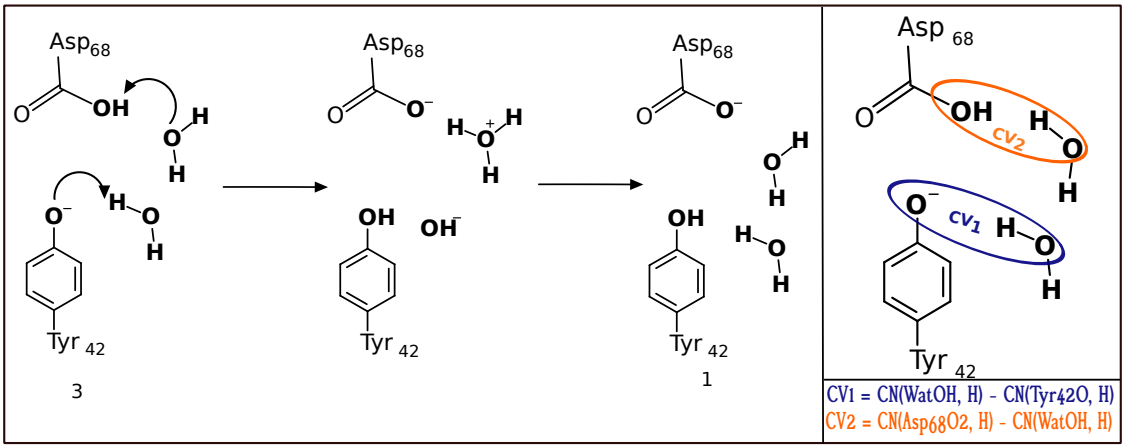


FIG. S28: (a) Reaction scheme for solvent assisted proton transfer between $Tyr_{42}O^-$, $Asp_{68}OH$ to form Tyr_{42} and Asp_{68}. (b) Definition of CV as per the scheme. CV1 shows abstraction of a proton by $Tyr_{42}O^-$ from water to form Tyr_{42} , OH^- and CV2 represents the donation of a proton to nearby water by $Asp_{68}OH$ to finally form Asp_{68}.

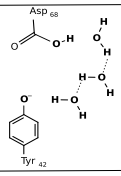
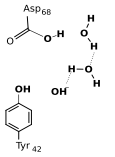
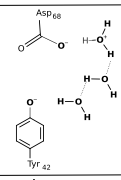
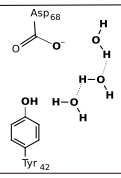
CV1	CV2	Structure
-0.5	-0.5	
+0.5	-0.5	
-0.5	+0.5	
+0.5	+0.5	

FIG. S29: Values expected for CV in the reactant and product states for solvent assisted proton transfer between $Tyr_{42}O^{-}$ and $Asp_{68}OH$.

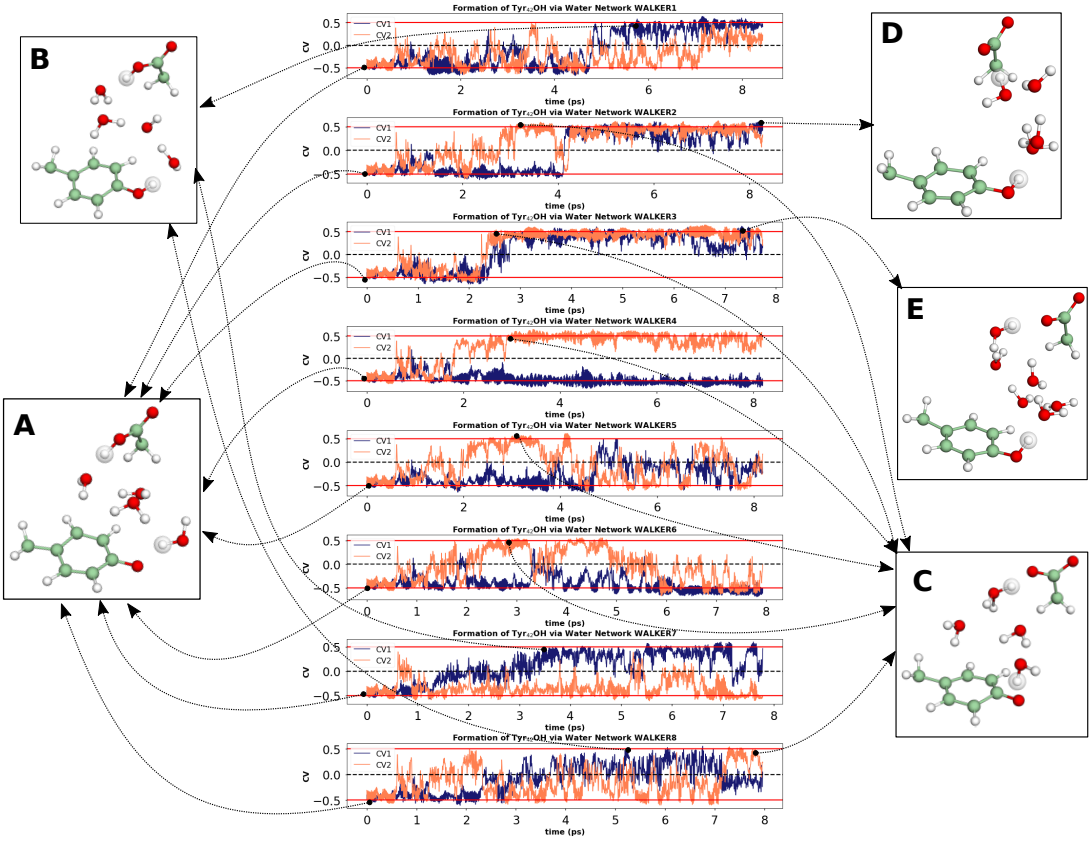


FIG. S30: Evolution of CVs' with time for $Tyr_{42}O^-$ stability. Blue lines represent CV1 and orange CV2. Structure for reactant(2) here is **A**. To distinguish hydrogen atoms involved in the CVs, they are highlighted. **B** shows the formation of OH^- with positive CV1. Here, walkers: 1, 7, and 8 form OH^- only. When CV2 alone is positive, it indicates the formation of H_3O^+ (**C**), which is formed by the majority of the walkers. As it can be seen for walkers: 2 and 3, both CV1 and CV2 have +0.5, forming product(1) (**D**) - four water network and (**E**) - six water network.

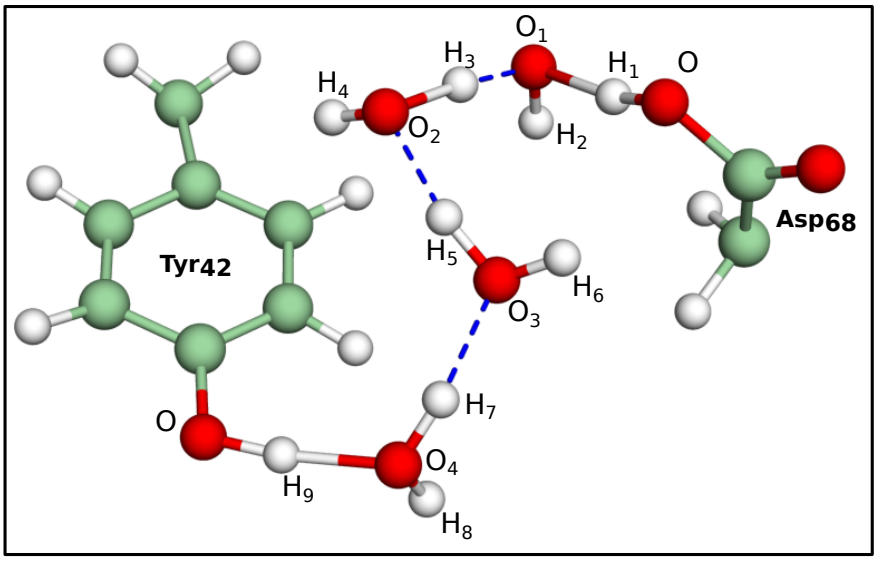


FIG. S31: Snapshot showing four water network.

S12. REFERENCES

[1] H. M. Berman, Nucleic Acids Res. **28**, 235 (2000).

[2] Schrödinger, LLC, "The PyMOL molecular graphics system, version 1.8," (2015).

[3] E. Lindahl, B. Hess, and D. van der Spoel, J. Mol. Model. **7**, 306 (2001).

[4] M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, and E. Lindahl, SoftwareX **1-2**, 19 (2015).

[5] W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, J. Chem. Phys. **79**, 926 (1983).

[6] K. Lindorff-Larsen, S. Piana, K. Palmo, P. Maragakis, J. L. Klepeis, R. O. Dror, and D. E. Shaw, Proteins: Struct. Funct. Bioinf. **78**, 1950 (2010).

[7] B. Hess, H. Bekker, H. J. Berendsen, and J. G. Fraaije, J. Comp. Chem. **18**, 1463 (1997).

[8] T. Darden, D. York, and L. Pedersen, J. Chem. Phys. **98**, 10089 (1993).

[9] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, J. Chem. Phys. **103**, 8577 (1995).

[10] W. Humphrey, A. Dalke, and K. Schulten, J Mol Graph. **14**, 33 (1996).

[11] D. Case, I. Ben-Shalom, S. Brozell, D. Cerutti, T. Cheatham, III, V. Cruzeiro, T. Darden, R. Duke, D. Ghoreishi, M. Gilson, H. Gohlke, A. Goetz, D. Greene, R. Harris, N. Homeyer, Y. Huang, S. Izadi, A. Kovalenko, T. Kurtzman, T. Lee, S. LeGrand, P. Li, C. Lin, J. Liu, T. Luchko, R. Luo, D. Mermelstein, K. Merz, Y. Miao, G. Monard, C. Nguyen, H. Nguyen, I. Omelyan, A. Onufriev, F. Pan, R. Qi, D. Roe, A. Roitberg, C. Sagui, S. Schott-Verdugo, J. Shen, C. Simmerling, J. Smith, R. SalomonFerrer, J. Swails, R. Walker, J. Wang, H. Wei, R. Wolf, X. Wu, L. Xiao, D. York, and P. Kollman, AMBER 2018, University of California, San Francisco (2018).

[12] C. U. Stirnimann, A. Rozhkova, U. Grauschopf, R. A. Böckmann, R. Glockshuber, G. Capitani, and M. G. Grütter, J. Mol. Biol. **358**, 829 (2006).

[13] G. Bussi, D. Donadio, and M. Parrinello, J. Chem. Phys. **126**, 014101 (2007).

[14] M. Parrinello and A. Rahman, Int. J. Appl. Phys. **52**, 7182 (1981).

[15] D. Marx and J. Hutter, Ab Initio Molecular Dynamics: Basic Theory and Advanced Methods (Cambridge University Press, 2009).

[16] J. Hutter, M. Iannuzzi, F. Schiffmann, and J. VandeVondele, Wiley Interdiscip. Rev. Comput. Mol. Sci. **4**, 15 (2013).

[17] B. G. Lippert, J. H. Parinello, and Michele, Mol. Phys. **92**, 477 (1997).

[18] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[19] C. Lee, W. Yang, and R. G. Parr, Phys. Rev. B **37**, 785 (1988).

[20] S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, J. Chem. Phys. **132**, 154104 (2010).

[21] S. Grimme, S. Ehrlich, and L. Goerigk, Journal of Computational Chemistry **32**, 1456 (2011).

[22] J. VandeVondele and J. Hutter, J. Chem. Phys. **127**, 114105 (2007).

[23] S. Goedecker, M. Teter, and J. Hutter, Phys. Rev. B **54**, 1703 (1996).

[24] C. Hartwigsen, S. Gœdecker, and J. Hutter, Phys. Rev. B **58**, 3641 (1998).

[25] J. Hostaš and J. Řezáč, J. Chem. Theory and Comput. **13**, 3575 (2017).

[26] P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka, and D. Marx, Nat. Chem. **9**, 164 (2017).

[27] P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka, and D. Marx, Angew. Chem. Int. Ed. **55**, 1304 (2016).

[28] P. Dopieralski, J. Ribas-Arino, P. Anjukandi, M. Krupicka, J. Kiss, and D. Marx, Nat. Chem. **5**, 685 (2013).

[29] J. A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K. E. Hauser, and C. Simmerling, J. Chem. Theory and Comput. **11**, 3696 (2015).

[30] F. Maseras and K. Morokuma, J. Comp. Chem. **16**, 1170 (1995).

[31] N. Reuter, A. Dejaegere, B. Maigret, and M. Karplus, J. Phys. Chem. A **104**, 1720 (2000).

[32] M. J. Field, P. A. Bash, and M. Karplus, J. Comp. Chem. **11**, 700 (1990).

[33] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, Journal of Chemical Theory and Computation **1**, 1176 (2005).

[34] T. Laino, F. Mohamed, A. Laio, and M. Parrinello, J. Chem. Theory and Comput. **2**, 1370 (2006).

[35] S. Nosé, J. Chem. Phys. **81**, 511 (1984).

[36] S. Nose, Mol. Phys. **52**, 255 (1984).

[37] A. Laio and M. Parrinello, Proc. Natl. Acad. Sci. **99**, 12562 (2002).

[38] P. Raiteri, A. Laio, F. L. Gervasio, C. Micheletti, and M. Parrinello, **110**, 3533 (2006).

[39] W. J. Hehre, R. Ditchfield, and J. A. Pople, J. Chem. Phys. **56**, 2257 (1972).

[40] F. Neese, WIREs Comput. Mol. Sci **8**, e1327 (2018).

[41] P. W. Haebel, D. Goldstone, F. Katzen, J. Beckwith, and P. Metcalf, EMBO J. **21**, 4774 (2002).

[42] C. W. Goulding, M. R. Sawaya, A. Parseghian, V. Lim, D. Eisenberg, and D. Missiakas, Biochemistry **41**, 6920 (2002).

[43] D. A. Mavridou, E. Saridakis, P. Kritsiligkou, A. D. Goddard, J. M. Stevens, S. J. Ferguson, and C. Redfield, J. Biol. Chem. **286**, 24943 (2011).

[44] L. S. Stelzl, D. A. Mavridou, E. Saridakis, D. Gonzalez, A. J. Baldwin, S. J. Ferguson, M. S. Sansom, and C. Redfield, Elife **9**, e54661 (2020).

[45] G. R. Grimsley, J. M. Scholtz, and C. N. Pace, Prot. Sci. **18**, 247 (2009).

[46] "Centre for Information Biology,Ochanomizu Universityaccessible surface area and accessibility calculation for protein,2012," `http://cib.cf.ocha.ac.jp/bitool/ASA/`, accessed:2022-02-5.

[47] E. Krissinel and K. Henrick, J. Mol. Biol. **372**, 774 (2007).

[48] M. Heinig and D. Frishman, Nucleic Acids Res. **32**, W500 (2004).

[49] R. Fraczkiewicz and W. Braun, J. Comp. Chem. **19**, 319 (1998).

[50] J. Sevcik, Z. Dauter, V. Lamzin, and K. Wilson, **52**, 327 (1996).

[51] R. Diamond, J. Mol. Biol. **82**, 371 (1974).

[52] T. S. Ulmer, A. Bax, N. B. Cole, and R. L. Nussbaum, J. Biol. Chem. **280**, 9595 (2005).

[53] S. Vijay-Kumar, C. E. Bugg, and W. J. Cook, J. Mol. Biol. **194**, 531 (1987).

[54] K. Edman, P. Nollert, A. Royant, H. Belrhali, E. Pebay-Peyroula, J. Hajdu, R. Neutze, and E. M. Landau, Nature **401**, 822 (1999).

[55] T. K. Harris and G. J. Turner, IUBMB Life **53**, 85 (2002).

[56] M.-F. Jeng, A. P. Campbell, T. Begley, A. Holmgren, D. A. Case, P. E. Wright, and H. J. Dyson, Structure **2**, 853 (1994).

[57] C. U. Stirnimann, A. Rozhkova, U. Grauschopf, R. A. Böckmann, R. Glockshuber, G. Capitani, and M. G. Grütter, J. Mol. Biol. **358**, 829 (2006).

[58] A. K. Shaytan, K. V. Shaitan, and A. R. Khokhlov, Biomacromolecules **10**, 1224 (2009).

[59] T. Meyer, G. Kieseritzky, and E.-W. Knapp, Proteins: Struct. Funct. Genet. **79**, 3320 (2011).