

University of Nottingham
UK | CHINA | MALAYSIA

Enhanced Sampling Methods for the Simulation of Slow Biomolecular Processes

Zhicheng He

School of Pharmacy

Thesis submitted to the University of Nottingham for the degree of

Doctor of Philosophy

September 2025

Supervisors: Charles Anthony Laughton, Shailesh Mistry, Mattea Finelli,

Phillip Micheal Williams

Abstract

Molecular dynamic (MD) simulations are a powerful tool to study biomolecular dynamics. However, many biologically important processes, such as ligand dissociation and protein unfolding, occur on timescales that exceed the capability of conventional MD simulations. This thesis applies enhanced-sampling methods to investigate (i) the dissociation pathways and kinetics of selective β-blockers at the β2-adrenergic receptor (β2-AR), and (ii) the effect of S-nitrosylation (SNO) on the FF↔LU conformational equilibrium of peroxiredoxin-2 (Prdx2).

For β2-AR, the detailed dissociation mechanisms of atenolol, bisoprolol, and cyanopindolol were previously unknown. Using τRAMD, we identify two principal egress routes: the extracellular vestibule (ECV) and a secondary "keyhole" pathway. Atenolol and cyanopindolol preferentially exit via the ECV, whereas bisoprolol dissociates via both pathways. Beyond conventional τRAMD analysis, a new force-extrapolation framework was established to estimate the zero-force dissociation rate (k_{off}^0). Across ligands, extrapolated k_{off}^0 values consistently exceed experiment by ~$10^{5.4}$, collapsing onto a calibration factor. Committor analysis locates mixed-outcome configurations along cyanopindolol pathways as candidate transition-state ensembles.

For Prdx2, although both the peroxidatic (C_P) and resolving (C_R) cysteines undergo SNO, its effect on FF↔LU conformational equilibrium was unknown. Principal-component analysis shows that the interconversion is too rare to be sampled by standard MD. By representing each conformational basin with its lowest-free-energy substate, alchemical free-energy perturbation reveals that SNO at C_R drives a stronger stabilization of the LU state than SNO at C_P. The SNO cumulatively shifts the equilibrium toward the LU state by +2.4 kcal/mol (~60-fold). We propose a mechanistic rationale that the C_R is S-nitrosylated by small NO^+ donor molecules. Finite-temperature string suggests the most probable FF→FU transition pathways under different modification states (FU is an intermediate state between FF and LU).

Together, these studies establish two frameworks: (i) a τRAMD-based protocol for ligand dissociation pathway and k_{off}^{0} estimation, and (ii) an alchemical free-energy strategy for quantifying the impact of post-translational modifications (PTMS) on protein dynamics. To our knowledge, this is the first application of free-energy perturbation to explore the conformational effects of site-specific PTMs. This thesis provides novel mechanistic insights into β-blocker dissociation at β2-AR and the role of SNO in Prdx2's dynamics.

Acknowledgements

I would like to express my sincere gratitude to my supervisors. I have always perceived Professor Charles Laughton as my role model. His deep understanding in the field, pragmatic spirit, honesty, and kindness exemplify me what a researcher should be. To Professor Phillip Williams, who opened the gate for me to approach biological processes from a physical perspective. To Dr Mattea Finelli, for giving me the valuable opportunity to write a publication review. I am also grateful to my assessor, Dr Naoto Hori, for his generous support and for recommending books.

My thanks also go to the members of A32, whose companionship has enriched my journey. I am especially grateful to my parents for their unconditional support. It is a privilege to purse a PhD without funding.

I would also like to express my special thanks to the difficulties I encountered in both research and life. It is the difficulties that make me stronger. Lastly, I want to thank myself, the one who is always confident and brave enough to keep denying myself and growing up from these denials.

Publications related to this thesis

Protein Oxidative Modifications in Neurodegenerative Diseases: From Advances in Detection and Modelling to Their Use as Disease Biomarkers

Authors: Sandra I Anjo[†], Zhicheng He[†], Zohaib Hussain, Aruba Farooq, Alan McIntyre, Charles A Laughton, Andreia Neves Carvalho, and Mattéa J Finelli

Status: Published in *Antioxidants*

(Anjo et al., 2024)[1]

Abbreviations

AC	adenylyl cyclase
AFM	atomic force microscopy
ASP	Active Site Pressurization
ATE	atenolol
BIS	bisoprolol
COM	centre of mass
COPD	chronic obstructive pulmonary disease
C_P	peroxidatic cystine
C_R	resolving cystine
CV	collective variable
CYA	cyanopindolol
ECV	extra cellular vestibule
FF state	fully folded state
FRET	Förster Resonance Energy Transfer
FTS	finite-temperature string
FU	intermediate state contributed from both states
GDP	Guanosine Diphosphate
GPCR	G protein-coupled receptor
GTP	Guanosine Triphosphate
H_2O_2	hydrogen peroxide
H4	Helix 4
HMW	High Molecular Weight
LU	local unflolding state
MBAR	Multi-State Bennett's Acceptance Ratio
NMDAR	N-methyl-D-aspartate receptor
NO	nitric oxide

NOSs	nitric oxide synthases
NOXs	NADPH oxidases
NPT	constant-pressure, constant-temperature ensemble
NVT	constant number, volume and temperature ensemble
PCA	Principal Component Analysis
PD	Parkinson's disease
PDB	Protein Data Bank
PKA	protein kinase A
PME	Particle Mesh Ewald
PTM	post-translational modification
RMSD	Root Mean Square Deviation
RNS	reactive nitrogen species
RONS	reactive oxygen and nitrogen species
ROS	reactive oxygen species
SH-Cys	thiol cysteine
SNO	S-nitrosylation
SNOC	S-nitrosocysteine
SO_2H	sulfinic acid cysteine
SOH	sulfenic acid cysteine
SS	disulfide bond
TI	Thermodynamic Integration
TM	transmembrane hehix
WHAM	weighted histogram analysis
β-AR	β-adrenergic receptor
ΔG_{FF}	ΔG bewteen two oxidation states in the FF state

ΔG_{LU}	ΔG bewteen two oxidation states in the LU state
ΔG_{SH}	ΔG_{SH} bewteen FF and LU
ΔG_{SNO}	ΔG_{SNO} bewteen FF and LU
ΔΔG	relative free energy change
$\Delta\Delta G^{equil}$	ΔG difference
τRAMD	τ-Random Acceleration Molecular Dynamics

List of Figures

Figure 1.1.1: Double-well energy schematic diagram along reaction coordinate x. The energy refers to the total energy approximating to the sum of potential and kinetic energy. Each well represents a state (A or B), with point 'a' indicating a single configuration in state A. The transition state C corresponds to the highest energy point between wells. The equilibrium constant (K_{eq}) between A and B is determined by the two-state energy gap. The forward reaction rate constant (k_f) is determined by the energy gap between A and C. The reverse reaction rate constant (k_r) is determined by the energy gap between B and C. 3

Figure 1.1.2: The structures of Salmeterol and Adrenaline 11

Figure 1.2.1: Thermodynamic cycle used to determine the effect of SNO on Prdx2's conformational equilibrium between the FF and LU state. The green circle represents the FF state, while the green loop represents the LU state. The SH-Cys is shown in blue, and the SNO-Cys in pink. By making use of the state function property of free energy, the ΔΔG can be calculated by deriving the difference between ΔG_{FF} and ΔG_{LU}. Source: Scarabelli *et al.*, 2021.....21

Figure 1.2.2: A diagram of the λ-dependent transformation from thiol to S-nitrosothiol. At λ = 0, the system is a fully interacting thiol (depicted in light blue), while at λ = 1, it is a fully interacting S-nitrosothiol (depicted in pink). The λ intermediate states exhibit hybrid properties and are illustrated with blended colors contributed from the two colors. 22

Figure 1.2.3: Diagram of the string optimization process. The red line indicates the string. The green circles mark the instantaneous positions of the images along the string. For each image, an unbiased MD simulation is performed to explore its local environment in phase space, illustrated by the large blue

circle. The red circle within each blue circle represents the averaged position of the image after sampling. The yellow circle denotes the saddle point on the underlying free-energy surface. Source: Kulshrestha *et al.*, 2022. 26

Figure 1.3.1: GPCR structure. The agonist binds to the extracellular region, and the $G_{\alpha\beta\gamma}$ complex is coupled with GPCR. Source: Congreve *et al.*, 2020. ..27

Figure 1.3.2: GPCR activation cycle. In the inactive state of the receptor (A), the GDP-bound Gα subunit is associated with $G_{\beta\gamma}$ dimer and they together are coupled to the GPCR. (B) Agonist binding induces a conformational change in the receptor, catalyzing the exchange of GDP for GTP on the G_{α} subunit. (C) The GTP-bound G_{α} subunit dissociates from the $G_{\beta\gamma}$ dimer, and both components activate downstream signaling pathways. (D) Intrinsic GTP hydrolysis by the G_{α} subunit—often accelerated by regulator of G protein signaling (RGS) proteins—converts GTP to GDP, enabling reassociation of G_{α} with $G_{\beta\gamma}$ and thereby terminating the signal. Source: Proft *et al*., 2014. 28

Figure 1.3.3: β-AR G protein signalling pathway. The β1-AR and β2-AR share the same pathway. Binding of epinephrine or norepinephrine to β-AR activates the associated Gα protein via GDP–GTP exchange. The dissociated Gα subunit stimulates AC, leading to increased cAMP production. Elevated cAMP primarily activates PKA, which phosphorylates downstream targets to elicit physiological responses. The specific responses depend largely on the tissue distribution of β1-AR versus β2-AR.....29

Figure 1.3.4: Structural features common to β-AR ligands and their conserved polar interactions, exemplified by adrenaline. The protonated amine group forms an ionic interaction with a conserved $Asp^{3.32}$, and hydrogen bond with a conserved $Asn^{7.39}$. The β-hydroxyl group also forms hydrogen bonds with $Asp^{3.32}$ and $Asn^{7.39}$, respectively.....31

Figure 1.3.5: Structural comparison between non-selective β-AR agonist and antagonist and selective β1-AR antagonist. Non-selective agonists, exemplified by adrenaline, typically feature two phenolic hydroxyl groups (colored by light orange), while non-selective inhibitors exhibit a large, rigid aromatic substituents such as propranalol and cyanopindolol. The selective β1-AR antagonists, such as atenolol and bisoprolol, possess an extended side-chain (highlighted in pink) that contributes to subtype selectivity. 32

Figure 1.3.6: Binding poses for (A) atenolol, (B) bisoprolol, and (C) cyanopindolol in complex with β2-AR. The protein is shown in cartoon style. Residues involving key polar interactions are shown in stick style. The hydrogen bonds are shown in yellow dashed lines. (D) The alignment of the three complexes, with atenolol colored green, bisoprolol cyan, and cyanopindolol pink.....34

Figure 1.3.7: Dissociation pathways of the ECV and keyhole, illustrated using bisoprolol-β2-AR complex. The ECV opens to the extracellular water, while the keyhole constitutes a fissure between transmembrane helices H4 and H5 linking the orthosteric binding site to the membrane bilayer. The long side chain of bisoprolol (pink) occupies the keyhole, while the remaining portion of the ligand is shown in green. 34

Figure 1.4.1: Oligomeric Conformations of Prdx2. (A) Homodimer of Prdx2 with one subunit coloured green and the other cyan. Sulphur atoms of the 51C_P and the 172C_R are shown as yellow spheres. (B) Decameric ring of Prdx2 with ten subunits shown in different colours. (PDB ID: 1QMV). (C) HMW of Prdx3. Adapted from Cao *et al.*, 2017. 38

Figure 1.4.2: The Prdx2 catalytic cycle. C_P is highlighted in red. 39

Figure 1.4.3: Conformations of the FF state (PDB: 1QMV) and the LU state (PDB: 5IJT). (A) Structure of the FF state in cartoon style with two crucial cysteines

(C_P and C_R) in stick representation. One chain is shown in yellow, while the other is colored green, cyan, and red. Wherein, the representation of red is the C-terminal and cyan is the α2 helix. (B) A diagram of the structure of the FF state, where the C-terminal is folded and interacts with the α2 helix located in domain I. The C_P and C_R orientate in the opposite direction of the position of each other. (C) Structure of the LU state. Same as (A), different regions are shown in various representation styles and colors. (D) A diagram of the LU state, the drawing style of which is the same as (B). Its C-terminal is released and has no interactions with the α2 helix. Moreover, the C_P and C_R orientate towards each other.....40

Figure 1.4.4: The active site of FF Prdx2 with a binding H_2O_2. Source: Hall *et al*, 2011.....41

Figure 1.4.5: Diagram for the Prdx2 conformational change during catalytic cycle. 43

Figure 1.4.6: Pathway linking Prdx2 SNO to Parkinson's disease. 46

Figure 2.1: Residence time prediction workflow. Equilibrations were run to relax the ligand-β2-AR system and generate various starting structures for subsequent dissociation simulations. The τRAMD method was employed to simulate the dissociation behaviour, and the resulting trajectories were grouped based on the egress route. Relative koff values were initially estimated, which allowed the calculation of absolute koff values, free from the custom force. 49

Figure 2.2: Visualization of the simulated system. The β2-AR is shown in cartoon style, while the mixed PA/PC lipid bilayer is shown in stick representation. Water molecules are displayed in line representation, and sodium (Na^+) and chloride (Cl^-) ions are shown as spheres. 50

Figure 2.3: Conformation of the FU state, consiting of one sub-FF state and the other sub-LU state. Two subunits are colored in green and cyan with C_P and C_R highlighted in stick style.....59

Figure 2.4: Stepwise alchemical pathway from SH-Cys to SNO-Cys. (A) The pathways correspond to the FF→LU conformational transition. In this process, all four cysteines including two C_P and two C_R residues are S-nitrosylated. The top pathway (1→2→3→4→5) mutates C_P first followed by C_R residues, and the bottom pathway (1→6→7→8→5) mutates C_R first followed by C_P residues. These two pathways together form a closed thermodynamic cycle, allowing error assessment for the calculation. (B) Alchemical pathways associated with the FF→FU conformational transition. The 9→10 pathway represents the C_P mutation in the sub-LU side, while 9→11 pathway represents that in the sub-FF side. For clarity, only the diagram of FU is shown.....63

Figure 3.1: Proportion of keyhole dissociation events for atenolol, bisoprolol, and cyanopindolol under varying magnitudes of the random force. The percentage of dissociation trajectories utilizing the keyhole pathway is plotted as a function of force magnitude. Atenolol and cyanopindolol exhibit force-dependent trends, with higher forces promoting greater usage of the keyhole route. In contrast, bisoprolol shows a relatively stable proportion across all force conditions.....72

Figure 3.2: Dissociation pathways for atenolol, bisoprolol and cyanopindolol at a random force magnitude of 764.5 pN. Red spheres represent the COM of the ligand sampled every 1 ps during the final 500 ps of each trajectory. For each ligand and pathway, 15 out of 45 trajectories were randomly selected for display. The spatial distribution of the spheres reflects the overlap of these trajectories, thereby indicating the dissociation route. The protein is shown in

cartoon representation, with helices H4 and H5 highlighted in yellow and the remaining residues in green. 73

Figure 3.3: Atenolol (A) and bisoprolol (B) in complex with β2-AR. Ligands are shown in cyan sticks, key residues involved in polar interactions in green sticks, and lipid molecules in pink sticks. The protein is depicted in green cartoon representation. Hydrogen bonds are shown as yellow dashed lines, and non-polar interactions between bisoprolol and lipid molecules as red dashed lines. 75

Figure 3.4: Unbinding time distributions at a random force magnitude of 764.5 pN for different ligand dissociation pathways: ATE_{ECV}, BIS_{ECV} and $BIS_{keyhole}$, CYA_{ECV}. 77

Figure 3.5: Derivation of relative k_{off} values at a certain magnitude random force for four ligand dissociation pathways: ATE_{ECV}, BIS_{ECV}, $BIS_{keyhole}$, and CYA_{ECV}. LEFT: Plot of the number of surviving ligand–β2-AR complexes (N) over time. At time 0, all complexes survive; unbinding events occur progressively until complete dissociation. RIGHT: A plot of $\ln(N)$ versus unbinding time and the dash line shows the linear regression excluding the three outliers (coloured orange). According to Eq. (2.4), the absolute value of the fitted slope corresponds represents the value of relative k_{off} at 764.5 pN. By using this protocol, k_{off} values can be estimated for each ligand dissociation pathway in various magnitudes of the random force. 79

Figure 3.6: Simulated k_{off} and k_{off}^0 values. (A) Simulated k_{off} values across all random force magnitudes (data from Table 3.1). The colorful lines show linear fits of the k_{off} values to the force magnitudes. According to Eq. (2.8), the extrapolated k_{off} in the absence of a custom force (k_{off}^0) can be reported as $10^{intercept}$. The resulting k_{off}^0 values for each ligand dissociation route are listed in Table 3.2. (B) Comparison between simulation and experimental k_{off}^0 values after normalization by ATE_{ECV}. An obvious strong correlation between

simulation and measured k_{off}^0 values can be observed among four ligand dissociation routes. 82

*ATE_{ECV} at 834 pN de-weighted. 82

Figure 3.7: Local energy-peak configurations along the cyanopindolol dissociation pathway. (A, B) Structures of two representative configurations. The protein is shown in green cartoon style, and the ligand cyanopindolol in cyan sticks. Conserved $Asp^{3.32}$ and $Asn^{7.39}$, together with residues forming hydrogen bonds, are displayed in stick. Hydrogen bonds are indicated as yellow dashed lines. Configuration 1 is shown in A and configuration 2 in B. (C, D) Evolution of the progress coordinate *s* over 10 ns standard simulations initiated from configuration 1 (C) and configuration 2 (D). The red triangle denotes the starting configuration, and the five replicates are shown in different colors. 88

Figure 4.1: Comparison of SNO-Prdx2 dynamics with the crystalized S-nitrosocysteine. Blue traces represent values obtained from dynamics, while red dashed lines indicate corresponding values from the X-ray crystal structure. (A) Structure of S-nitrosocysteine with atom names indicated. (B) Distribution of O2-N2-SG-CB dihedral angle. (C) N2-O2 bond length. (D) SG-N2 bond length. (E) CB-SG bond length. 92

Figure 4.2: FF-state conformational sampling. Trajectories for SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P are projected onto a common PC1–PC2 subspace derived from all Cα coordinates. The grey background represents the aggregated explored space from all PTMs, while red marks the sampling of the PTM highlighted in that panel. 94

Figure 4.3: Visualization of FF-state motions along PC1 and PC2 exemplified by SH_C_{PR}. The two monomers are colored in green and cyan. The maximal

displacements along the principal components are represented by white arrow. 95

Figure 4.4: LU-state conformational sampling. Trajectories for SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P are projected onto a common PC1–PC2 subspace derived from the Cα coordinates of residues 1–168 and 197–364 (C-termini excluded). 97

Figure 4.5: Visualization of LU-state motions along PC1 (left panel) and PC2 (right panel) from SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P simulations. The two monomers are colored in green and cyan. The maximal displacements along the principal components are represented by white arrow. 98

Figure 4.6: Determination of equilibration time. (A) Stability of the calculated ΔG_{FF} for 1→2 transition was assessed by changing discarded simulation time. Data prior to the discarded time point are excluded. The remaining trajectory is split into two halves, and ΔG_{FF} is calculated independently for each half; the $\Delta\Delta G_{FF}^{equil}$ reflects the variation between the two halves. At a discarded time of 2.08 ns, $\Delta\Delta G_{FF}^{equil}$ falls below 0.1 kcal/mol, suggesting that the system has reached relative equilibrium. The obvious severe fluctuation of $\Delta\Delta G_{FF}^{equil}$ after 8 ns derives from an insufficient sampling. (B) Distribution of equilibration times across all performed FEP simulations. The red dashed line refers to a fitted Gaussian curve. The mean equilibration time (μ) is 3.58 ns, indicating that approximately half of the simulations reach equilibrium by this point. To ensure consistency in sampling duration and to maximize the extent of equilibrium for all systems, 3.58 ns is adopted as a uniform equilibration time cutoff for all FEP simulations. All the calculated ΔG values are listed in the Table 4.1. 104

Figure 4.7: Potential of mean force with respect to λ for TI and MBAR. The ΔG_{FF} along 1→2 transition was estimated by using TI and MBAR, respectively. 105

Figure 4.8: Diagram of a conformational energy landscape explaining the relation and difference between substate and state. The horizontal axis denotes the conformational coordinate, while the vertical axis H(x) represents the free energy. Two major conformational states (State 1 and State 2) are separated by a high-energy barrier, with interconversion occurring on the microsecond-to-millisecond timescale. Each state comprises several local minima (Substates 1–5), which are separated by lower barriers, allowing transitions on the microsecond timescale. This hierarchical landscape illustrates the coexistence of fast local motions between sub-state and slower global transitions between states. 107

Figure 4.9: Comparison of initial SNO-C_P poses in the FF state. (A) Alignment of all initial SNO-C_P poses. (B) Comparison of the two initial poses: $9 \rightarrow 10^{(pose1)}$ shown in green and $9 \rightarrow 10^{(pose2)}$ shown in dark green. 109

Figure 4.10: Proposed mechanism of C_R SNO by small-molecule NO^+ donors. In an intermediate state where C_R becomes solvent-exposed, RNS can directly modify C_R 113

Figure 4.11: Proposed mechanism of C_R SNO by SNO-C_P. 115

Figure 4.12: Optimization process of the FF→FU transition pathway for the SNO_C_P system. LEFT: Successive iterations of the string, projected onto the PC1–PC2 subspace. Strings are colored from purple (initial) to bright red (final), showing convergence toward the most probable pathway. PC1 accounts for 43% of the variance and PC2 for 6%. RIGHT: Structural overlap of the intermediate states, with the two subunits colored yellow and green. C_P and C_R are highlighted in sphere style. 117

Figure 4.13: Comparison of optimized FF→FU transition pathways under different PTMs. Final strings for SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P systems are projected onto the PC1–PC2 subspace. PC1 and PC2 account for 45% and 8% of the variance, respectively. 118

List of Tables

Table 1.1: Physical energy-biological measurement connection 8

Table 1.2: τRAMD applications and their correlation of k_{off} with experimental data. 17

Table 2.1: Used software and tools. 48

Table 2.2: Crystal structures of human source Prdx2. 57

Table 2.3: The built systems of Prdx2, each representing a unique combination of PTM location and conformational state. 59

Table 3.1: Estimated k_{off} values and corresponding error at various magnitudes of the random force for four ligand dissociation pathways: ATE_{ECV}, BIS_{ECV}, $BIS_{keyhole}$, and CYA_{ECV}. 80

Table 3.2: Simulation and experimental k_{off}^0 values before and after normalization by ATE_{ECV}. 85

Table 4.1: Computed free energy changes in selected alchemical transition. ΔG_{FF} and ΔG_{LU} represent the free energy changes associated with given alchemical transition, in FF and LU conformational state, respectively. These alchemical transitions are categorized into two groups based on their perturbation sites: CP and CR. In each group, the variation in calculated ΔG values derives from different sampling configurational spaces. For 9→11 transition, the sub-FF side CP is mutated in the FU conformational state; thus, its calculated ΔG values correspond to the FF state. 102

Table 4.2: Effective free energy differences and relative free energy difference (ΔΔG). 111

Table of Contents

ABSTRACT ..... **I**

ACKNOWLEDGEMENTS ..... **III**

PUBLICATIONS RELATED TO THIS THESIS ..... **IV**

ABBREVIATIONS ..... **V**

LIST OF FIGURES ..... **VIII**

LIST OF TABLES ..... **XVIII**

TABLE OF CONTENTS ..... **XIX**

1. INTRODUCTION **1**

1.1 ENERGY, EQUILIBRATION CONSTANT, DISSOCIATION CONSTANT, PROBABILITY AND THEIR INTERCONNECTION 1

1.1.1 Physics Underlying Theory *2*

1.1.2 Dissociation Constant *8*

1.1.3 Thermodynamic and Kinetics of Protein Conformational Change *12*

1.2 ENHANCED SAMPLING METHODS 15

1.2.1 τ-Random Acceleration Molecular Dynamics *15*

1.2.2 Committor Analysis *18*

1.2.3 Free Energy Perturbation *19*

1.2.4 Finite Temperature String Method *24*

1.3 BETA-ADRENOCEPTORS AND THEIR BINDERS 27

1.3.1 The Structure and Function of GPCR *27*

1.3.2 β-AR G Protein Signalling Pathway *28*

1.3.3 Selective β1-AR Inhibitors *30*

1.4 PEROXIREDOXIN 2 36

1.4.1 Overview on Peroxiredoxin Family *36*

1.4.2 Peroxiredoxin 2 and Its Catalytic Cycle *37*

1.4.3 FF and LU state – Two Main Conformational States of Prdx2 ..... *39*

1.4.4 FF state—a structure basis of antioxidant and chaperone activity *41*

1.4.5 Conformational Change During Catalytic Cycle *42*

1.4.6 SNO of Prdx2 and Its Pathological Consequences in Parkinson's Disease *44*

1.5 AIMS AND OBJECTIVES 47

2. METHODS 48

2.1 LIGAND RESIDENCE TIME PREDICTION 48

2.1.1 System Preparation *49*

2.1.2 τRAMD protocols *52*

2.1.3 Derivation of Relative k_{off} *53*

2.1.4 Derivation of Absolute k_{off}^{0} *55*

2.1.5 Transition State Confirmation *56*

2.2 PREDICTION AND ANALYSIS OF PRDX2 CONFORMATIONAL TRANSITIONS 57

2.2.1 System buildup *57*

2.2.2 Parameterization *59*

2.2.3 MD Details *60*

2.2.4 Principal Component Analysis *61*

2.2.5 Free Energy Perturbation *61*

2.2.6 Finite Temperature String Method *67*

3. LIGAND DISSOCIATION MECHANISMS FOR B-AR INHIBITORS FROM B2-AR 69

3.1 DISSOCIATION PATHWAY ANALYSIS 69

3.2 ESTIMATION OF K_{OFF} IN VARIOUS MAGNITUDES OF THE RANDOM FORCE 74

3.2.1 Distribution of Unbinding Times *76*

3.2.2 Estimation of k_{off} in various magnitudes of the random force *77*

3.3 ESTIMATION OF K_{OFF}^{0} 83

3.4 TRANSITION STATE CONFIRMATION 86

3.5 CONCLUSION 89

4. EFFECT OF S-NITROSYLATION ON PRDX2 CONFORMATIONAL EQUILIBRIUM 91

4.1 STANDARD MD SIMULATIONS 91
4.1.1 Parameterization Validation *91*
4.2.2 Principal Component Analysis *93*
4.2 FREE ENERGY PERTURBATION 99
4.2.1 Independence of Alchemical Transitions *99*
4.2.2 Equilibration Time Determination *102*
4.2.3 Error Assessment *104*
4.2.4 Substate Sampling and Approximation of State Free Energy *106*
4.2.5 Effect of SNO on Prdx2 Conformational Equilibrium *109*
4.2.6 How C_R is S-nitrosylated? *111*
4.3 TRANSITION PATHWAY IDENTIFICATION 116
4.4 CONCLUSION 119
REFERENCE 121

1. Introduction

1.1 Energy, Equilibration Constant, Dissociation Constant, Probability and Their Interconnection

Molecular dynamics (MD) simulations model a system of interest at the atomic level, predicting the motion of each particle according to classical laws of physics[1-4]. Despite the simplicity of the idea, MD simulation has been proven to be a powerful tool to predict the dynamics of biological molecules, including proteins, DNA, and RNA[5-7]. However, the accessible simulation timescales are limited by computational power. Many biomolecular processes, such as protein folding, ligand binding, and ion permeation through membranes, involve large conformational rearrangements. These events typically span a broad range of timescales, from microseconds to seconds or longer, depending on the system and the underlying energy barriers[3, 8]. By contrast, standard MD simulations are usually limited to the nanosecond to low-microsecond regime, where they can primarily capture only local conformational fluctuations for example, loop motions, side-chain rearrangements, or small helical breathing motions[2, 9].

It is important to note that a slow biological process with a long macroscopic timescale does not imply that each individual event runs slowly over that entire period. Taking ligand dissociation as an example, a long macroscopic timescale reflects the rarity of unbinding: the probability of dissociation at any given moment is low, so one may need to wait a long time before observing it. When the event finally occurs, it typically completes on a much shorter timescale—often milliseconds or less. Thus, the macroscopic timescale extracted from experiments is not the duration of the event itself, but instead a statistical reflection of the underlying microscopic probability distribution governing its occurrence[10-12].

To access biomolecular processes occurring on timescales beyond microseconds, enhanced sampling methods have been developed. Two representative slow processes are used as illustrative examples[13-17]:

1. Ligand dissociation from a bound protein.
2. Protein conformational change between folded and unfolded states.

A biomolecular process can be characterised by a change in the state of the system. For example, ligand dissociation corresponds to a change from the bound to the unbound state[10], while a protein conformational transition can be described as a change between folded and unfolded states[2, 18]. Physics provides a rigorous framework for defining such states and for developing methods to characterize their thermodynamic and structural properties[19]. The following section introduces the theoretical foundations underpinning these approaches.

1.1.1 Physics Underlying Theory

A system may possess various states, each characterized by a specific probability of occurrence. The knowledge of the states and their associated probabilities is fundamental to gaining insights into the system's behaviour. For example, a protein exists in either a folded or unfolded state. Assuming a correlation between protein conformation and function, knowledge of the equilibrium distribution between the two states can provide valuable information about the protein's functional properties[8, 20]. Similarly, a drug molecule may either bind to its target protein to form a complex or keep unbound in solvent. Quantifying the fraction of bound molecules enables the estimation of drug affinity[21].

To describe state probabilities, physicists often employ energy-based frameworks, whereas chemists and biologists may prefer to use concentration often derived from experimental observations. Thereby, probability unifies energy and concentration[22]. In this context, we aim to elucidate commonly used physical and biological terms and their interconnections. This dual perspective enables physicists to interpret how abstract

energy terms translate into measurable biological outcomes, while providing biologists with theoretical grounding for experimental observations. Moreover, it allows for cross-validation of a single phenomenon through the independent application of physical and biological methodologies[4, 19].

To begin a discussion on energy, it is first necessary to define the concepts of *configuration* and *state*. A configuration refers to a single conformation of a system[23]. In contrast, a state is defined as a collection of configurations that share similar structural features and reside within the same energy well on the energy landscape[19, 24]. As shown in Fig 1.1.1, point 'a' corresponds to an individual configuration belonging to state A. Configurations spanning from $x = 0$ to 2 collectively comprise state A. These configurations share similar structural characteristics and occupy the same energy well.

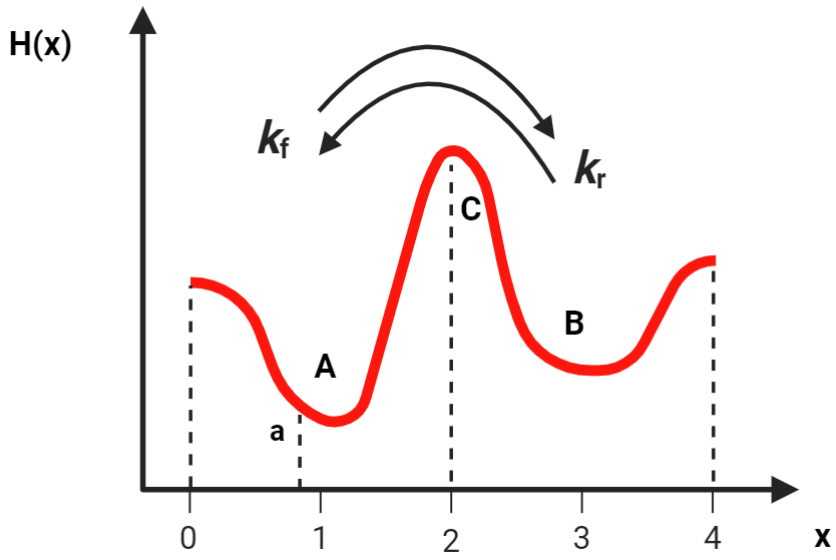


Figure 1.1.1: Double-well energy schematic diagram along reaction coordinate x. The energy refers to the total energy approximating to the sum of potential and kinetic energy. Each well represents a state (A or B), with point 'a' indicating a single configuration in state A. The transition state C corresponds to the highest energy point between wells. The equilibrium constant (K_{eq}) between A and B is determined by the two-state energy gap. The forward reaction rate constant (k_f) is determined by the energy gap between A and C. The reverse reaction rate constant (k_r) is determined by the energy gap between B and C.

Once the energy of a particular configuration is known, its probability of occurrence can be estimated using the Boltzmann factor (assuming the system is in thermodynamic equilibrium)[23, 25]. Using point 'a' as an example:

$$P_a \propto \exp[-H(x = a)/k_B T] \tag{1.1}$$

where P_a is the probability of the system being in configuration 'a', exp is the exponential function, $H(x = a)$ is the total energy of that configuration, and $k_B T$ is the product of the Boltzmann constant k_B and absolute temperature T. As Eq. (1.1) suggests, the higher a configuration's total energy, the lower its probability of occurrence[18, 19].

Given that the probability of any single configuration is typically very small, and that configurations within the same state often share similar energetic and structural characteristics, it is generally more meaningful to consider the cumulative probability of an entire state rather than that of an individual configuration[19, 24].

In a system at thermodynamic equilibrium, energy and the probability distribution across states reach a steady condition, exhibiting no significant changes and thus can be approximated to constants[25]. Under such thermodynamic equilibrium, the total probability of state A—comprising configurations from x = 0 to 2—is calculated by integrating the Boltzmann-weighted probabilities of all included configurations[23, 26]. This accumulation is the definition of *partition function*, which can be expressed as:

$$Z = \int dx \exp[-H(x)/k_B T] \tag{1.2}$$

The probability and partition function of state A, Z_A, can be described as Eq. (1.3):

$$Z_A = \int_0^2 dx \exp[-H(x)/k_B T] \tag{1.3}$$

However, this accumulated probability exhibits no meaning without comparing with other states. One way to solve this problem is to sample all the possible states and explore their corresponding energies. By sampling all the states, the total partition

function of the system, denoted as Z_{all}, can be calculated. This enables normalization, allowing the absolute probability of state A to be determined as[27]:

$$P_A = \frac{Z_A}{Z_{\text{all}}} \tag{1.4}$$

Nevertheless, in practice, it is often difficult to sample all the states and determine their energies, especially the ones with high energies and low probability[27, 28]. In such cases, a more realistic alternative is to calculate the relative probability between some of the states but not all states. For example, the relative probability between state A and B can be expressed as[27, 28]:

$$\frac{P_A}{P_B} = \frac{Z_A}{Z_B} \tag{1.5}$$

Eq. (1.5) enables a direct comparison between state A and B, facilitating the evaluation of their relative stabilities or population distributions without a complete sampling of the entire system. By knowing the property of one state, the corresponding characteristics of the other can be calculated.

The energies discussed thus far refer to the *total energy* of a configuration, the sum of potential and kinetic energies[27]. Here, a new energy concept, *Gibbis free energy* (G), is introduced. Unlike the total energy, which characterizes individual configurations, G is a macroscopic property that describes an entire state (a collection of configurations)[27, 28]. Once the partition function of a state is known, its G can be derived using the Boltzmann factor[23]:

$$G = -k_B T \ln Z \tag{1.6}$$

Eq. (1.6) defines G as a state-level energy that can be used to directly calculate the state's partition function and probability. By combining Eq. (1.5) and (1.6), the relative probability between two states, A and B, can be derived:

$$\frac{P_A}{P_B} = \frac{Z_A}{Z_B} = \frac{\exp(-G_A/k_B T)}{\exp(-G_B/k_B T)} = \exp(-\Delta G_{AB}/k_B T) \tag{1.7}$$

where G_A and G_B denote the G of states A and B, respectively, and ΔG_{AB} represents the energy difference between two states. As shown in Eq. (1.7), a knowledge of the free energy gap between two states directly derives their relative probability[24].

In biological systems, the state probability is often manifested through the experimental observation, namely concentration. To illustrate this relationship, consider a system comprising a large number of identical protein molecules in solution. This protein can adopt two interconvertible conformational states, A and B[19, 27]:

$$\mathrm{A} \underset{k_r[\mathrm{B}]}{\overset{k_f[\mathrm{A}]}{\rightleftharpoons}} \mathrm{B} \tag{1.8}$$

Under the thermodynamic equilibrium condition, the concentration of each state keeps constant and reflects its relative probability. The ratio of these concentrations thus provides a measure of the relative probability between the two states[28]. This relationship is expressed by the equilibrium constant (K_{eq}):

$$K_{eq} = \frac{[B]}{[A]} = \frac{P_B}{P_A} \tag{1.9}$$

where K_{eq} quantifies the equilibrium between the two conformational states. If $K_{eq} = 1$, the system exists in a 1:1 ratio between states A and B. If $K_{eq} > 1$, state B is more populated than A, whereas $K_{eq} < 1$ indicates a higher population of state A[24].

Importantly, the constancy of concentration does not imply an absolutely stationary system, but an equality between forward and reverse conformational transition rates[19, 27, 28]. For this simple two-state system, the forward transition rate is given by the product of the forward transition rate constant (k_f) and the concentration of A, while the reverse rate is the product of the reverse transition rate constant (k_r) and the concentration of B:

$$k_f[A] = k_r[B] \tag{1.10}$$

and imposing Eq. (1.9), giving

$$K_{eq} = \frac{k_f}{k_r} \tag{1.11}$$

From Eq. (1.11), a direct relationship is established between K_{eq}, and kinetic rate constants: k_f and k_r. These three constants are intrinsic properties of the protein and are ultimately governed by the protein's underlying free energy landscape[4, 18, 24]. By combining Eq. (1.7) and (1.9), giving

$$K_{eq} = \frac{Z_B}{Z_A} = \exp(-\Delta G_{BA}/k_B T) \tag{1.12}$$

From Eq. (1.12), we can see how probability unifies free energy and equilibrium constant. This unity indicates that the knowledge of either K_{eq} or ΔG_{AB} allows a direct calculation of the other. As for the kinetic rate constant, it is determined by the *energy barrier* ($G_{barrier}$) which is the energy difference between the initial stable state and the *transition state*[19, 29].

$$k \propto \exp[-G_{barrier}/k_B T] \tag{1.13}$$

The transition state is defined as the conformational state with the highest energy along the reaction coordinate. For a simple two-state system such as Eq. (1.8), the state C having the highest energy represents the transition state (Fig 1.1.1). As for the stable state, its definition depends on the initial state. If the system begins in state A, to transit to state B, the system has to overcome the energy barrier ΔG_{AC} which governs k_f; conversely, if the system begins in state B, the ΔG_{BC} determines k_r. These relationships can be expressed as[4, 19]:

$$k_f \propto \exp[-\Delta G_{AC}/k_B T] \tag{1.13}$$

$$k_r \propto \exp[-\Delta G_{BC}/k_B T] \tag{1.14}$$

where ΔG_{AC} and ΔG_{BC} represent the energy barrier from stable states A and B to the state C, respectively. Therefore, given knowledge of the energy difference between the stable state and the transition state, the rate constant for leaving that stable state can be quantitatively estimated—bridging kinetics with free energies[19]. By combining Eq.

(1.13) and Eq. (1.14), we can also derive the relation of K_{eq}, k_f and k_r from the perspective of energy[4]:

$$K_{eq} = \frac{k_f}{k_r} = \frac{\exp(-\Delta G_{AC}/k_B T)}{\exp(-\Delta G_{BC}/k_B T)} = \exp(-\Delta G_{BA}/k_B T) \tag{1.15}$$

In summary, probability bridges microscopic configurations with macroscopic observables by linking energy and concentration. Through the Boltzmann distribution, partition functions, and free energy landscapes, we can quantitatively relate the physical behaviour of biomolecules to experimentally measurable quantities such as binding affinity and kinetic rates. Their relations are shown in Table 1.1.

Physical Concept	Biological Observation	Connection via Probability
Energy of a configuration (E)	Not directly observable	
Partition function (Z)	Not directly observable	
Free energy difference (ΔG)	Equilibrium constant (K_{eq})	$K_{eq} = \frac{Z_B}{Z_A} = \exp(-\Delta G_{BA}/k_B T)$
	Reaction kinetics (k)	$k \propto \exp[-G_{barrier}/k_B T]$

Table 1.1: Physical energy-biological measurement connection

1.1.2 Dissociation Constant

The prerequisite of drug action (in most cases) is the formation of the binary complex between the drug and the receptor. The longer the complex survives, the stronger the drug effects[30]. To quantify the drug effect duration, researchers have traditionally relied on drug affinity, measured as the equilibrium dissociation constant (K_d), a specific application of K_{eq}[10, 21]. The K_d is a measure of the tendency of the ligand-receptor complex to dissociate at equilibrium. In a one-step binding event,

$$LR \underset{k_{on}[L][R]}{\overset{k_{off}[LR]}{\rightleftharpoons}} L + R \tag{1.16}$$

where LR represents the ligand-protein complex. The L and R represent ligand and protein, respectively, the K_d can be defined as:

$$K_d = \frac{[L][R]}{[LR]} \tag{1.17}$$

In *in vitro* experiments, a closed system eventually reaches thermodynamic equilibrium, where the concentrations of the ligand and its target remain constant. In this equilibrium, K_d well indicates the proportion of protein bound to ligands, and therefore strongly correlates the drug effect duration[10]. However, *in vivo* drug concentration fluctuates over time after dosing due to various physiological processes including circulation, distribution, metabolism and diffusion. As a result, this open system is far from the equilibrium condition, rendering K_d a limited predictor of pharmacological efficacy[10, 21].

To address this, researchers have proposed that the *in vivo* drug efficacy is more relevant to drug-target binding kinetics—the association rate constant (k_{on}) and, most critically, dissociation rate constant (k_{off})[10-12, 31, 32]. The relation between K_d, k_{on} and k_{off} is given by Eq. (1.11). While both rate constants contribute to the drug-target complex duration, improving k_{on} is often restricted by the concentration of ligand, and the effect of absorption, distribution, and ultimately limited by the diffusion rate of the binding partners[10, 11]. Therefore, modulating k_{on} is complicated and less effective. In contrast, k_{off}—which determines how quickly a drug dissociates from its target after binding—is a more relevant indicator of *in vivo* drug efficacy. The residence time (τ)—defined as the inverse of the dissociation rate ($\tau = 1/ k_{off}$)—more directly reflects how long a drug remains bound to its receptor[10, 31].

The k_{off} has entire relevance to the interactions between the ligand and its receptor along the dissociation pathway. Medicinal chemists can influence k_{off} through structure–activity relationship (SAR) studies aimed at optimizing these interactions[31, 33]. The strategy of optimizing k_{off} has gained significant attraction. For example, Hyre et al. studied the kinetics of biotin to wild-type streptavidin and its three mutants, observing less than a twofold difference in k_{on} values, but over seven orders of magnitude variation in k_{off}[34]. Similarly, Lu developed a series of chemically related inhibitors targeting FabI enoyl reductase and found that the 10-day survival percent of

infected animals correlated linearly with the inhibitors' residence times, not their K_d values[35]. Guo and colleagues also investigated a set of A_2A receptor agonists and reported no correlation between efficacy and K_d[36], but a significant linear relationship between efficacy and residence time.

In addition to being a good indicator of drug efficacy, drug-target binding kinetics also offers a valuable framework for enhancing drug selectivity[21, 32, 37-39]. The differential binding and unbinding rates of a ligand to and from receptors can lead to the selectivity to one over the other receptors. If a drug's k_{off} from the desired receptor can be decreased, and the survival time of its complex can be increased, meanwhile increasing the k_{off} from other undesired receptors, the selectivity of the drug can be improved. It is generally considered advantageous when the dissociation rate from the target is lower than that from non-target receptors.

For example, Salmeterol is a (β2-AR) selective agonist, while Adrenaline is a non-selective adrenergic agonist[40]. Compared to Adrenaline, Salmeterol possess a long aryloxyalkyl tail allowing additional favourable interactions with β2-AR (Fig 1.1.2)[41]. These additional interactions are very likely to raise the energy barrier for Salmeterol exiting from β2-AR. Despite the structural flexibility of the long tail which incurs an entropic penalty upon binding, this penalty is effectively compensated in β2-AR by its exosite that stabilizes the bound state, rendering the entropic contribution relatively minor compared with the enthalpic gain[42]. According to Eq. (1.13), the raised energy barrier leads to a reduced k_{off}. Indeed, the k_{off} value of Salmeterol is 7-fold smaller than that of Adrenaline[43]. With the same dosing, the Salmetrerol-β2-AR complexes survive longer than Adrenaline-β2-AR complexes. This reduced k_{off} contributes to the selectivity of Salmeterol to β2-AR.

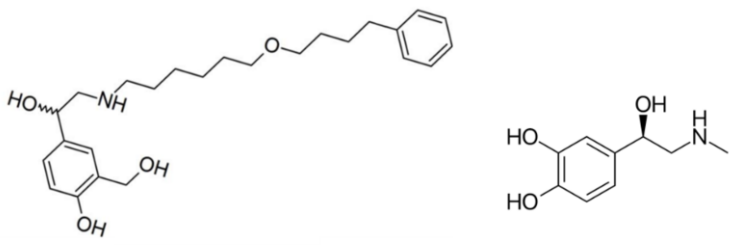


Figure 1.1.2: The structures of Salmeterol and Adrenaline

To determine the k_{off} value of a ligand from its target receptor, one may do experiments and simulations. The current available experimental methods for measuring k_{off} typically rely on monitoring a time-dependent signal that reflets the unbinding process. These methods can be broadly classified into three classes: (a) label-based[44-46], (b) label-free[47-49], and (c) enzyme activity-based assays[50]. Despite their methodological differences, they share a common principle—tracking the dynamic change of unbound ligand directly or indirectly. Thus, by analysing the ligand concentration changing rate and applying appropriate mathematical models, the k_{off} can be inferred. Our analysis framework for estimating relative k_{off} values through simulation is developed based on insights from this.

In label-based methods, the signal produced by labels changes depending on whether the ligand is is free or bound[44, 46]. By monitoring the time-dependent change in signal, one can infer how quickly the ligand dissociates from the receptor. Typically, k_{off} is extracted by fitting the signal decay to an exponential function over time. Specifically, Fluorescence Polarization (FP) and Radioligand Binding methods label on the ligand, thus directly monitoring the concentration of unbound ligand as dissociation proceeds[51, 52]. In contrast, Förster Resonance Energy Transfer (FRET) labels both ligand and acceptor, therefore detecting the distance between them and tracking the intact binding complex[53].

Most of label free techniques measure intrinsic biophysical changes at the binding interface. Surface Plasmon Resonance (SPR) detects the refractive index which

fluctuates with ligand binding[47, 48]. A more advanced method: Isothermal Titration Calorimetry (ITC) measures heat changes during molecular interaction which changes following binding and unbinding event[54].

Enzyme activity-based techniques rely on that the ligand binding alters the catalytic activity of the enzyme. As the ligand dissociates, the enzyme activity changes accordingly, allowing k_{off} to be inferred from the time-dependent enzymatic response[21].

Specially, different from the above three assays which detect the unbound ligand concentration, atomic force microscopy (AFM) follows a different protocol to measure the k_{off}. In this approach, a cantilever tip is chemically attached to the ligand. By lowering the height of the cantilever, the ligand can gradually approach and bind to its target protein. Then the cantilever is gradually retracted, resulting in an increasing force to pull the ligand from the binding site until the ligand-protein interaction breaks. Therefore, the strength of the ligand-receptor bound can be measured. Moreover, in dynamic force spectroscopy, measurements are performed over a range of force loading rates, enabling extrapolation of the intrinsic k_{off} in the absence of an external force[55, 56].

The simulation methods for k_{off} prediction are detailed in the Section 1.2.1, with their applications illustrated in Chapter 3.

1.1.3 Thermodynamic and Kinetics of Protein Conformational Change

There is a strong correlation between a protein's conformation and its function—an idea central to the structure–function paradigm. For many proteins, static structures have been resolved, providing atomistic detail that helps uncover the chemical basis of their biological activity[57]. SAR theory builds on these static models, typically treating the receptor as a rigid entity. SAR seeks to understand how modifications to a ligand's shape, functional groups, or stereochemistry—based on the fixed structure of the receptor binding site—affect potency, selectivity, or mechanism of action. By

examining the atomic features of the receptor's active site, SAR can help explain why only specific substrates are recognized and bound[58].

For example, peroxiredoxin 2 (Prdx2) is an antioxidant enzyme that rapidly reduces hydrogen peroxide (H_2O_2). In the crystal structure of its fully folded (FF) state, the peroxidatic cysteine (C_P) is positioned adjacent to an arginine residue, which lowers its pK_a and stabilizes the deprotonated thiolate form (Fig 1.4.4). In addition, a nearby threonine residue contributes to substrate destabilization through polar interactions. This acid–base catalytic motif helps explain the high reactivity of Prdx2 toward H_2O_2, offering a clear structure-based rationale for its enzymatic efficiency[59].

However, all molecules are not rigid but keep moving. Protein functions are often ultimately governed by their dynamic character rather than static structures alone[8, 60]. Using Prdx2 as an example again, two conformational states have been characterized: the FF state and the locally unfolded (LU) state[61, 62]. The FF state is responsible for the reduction of H_2O_2, while the LU state facilitates functional regeneration by facilitating redox recovery following catalysis. Post-translational modifications (PTMs)—covalent changes to proteins occurring after biosynthesis—can modulate Prdx2's structure and function[63]. Simply examining the static structures of these two states is insufficient to fully understand how different PTMs trigger conformational transitions and drive the functional switch.

A free energy landscape represents the free energy throughout the configurational or conformational space of a biomolecule[8, 64]. A knowledge of the dynamic pathway between conformational states and the corresponding energy landscape along the pathway can answer how likely each state is (thermodynamics) and how fast they interconvert (kinetics). For example, if we obtain the free energy landscape of Prdx2 transiting from FF to LU, the free energy difference between the two states can be calculated, and thus according to Eq. (1.7), their relative probability can be also calculated. Since a complete energy landscape includes the sampling of the transition state, the kinetic transition rates between FF and LU can be calculated according to Eq.

(1.13). By constructing such an energy landscape, we can understand the underlying mechanism of PTM modulation on Prdx2's conformational transition and functional switch.

1.2 Enhanced Sampling Methods

1.2.1 τ-Random Acceleration Molecular Dynamics

The τ-Random Acceleration Molecular Dynamics (τRAMD) method is utilized to simulate ligand dissociation[65]. We first distinguished two related but distinct concepts: *unbinding time* and *residence time*. The unbinding time refers to the duration required for a ligand to dissociate from its receptor along a given trajectory. This time is trajectory-dependent and therefore varies from case to case. While the residence time is a constant, which reflects the average lifetime of the ligand–receptor complex and is only determined by the property of the ligand dissociation route. By sampling sufficient unbinding events, the residence time can be extrapolated theoretically.

The ligand dissociation is a slow process. By running an extremely long standard MD simulation for a ligand-protein complex, an unbinding event with atomistic details can be captured eventually. However, the residence time of drugs with an interesting pharmacokinetic often span from minutes to hours-timescales which is far exceed current computational ability[66]. It is extremely unlikely to observe the dissociation event in microsecond timescale for a ligand with minute/hour-long residence time[67]. Moreover, because dissociation is stochastic and the experimental residence time represents an ensemble average, a single simulated trajectory cannot provide meaningful kinetic data. To overcome these challenges, researchers have developed many different enhanced sampling methods to computationally characterise ligand-receptor association/dissociation pathways, including Weighted Ensemble[68], Metadynamics[69], Milestoning[70], Steered Molecular Dynamics[71] and Markov state model[24]. It is a feature, and sometimes a limitation, of many of these methods that they require a prior knowledge of the dissociation pathway, and a well-designed collective variable along this dissociation pathway for the sampling process.

An alternative method is τRAMD, developed by Rebecca C. Wade's group[65]. Unlike methods above, this method does not require a prior knowledge of the

dissociation pathway. Similar to AFM, in which a force is applied to pull a ligand out of its target, the underlying rationale of τRAMD is applying a custom force to the ligand centre of mass (COM) to pull the ligand, thereby facilitating the dissociation process without predefining a dissociation pathway. The 'random' refers to the force orientation. The orientation changes as the COM of the ligand moves less than a threshold distance in every time interval. It implies that only force orientations resulting in productive ligand movement are maintained. These effective movements would eventually pull the ligand out of the binding site within an acceptable computational time.

The most important parameter is the magnitude of the random force which mainly affects the accuracy and efficiency of the residence time prediction[65, 72-76]. A weaker random force tends to result in a longer observed unbinding time, more closely approximating the intrinsic unbinding time in the absence of acceleration, therefore predicts more accurately, however, at the cost of a larger computational power. This reflects an inherent accuracy–efficiency trade-off in the method.

Within an appropriate range of force magnitudes, τRAMD is able to distinguish residence times of different systems (Table 1.2). The tested systems include complicated ones such as HSP90α with a highly flexible binding site and membrane-buried GPCRs (β2-AR and ACh2R M2). The method has also been applied under varied conditions, including different mutations, binding sites, and temperatures. Collectively, these results demonstrate that τRAMD is broadly applicable and yields strong correlations between computed and experimental k_{off} values.

System	No. of complex	Pearson coefficient of determination R^2	reference
HSP90α	70	0.86	[77]
β2-AR	1	0.94	[78]
ACh2R M2	3		
InhA	6	n.d.	[79]
Four T4L mutants	6	0.78	[80]
Wild type and two mutants of haspin	6	n.d.	[81]
FAK	12	0.95	[82]
PYK2	10	0.63	

Table 1.2: τRAMD applications and their correlation of k_{off} with experimental data.

In all of these τRAMD works, the relative residence times have been approximated to the average of simulated unbinding times. To give a trustable average and corresponding uncertainty, commonly 5-8 replicas were performed[65, 72, 73, 75, 76]. This approach provides a credible estimate of residence times without incurring excessive computational cost.

This approximation is based on a correlation between unbinding time and residence time. Give a specific ligand-protein complex, an intrinsically long residence time generally results in a delayed observation of ligand dissociation. Despite the correlation, the unbinding time and residence time are two distinct concepts and can not be regarded as equivalent mathematically. Consequently, the conventional τRAMD analysis method does not allow a direct mathematical mapping between individual unbinding times and the absolute residence time.

To address this limitation, we developed an analysis framework that enables rigorous computation of relative k_{off} under a defined random force, as well as the absolute k_{off} without the random force (k_{off}^0) in a mathematically rigorous manner, inspired by the experimental methods: label assay and AFM[55]. The former measures the

k_{off} by monitoring the bound ligand concentration over time[44-47], and the latter measures it by applying various force loading rate[55].

1.2.2 Committor Analysis

Verification of the transition state is informative to understand a biological process in the aspect of kinetics[18, 19]. The transition state ensemble (TSE), however, is difficult to characterize. By definition, transition states correspond to configurations at the top of the free-energy barrier, which are sampled with exceedingly low probability in equilibrium simulations[19, 28]. Even if a trajectory covers an entire biological process—such as ligand dissociation—the TSE is not immediately evident within it.

A straightforward but often impractical approach to identify the TSE is to construct the full free-energy landscape along the dissociation pathway[83, 84]. Enhanced sampling strategies, such as weighted ensemble, metadynamics, milestoning, have been widely employed to achieve this for slow processes. These methods provide a global view of the energy landscape but typically require substantial computational effort and depends on chosen collective variables (CVs). A CV is a low-dimensional quantity that captures the essential slow motions of a complex system[85].

Committor analysis offers a more direct and efficient approach to TSE identification. The committor for a configuration x, denoted $p_B(x)$, is defined as the probability that a trajectory initiated from x with randomized velocities reaches the product basin B before returning to the reactant basin A[86, 87]:

$$p_B(x) = P\{\text{reach } B \text{ before } A \mid x\}$$

In practice, basins A and B are specified, a commitment time is chosen to suppress rapid recrossing, and many short unbiased "shooting" trajectories are launched from candidate configurations with randomized velocities. Randomizing velocity orientations is essential to avoid bias from inertia. In the case of ligand dissociation, the basins correspond to bound and unbound states, and each trajectory is classified as forward (committing to B) or backward (committing to A). Configurations yielding p_B

≈ 0.5 constitute the transition state ensemble. Because this definition depends only on dynamical probabilities, committor analysis is often regarded as the gold standard for transition state validation[86-88].

A key requirement for committor analysis is the definition of a suitable progress coordinate that allows an effective representation of the process. Biomolecular transitions generally occur in a high-dimensional conformational space, making direct analysis intractable[24]. To address this, dimensionality reduction techniques are employed to define the process as a one-dimensional 'string'. Therefore, a complex multidimensional pathway can be mapped onto a simplified 1D representation. Depending on the system, structural measures such as the Root Mean Square Deviation (RMSD) of selected atoms, interatomic distances, or projections along principal components may serve as effective reduced coordinates[3].

Committor analysis has been successfully applied across a broad spectrum of molecular processes**.** In protein–ligand unbinding, it pinpoints the kinetic bottlenecks that govern dissociation rates[87, 88]. In protein folding and conformational switching, it identifies critical intermediate states[24, 88]. In enzymatic and allosteric transitions, it reveals how structural changes couple to function. Across these diverse applications, it provides a rigorous framework for confirming the TSE and thereby deepening mechanistic understanding[88].

1.2.3 Free Energy Perturbation

Free energy perturbation (FEP) is a computational method used to quantify the effect of a perturbation on the equilibrium constant between two conformational states[14, 89-91]. Typical applications include evaluating differences in ligand binding affinities to the same protein target[17, 92, 93], assessing the impact of a single-residue mutation on conformational equilibria[94, 95], and here investigating the influence of PTMs on protein dynamics (to the best of our knowledge, it has not been previously investigated). Since

the equilibrium constant between two states is directly related to their free energy difference (ΔG), the perturbation effect can be expressed as a change in ΔG (ΔΔG)[14].

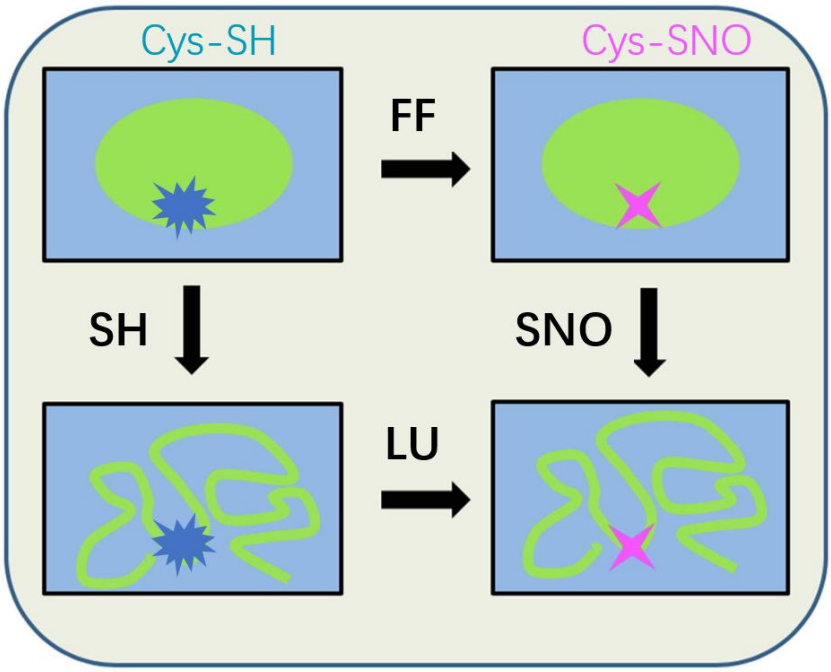
A straightforward approach is to calculate the two ΔG values before and after the perturbation, and then comparing the results. The resulting ΔΔG quantifies the perturbation's influence on the equilibrium constant. For illustration, we consider the effect of cysteine SNO on the conformational equilibrium between the FF and LU states of Prdx2.

According to Eq. (1.12), estimation of ΔG_{SH} (between FF and LU in the thiol state) requires extensive sampling of all possible configurations in both conformational states while the cysteine in its reduced form (–SH) (Fig 1.2.1). This would, in principle, involve sufficiently long MD simulations to capture the interconversion between FF and LU until equilibrium is reached, where the populations of each state remain constant over time[19, 23]. In practice, such sampling is challenging due to the limited timescales accessible to standard MD. The calculation of ΔG_{SNO} for the system after cysteine modification follows the same principle, and the difference between the two values yields ΔΔG[19].

$$\begin{aligned} \Delta\Delta G &= \Delta G_{SNO} - \Delta G_{SH} \\ &\approx -k_B T \left(\ln \frac{Z(FF,SNO)}{Z(LU,SNO)} - \ln \frac{Z(FF,SH)}{Z(LU,SH)} \right) \end{aligned} \tag{1.18}$$

Considering the difficulty of directly calculating the partition function ratios $\frac{Z(FF,SNO)}{Z(LU,SNO)}$ and $\frac{Z(FF,SH)}{Z(LU,SH)}$, the relative FEP method circumvents the need to simulate the interconversion of two conformational states by exploiting the fact that the free energy is a state function and exploiting the thermodynamic cycle illustrated in Fig 1.2.1[96]. This is evident after rewriting Eq. (1.18):

$$\begin{aligned} \Delta\Delta G &\approx -k_B T \left(\ln \frac{Z(FF,SNO)}{Z(LU,SNO)} - \ln \frac{Z(FF,SH)}{Z(LU,SH)} \right) \\ &= -k_B T \left(\ln \frac{Z(FF,SNO)}{Z(FF,SH)} - \ln \frac{Z(LU,SNO)}{Z(LU,SH)} \right) \\ &= \Delta G_{FF} - \Delta G_{LU} \end{aligned} \tag{1.19}$$



$$\Delta\Delta G = \Delta G_{SH} - \Delta G_{SNO} = \Delta G_{FF} - \Delta G_{LU}$$

Figure 1.2.1: Thermodynamic cycle used to determine the effect of SNO on Prdx2's conformational equilibrium between the FF and LU state. The green circle represents the FF state, while the green loop represents the LU state. The SH-Cys is shown in blue, and the SNO-Cys in pink. By making use of the state function property of free energy, the ΔΔG can be calculated by deriving the difference between ΔG_{FF} and ΔG_{LU}. Source: Scarabelli *et al.*, 2021.

According to Eq. (1.19), only $\Delta G_{FF/LU}$, the free energy change of transferring -SH to -SNO in the FF/LU state, is calculated. In practice, the calculation is facilitated by introducing one or more non-physical intermediate states parameterized by a coupling variable λ. The parameter λ interpolates the interactions between the two chemical states by modulating their potential energy functions[14, 89, 90]. For example, in the case of a bond potential, the conventional expression is given in Eq. (1.20)[97]:

$$V_b = \frac{1}{2}k\left(b - b_0\right)^2 \tag{1.20}$$

where b represents the bond length, b_0 is the equilibrium bond distance, and k is the force constant. In FEP simulations, this potential is modified according to λ, as expressed in Eq. (1.21)[97]:

$$V_b = \frac{1}{2}\left[(1-\lambda)k^{SH} + \lambda k^{SNO}\right]\left[b - (1-\lambda)b_0^{SH} - \lambda b_0^{SNO}\right]^2 \quad (1.21)$$

such that when λ = 0 the system is a fully interacting thiol, while at λ = 1 it represents a fully interacting S-nitrosothiol. At intermediate λ values, the system exhibits a hybrid character, with partial contributions from both states. The introduction of these λ windows allows a smooth transformation from thiol to S-nitrosothiol (Fig 1.2.2), where increasing the number of windows yields a smoother transition[14].

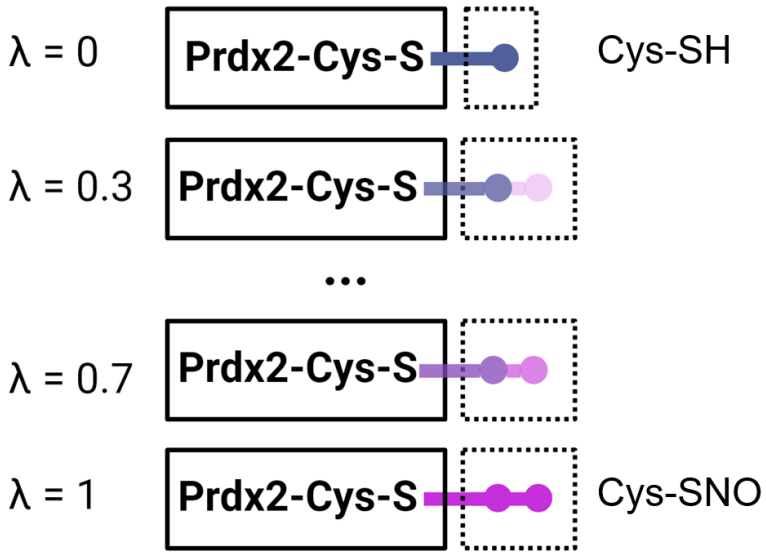


Figure 1.2.2: A diagram of the λ-dependent transformation from thiol to S-nitrosothiol. At λ = 0, the system is a fully interacting thiol (depicted in light blue), while at λ = 1, it is a fully interacting S-nitrosothiol (depicted in pink). The λ intermediate states exhibit hybrid properties and are illustrated with blended colors contributed from the two colors.

Although, in principle, ΔG could be estimated directly from samples obtained only at the two end states (-SH and -SNO), the efficiency of the analysis methods decreases markedly as the phase-space overlap between two states decreases[14]. The phase-space

overlap reflects the extent to which high-probability configurations in one state are also of high-probability in the other. When sampling is restricted to the two endpoints, the overlap is typically limited, resulting in poor convergence and large statistical uncertainties[14]. By introducing intermediate λ-windows, however, neighboring states share a much higher degree of overlap, thereby facilitating reliable reweighting between adjacent windows and improving both the accuracy and efficiency of free energy estimation[14].

As we illustrated earlier, the relative FEP is aimed to how a perturbation shifts the conformational equilibrium between two states, thereby providing a broad range of applications. These applications include, but are not limited to, the evaluation of ligand binding potency, the effect of protein mutations, and solvation or transfer processes.

One of the most common uses is in the study of ligand binding behaviour. In this context, the two states under comparison are typically the bound and unbound forms of the ligand. Relative FEP is particularly effective when the ligands share a common structural scaffold but differ in specific substituents, as these subtle variations are well-suited to alchemical transformations[14, 17, 92, 93]. By calculating the ΔG in bound/unbound state and subsequently ΔΔG, the ligand binding strengths can be quantitatively compared among these substituents.

The perturbation can also be introduced as a residue mutation in the protein, providing insight into how a point mutation influences ligand affinity or protein conformational equilibrium[94, 95]. From current knowledge, however, no studies have used this FEP method to learn how PTMs modulate protein conformational equilibrium. Since both residue mutations and PTMs involve chemical alterations of side chains, the alchemical protocols are conceptually similar. In this study, we apply relative FEP to investigate how SNO modulates the conformational equilibrium of Prdx2.

1.2.4 Finite Temperature String Method

Understanding the transition pathway that connects protein conformations is crucial for interpreting their dynamics. Such a pathway clarifies the sequence of structural rearrangements. However, direct identification of transition pathways is computationally challenging because the underlying conformational changes are rare events[16]. Enhanced-sampling approaches such as umbrella sampling, metadynamics, or steered molecular dynamics can recover free-energy surfaces along predefined CVs. These methods rely heavily on the suitability of the chosen CVs and are often impractical without prior knowledge of the pathway.

Whereas these enhanced-sampling methods aim to sample all relevant rare events and reconstruct the full free-energy landscape, the finite-temperature string (FTS) method provides a powerful alternative to rapidly identifying the most probable pathway[98-100]. Because FTS focuses exclusively on the dominant pathway rather than sampling the entire configurational space, including rare events, it is computationally less demanding than exhaustive free-energy mapping. The method requires two endpoint structures and a CV that capture the essential progress of the transition. The CV must be sufficiently descriptive to distinguish between the end states and the intermediates; in practice, they may consist of distances, dihedral angles, inter-domain vectors, or low-dimensional projections such as principal components[98].

In its standard implementation (Fig 1.2.3)[98], FTS begins with the construction of an initial string—typically a set of images generated by linear interpolation between the endpoints in CV space. For each image, short unrestrained molecular dynamics simulations are carried out to sample configurations in its neighborhood. From these simulations, the average drift is calculated and projected to the string. Then the images are subsequently re-parametrized to calculate the length along the pathway. This cycle of local sampling, average calculation, projection, and re-parametrization is repeated until the path converges. The resulting converged string defines the most probable

pathway and provides a mechanistic description of how the system is likely to progress between the two states.

The pathway identified by FTS not only enhances understanding of the transition between two states but also helps the refinement of CV. A well-defined pathway and optimized CV provide a solid foundation for subsequent enhanced-sampling methods[98, 99]. While FTS yields the structural pathway, it does not in itself provide the associated free-energy profile or transition kinetics. To extract thermodynamic information, the converged string is typically combined with umbrella sampling, weighted ensemble, or other biased-sampling strategies. When the string is generated, umbrella windows can be positioned around each image, and simulations carried out to generate overlapping phase space[98, 101]. These data are then analysed using weighted histogram analysis (WHAM)[102] or the Multi-State Bennett's Acceptance Ratio (MBAR)[103] to reconstruct the free-energy landscape along the pathway. Kinetic information, in turn, requires further methods: partitioning configuration space along the string can serve as the basis for Weighted Ensemble simulations[104]. Together, these approaches can provide both mechanistic insight into the transition pathway and quantitative estimates of free-energy barriers and rates.

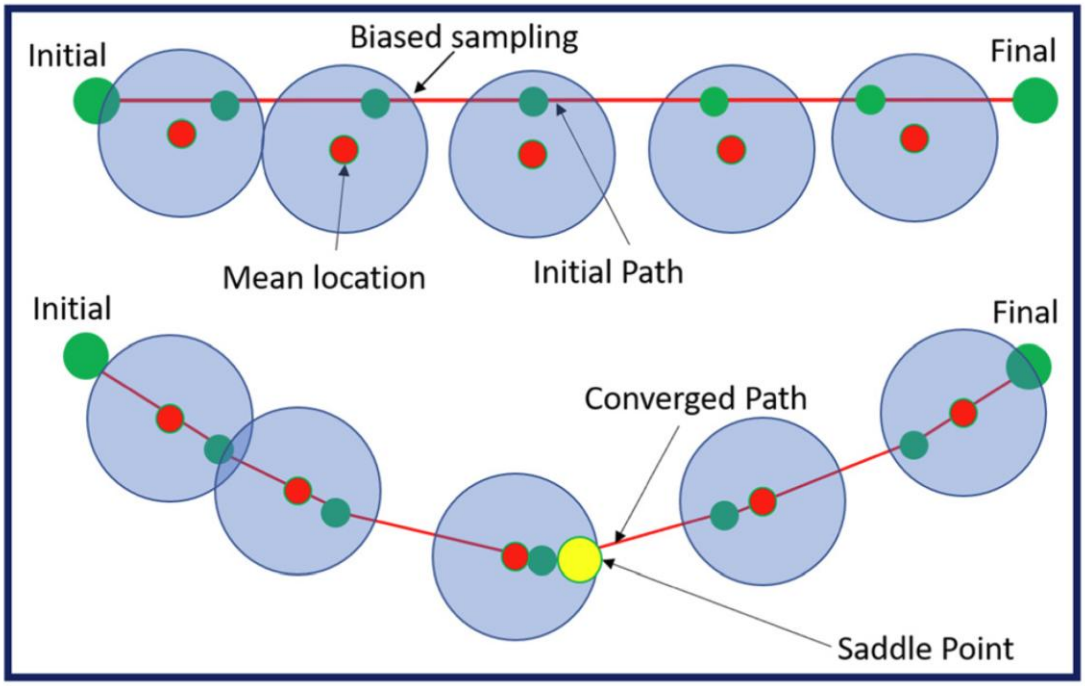


Figure 1.2.3: Diagram of the string optimization process. The red line indicates the string. The green circles mark the instantaneous positions of the images along the string. For each image, an unbiased MD simulation is performed to explore its local environment in phase space, illustrated by the large blue circle. The red circle within each blue circle represents the averaged position of the image after sampling. The yellow circle denotes the saddle point on the underlying free-energy surface. Source: Kulshrestha *et al.*, 2022.

1.3 β-Adrenoceptors and Their Binders

1.3.1 The Structure and Function of GPCR

β-ARs are part of the G protein-coupled receptor (GPCR) family, which play a central role in cellular signal transduction[105, 106]. GPCRs are embedded in the cell membrane and are structurally characterized by seven transmembrane helices. The extracellular regions detect external signals, such as hormones or neurotransmitters, while the intracellular loops interact with a heterotrimeric G protein complex composed of α, β, and γ subunits. The structure of GPCR is shown in Fig 1.3.1[107].

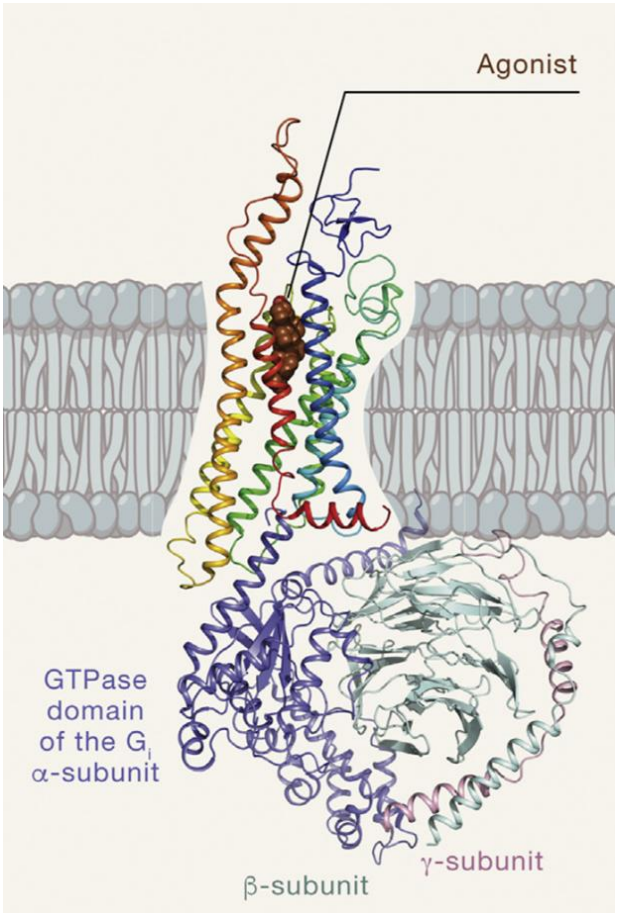


Figure 1.3.1: GPCR structure. The agonist binds to the extracellular region, and the $G_{\alpha\beta\gamma}$ complex is coupled with GPCR. Source: Congreve *et al.*, 2020.

The structures of GPCR underpin their function. The binding of an agonist ligand to the extracellular region induces a conformational change in the receptor, particularly

in its intracellular loops. This change promotes the exchange of Guanosine Diphosphate (GDP) for Guanosine Triphosphate (GTP) on the G_α subunit. As a result, the G_α-GTP dissociates from the $G_{\beta\gamma}$ dimer, and both components go on to regulate their respective downstream effectors. Eventually, the intrinsic GTPase activity of the G_α subunit hydrolyzes GTP to GDP, allowing G_α to reassociate with $G_{\beta\gamma}$, thereby restoring the inactive heterotrimeric G protein complex. The complete GPCR activation cycle is shown in Fig 1.3.2[108].

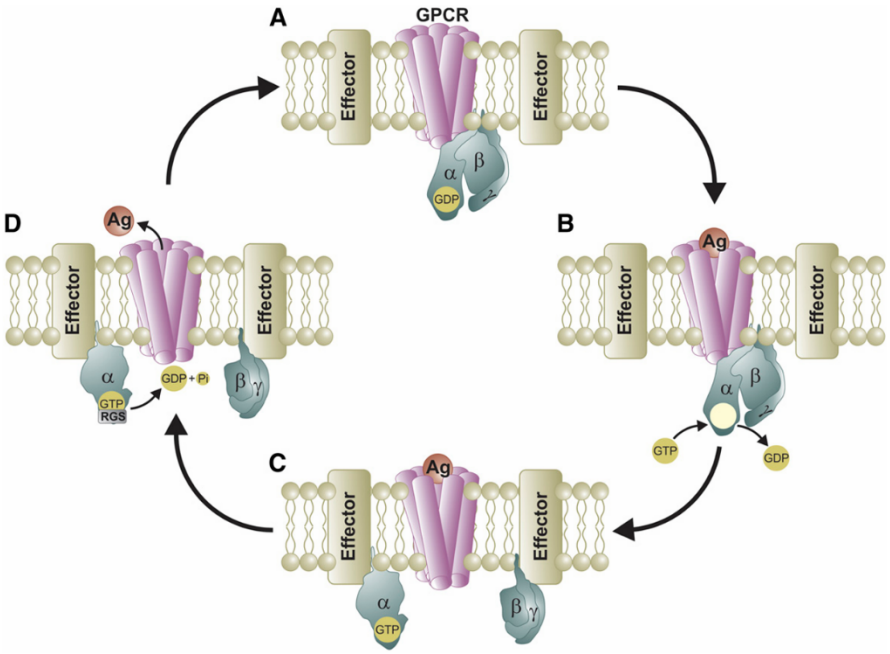


Figure 1.3.2: GPCR activation cycle. In the inactive state of the receptor (A), the GDP-bound Gα subunit is associated with $G_{\beta\gamma}$ dimer and they together are coupled to the GPCR. (B) Agonist binding induces a conformational change in the receptor, catalyzing the exchange of GDP for GTP on the G_α subunit. (C) The GTP-bound G_α subunit dissociates from the $G_{\beta\gamma}$ dimer, and both components activate downstream signaling pathways. (D) Intrinsic GTP hydrolysis by the G_α subunit—often accelerated by regulator of G protein signaling (RGS) proteins—converts GTP to GDP, enabling reassociation of G_α with $G_{\beta\gamma}$ and thereby terminating the signal. Source: Proft *et al.*, 2014.

1.3.2 β-AR G Protein Signalling Pathway

The β1-AR and β2-AR primarily signal through the same G protein-coupled pathway[108-110]. Their endogenous ligands are epinephrine and norepinephrine. Upon

ligand binding, the receptor activates the heterotrimeric G protein, leading to the exchange of GDP for GTP on the G_α subunit. This causes G_α to dissociate from the $G_{\beta\gamma}$ dimer and activate the primary effector, adenylyl cyclase (AC). Activated AC converts ATP into cyclic AMP (cAMP), which in turn activates protein kinase A (PKA). PKA phosphorylates specific target proteins, eliciting physiological responses. In some tissues, cAMP can also activate cAMP-gated ion channels, contributing to cellular excitability (Fig 1.3.3)[108-110].

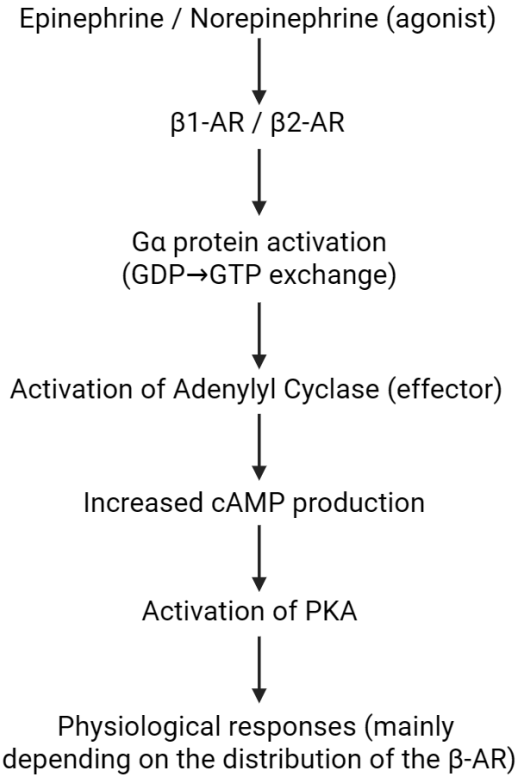


Figure 1.3.3: β-AR G protein signalling pathway. The β1-AR and β2-AR share the same pathway. Binding of epinephrine or norepinephrine to β-AR activates the associated Gα protein via GDP–GTP exchange. The dissociated Gα subunit stimulates AC, leading to increased cAMP production. Elevated cAMP primarily activates PKA, which phosphorylates downstream targets to elicit physiological responses. The specific responses depend largely on the tissue distribution of β1-AR versus β2-AR.

The primary reason for the different physiological responses of β1-AR and β2-AR lies in their distinct tissue distributions. The β1-ARs are predominantly distributed in

the heart and kidneys; thus, their activation primarily leads to increased cardiac output and elevated blood pressure. In contrast, the β2-ARs are mainly found in the lungs and vascular or skeletal muscle, where their activation induces bronchodilation and vasodilation, contributing to decreased peripheral resistance and facilitating airflow (Fig 1.3.3)[111, 112].

1.3.3 Selective β1-AR Inhibitors

Non-selective β-AR antagonists are widely used to treat cardiovascular conditions including heart failure, myocardial infarction, angina, and hypertension. Their therapeutic effect is primarily mediated through inhibition of β1-ARs which are predominantly distributed in the heart[113]. However, non-selective β-AR antagonists also block β2-ARs, which are mainly distributed in lung cells, leading to bronchoconstriction and other respiratory adverse effects. These side-effects prevent patients with chronic obstructive pulmonary disease (COPD) from receiving non-selective β-AR antagonists. Therefore, β1-selective blockers can avoid these side-effects and be more preferred for such patients[114-117].

Structurally, β-AR ligands—whether agonists or antagonists—commonly share a phenylethanolamine scaffold (Phe–CH(OH)–CH_2–NH–), comprising an aromatic ring attached to a hydroxy-substituted ethylamine chain[118]. At physiological pH, the amine group is protonated, enabling a strong ionic interaction with a conserved $Asp^{3.32}$ residue in transmembrane helix 3 (TM3) and an additional hydrogen bond with a conserved $Asn^{7.39}$ in TM7. The β-hydroxyl group also form two hydrogen bonds with both residues. Together, the amine and hydroxyl moieties provide four polar interactions, anchoring the ligand firmly within the orthosteric binding site (Fig 1.3.4). The aromatic ring is accommodated within a hydrophobic pocket, contributing to binding stability through van der Waals contacts[119-122]. Substituents on the aromatic ring influence receptor subtype selectivity and the functional outcome of binding.

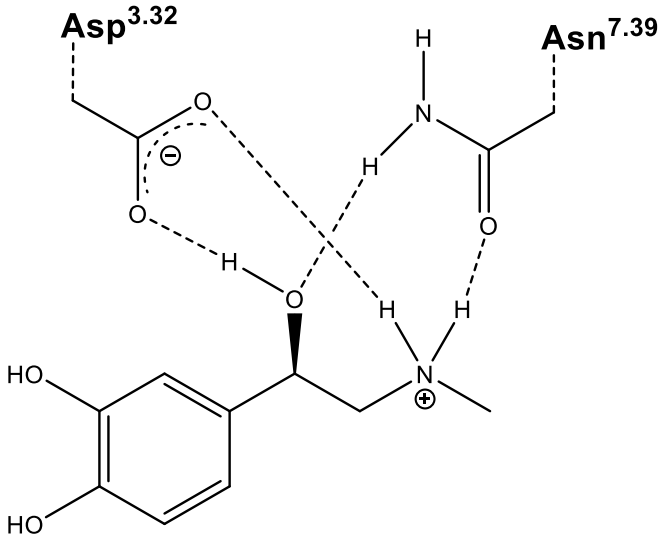


Figure 1.3.4: Structural features common to β-AR ligands and their conserved polar interactions, exemplified by adrenaline. The protonated amine group forms an ionic interaction with a conserved $Asp^{3.32}$, and hydrogen bond with a conserved $Asn^{7.39}$. The β-hydroxyl group also forms hydrogen bonds with $Asp^{3.32}$ and $Asn^{7.39}$, respectively.

In non-selective agonists such as epinephrine and norepinephrine, two phenolic hydroxyl groups enable hydrogen bonding with two conserved Ser residues in TM5, an interaction essential for β-AR activation[120]. Non-selective antagonists typically lack these phenolic hydroxyl groups; instead, the presence of large, rigid aromatic substituents can sterically hinder the outward movement of TM6[119, 123-125]—a conformational change critical for receptor activation—thereby stabilizing the inactive state of the receptor and producing inhibitory effects exemplified by propranolol and cyanopindolol (Fig 1.3.5).

Different from the rigid aromatic ring characteristic of non-selective antagonists, selective β1-AR inhibitors, such as atenolol and bisoprolol, feature an extended side chain attached to the ring of phenylethanolamine[42, 126, 127]. However, for all the published X-ray crystal structures of complex of β-ARs, their binding ligands all lack extended side chains (Fig 1.3.5). Simulation directly docking "extended" ligands to β-ARs fails to reproduce the canonical binding pose[128], in which the ethanolamine core engages in four polar interactions with the conserved residues $Asp^{3.32}$ and $Asn^{7.39}$—a

hallmark of β-AR ligands. Furthermore, no apparent static conformation of the ligand-binding pocket appears capable of accommodating these extended moieties. It implies that the β1-AR may undergo conformational rearrangement upon the binding of these selective β1-AR antagonists, thus the extended side chain can be accommodated (Fig 1.3.6)[128].

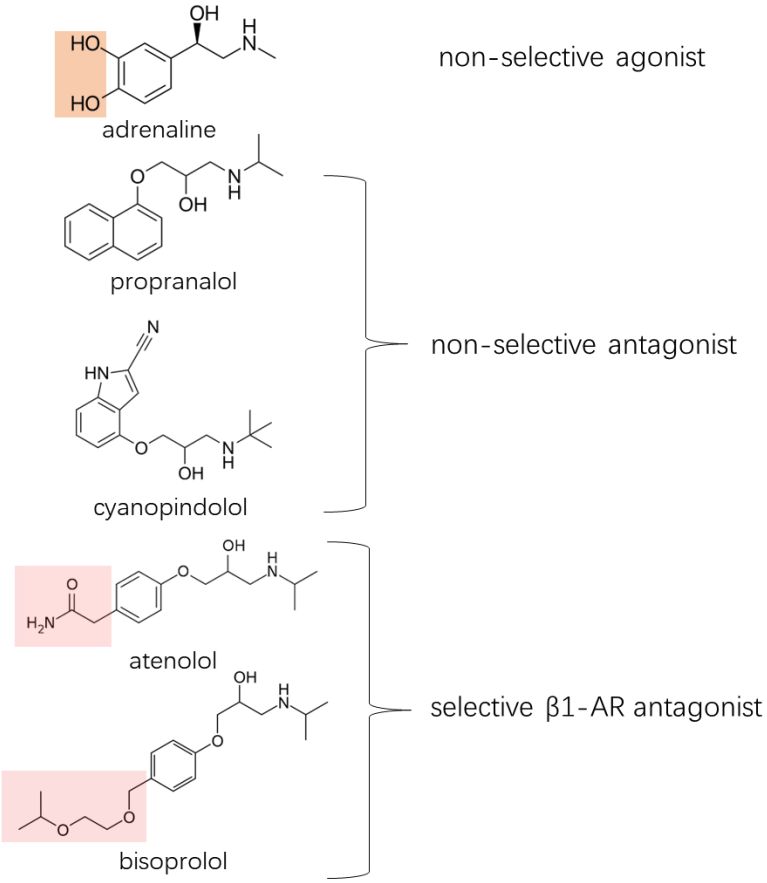


Figure 1.3.5: Structural comparison between non-selective β-AR agonist and antagonist and selective β1-AR antagonist. Non-selective agonists, exemplified by adrenaline, typically feature two phenolic hydroxyl groups (colored by light orange), while non-selective inhibitors exhibit a large, rigid aromatic substituents such as propranalol and cyanopindolol. The selective β1-AR antagonists, such as atenolol and bisoprolol, possess an extended side-chain (highlighted in pink) that contributes to subtype selectivity.

Assuming that the flexible region of the protein is responsible for accommodating extended substituents, the conformational flexibility of β-ARs was investigated using

Active Site Pressurization (ASP)—a computational approach that continuously injects Lennard–Jones particles into the binding pocket to probe structural plasticity[129, 130]. During the ASP simulations, a fissure between transmembrane helices H4 and H5 was found in both subtypes. This fissure is termed ‘keyhole’ (Fig 1.3.7). Docking ligands with the extended side-chain into the keyhole conformation successfully generates a pose in which the keyhole ideally accommodates the extended-side chain, while the rest of their structure interacts with the orthosteric binding site in the canonical pose (Fig 1.3.6)[128].

Interestingly, the keyhole may have a crucial relevance to subtype selectivity. It is proposed that the slight difference in how extensive the side chain occupies the keyhole, enforced by structural constraints given by similar but different conformation of two subtypes, confers different effectiveness for the tail extensions resting and interacting in and with the keyhole (Fig 1.3.6). This difference derives from the non-conserved Val/Thr$^{4.56}$ in β1- and β2-AR, respectively[131]. The polar interaction between Thr$^{4.56}$ in β2-AR and water molecules may occlude the keyhole, limiting access of the ligand’s extended side chain. Therefore, the keyhole is more readily formed in β1-AR than in β2-AR, explaining why selective β1-AR antagonists preferentially induce and exploit this conformation[131].

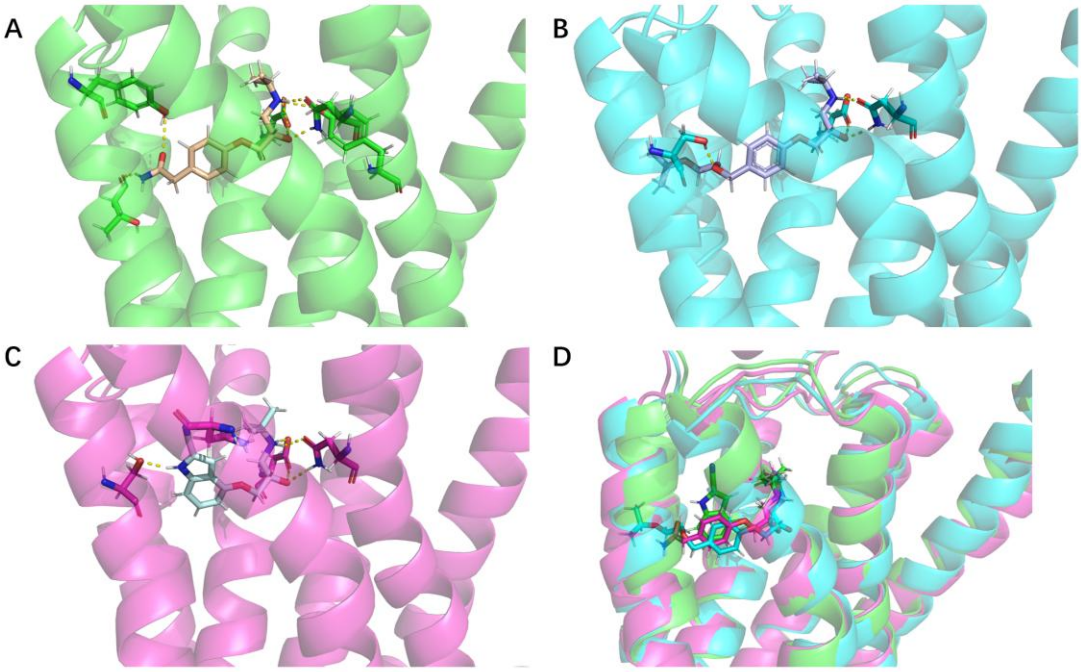


Figure 1.3.6: Binding poses for (A) atenolol, (B) bisoprolol, and (C) cyanopindolol in complex with β2-AR. The protein is shown in cartoon style. Residues involving key polar interactions are shown in stick style. The hydrogen bonds are shown in yellow dashed lines. (D) The alignment of the three complexes, with atenolol colored green, bisoprolol cyan, and cyanopindolol pink.

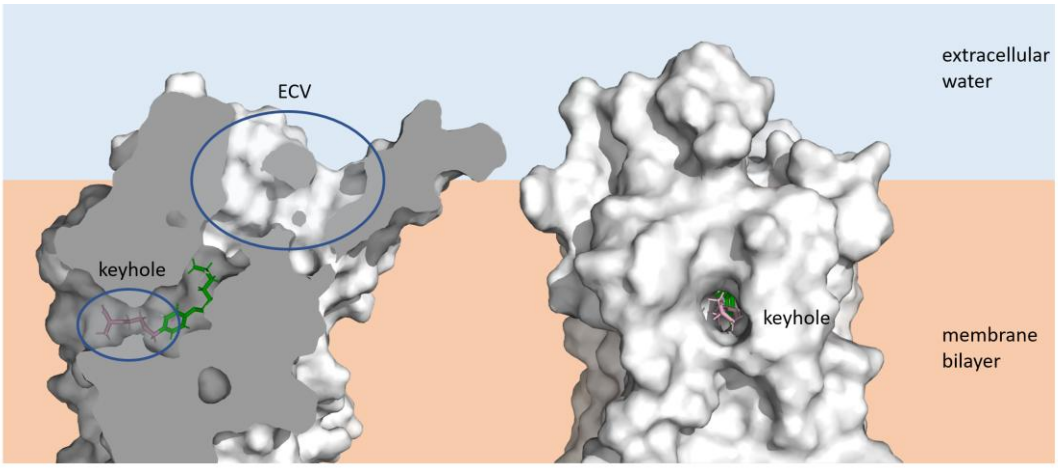


Figure 1.3.7: Dissociation pathways of the ECV and keyhole, illustrated using bisoprolol-β2-AR complex. The ECV opens to the extracellular water, while the keyhole constitutes a fissure between transmembrane helices H4 and H5 linking the orthosteric binding site to the membrane bilayer. The long side chain of bisoprolol (pink) occupies the keyhole, while the remaining portion of the ligand is shown in green.

The above findings explain the β1-AR selectivity in the perspective of structure. As illustrated in Section 1.1.2, the drug selectivity may have high relevance on its kinetics. An accurate prediction of dissociation pathway and rate is crucial for drug selectivity optimization. The k_{off} is mainly determined by the ligand-protein interactions along the pathway. The keyhole accommodates the extended side-chain, it may potentially also act as a dissociation pathway. To prove that not only conventional extra cellular vestibule (ECV) as the dissociation pathway, keyhole also can be a dissociation pathway for atenolol and bisoprolol (Fig 1.3.7), dissociation simulations are run via τRAMD.

Although the primary pharmacological targets of atenolol and bisoprolol are β1-ARs, no experimentally measured K_d and k_{off} values are available, so instead we focus here on their binding to β2-AR as the model system, where experimental data for both ligands is available, as well as for cyanopindolol. In contrast to atenolol (0.02 ± 0.005 s^{-1}) and bisoprolol (0.1 ± 0.03 s^{-1}), cyanopindolol ($3\times10^{-5} \pm 1.6\times10^{-5}$ s^{-1}) is a non-selective β-AR antagonist lacking an extended side chain, and it exhibits a markedly lower k_{off}[132]. The three ligands therefore form an interesting test set for the comparison of dissociation mechanisms between selective β1-AR antagonists and non-selective β-AR antagonists.

1.4 Peroxiredoxin 2

1.4.1 Overview on Peroxiredoxin Family

Peroxiredoxins (Prxs) constitute a large and highly conserved family of thiol-dependent peroxidases[133-136]. Their catalytic activity relies on a conserved cysteine residue, termed the "peroxidatic" cysteine (C_P), which undergoes oxidation by peroxides. During catalysis, the thiol cysteine (SH-C_P) is oxidized to a sulfenic acid cystine (SOH-C_P), which subsequently reacts with a second cysteine, known as the "resolving" cysteine (C_R), to form a disulfide bond[135]. The classification of Prxs is based on the presence and location of the C_R: (i) typical 2-Cys Prxs (Prdx1, Prdx2, Prdx3, Prdx4) form intersubunit disulfide bonds, (ii) atypical 2-Cys Prxs (Prdx5) form intrasubunit disulfide bonds, and (iii) 1-Cys Prxs (Prdx6) lack a C_R and instead form disulfide bonds with other proteins or low-molecular-weight thiols[63, 135, 136].

Prxs are expressed at high levels across virtually all species and are ubiquitously distributed in diverse subcellular compartments, including mitochondria, chloroplasts, peroxisomes, and, most abundantly, the cytosol[137]. In *Escherichia coli*, Prxs rank among the ten most abundant proteins[138], whereas in human erythrocytes, they represent the third most abundant protein. Their widespread distribution and abundance underscore their central role in the reduction of peroxides and maintenance of redox homeostasis[135].

Beyond their role in peroxide reduction, Prxs participate in the regulation of cell signaling. By modulating the level of local H_2O_2 concentration, they indirectly engage in redox-sensitive signaling pathways: their activation decreases local H_2O_2 levels, while their inactivation increases, thereby triggering stress responses[134, 136, 139]. In addition, via redox-switch, Prxs can play a more direct role in signaling. For example, oxidized and overoxidized Prxs can be repaired by Thioredoxin (Trx)[140] and Sulfiredoxin (Srx)[141], respectively (Fig 1.4.2). These repair processes not only restore Prx antioxidant activity but also alter the redox state of Trx and Srx, potentially leading

to related subcellular relocalization, and functional modulation of their target proteins[142]. Furthermore, certain Prxs display molecular chaperone activity, facilitating the folding and unfolding of client proteins and protecting them from aggregation. Notably, hyperoxidation of Prxs, which impairs their peroxidase function, is typically accompanied by an enhancement of their chaperone activity, suggesting a functional compensation mechanism[134, 143, 144].

1.4.2 Peroxiredoxin 2 and Its Catalytic Cycle

Peroxiredoxin 2 (Prdx2) is a member of the typical 2-Cys Prx subfamily and exhibits both peroxidase and chaperone activities. The intracellular concentration of Prdx2 reaches ~5.6 mg mL^{-1}, making it the third most abundant protein in erythrocytes[145]. Its catalytic mechanism involves the formation of an intersubunit disulfide bond between 51C_P of one and 172C_R of the other subunit within the homodimer (Fig 1.4.1A). Each subunit of Prdx2 consists of 198 amino acid residues with a molecular mass of ~22 kDa. While the functional unit is a homodimer (Fig 1.4.1A), Prdx2 can also assemble into higher-order oligomers, including decamers (Fig 1.4.1B) and high-molecular-weight (HMW) complexes that typically form hollow cylindrical structures (Fig 1.4.1C)[134, 146-148].

The catalytic cycle of Prdx2 is shown in Fig 1.4.2. The SH-C_P first reacts with H_2O_2 at an exceptional reaction rate, resulting in its oxidation to SOH-C_P. This intermediate can follow two distinct fates: it may (i) react with the SH-C_R to form an intersubunit disulfide bond, which is subsequently reduced by Trx, or (ii) undergo further oxidation to sulfinic acid cysteine (SO_2H-C_P) upon reaction with an additional H_2O_2 molecule. The latter modification is reversible through the assistance of Srx, which reduces SO_2H-C_P back to SH-C_P. The decision between disulfide bond formation and hyperoxidation is governed by the relative kinetics of these competing pathways and the local concentration of H_2O_2. When disulfide bond formation is faster than hyperoxidation, SOH-C_P preferentially reacts with SH-C_R; conversely, at elevated H_2O_2 levels, hyperoxidation predominates. Under conditions of extreme oxidative stress, C_P

can be further oxidized to the irreversibly inactivated cysteine sulfonic acid (SO_3H–C_P)[133].

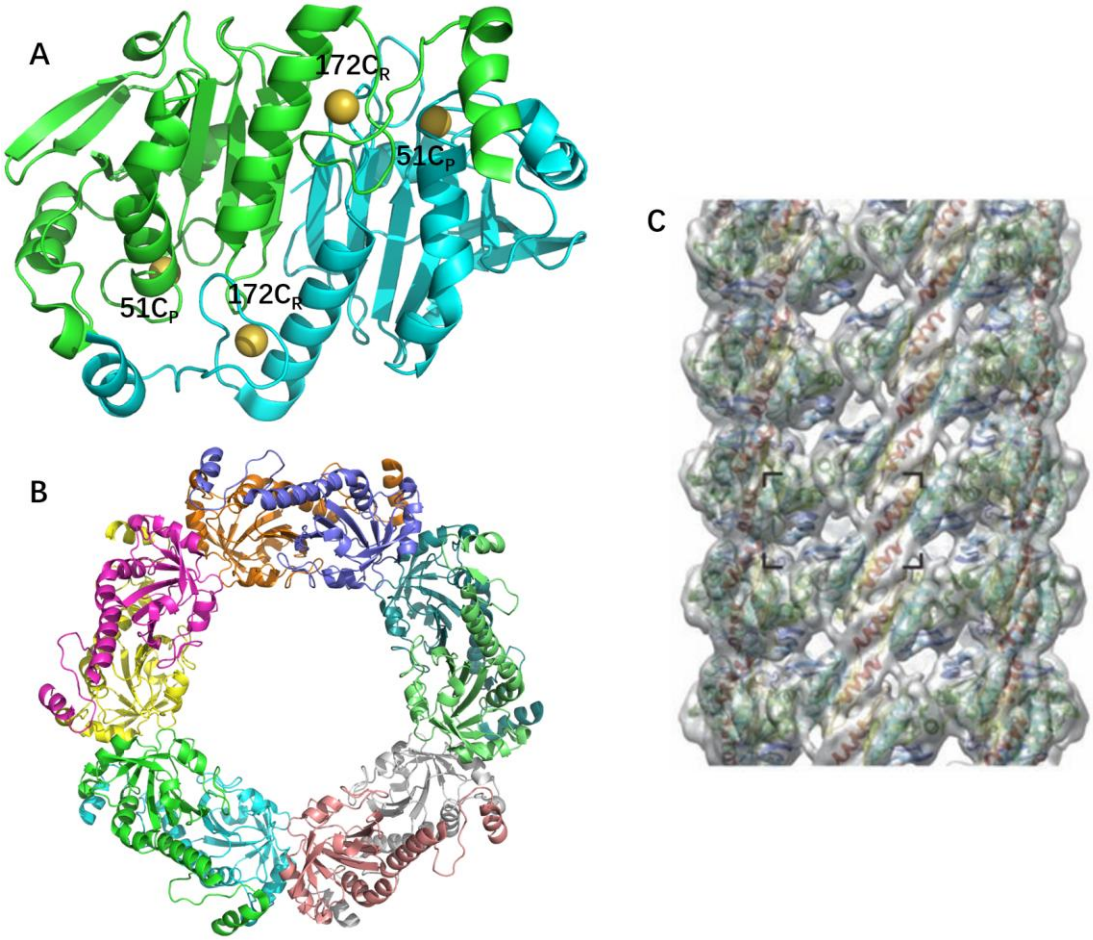


Figure 1.4.1: Oligomeric Conformations of Prdx2. (A) Homodimer of Prdx2 with one subunit coloured green and the other cyan. Sulphur atoms of the $51C_P$ and the $172C_R$ are shown as yellow spheres. (B) Decameric ring of Prdx2 with ten subunits shown in different colours. (PDB ID: 1QMV). (C) HMW of Prdx3. Adapted from Cao *et al.*, 2017.

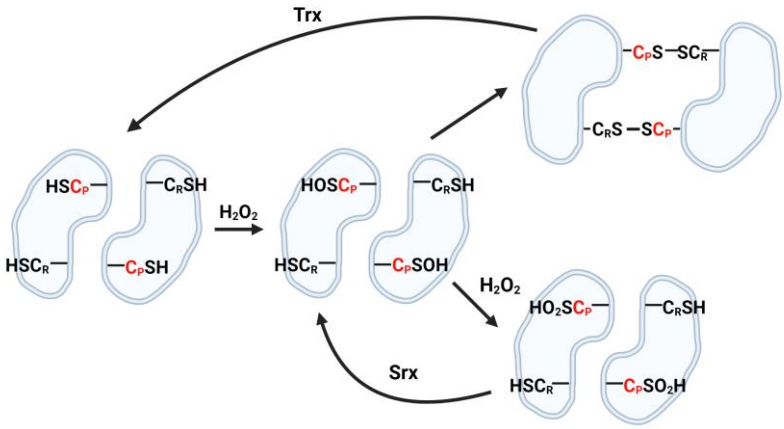


Figure 1.4.2: The Prdx2 catalytic cycle. C_P is highlighted in red.

1.4.3 FF and LU state – Two Main Conformational States of Prdx2

To begin with, we introduce post-translational modifications (PTMs). PTMs are covalent modifications of one or more amino acid residues that occur after protein biosynthesis. They serve as a crucial regulatory layer that greatly expands the functional diversity of the proteome beyond what is encoded at the genomic level. Examples include phosphorylation, acetylation, methylation, and various redox-based modifications, such as those that occur during the catalytic cycle of Prdx2[1, 149]. From this perspective, the Prdx2 catalytic cycle can be regarded as a paradigm for how PTMs dynamically regulate protein conformation and activity.

During the catalytic cycle, Prdx2 undergoes conformational transitions that are closely coupled to its PTMs. X-ray crystallography has revealed two principal conformational states: the FF state[145] and the LU state[147] (Fig 1.4.3). These conformations are distinguished primarily by the structural arrangement of the C-terminal region (highlighted in red in Fig 1.4.3A and C) and the relative positioning of the C_P and the C_R (depicted as sticks in Fig 1.4.3A and C).

In the FF state (Fig 1.4.3A and B), the C-terminal adopts a well-ordered α-helix that folds back onto the protein core and makes stable contacts with the α2 helix (cyan). In this conformation, C_P and C_R are oriented away from each other. In contrast, in the LU state (Fig 1.4.3C and D), the C-terminal helix becomes unwound and displaced,

adopting a more flexible and solvent-exposed conformation that no longer interacts with the α2 helix. This conformational change allows C_P and C_R to reorient toward each other, thereby reducing the distance between them and facilitating the formation of an intersubunit disulfide bond.

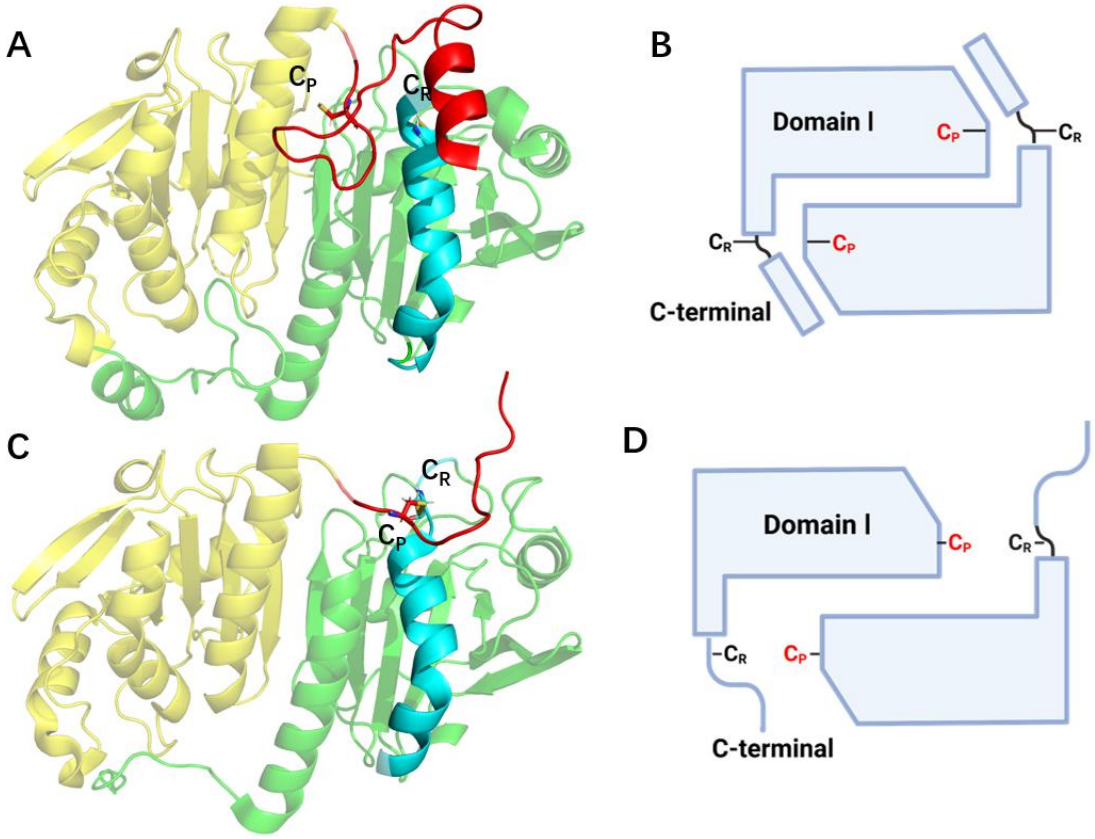


Figure 1.4.3: Conformations of the FF state (PDB: 1QMV) and the LU state (PDB: 5IJT). (A) Structure of the FF state in cartoon style with two crucial cysteines (C_P and C_R) in stick representation. One chain is shown in yellow, while the other is colored green, cyan, and red. Wherein, the representation of red is the C-terminal and cyan is the α2 helix. (B) A diagram of the structure of the FF state, where the C-terminal is folded and interacts with the α2 helix located in domain I. The C_P and C_R orientate in the opposite direction of the position of each other. (C) Structure of the LU state. Same as (A), different regions are shown in various representation styles and colors. (D) A diagram of the LU state, the drawing style of which is the same as (B). Its C-terminal is released and has no interactions with the α2 helix. Moreover, the C_P and C_R orientate towards each other.

1.4.4 FF state—a structure basis of antioxidant and chaperone activity

Only the FF state of Prdx2 provides the catalytic environment necessary for C_P to reduce peroxides at an exceptionally high rate (Fig 1.4.4). The 51C_P is embedded within a highly conserved microenvironment formed by 44Pro, 48Thr, and 127Arg, residues that are also preserved across other Prx subtypes. The positively charged guanidinium group of 127Arg stabilizes the thiolate form of C_P (S^--C_P) by lowering its pK_a, thereby enhancing its nucleophilicity. The hydroxyl group of 48Thr forms a hydrogen bond with the incoming peroxide substrate, contributing to its proper orientation in the active site. Meanwhile, 44Pro helps maintain the structural geometry of the active-site pocket, ensuring the correct positioning of the catalytic residues. Together, this acid–base catalytic motif generates an activated S^--C_P that attacks the electrophilic oxygen atom of H_2O_2, thereby facilitating its reduction. This unique structural arrangement accounts for the exceptional reactivity of Prdx2 in the FF state[147, 150].

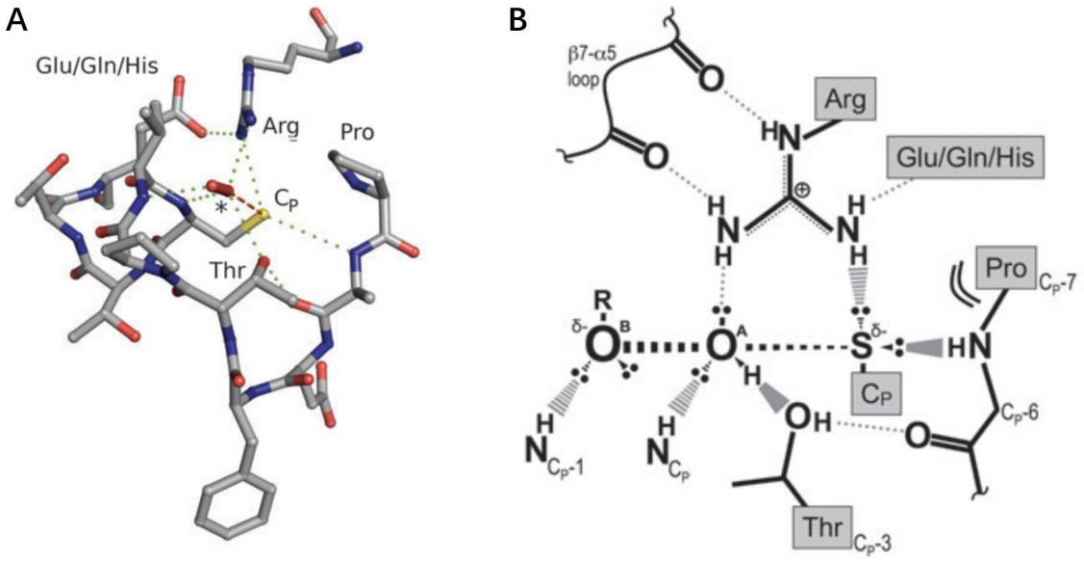


Figure 1.4.4: The active site of FF Prdx2 with a binding H_2O_2. Source: Hall *et al*, 2011.

The FF state is also fundamental to the chaperone activity of Prdx2. In this conformational state, Prdx2 exhibits a strong tendency to assemble into decamers and, in some cases, into HMW oligomers composed of stacked decamers. By contrast, the LU state is predominantly dimeric and lacks the capacity for higher-order assembly. The structure of the FF state therefore provides the molecular basis for decamer and HMW formation, which are essential for chaperone function. Nevertheless, this does not imply that the LU state is incapable of decamer formation; rather, it tends to form less stable decamers. The crystallized decameric structures derived from the LU conformation demonstrates that decamerization, although weakened, remains possible in this state[59, 134, 146, 148].

The link between conformational state and oligomerization explains why the FF state is consistently associated with chaperone activity. The transition from FF to LU destabilizes decamers, leading to their dissociation into dimers and the loss of chaperone function. Conversely, the reverse transition from LU to FF promotes the reassembly of decamers and the recovery of chaperone activity.

In summary, the FF state not only furnishes the catalytic environment for antioxidant activity but also provides the structural framework for higher-order oligomerization, thereby enabling the chaperone function of Prdx2.

1.4.5 Conformational Change During Catalytic Cycle

The catalytic cycle of Prdx2 involves the switch between FF and LU states, driven by redox modifications of the two cysteines[136, 146, 147]. The diagram for the Prdx2 conformational change during catalytic cycle is shown in Fig 1.4.5. Although the precise dynamics of the transition between the FF and the LU state remained unknown, an SH-C_P equilibrium favouring the FF state has been certified[147]. Upon oxidation of SH-C_P to SOH-C_P, the protein tends to shift toward the LU state. However, the FF→LU transition requires time, creating a window in which two competing fates for SOH-C_P can occur.

At physiological H_2O_2 concentrations, the conformational transition to the LU state proceeds successfully, bringing SOH-C_P and SH-C_R into close proximity and enabling disulfide bond formation. This locks the protein in the LU state. By contrast, under elevated H_2O_2 conditions, SOH-C_P can be further oxidized to sulfinic acid (SO_2H-C_P) before the FF→LU transition is complete[144]. Unlike the reduced form, SO_2H-C_P exhibits an even stronger preference for the FF state, thereby promoting stable decamer and HMW oligomer formation[143, 151]. Importantly, the oxidation of SOH-C_P to SO_2H-C_P is reversible through the action of Srx, which regenerates SOH-C_P in the LU state, allowing subsequent disulfide bond formation with SH-C_R[148, 152]. By contrast, once a disulfide bond between C_P and C_R (C_P-S-S-C_R) is established, the protein remains locked in the LU state until the bond is reduced by Trx[136, 148]. This reduction restores SH-C_P in the LU state, which can then rapidly transition back to the FF state.

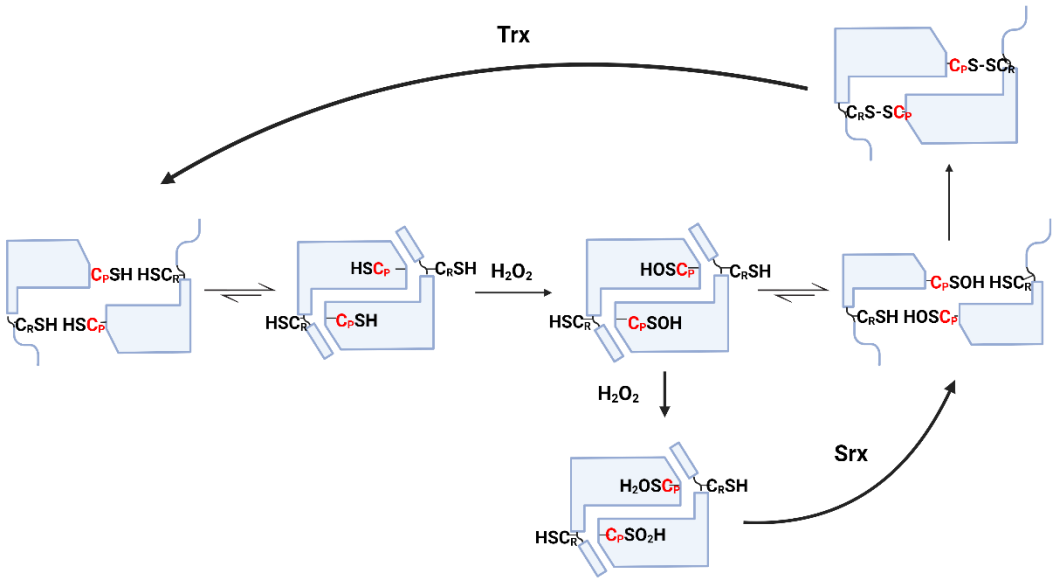


Figure 1.4.5: Diagram for the Prdx2 conformational change during catalytic cycle.

As illustrated before, Prdx2 not only exhibits antioxidant activity but also chaperone activity, and the latter is strongly correlated with the ability of the protein to assemble into HMW oligomers. The influence of PTMs on this function is particularly evident. Although both SH-C_P and SO_2H-C_P retain chaperone activity, the latter

consistently displays stronger activity. This distinction arises from the conformational preferences imposed by the modifications.

In the SH-C_P state, Prdx2 exists in an equilibrium between the FF and LU states. Because LU destabilizes decameric assemblies, the equilibrium limits oligomer stability, allowing only relatively weak decamers[147] to form and thereby supporting basal levels of chaperone activity. By contrast, in the SO_2H-C_P state, the protein is conformationally locked in the FF state[145], which favors stable decamerization and further the assembly of HMW complexes. This structural stabilization underlies the markedly enhanced chaperone activity of SO_2H-C_P[153, 154]. In contrast, when a C_P-S-S-C_R is formed, the protein adopts the LU state exclusively, which disrupts decamer stability and abolishes chaperone function[155].

Taken together, these observations establish a clear pattern: regardless of the specific PTM, it is the FF state that provides the structural foundation for decamer and HMW formation, and thereby confers chaperone activity to Prdx2. PTMs influence chaperone function primarily by modulating the conformational equilibrium between FF and LU states—stabilizing the FF state enhances chaperone activity, whereas favoring the LU state diminishes or eliminates it.

1.4.6 SNO of Prdx2 and Its Pathological Consequences in Parkinson's Disease

The catalytic cycle of Prdx2 includes multiple redox-modifications under physiological conditions (Fig 1.4.2). They together modulate the function of Prdx2. However, aberrant redox modifications frequently occur under pathological conditions and can impair the activity of Prdx2. For instance, a significantly increased level of SNO-Prdx2 has been reported in Parkinson's disease (PD), where both C_P and C_R are S-nitrosylated[156, 157]. Upon SNO, the C_P can no longer react with H_2O_2 effectively, thereby impairing the antioxidant activity of Prdx2.

A direct result of the dysfunction of Prdx2 is the accumulation of its substrates-reactive oxygen and nitrogen species (RONS). RONS are broadly classified into two categories: (i) radical species, such as hydroxyl (•OH), superoxide ($O_2^{\bullet -}$), and nitric oxide (NO•), and (ii) non-radical species, including H_2O_2, nitrosyl cation (NO^+), and dinitrogen trioxide (N_2O_3)[158, 159]. Typically, the former exhibits higher reactivity than the latter; however, under certain conditions, non-radical RONS can transform into radical RONS, such as the homolytic cleavage of H_2O_2 resulting the formation of two $OH^{\bullet}$s[160, 161]. The endogenous source of radicals primarily stem from metabolism and the functions of enzymes: NADPH oxidases (NOXs), myeloperoxidase, and nitric oxide synthases (NOSs)[162, 163]; exogenous sources like air pollution, alcohol, and pesticide also contribute to the generation of radicals[163]. Under normal physiological conditions, a fair level of RONS plays an important role in intra- and intercellular signalling pathways. However, under pathological conditions, excessive RONS accumulation occurs. Due to their high chemical reactivity, these species readily attack lipids, nucleic acids, and proteins, leading to structural damage and loss of cellular function[160, 164].

An abnormal high level of RNOS is present in aging and various diseases such as cancer, diabetes mellitus, inflammation, liver damage, asthma, cataract, cardiovascular diseases, and neurodegenerative diseases[165-170]. Although the precise mechanism of how RNOS interact with the above diseases is complicated and not yet fully understood but from the existing knowledge, the pathological mechanism primarily falls into the alteration of DNA, RNA, lipid, and protein along with aberrant cell signalling by radicals[159-161, 169]. For example, by means of radicals, an elevated growth signalling promotes the proliferation of cancer cells, while the mutations of DNA and subtle modifications in key cancer-associated proteins impede the DNA repair and cell apoptosis. Age-related degeneration, on the other hand, partly results from free radical damage, and decreasing the radical level has been found to prolong the lifespan of cells.

The PD pathway in detail is as followed (Fig 1.4.6). An increased extracellular glutamate (Glu) level is observed. Glu target protein N-methyl-D-aspartate receptor (NMDAR) is hyper-activated, leading to an increase in the free radical and nitric oxide (NO) level, and excessive intracellular calcium (Ca^{2+}) level[171-173]. The excessive NO reacts with a specific group of proteins and S-nitrosylates specific cysteines from proteins, including Prdx2. The SNO-Prdx2 results in the loss of antioxidant activity of Prdx2, further, in turn, worsens the radical stress. Dopaminergic neurons are vulnerable to oxidative stress. The loss of dopaminergic neurons is the characterization of PD, consequently, the SNO of Prdx2 would lead to or worsen the pathology of PD[167, 171-173]. On the other hand, the repair of SNO-Prdx2 to a reduced state is expected to alleviate PD.

Also, the hyperactivation of NMDAR gives rise to an elevated intracellular Ca^{2+} level, which brings about hypersynchrony, hyperexcitability, synaptic damage, and excitotoxicity. These abnormal actions are related to neurodegenerative diseases.

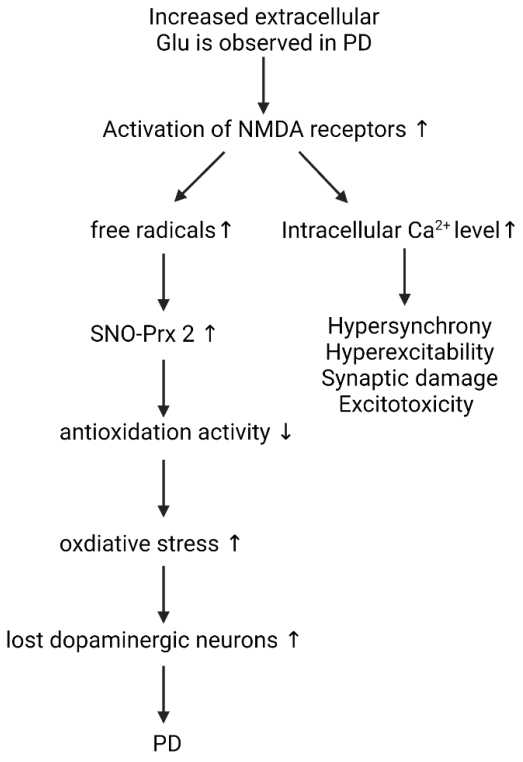


Figure 1.4.6: Pathway linking Prdx2 SNO to Parkinson's disease.

1.5 Aims and Objectives

- Propose a new analysis framework for predicting ligand residence times in the absence of a custom force based on τRAMD simulations.
- Investigate the dissociation pathways and rates of atenolol, bisoprolol, and cyanopindolol from β2-AR, thereby elucidating the mechanistic differences in dissociation between β1-AR–selective antagonists and non-selective antagonists.
- Estimate TSE along dissociation pathways by using committor analysis, distinguish the structural features of the three ligands that contribute to differences in k_{off}.
- Apply PCA to evaluate how different PTMs influence the dynamics of Prdx2 in both FF and LU conformational states over 40-ns trajectories.
- Establish a FEP protocol to investigate the effects of PTMs on protein conformational equilibrium.
- Quantify the impact of S-nitrosylation on the FF and LU conformational equilibrium with FEP, and further propose a mechanistic model for S-nitrosylation process of Prdx2.
- Utilize the FTS method to characterize the most probable transition pathway between the FF and LU states.

2. Methods

All software and computational tools used in the thesis are listed in Table 2.1.

Software or tool	Version
GROMACS	2021.6
Amber	24
Schrödinger	2025.3
Antechamber	24
BisoSimSpace	2024.3.0
PropKa	2.0
CHARM-GUI	23
Open Babel	2023
MDTraj	1.10.2
ParMed	4.3.0
VMD	1.9
PyMOL	2.6
Chimera	1.19
MDplus	0.1.1
Matplotlib	3.10.0
Origin	2018

Table 2.1: Used software and tools.

2.1 Ligand Residence Time Prediction

Through enough τRAMD simulations and our novel data analysis framework, the relative and, notably, absolute k_{off} values can be predicted. Different from conventional approaches that average unbinding times to represent the residence time, we monitored the temporal evolution of the survival number of the ligand–receptor complex and extrapolated both relative and absolute k_{off} from it[72-75]. The complete residence time

prediction workflow including system preparation, dissociation simulation, egress route check and k_{off} analysis was illustrated in this section and summarized in Fig 2.1.

With the generated ligand dissociation pathways, an attempt was made to apply committor analysis to identify the transition state[86].

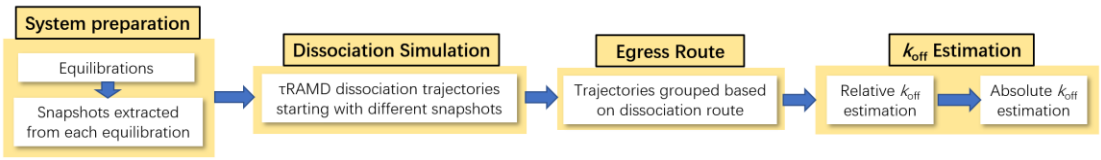


Figure 2.1: Residence time prediction workflow. Equilibrations were run to relax the ligand-β2-AR system and generate various starting structures for subsequent dissociation simulations. The τRAMD method was employed to simulate the dissociation behaviour, and the resulting trajectories were grouped based on the egress route. Relative koff values were initially estimated, which allowed the calculation of absolute koff values, free from the custom force.

2.1.1 System Preparation

Thanks to earlier researchers in our group, the structures of atenolol, bisoprolol, and cynapindolol-β2-AR complexes had already been constructed at 2017[131]. The β2-AR structure was homology modelled by using 2RH1 as template which is a human sourced β2-AR crystal structure with three stabilizing mutations[119]. With the improvement of the technologies, increasing β2-AR structures have been resolved by using X-ray diffraction and electron microscopy[174-176]. However, all of them are not strictly wild type. To stabilize the structures, some contain mutations, and the others contain fusion partners. Therefore, we consistently used the homology modelled β2-AR structure in this study. The agreement between our estimated k_{off} values and experimental data indirectly supports that this homology model is accuracy enough for this study (Table 3.2).

The simulation system comprised a single copy of β2-AR embedded in a mixed phosphatidic acid (PA) / phosphatidylcholine (PC) lipid bilayer containing 1,380 PA

and 570 PC molecules. The bilayer was sandwiched between two water layers modelled with TIP3P water molecules. Sodium and chloride ions were added to neutralize the system and achieve an ionic strength of approximately 0.15 M. The periodic simulation box dimensions were 7.60 × 7.60 × 10.16 nm^3 (Fig 2.2).

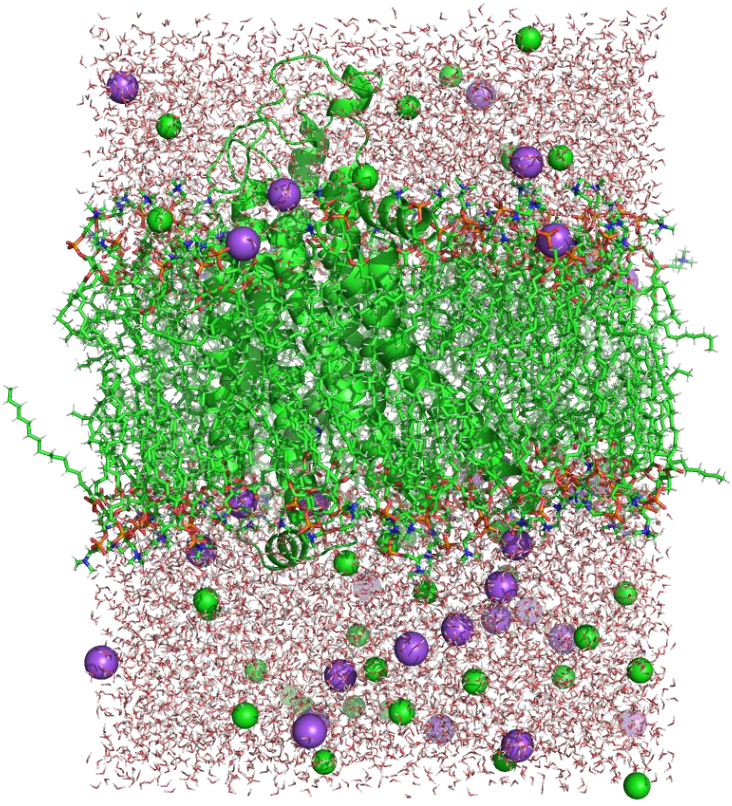


Figure 2.2: Visualization of the simulated system. The β2-AR is shown in cartoon style, while the mixed PA/PC lipid bilayer is shown in stick representation. Water molecules are displayed in line representation, and sodium (Na^+) and chloride (Cl^-) ions are shown as spheres.

During the ASP simulations, the keyhole conformations were generated. Docking the three ligands into the keyhole conformation successfully reproduces the canonical binding pose for β-AR ligands. These studies were performed previously by other members of our research group. I began our study from the preformed ligand-β2-AR complexes.

These complex structures generated by docking were in AMBER format[97]. To simulate the ligand dissociation, the GROMACS plugin-τRAMD was employed[72-74, 177].

Therefore, coordinate, topology, and trajectory files were converted from AMBER to GROMACS format by using MDtraj and Parmed[178, 179]. Then, simulations can be run in GOMACS. For all simulations, periodic boundary conditions were applied. A nonbonded cutoff of 10 Å was used for short-range Coulombic interactions, and long-range electrostatics were treated using the Particle Mesh Ewald (PME) method. A 2 fs time step was employed with bonds to hydrogen atoms constrained using the SHAKE algorithm.

These converted structures were not suitable for production MD simulations directly. A series of preparatory MD simulations were conducted to ensure the system reached the simulated condition gradually. The preparation consisted of energy minimization, heating, and equilibration. Minimization was done in stages with gradually decreasing position restraints on solute atoms until they reached zero. Then the system was gradually heated to 300K in the constant number, volume and temperature ensemble (NVT) with restraints of 50 kcal mol^{-1} $Å^{-2}$ to non-hydrogen solute atoms. This was followed by a series of 1 ns equilibration steps in a constant-pressure, constant-temperature (NPT) ensemble, during which the restraints were gradually reduced from 50 to 0.5 kcal·mol^{-1}·$Å^{-2}$. Finally, a 10 ns unrestrained NPT equilibration was performed to fully relax the systems (Langevin thermostat at 300 K and Parrinello–Rahman barostat at 1 bar). Upon completion of the whole suite of preparation steps, the systems were deemed ready for any further production simulations[180].

The ligand dissociation behaviour may depend on the initial ligand-protein interactions. Consistently using a single starting structure may result in a bias in the residence time prediction. To guarantee the variety of initial structures for the same system, snapshots were extracted from longer simulations as the starting structures for dissociation simulations. Specifically, each complex system underwent a 200 ns NPT molecular dynamics simulation in GROMACS[181] (Langevin thermostat at 300 K and Parrinello–Rahman barostat at 1 bar), which was repeated three times to generate

independent conformational ensembles. From each trajectory, five representative snapshots were extracted at 40 ns intervals. The atomic coordinates and velocities of these snapshots were used as starting conformations for the τRAMD dissociation simulations, resulting in 15 initial structures per ligand.

2.1.2 τRAMD protocols

In τRAMD[74], a custom force—termed a random accelerated force—is applied to the COM of the ligand to promote its unbinding from the receptor. The direction of the applied force was randomized and updated unless the displacement of the ligand COM relative to the protein COM is higher than 0.025 Å over 100 fs.

Each τRAMD simulation was terminated when the COM distance between the ligand and the target exceeded a predefined maximum displacement threshold, ranging from 30 to 60 Å depending on the ligand. Atenolol and bisoprolol possess an extended chain compared to the cyanopindolol. For atenolol and bisoprolol, a smaller displacement threshold (e.g., 30 Å) often resulted in incomplete dissociation, where part of the ligand remained in the protein pocket when simulation ended. To avoid this, a larger maximum displacement of 50–60 Å was set for their simulations, whereas a 30 Å threshold was sufficient for cyanopindolol. Reaching this threshold distance would be considered that the ligand has dissociated. Trajectory coordinates were recorded every 10 ps, and the COM displacement was also calculated to track the dissociation progress.

To evaluate the effect of force magnitude on the dissociation rate constant and taken pathway in the β2-AR system, seven different force magnitudes—11.0, 11.5, 12.0, 12.5, 13.0, 13.5, and 14.0 kcal·mol^{-1}·Å^{-1}—were applied independently while keeping all other τRAMD parameters the same. These values correspond to 764.5, 799.25, 834.0, 868.75, 903.5, 938.25, and 973.0 pN, respectively, using the conversion factor of 1 kcal·mol^{-1}·Å^{-1} = 69.48 pN. This unit conversion simplified the forces calculation in the further analysis.

Except for the above τRAMD-specific settings, all remaining simulation parameters were consistent with those used in the standard MD simulations, including the 200 ns NPT equilibration runs described previously.

As previously described, 15 snapshots were selected as starting conformations for each system. Each of these starting structures was subjected to three independent τRAMD simulations, resulting in a total of 45 dissociation trajectories per force magnitude for each system.

2.1.3 Derivation of Relative k_{off}

In previous τRAMD works, the relative residence times were all approximated to the average of simulated unbinding times[72-75]. Our analysis framework monitored the bound ligand number change over time. This approach is conceptually similar to the analysis strategies employed in both label-based and label-free binding assays. Specifically, for each magnitude and ligand-receptor system, 45 simulations were run. This relatively large population of trajectories enables monitoring of the number of surviving complexes as a function of time. To achieve this, the following five assumptions have to be made:

1. The removal of the ligand from the binding pocket is a thermally driven diffusion process.
2. A single dominant energy barrier is present to unbind.
3. The single barrier to thermally-driven unbinding gives a single dissociation rate constant, k_{off}.
4. The presence of the random force weakens the energy barrier causing an exponential increase in k_{off}.
5. No rebinding events occur following unbinding, and the k_{off} is consistent with 1st order kinetics.

Assumption 1, 2 and 3 together imply that different ligand dissociation pathways correspond to different k_{off} values. Assumption 4 implies that different magnitudes of

random force result in different relative k_{off} values. Assumption 5 suggests that the number of survival complex decreases overtime under 1st order kinetics and does not increase due to rebinding. Based on these considerations, dissociation simulations were grouped according to the ligand dissociation pathway and the random force magnitude. In each group, the population of the ligand bound complexes, *N*, decreases as a simulation ends with its ligand unbinding. When all simulations have finished, the population reaches zero. Under the assumptions described previously, the unbinding reaction can be represented by the following equation:

$$[LR] \xrightarrow{k_{\mathrm{off}}[LR]} [L] + [R] \tag{2.1}$$

where $[LR]$, also denoted as N, is the concentration of the formed complex. The $[L]$ and $[R]$ are the free ligand and macromolecule concentrations, respectively. The rate of the unbinding reaction (v) is in the 1st order kinetics described by:

$$v = {}^{dN}/_{dt} = -k_{\mathrm{off}}N \tag{2.2}$$

Solving this first-order differential equation yields an exponential decay function for the number of the ligand bound complex over time:

$$N_{(t)} = N_0 \exp(-k_{\mathrm{off}}\, t) \tag{2.3}$$

where *t* represents the elapsed time, N_0 denotes the initial concentration of the formed complex at $t = 0$, and $N_{(t)}$ represents the concentration at any subsequent time *t*. Taking the natural logarithm to both sides:

$$\ln\left(N_{(t)}\right) = \ln(N_0) - k_{\mathrm{off}}t \tag{2.4}$$

Thus, plotting $\ln\left(N_{(t)}\right)$ versus time *t* results in a straight line, and the slope of this line corresponds to $-k_{\mathrm{off}}$. Consequently, the absolute value of the slope gives the k_{off} value under the specified random force, as shown in Fig 3.5, right.

To estimate the error in k_{off}, each simulation group was divided into three subgroups, and k_{off} was independently calculated for each subgroup using the same

procedure. The standard deviation among the three values was then taken as the uncertainty of the estimated k_{off}. The calculated k_{off} values and their corresponding deviations are presented in Table 3.1.

All values reported in the text and tables represent mean ± standard error of the mean (S.E.M.) and are shown graphically as mean ± range. Linear regression was performed using either Excel or SciPy, with or without weighting, as appropriate[182].

2.1.4 Derivation of Absolute k_{off}^0

The k_{off} in the absence of the custom force corresponds to the absolute k_{off} (k_{off}^0), which is expected to correlate with experimental measurements. The absolute k_{off} can be obtained by extrapolating relative k_{off} values as a function of the applied force magnitude. This approach parallels the analysis commonly employed in AFM experiments[55], where force-dependent dissociation rates are measured and used to infer the force-free k_{off}. According to the fourth assumption, the relationship between the dissociation rate under an applied random force, $k_{\mathrm{off}}(f)$, and the force-free dissociation rate, k_{off}^0, is described by the following equation:

$$k_{\mathrm{off}}(f) = k_{\mathrm{off}}^0 \exp\left(f \Big/ f_\beta\right) \tag{2.5}$$

where f represents the magnitude of the random force, and f_β represents a "force scale" – indicating the barrier effect of force on the dissociation. A larger f_β value indicates that the unbinding rate is less affected by the force. The f_β is an intrinsic property of the system, and does not change with the magnitude of the random force.

Using the k_{off} values obtained under different magnitudes of random force, the data were plotted as a semi-logarithmic plot (logarithmic scale for k_{off}, linear scale for f), and a weighted linear fit was applied (Fig. 3.5A). In this representation, the intercept yields the force-free dissociation rate k_{off}^0, while the slope of the fitted line reflects the force sensitivity of dissociation and allows the determination of the force scale f_β.

2.1.5 Transition State Confirmation

Committor analysis was employed to identify the TSE along ligand dissociation pathways by using an in-house Python script. Key frames were first extracted to define the dissociation pathway. A one-dimensional progress coordinate *s* along the dissociation pathway was then defined as a cumulative RMSD values between successive intermediate structures. Specifically, the first structure was assigned $s = 0$; the RMSD between the first and second frames was assigned to obtain *s* for the second structure. For all subsequent structures, the RMSD was calculated relative to the preceding structure rather than the first, and the incremental RMSD changes were accumulated along the trajectory. This approach ensured that the RMSD values monotonically increased along the dissociation pathway. RMSDs were evaluated over four atoms: the nitrogen and oxygen atoms within the ligand's hydroxy-substituted ethylamine chain, and the C_γ atoms of the conserved $Asp^{3.32}$ and $Asn^{7.39}$. Finally, *s* was normalized to a range between 0 and 1, where 0 corresponds to the initial bound state, 1 to the fully dissociated state, and values between 0 and 1 represent intermediate conformations.

Along the dissociation trajectory, multiple short unbiased MD simulations from different candidate configurations were performed. With the defined dissociation pathway and its progress coordinate, the positions of the selected configurations before and after the short unbiased MD simulations were analyzed to determine whether the ligand moved forward (toward dissociation) or backward (toward rebinding). The configurations which have a nearly 1/2 probability of going back and forward were identified.

2.2 Prediction and analysis of Prdx2 Conformational Transitions

There is a strong correlation between protein's dynamic and function. The PTMs regulate Prdx2's conformational transition between FF and LU state, therefore, regulate Prdx2's function. We learn how likely each state is by using FEP, and the transition pathway by using FTS.

2.2.1 System buildup

Our work focuses on the conformational transition between FF and LU state. The very first step is to find a good structure of each state as template. The crystal structures of human source Prdx2 were listed in Table 2.2. The 5IJT is the only LU state, therefore used as the template for the LU state. The structures of 1QMV, 7KJ0, and 7KIZ are all in FF state. Despite that we aimed to learn the impact of both SO_2^- and SH on the FF-LU conformational equilibrium, we utilized SO_2^- as the FF state template due to that SO_2^- is more bulky than SH, resulting in an easier covalent bond change from SO_2^- to SH than the reverse in system buildup step. 1QMV has a higher resolution than 7KJ0, therefore used as the FF template. What's more, all the crystal structures are formed in a decamer ring. Our research focused on the dimer. For each structure, five dimers can be chosen, and we consistently chose the dimer with the most complete amino acid sequence.

PDB ID	Oxidation state	Conformational State	Resolution	Released Year	Reference
1QMV	SO_2^-	FF	1.7 Å	2000	[145]
5IJT	SS	LU	2.15 Å	2017	[147]
7KJ0	SO_2^-	FF	2.29 Å	2021	[183]
7KIZ	SH	FF	1.7 Å	2021	[183]
7KJ1	C172S mutant	FF	2.15 Å	2021	[183]

Table 2.2: Crystal structures of human source Prdx2.

In the LU state, the C-terminus of the protein is flexible, and therefore not completely resolved in PDB structure 5IJT. The program CHARM-GUI[184] was utilized

to complete the missing residues to ensure that the FF state share the same amino acid sequence with the LU state (the application of FTS requires the two end structures: FF and LU states share the same amino acid sequence).

Though the available crystal structures have both half sites in the same conformational state (FF or LU), there is no experimental data to suggest that their dynamics is actually coupled. Therefore, we manually created a model for what we term the 'FU structure' representing an intermediate structure during the conformational transition from FF to LU, in which one side of the protein adopts the FF conformation and the other adopts a sub-LU conformation (Fig 2.3). Note that a "side" does not correspond to a single monomer; rather, it represents a structural unit formed by contributions from both monomers.

To generate the FU structure, molecular alignment between FF and LU was first performed using Chimera[185]. Then, one C_P/C_R site was designated as the sub-FF side, and the other as sub-LU side. For each atom, distances to the C_P/C_R sites of both FF and LU were computed. Atoms near the FF C_P/C_R site stay in FF coordinates, while those close to the LU C_P/C_R site adopt LU coordinates. Atoms in the middle were assigned averaged coordinates between FF and LU. This procedure resulted in the FU structure (Fig 2.3).

With the FF, LU and FU structure templates, the desired oxidation states of cysteines, including -SH, -SNO, and $-SO_2^-$, were manually assigned using the 'buildup' tool in Schrodinger[186]. For example, the buildup operation from SH to SNO requires the atom type change from 'H' to 'N', and an atom addition of 'O' connecting to the newly introduced 'N'. Then, the newly built mutated cysteine underwent local energy minimization to a reasonable position. By introducing SH, SNO, and SO_2^- to C_P and C_R in FF, LU and FU state, twelve systems were generated, each representing a unique combination of PTM and conformational state (Table 2.3).

	SH-C_P-C_R	SNO-C_P	SNO-C_R	SNO-C_P-C_R	SO_2^--C_P
FF	FF_SH_C_{PR}	FF_SNO_C_P	FF_SNO_C_R	FF_SNO_C_{PR}	FF_SO_2^-_C_P
LU	LU_SH_C_{PR}	LU_SNO_C_P	LU_SNO_C_R	LU_SNO_C_{PR}	LU_SO_2^-_C_P
FU		FU_SNO_C_P	FU_SNO_C_R		

Table 2.3: The built systems of Prdx2, each representing a unique combination of PTM location and conformational state.

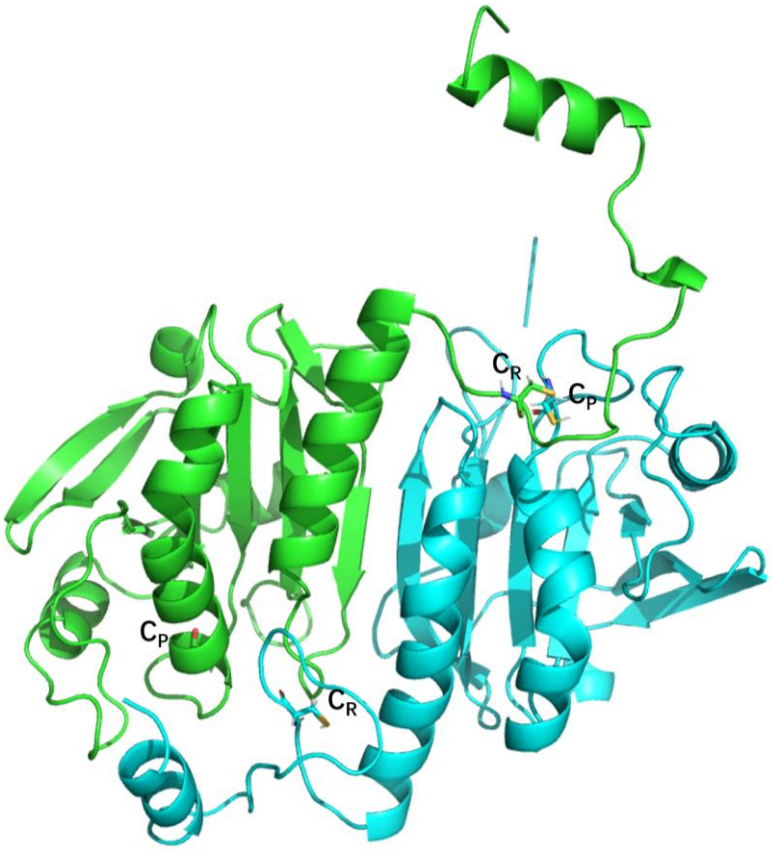


Figure 2.3: Conformation of the FU state, consiting of one sub-FF state and the other sub-LU state. Two subunits are colored in green and cyan with C_P and C_R highlighted in stick style.

2.2.2 Parameterization

In addition to the atom coordinates of the systems, to run MD simulations, a step of parameterization is required to describe all interactions between atoms. The

parameterization was performed using AMBER[187]. The *protein.ff14SB* and *water.opc* force fields were employed to parameterize protein residues and explicit water molecules, respectively.

The non-standard residues: SNO-Cys and SO_2^--Cys are not included in standard protein force fields and therefore require an additional unique parameterization. The 3D structures of non-standard amino acid SNO-Cys and SO_2^--Cys were produced by using the software Open Babel[188]. Subsequently, AMBER tools were used to parameterize the two custom residues. The *antechamber* module was utilized to derive charges and atom type of the residues with the *bcc* charge[189]. The charge computation for the *antechamber* requires the presence of the complete molecule. In the next, to ensure the connectivity of SNO-Cysteine and SO_2^--Cysteine to the next amino acids, the unneeded ("capping") atoms at two ends were removed using *prepgen*. Following this step, the information of atom charge and type was available but that of bonds, angles, and dihedrals in the modified amino acids remained unknown. To address this, two force fields, namely *parm10.dat* and *gaff.dat*, were introduced in *parmchk2*, generating parameter files. The accuracy of the parameterizations was validated by comparing the obtained 40 ns long simulation results with the crystal structure of S-nitrosocysteine (CID: 9793848) derived from the PubChem[190]. Specifically, the O2–N2–SG–CB dihedral angle and the CB–SG, SG–N2, and N2–O2 bond lengths were compared.

2.2.3 MD Details

With the coordinate and parameterization, MD simulations can be run. All standard MD simulations were performed using the AMBER software package. The systems were solvated in an octahedral periodic box, and long-range electrostatics were treated with periodic boundary conditions. Simulations were run with a time step of 2 fs, all bonds involving hydrogen atoms were constrained using the SHAKE algorithm, a nonbonded cutoff of 10 Å was applied, the Langevin collision frequency was set to 1 ps^{-1}.

The equilibration procedure followed standard protocols. First, to remove any steric clashes in the system, energy minimization was initially carried out using the steepest descent method for 500 steps, followed by an additional 1000-step minimization without restraints. The system was then gradually heated from 0 K to 300 K over 50 ps under constant volume conditions with a positional restraint of 50 kcal mol^{-1} $Å^{-2}$, using the Berendsen thermostat. This was followed by a 20 ps NVT equilibration at 300 K. Subsequently, four successive NPT equilibration phases were conducted: 10,000 steps each with positional restraints of 50, 10, and 2 kcal mol^{-1} $Å^{-2}$, respectively, followed by 100,000 steps without restraints. After the above processes, the structures were equilibrated and ready for further production simulations.

To evaluate the influence of PTMs on the dynamics of Prdx2, standard MD simulations were performed for all systems. All the production simulations were run in a consistent way, each lasting 20 ns, and conducted under NPT condition. Four replicate simulations were performed on each system, differing only in the random number seed used to assign initial velocity distributions.

2.2.4 Principal Component Analysis

For each of the five systems with different PTMs, four independent 40-ns molecular dynamics simulations were performed. Coordinates of Cα atoms from all trajectories served as input to PCA, which was conducted using the Python package *mdplus*[191]. To allow direct comparison among PTMs, trajectories from all PTM conditions were combined and subjected to a single PCA, yielding a common subspace defined by PC1 and PC2. For the FF state, all Cα atoms were included in the analysis, whereas for the LU state, the C-terminal segments were excluded, and only residues 1–168 and 197–364 were analyzed.

2.2.5 Free Energy Perturbation

FEP is a tool to evaluate the influence of a perturbation on the conformational equilibrium between two states by calculating the free energy change. In our case, the

perturbation is different PTMs, and the conformational equilibrium refers to that of FF and LU state. Therefore, FEP was utilized to estimate the relative probability change of FF and LU state under different PTMs.

2.2.5.1 Alchemical Pathway

In neurodegeneration diseases, a high level of S-nitrosylated Prdx2 was observed. To evaluate the effect of SNO to the FF and LU conformational equilibrium for Prdx2, the alteration in the free energy difference between FF and LU after the SNO process was estimated. Two stepwise alchemical pathways were defined as shown in Fig 1A, to represent the complete SNO process, which involves the oxidation of two C_P and two C_R residues from SH-Cys to SNO-Cys. In the pathways, alchemical state 1 represents the starting reduced state with four thiol cysteines, and state 5 is the end state with four SNO-Cys residues.

The first pathway (1→2→3→4→5) mutates C_P first followed by C_R residues, and the second (1→6→7→8→5) mutates C_R first followed by C_P residues. Free energy is a state function regardless of pathway. Thus, the computed free energy value obtained via 1→2→3→4→5 is expected to closely converge with that derived from 1→6→7→8→5, considering their same start and end states. This comparison between two pathways offers an opportunity assessing the error of the free energy calculation method.

Specially, an intermediate state—designated FU—is postulated, in which one side of the protein adopts a sub-FF conformation while the other is in a sub-LU conformation. Given that no available evidence supports that the two sides of the Prdx2 dimer unfold simultaneously, an attempt to evaluate the effect of SNO-C_P to the FF→FU transition is made. In the alchemical pathway of 9→10, the C_P located in the sub-LU side where unfolding event occur is mutated, and in 9→11, the C_P in the sub-FF side is mutated (Fig 2.1). These two transitions enable the effects of a side-specific evaluation of the impact of C_P SNO on the FF→FU conformational change.

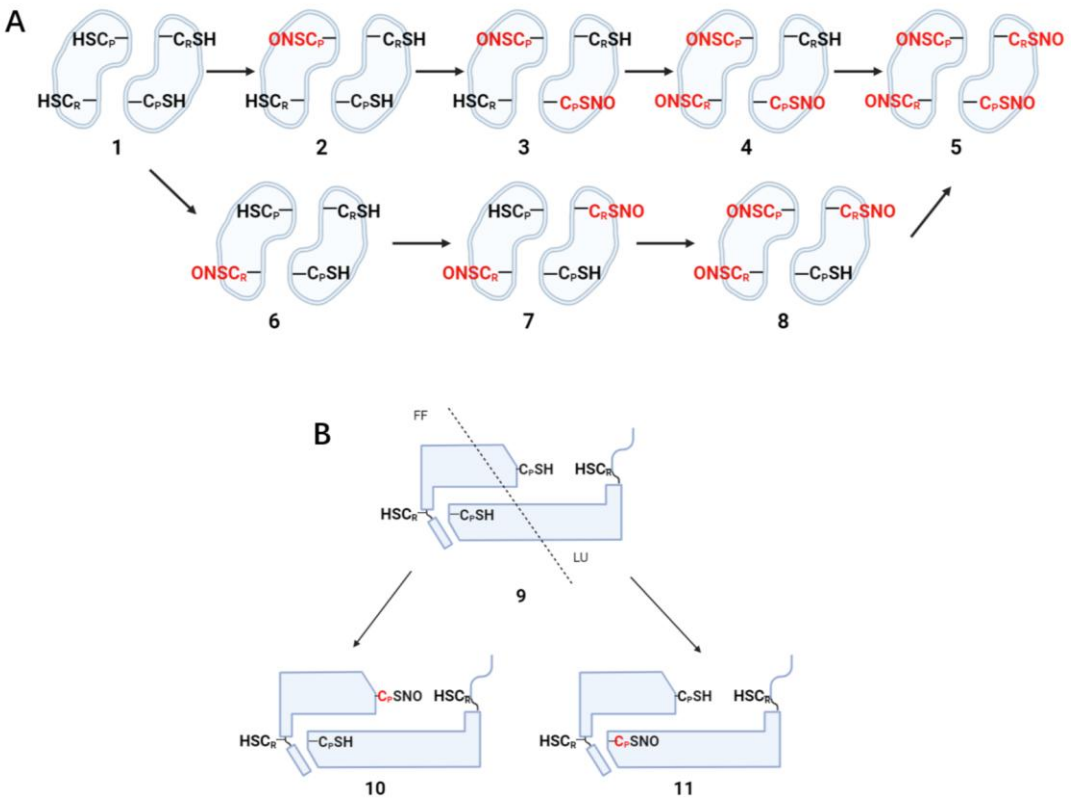


Figure 2.4: Stepwise alchemical pathway from SH-Cys to SNO-Cys. (A) The pathways correspond to the FF→LU conformational transition. In this process, all four cysteines including two C_P and two C_R residues are S-nitrosylated. The top pathway (1→2→3→4→5) mutates C_P first followed by C_R residues, and the bottom pathway (1→6→7→8→5) mutates C_R first followed by C_P residues. These two pathways together form a closed thermodynamic cycle, allowing error assessment for the calculation. (B) Alchemical pathways associated with the FF→FU conformational transition. The 9→10 pathway represents the C_P mutation in the sub-LU side, while 9→11 pathway represents that in the sub-FF side. For clarity, only the diagram of FU is shown.

2.2.5.2 Substate Generation

Limited by the sampling ability of standard MD algorithm which FEP follows, it is difficult to explore the entire configurational space of a state within a reasonable simulation time. To encourage sampling, multiple starting conformations of SNO-Cys were generated.

The construction of S-nitrosylated Prdx2 structures in Maestro follows a nondeterministic procedure, such that the position of the added nitrogen atom varies between runs. To further diversify conformations, the C-S-N-O dihedral angle was manually adjusted to generate additional side-chain poses, each of which was subjected to local energy minimization. This combination of manual intervention and nondeterministic protocols (atom addition and minimization) resulted in variable conformations of SNO-Cys and its interacting residues. As expected, repeating this process yielded multiple distinct SNO-Cys poses. In this study, one pose was generated for each alchemical state in the FF→LU transition (Fig 2.4A), and two poses were generated for each state in the FF→FU transition (Fig 2.4B).

These initial poses were not equilibrated and were therefore refined through short MD simulations, including energy minimization, heating, and equilibration, following the same protocol described in Section 2.1.1. The equilibrated structures obtained from these simulations were then used as representative SNO-Cys conformations in the FF, LU, or FU states for subsequent FEP calculations.

2.2.5.3 Equilibrium

In FEP simulations, free energy estimates ΔG are only reliable once the system has reached equilibrium[14]. Consequently, each simulation requires a finite period to achieve equilibration, and this duration must be determined. Since ΔG is value of interest, the ΔG is used as the metric to evaluate whether the system has equilibrated. Specifically, to identify the equilibration time, we plotted ΔG difference ($\Delta\Delta G^{equil}$) versus discarded time. At each time point, all prior data were discarded, and the remaining trajectory was divided into two halves. The free energy differences were calculated independently for each half, and the difference between them—denoted as $\Delta\Delta G^{equil}$—was evaluated.

We defined the onset of production sampling as the earliest time point where $\Delta\Delta G^{equil}$ first drops below 0.1 kcal/mol. Using this protocol, we obtained a distribution of equilibration times across all systems and fitted the distribution with a Gaussian

function. To ensure both consistency of sampling and near-complete equilibration across systems, the mean equilibration time from this distribution was adopted as a uniform equilibration period. Production sampling for all systems was then initiated at this standardized time point, thereby ensuring comparability of ΔG estimates.

2.2.5.4 Lambda Schedules

With the equilibrated structures, the assembly of the FEP system can proceed. However, building the FEP system is a complex and meticulous process. Using BioSimSpace[192], a Python package, simplified the system generation process in GROMACS format. A single topology was used during the transformation process, which gradually converts the atom types from the initial state to the final state by interpolating partial charges and bond parameters across two end redox states. The conversion process was divided into 13 windows, each with different lambda values ranging from 0 to 1, corresponding to the initial and final states. These lambda values serve as interpolation parameters, and at each lambda state, both van der Waals and electrostatic interactions are modified simultaneously. To mitigate the inherent singularity problem in Lennard-Jones (LJ) interactions at lambda = 0 or 1, the Beutler soft-core function was applied to modify the LJ interactions[14].

For each lambda window, the systems underwent a 10,000-step minimization, followed by heating from 0.01 K to 300 K over 100 ps. Subsequently, a 50 ps NPT simulation was performed to first roughly equilibrate the system, and a final production run of 10 ns was conducted. Notably, in the ‘production run’, only the data sampled after the ‘mean equilibrium time’ was used for ΔG calculation.

2.2.5.5 Estimation of Free Energy and Statistical Uncertainty

Two estimators, Thermodynamic Integration (TI) and MBAR[103, 193], were employed independently to analyze the same trajectory and estimate the corresponding ΔG value. Both estimators require the collected data to be uncorrelated. However, since MD simulations inherently generate time-dependent data, particularly when data are

sampled at short intervals, sampling was performed at relatively long intervals (0.4 ps) to reduce correlation and ensure approximate normal distribution.

TI estimates the ΔG by computing the numerical integration of the ensemble average d*u*/dλ over λ as expressed in Eq. (2.9):

$$\Delta f = \int_0^1 \left\langle \frac{du(\vec{\lambda})}{d\vec{\lambda}} \right\rangle_{\vec{\lambda}} d\vec{\lambda} \tag{2.9}$$

where λ represents different alchemical transformation states. The initial state corresponds to SH-Cys with λ = 0, and the final state corresponds to SNO-Cys with λ = 1.

MBAR, on the other hand, requires a complete set of potential energy differences across all λ windows and a sufficient overlap in phase space between neighboring λ windows. The ΔG between two end windows is computed by solving Eq. (2.10):

$$\Delta G = -\ln \left\langle \frac{\exp[-u_i(x_n)]}{\sum_{k=1}^{K} \frac{N_k}{N} \exp[f_k - u_k(x_n)]} \right\rangle \tag{2.10}$$

where N_k represents the number of samples from each of the K alchemical transformation states, with $\sum_{k=1}^{K} N_k = N$ the total number of samples[14].

These two estimators are expected to yield consistent results, thereby providing mutual validation of their reliability. MBAR is generally expected to provide greater accuracy due to its efficient use of all collected data. Consequently, MBAR was primarily used for data analysis and final ΔG calculations.

With either of the estimators, the ΔG between SH-Cys and SNO-Cys in different conformational states, ΔG_{FF} and ΔG_{LU}, respectively, can be calculated. The relative unfolding free energies (ΔΔG) were finally obtained by subtracting ΔG_{FF} from ΔG_{LU}.

$$\Delta\Delta G = \Delta G_{FF} - \Delta G_{LU} \tag{2.11}$$

The ΔΔG corresponds to the free energy change associated with transferring Prdx2 from the FF to the LU conformational state following SNO.

To assess statistical uncertainty, three independent replicas were simulated. The mean values were reported as ΔG, with standard deviations used as error estimates (Table 4.1). The error in ΔΔG derives from the error accumulation in ΔG (Table 4.2). The free energy computation protocol was validated by using thermodynamic cycle closure calculations, the free energy change along a close alchemical path 1→2→3→4→5→8→7→6→1 should be zero (Fig 2.4A).

2.2.6 Finite Temperature String Method

FTS was employed to explore the most probable pathway connecting the FF and FU states using an in-house Python script. This method requires that the two endpoint systems contain an identical set of atoms. To meet this requirement, CHARMM-GUI was used to add the missing residues in the FU state, ensuring that both conformations had the same number of protein atoms. Because it is challenging to preserve identical solvent atoms in both systems, an implicit solvent model was adopted.

As in committor analysis, the transition pathway was represented as a one-dimensional string defined by the cumulative RMSD between successive intermediate structures, evaluated over all protein atoms. Using RMSD as a descriptor, the pathway was initially interpolated into 20 images, including the two endpoint structures. Directly interpolated structures contained severe atomic clashes and were therefore unsuitable for MD simulations. To address this, the interpolated structures were used as templates, and targeted MD simulations (1 ns each) were performed using the two endpoint structures as starting conformations.

At the beginning of each FTS cycle, short restrained MD simulations (0.5 ns) were carried out for each image. In the first cycle, each image used itself as the restraint reference. In subsequent cycles, the interpolated structures from the previous iteration served as references. From these simulations, average structures were extracted to construct an updated string, which was then re-interpolated. This cycle of restrained simulations, averaging, and re-interpolation was repeated until convergence.

Convergence of the string was assessed using the overall string length and the positions of images along the pathway. Once these quantities stabilized with no significant fluctuations over recent cycles, the string was considered converged. PCA was used for visualization. Trajectories from all string images were combined and projected into a common subspace defined by PC1 and PC2, calculated over all Cα atoms.

3. Ligand Dissociation Mechanisms for β-AR Inhibitors from β2-AR

3.1 Dissociation Pathway Analysis

3.1.1 Selection of Dissociation Pathway

The dissociation pathways for atenolol, bisoprolol and cyanopindolol under a random force of 764.5 pN are shown in Fig 3.2. The magnitude of the applied random force can influence the prediction of dissociation pathway and rate. Among the forces tested, 764.5 pN represented the lowest magnitude, and its corresponding dataset was selected for pathway analysis. The specific impact of force magnitude on pathway and rate predictions is discussed later.

For atenolol (Fig 3.2) and bisoprolol (Fig 3.2), the extended side chains—particularly pronounced in bisoprolol—occupy the keyhole region, which can also serve as a dissociation pathway. Atenolol mainly dissociates via ECV (~90%) but also via keyhole (~10%). Bisoprolol dissociates through the ECV (~60%) and the keyhole (~27%), with an additional pathway (~13%) through a fissure between helices 5 and 6 (H5/H6). Both the keyhole and H5/H6 fissure link to the membrane system. Cyanopindolol (Fig 3.2) only dissociates through the ECV (100%).

The ligand k_{off} is primarily governed by the favourable interactions along the dissociation pathway[29]. Therefore, the k_{off} is closely linked to the specific egress route utilized during unbinding. Here, unbinding times are used to calculate the k_{off}. To minimize variability and improve the accuracy to predict the k_{off}, unbinding times are categorized based on the dissociation pathways.

In our analysis framework for relative k_{off}, a reliable estimation requires a sufficient sampling of unbinding times. Here, 45 dissociation trajectories were sampled for each ligand at a specific magnitude of the random force. The H5/H6 fissure was

never observed as an egress route for atenolol or cyanopindolol at any magnitude. For bisoprolol, at most 6 out of 45 trajectories followed the H5/H6 route across all tested magnitudes—a sample size inadequate for a reliable estimate of k_{off}, let alone smaller counts. Therefore, it is unpractical to calculate the k_{off} for the H5/H6 fissure in isolation. Since both the keyhole and H5/H6 fissure pathways are membrane-associated (Fig 3.2), their unbinding times were grouped together under the "keyhole" category. Then, all unbinding times are classified as either ECV or keyhole pathways (the latter inclusive of H5/H6), regardless of minor differences within each category.

The magnitude of the force has influence on the dissociation pathway selection. The trend observed as the force magnitude decreases may therefore reflect the dissociation preference in the absence of the random force.

For atenolol, reducing the force from 973 to 868.75 pN lowers the keyhole fraction from ~24% to ~11%, and further reduction to 764.5 pN maintains a plateau near 10%. It suggests that decreasing the force from 868.75 to 764.5 pN has little impact on the dissociation pathway selection (Fig 3.1). Consequently, it is reasonable to expect that further reducing the force to zero would not alter the pathway distribution, maintaining a similar plateau level—or potentially lowering it slightly—but is unlikely to increase the keyhole fraction. Under natural condition with zero force, atenolol thus predominantly dissociates via the ECV (> 90%) with a minor contribution from the keyhole (< 10%). When considering the overall ligand k_{off} regardless of pathway (as experimental measurements do not distinguish between pathways), the ECV overwhelmingly dominates. Therefore, due to the small contribution of the keyhole pathway, we will calculate the k_{off} solely for the ECV and use it to represent the overall k_{off}.

For bisoprolol, no obvious correlation is observed between the applied force magnitude and the selected dissociation pathway. Across all force conditions tested, the proportion of keyhole pathway fluctuates around 40% (Fig 3.1). Therefore, we infer that further reducing the force toward zero would not substantially alter the pathway

distribution, and the keyhole contribution would likely remain near 40%. These observations suggest that, under natural conditions, bisoprolol can dissociate through both the ECV (~60%) and keyhole (~40%) pathways.

For cyanopindolol, an obvious tendency is observed that as the reduction of the force magnitude from 973 to 764.5 pN, the portion of keyhole steadily declines from ~16% to 0% (Fig 3.1). This tendency strongly suggests that in the natural condition where the random force is absent, cyanopindol dissociates solely via the ECV.

The variation of pathway selection among the three ligands under the natural condition may derive from differences in their side chains. Bisoprolol possesses the longest extended side chain, shorter in atenolol, and absent in cyanopindolol (Fig 1). These side chains can interact with the keyhole region. The longer the side chain is, the stronger these interactions become. Consequently, bisoprolol exhibits the strongest association with the keyhole and the highest probability of dissociation via the keyhole, followed by atenolol. In contrast, the absence of a side chain in cyanopindolol results in no observed dissociation events through the keyhole route.

However, the observed pathway selection tendency is limited to the force magnitude ranging from 973 to 764.5 pN. Further simulations employing weaker force magnitudes are required to more precisely predict the ligands' unbinding pathways under the natural condition. In the aim of calculating the ligands' k_{off} values, the dissociation pathways and their corresponding unbinding times were eventually grouped into four categories after comprehensive consideration: atenolol ECV (ATE_{ECV}), bisoprolol ECV (BIS_{ECV}), bisoprolol keyhole ($BIS_{keyhole}$) and cyanopindolol ECV (CYA_{ECV}) pathways.

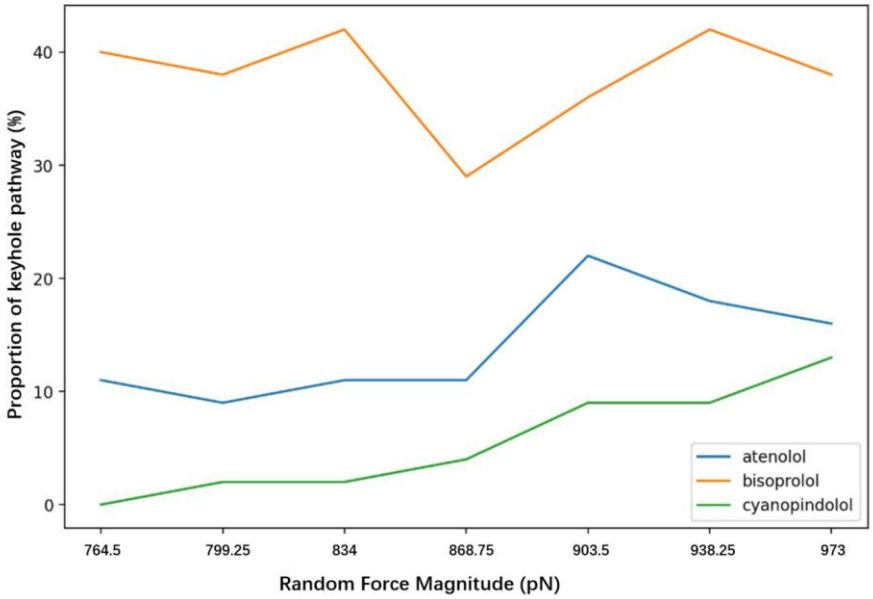


Figure 3.1: Proportion of keyhole dissociation events for atenolol, bisoprolol, and cyanopindolol under varying magnitudes of the random force. The percentage of dissociation trajectories utilizing the keyhole pathway is plotted as a function of force magnitude. Atenolol and cyanopindolol exhibit force-dependent trends, with higher forces promoting greater usage of the keyhole route. In contrast, bisoprolol shows a relatively stable proportion across all force conditions.

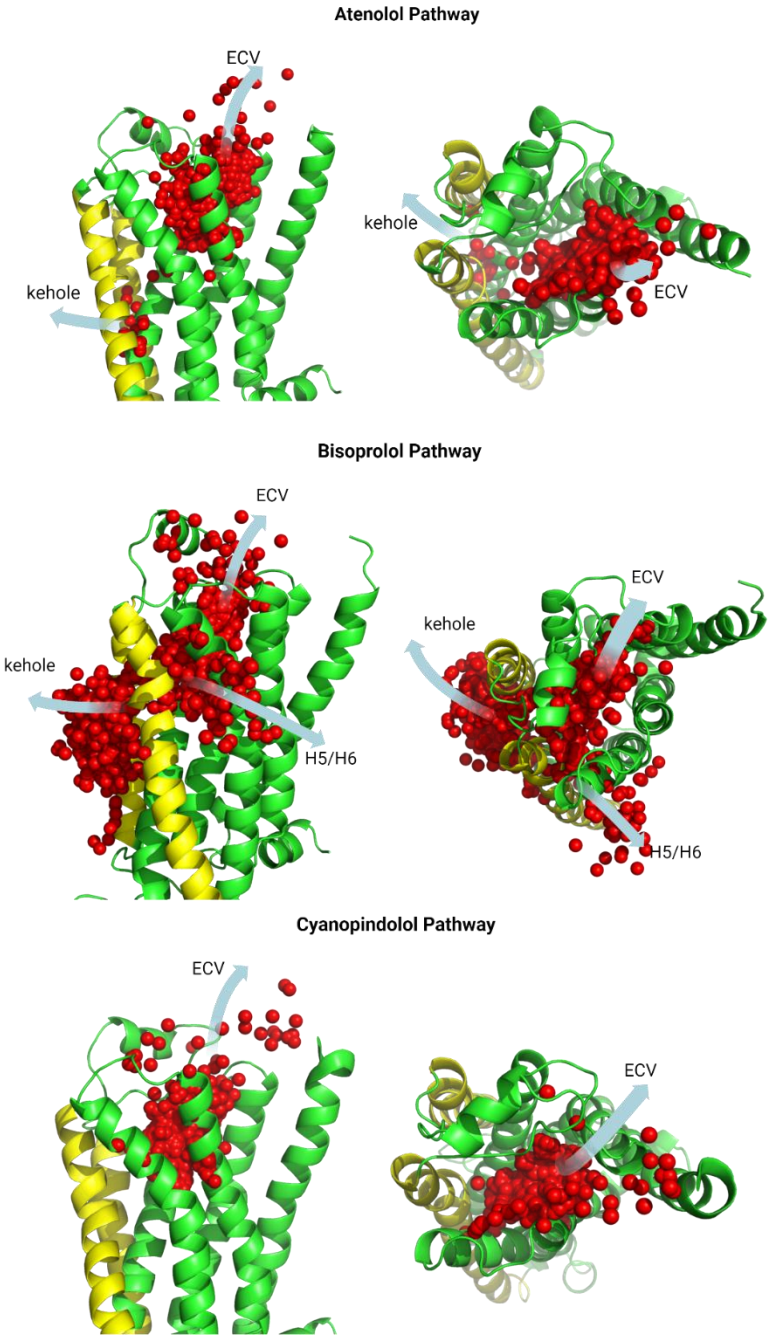


Figure 3.2: Dissociation pathways for atenolol, bisoprolol and cyanopindolol at a random force magnitude of 764.5 pN. Red spheres represent the COM of the ligand sampled every 1 ps during the final 500 ps of each trajectory. For each ligand and pathway, 15 out of 45 trajectories were randomly selected for display. The spatial distribution of the spheres reflects the overlap of these trajectories, thereby indicating the dissociation route. The protein is shown in cartoon representation, with helices H4 and H5 highlighted in yellow and the remaining residues in green.

3.1.2 Mechanism of Dissociation Pathway Selection

Atenolol and bisoprolol share a common structural feature in that both possess an extended side chain that occupies the keyhole region, and the extended side chain forms a hydrogen bond with residues within this site. The keyhole is defined as a cluster of residues including $Thr^{4.56}$, $Ser^{4.57}$, $Pro^{4.60}$, $Ser^{5.42}$. In atenolol, the acetamide nitrogen forms a hydrogen bond with $Thr^{4.56}$ (Fig 3.3A), whereas in bisoprolol, the ether oxygen proximal to the aromatic ring forms a hydrogen bond with $Ser^{5.42}$ (Fig 3.3B). Occupation of the keyhole by the extended side chain, together with the formation of these hydrogen bonds, likely underlies the observation that dissociation events via the keyhole pathway occur for both ligands. By contrast, cyanopindolol lacks an extended side chain and correspondingly does not exhibit dissociation events via the keyhole pathway at forces low to 764.5 pN (Fig 3.1).

Bisoprolol preferentially adopts the keyhole as its dissociation pathway compared with atenolol, a difference that arises from the distinct characteristics of their extended side chains. In bisoprolol, the longer side chain terminates in an isopropyl group, allowing the ligand to penetrate more deeply into the keyhole region such that the isopropyl head becomes geometrically proximal to the membrane interface. This proximity enables favorable non-polar contacts between the isopropyl group and lipid molecules (Fig 3.3B), which stabilizes membrane-adjacent dissociation intermediates and biases unbinding toward the keyhole pathway. By contrast, the side-chain terminus of atenolol consists of a polar acetamide group (Fig 3.3A). Due to its polarity, this group is energetically unfavorable in the hydrophobic membrane environment, reducing the likelihood of stable membrane-associated intermediates and thereby diminishing the probability of dissociation via the keyhole pathway.

The above focuses on enthalpic contributions. From an entropic perspective, bisoprolol is also more favorable to dissociate via the keyhole pathway. Given that atenolol and bisoprolol share the same core scaffold, the core is expected to contribute similarly to the entropy of the still-bound state immediately prior to dissociation.

Consequently, the entropy difference between the two ligands mainly originates from their extended side chains. The longer side chain of bisoprolol allows greater geometrically accessible conformational freedom near the membrane interface while the orthosteric head remains bound. This expanded conformational space increases the number of accessible microstates for membrane-adjacent pre-dissociation configurations, thereby stabilizing these states through an entropic contribution and lowering their free energy. As a result, dissociation via the keyhole pathway is entropically favored for bisoprolol relative to atenolol.

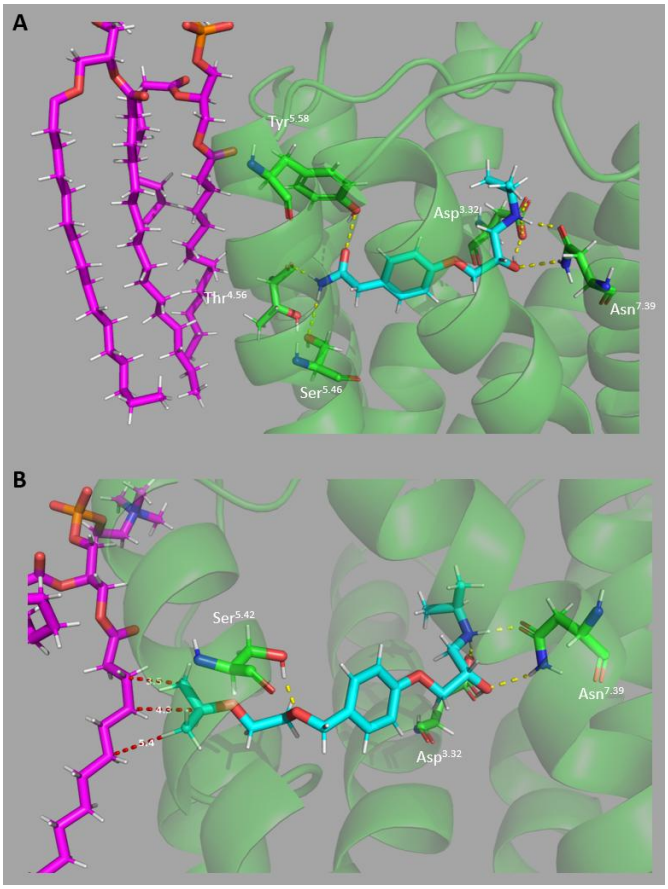


Figure 3.3: Atenolol (A) and bisoprolol (B) in complex with β2-AR. Ligands are shown in cyan sticks, key residues involved in polar interactions in green sticks, and lipid molecules in pink sticks. The protein is depicted in green cartoon representation. Hydrogen bonds are shown as yellow dashed lines, and non-polar interactions between bisoprolol and lipid molecules as red dashed lines.

3.2 Estimation of k_{off} in various Magnitudes of the Random Force

3.2.1 Distribution of Unbinding Times

After grouping the unbinding times based on ligand dissociation pathway and the magnitude of the force, the unbinding time distribution of each group was visualized. While the force magnitude is at 764.5 pN, the unbinding time distributions of ATE_{ECV}, BIS_{ECV} and $BIS_{keyhole}$, CYA_{ECV} are shown in Fig 3.4. An obvious observation is that the distributions for all the four ligand pathways all converge toward shorter times, especially for ATE_{ECV} and CYA_{ECV}; however, they do not follow a normal (Gaussian) distribution. A similar overall shape in the distributions was also observed at other magnitudes of the force.

This non-Gaussian behaviour may be mainly derived from that the sample size is progressive reduced along the unbinding time, therefore the shorter times are easier to be sampled. As illustrated in Fig 3.5 (left), using the ATE_{ECV} at 764.5 pN random force as an example, all 40 complexes are initially bound at time zero, and data from all 40 simulations contribute to the sampling. However, once one ligand dissociates, only 39 simulations remain to sample the next unbinding event. When the unbinding time comes to 100 ns, only 11 simulations are contributing. The points after 100 ns tend to spread out more. This progressive reduction in the sampling population continues until all ligands dissociate.

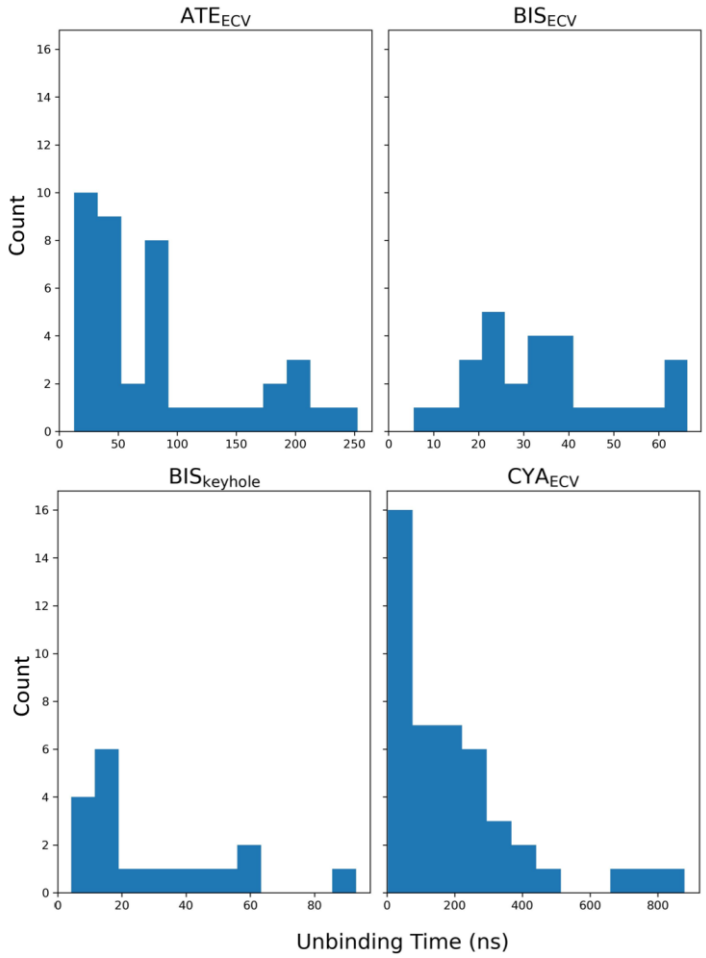


Figure 3.4: Unbinding time distributions at a random force magnitude of 764.5 pN for different ligand dissociation pathways: ATE_{ECV}, BIS_{ECV} and $BIS_{keyhole}$, CYA_{ECV}.

3.2.2 Estimation of k_{off} in various magnitudes of the random force

In previous τRAMD works, the k_{off} values were all approximated to the inverse of the average of simulated unbinding times. However, the unbinding time and residence time are two distinct concepts. They cannot be regarded as equivalent mathematically. Here, our analysis approach monitors the surviving complex number change over time enlightened by some experimental methods which monitor change in the surviving bound complex concentration. Both number and concentration of complex evaluate the quantity of surviving complex.

In the experiments, ligand dissociation events are independent of one another. The same principle applies to the simulation, as each trajectory is run separately. Therefore, all the simulations can be considered as a single integrated dataset. Initially, all ligands are bound; as time progresses, dissociation events occur, leading to a progressive decrease in the number of surviving bound complexes (Fig 3.5, left).

Since the assumption that no rebinding events occur following unbinding, and the k_{off} is consistent with 1st order kinetics is made, the number of surviving complexes is assumed to decrease overtime under 1st order kinetics and never increase due to no rebinding. Therefore, a plot of $\ln(N)$ versus unbinding time is constructed (Fig 3.5, right). According to Eq. 2.2-2.4, in these plots, the absolute value of the slope from the fitted line corresponds to the value of relative k_{off} at 764.5 pN random force. For example, the absolute value of the slope in ATE_{ECV} is 0.012, therefore the relative k_{off} of ATE_{ECV} at 764.5 pN random force is 0.012 ns^{-1}. Specifically, the last three points colored orange are outliers and excluded in the linear regression due to increasing variance at longer timescales. By using this protocol, relative k_{off} values can be estimated for each ligand dissociation pathway in various magnitudes of the random force, as listed in Table 3.1. The detailed procedure of relative k_{off} and corresponding error estimation is illustrated in the Section 2.1.3.

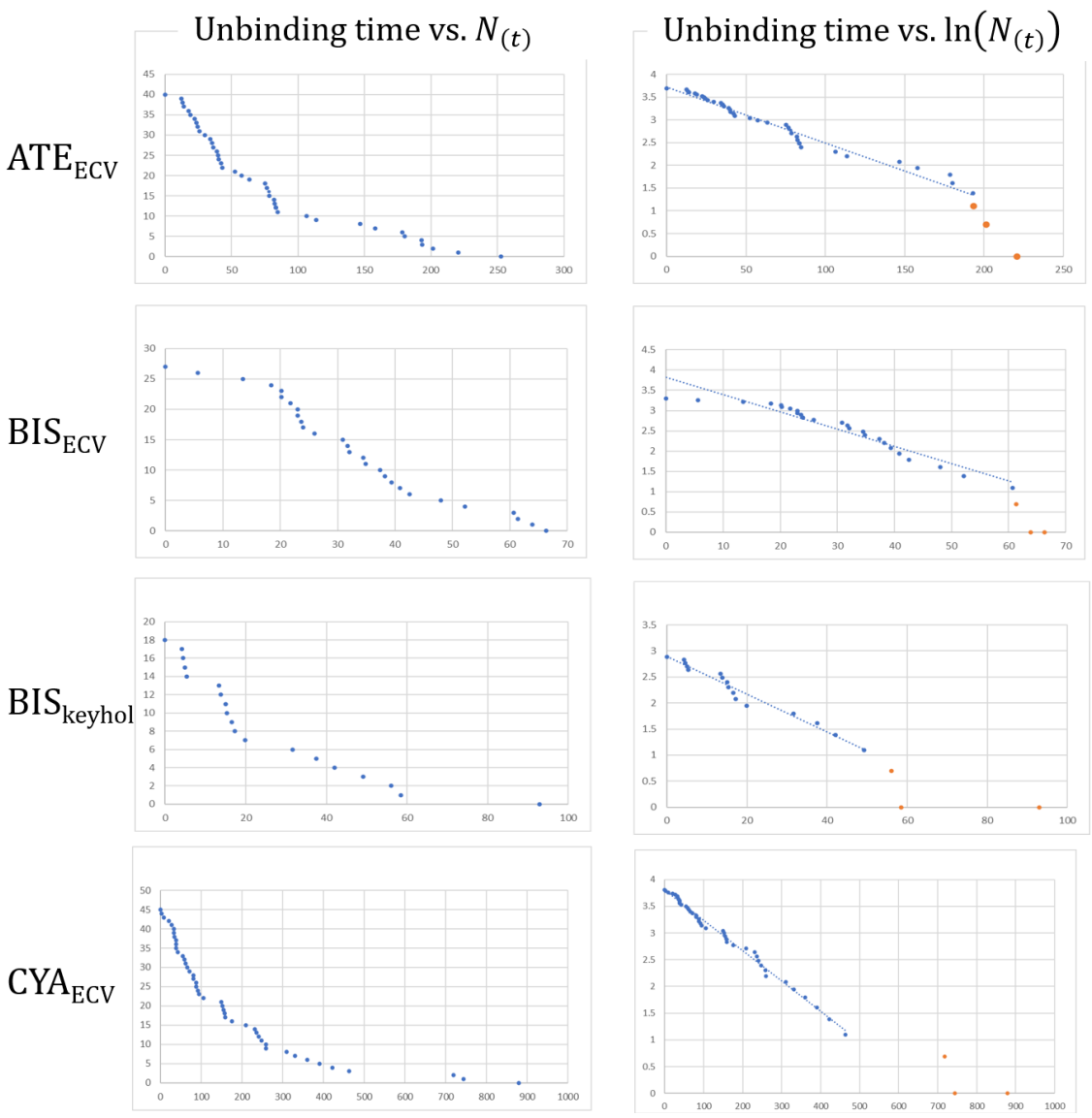


Figure 3.5: Derivation of relative k_{off} values at a certain magnitude random force for four ligand dissociation pathways: ATE_{ECV}, BIS_{ECV}, $BIS_{keyhole}$, and CYA_{ECV}. LEFT: Plot of the number of surviving ligand–β2-AR complexes (N) over time. At time 0, all complexes survive; unbinding events occur progressively until complete dissociation. RIGHT: A plot of $\ln(N)$ versus unbinding time and the dash line shows the linear regression excluding the three outliers (coloured orange). According to Eq. (2.4), the absolute value of the fitted slope corresponds represents the value of relative k_{off} at 764.5 pN. By using this protocol, k_{off} values can be estimated for each ligand dissociation pathway in various magnitudes of the random force.

Force (pN)	ATE_{ECV}		BIS_{ECV}		$BIS_{keyhole}$		CYA_{ECV}	
	k_{off} (ns^{-1})	s. d.	k_{off} (ns^{-1})	s. d.	k_{off} (ns^{-1})	s. d.	k_{off} (ns^{-1})	s. d.
764.5	0.012	0.0049	0.042	0.019	0.036	0.0096	0.0057	0.0012
799.25	0.016	0.0039	0.053	0.032	0.034	0.0087	0.011	0.0035
834	0.06	0.0053	0.073	0.018	0.11	0.029	0.022	0.0083
868.75	0.034	0.016	0.094	0.1	0.071	0.15	0.041	0.015
903.5	0.07	0.016	0.19	0.041	0.14	0.031	0.068	0.031
938.25	0.073	0.026	0.18	0.12	0.19	0.21	0.15	0.021
973	0.11	0.0052	0.26	0.064	0.27	0.2	0.17	0.062

Table 3.1: Estimated k_{off} values and corresponding error at various magnitudes of the random force for four ligand dissociation pathways: ATE_{ECV}, BIS_{ECV}, $BIS_{keyhole}$, and CYA_{ECV}.

In Fig 3.6A, the simulated k_{off} values are plotted as a function of the random force magnitude for all ligand dissociation pathways. The experimentally measured k_{off} values follow the rank order: BIS > ATE > CYA, with experimental pathways indistinguishable (Table 3.2). The computed relative k_{off} values are expected to reproduce this experimental ranking.

For atenolol and bisoprolol, the computed k_{off} values for BIS_{ECV} and $BIS_{keyhole}$ consistently exceed those of ATE_{ECV} across all tested force magnitudes, indicating that the τRAMD protocol is capable of resolving the kinetic differences between ATE and BIS across all tested force magnitudes. However, as the applied force increases beyond 868.75 pN, the relative ranking begins to deviate from experimental trends for cyanopindolol. In particular, the computed k_{off} for CYA_{ECV} gradually exceeds that of ATE_{ECV} and approaches the values of BIS_{ECV} and $BIS_{keyhole}$. This suggests that for such a complicated system as ligand-β2-AR–interactions, an upper threshold for the applied random force exists—beyond which τRAMD may result in an incorrect relative k_{off} rank. In this case, the threshold appears to be approximately 868.75 pN and may vary depending on the specific ligand–receptor system.

Interestingly, within the error, the two k_{off} values for BIS_{ECV} and $BIS_{keyhole}$ show no significant difference across all tested force magnitudes (Fig 3.6A). This observation suggests that the energy barriers associated with the ECV and keyhole pathways are of similar height for bisoprolol (Eq. (1.13)). If a substantial difference existed between the two, bisoprolol would preferentially dissociate through the pathway with the lower energy barrier, and transitions via the higher-barrier route would be rare or absent. Therefore, the comparable k_{off} values imply that both pathways are energetically accessible and contribute similarly to the overall unbinding process. Given the limited sampling and statistical population, it is expected that the two k_{off} values would appear equivalent or closely aligned.

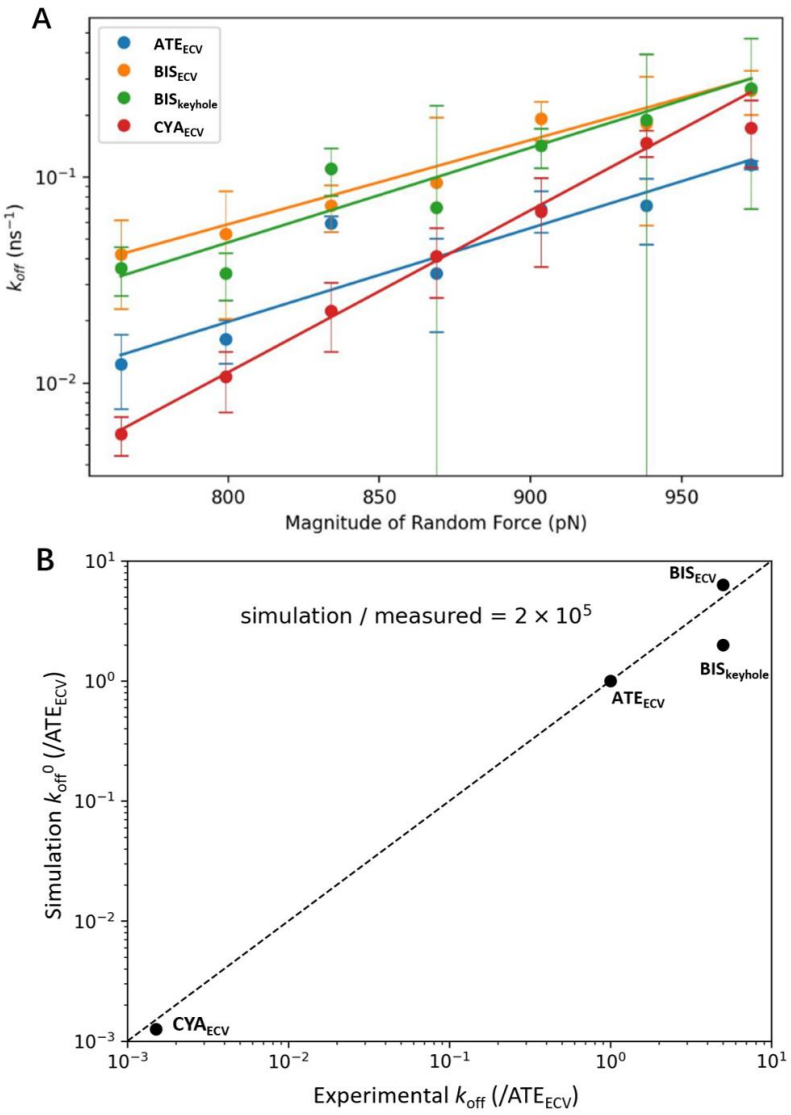


Figure 3.6: Simulated k_{off} and k_{off}^0 values. (A) Simulated k_{off} values across all random force magnitudes (data from Table 3.1). The colorful lines show linear fits of the k_{off} values to the force magnitudes. According to Eq. (2.8), the extrapolated k_{off} in the absence of a custom force (k_{off}^0) can be reported as $10^{intercept}$. The resulting k_{off}^0 values for each ligand dissociation route are listed in Table 3.2. (B) Comparison between simulation and experimental k_{off}^0 values after normalization by ATE_{ECV}. An obvious strong correlation between simulation and measured k_{off}^0 values can be observed among four ligand dissociation routes.

*ATE_{ECV} at 834 pN de-weighted.

3.3 Estimation of k_{off}^0

Instead of focusing solely on the rank of computed relative k_{off} value, an attempt to compute the estimated absolute k_{off} in absence of custom force (k_{off}^0) is made as detailed in the Section 2.1.1. This analysis framework is inspired from AFM experiments, in which a pulling force is applied to a ligand to facilitate dissociation from its target protein[55]. In dynamic force spectroscopy, measurements are performed over a range of force loading rates, enabling extrapolation of k_{off}^0 values[56]. In τRAMD simulations, the random force also pulls the ligand out of the binding site. This similarity may allow an application of the AFM analysis framework to τRAMD data. The AFM framework assumes that the dissociation occurs in the thermally activated regime, with the applied force lowering the free-energy barrier. We adopt the analogous assumption that the random force employed in τRAMD reduces the dissociation barrier in a comparable manner, thereby allowing k_{off}^0 to be extrapolated from force-dependent k_{off} values obtained at different force magnitudes.

Specifically, according to Eq. (2.5), the estimated k_{off}^0 for each ligand dissociation pathway is obtained as intercept from the linear fits in Fig 3.6A, while a relationship between the effective force factor (f_β) and slope can be built. The f_β represents a "force scale" – indicating the effect of force on the dissociation. The larger f_β, the less effect force has on the dissociation rate. Applying this procedure yields k_{off}^0 and f_β values, summarized in Table 3.2. The resulting rank of estimated k_{off}^0 values, BIS_{ECV} > $BIS_{keyhole}$ > ATE_{ECV} > CYA_{ECV}, aligns with that of experimental[132].

To facilitate comparison, all the simulated k_{off}^0 values were normalized to the value of simulated ATE_{ECV}, and the corresponding experimental values were normalized in the same manner. For CYA_{ECV}, the normalized simulated k_{off}^0 and measured k_{off} are 0.0013 and 0.0015, respectively (see in Table 3.2). The normalized simulated k_{off}^0 values of BIS_{ECV} and $BIS_{keyhole}$ are about 6.3 and 2, respectively, while the experimental k_{off} is around 5. After normalization, the simulated and experimental k_{off} values are of

the same order of magnitude across all the three systems. As illustrated in Fig 3.6B, therefore, despite only three data points (exclude ATE_{ECV} which is used for normalization), a correlation is observed between the simulated and experimental data for CYA_{ECV}, BIS_{ECV} and $BIS_{keyhole}$.

Notably, the simulated values systematically overestimate the experimental k_{off} values by approximately a factor of $10^{5.4}$ across the three data points examined (Table 3.2). To assess whether this calibration factor is transferable to other ligands of the β2-AR complex or to different protein systems, additional experimental measurements will be required.

Nevertheless, the significance of the five-order-of-magnitude gap simply implies that the dissociation dynamics are not fully governed by the assumptions used in our model. In AFM-based dynamic force spectroscopy, ligand dissociation occurs in a diffusive regime and is primarily thermally activated. The externally applied force acts as a perturbation that lowers the free-energy barrier along the dissociation pathway, without fundamentally altering the pathway selection. Within this framework, the barrier reduction is given by $Fx^{\ddagger}$, where F denotes the magnitude of the pulling force, and $x^{\ddagger}$ denotes the distance from the bound state to the transition state along the reaction coordinate. The value of $x^{\ddagger}$ can be extracted according to the Bell–Evans relationship, $x^{\ddagger} = k_B T / f_\beta$[56].

Applying this analysis to our simulated f_β values yields a maximum estimated $x^{\ddagger}$ of approximately 0.7 Å for CYA_{ECV}. Such a small value is inconsistent with a thermally driven diffusive escape process, for which transition-state distances are typically on the order of several ångströms. In conjunction with the significant difference between simulated and experimentally measured k_{off}^0 values, this observation indicates that τRAMD simulations do not operate within the diffusive regime assumed by AFM-based models.

Instead, the dissociation dynamics in τRAMD are more consistent with a friction-dominated regime, in which ligand unbinding is primarily driven by the external force dragging effect, while thermal fluctuations play a less important role. Under such conditions, the applied force does not merely lower the free-energy barrier but reshapes the dissociation pathway, biasing it along the direction of the applied force.

Despite these limitations, our analysis remains practically valuable. Across all three tested systems, it consistently yields a calibration factor of approximately five orders of magnitude, underscoring its robustness for predicting k_{off}^0 values.

Ligand dissociation route	**Simulation k_{off}^0 (s^{-1})**	**Experimental k_{off} (s^{-1})**	**Force scale $f\beta$ (pN)**	**Simulation k_{off}^0/ATE_{ECV}**	**Experimental k_{off}/ATE_{ECV}**	**Simulation/ Experimental**
ATE_{ECV}	$10^{3.7\pm0.62}$	0.02 ± 0.005	95 ± 15	1	1	$10^{5.4}$
BIS_{ECV}	$10^{4.5\pm0.41}$	0.1 ± 0.03	106 ± 12	6.3	5	$10^{5.5}$
$BIS_{keyhole}$	$10^{4.0\pm0.67}$		94 ± 16	2		$10^{5.0}$
CYA_{ECV}	$10^{0.8\pm0.27}$	3×10^{-5} ± 1.6×10^{-5}	55 ± 4	0.0013	0.0015	$10^{5.3}$

Table 3.2: Simulation and experimental k_{off}^0 values before and after normalization by ATE_{ECV}.

3.4 Transition State Confirmation

According to the Eq. (1.13), k_{off} is determined by the free energy difference between the bound state and transient state. Therefore, the identification of the TSE can be informative for k_{off} lead drug optimization. With τRAMD simulations, complete dissociation trajectories were available. However, the position of the TSE is not automatically evident in these trajectories. We therefore employed committor analysis to identify the TSE[194].

The application of committor analysis requires a predefined pathway and its associated progress coordinate. We defined each dissociation trajectory into a "string" and a one-dimensional progress coordinate *s* from RMSD. Specifically, the first structure was assigned $s = 0$; the RMSD between the first and second frames was assigned to obtain *s* for the second structure. For all subsequent structures, the RMSD was calculated relative to the preceding structure rather than the first, and the incremental RMSD changes were accumulated along the trajectory. This approach ensured that *s* monotonically increased along the dissociation pathway, yielding a pathway-dependent progress coordinate.

Because the canonical β-binder pose is stabilized by four polar interactions whose breakdown plays a key role in the ligand dissociation event, *s* was computed as the RMSD of the relative geometry of the atoms defining these contacts. This string representation compresses the inherently multidimensional reaction coordinate into a simple, effective 1D metric suitable for committor tests.

For cyanopindolol, multiple short unbiased MD simulations from configurations distributed along *s* were performed. The predefined pathway and its associated progress coordinate allow to classify outcomes as advancing toward dissociation or reverting toward rebinding. Two configurations (Fig 3.7A and B) display mixed outcomes. For the first, 3 three trajectories return and two proceed (Fig 3.7C). For the second

configuration, four return and one proceeds (Fig 3.7D). Both configurations possess local highest energy and represent the bottleneck for unbinding.

However, these configurations do not necessarily constitute the TSE which possess the global highest energy. To verify whether the two configurations are local or global energy peak, a construction of the entire free-energy landscape along the dissociation pathway is required. In practice, this involves many longer trajectories from densely spaced intermediates, and these trajectories are expected to have a sufficient phase-space overlap. Moreover, to identify structural determinants of k_{off} differences across the three ligands, a comprehensive comparison between their TSE is essential.

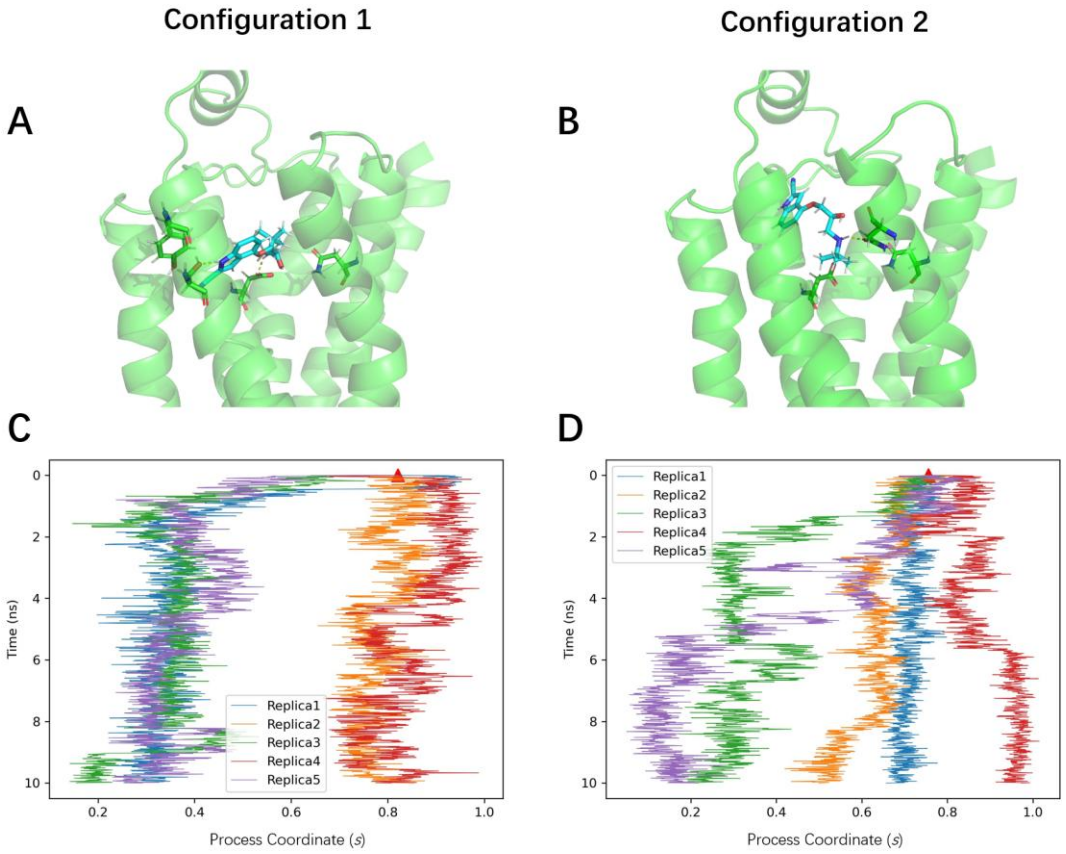


Figure 3.7: Local energy-peak configurations along the cyanopindolol dissociation pathway. (A, B) Structures of two representative configurations. The protein is shown in green cartoon style, and the ligand cyanopindolol in cyan sticks. Conserved $Asp^{3.32}$ and $Asn^{7.39}$, together with residues forming hydrogen bonds, are displayed in stick. Hydrogen bonds are indicated as yellow dashed lines. Configuration 1 is shown in A and configuration 2 in B. (C, D) Evolution of the progress coordinate *s* over 10 ns standard simulations initiated from configuration 1 (C) and configuration 2 (D). The red triangle denotes the starting configuration, and the five replicates are shown in different colors.

3.5 Conclusion

This chapter applies τRAMD simulations to characterize the dissociation pathways and kinetics of atenolol, bisoprolol, and cyanopindolol bound to β2-AR. A central contribution is the development of a new analysis framework that goes beyond conventional τRAMD practice. Instead of averaging unbinding times, the framework monitors the temporal evolution of the bound ligand population to derive relative k_{off} values. By analogy with AFM experiments, it also extrapolates absolute k_{off}^0 values in the absence of the applied random force, enabling direct comparison with experimental kinetics.

The approach provides both mechanistic and kinetic insights. Atenolol and cyanopindolol mainly dissociate through the ECV at low force magnitudes, suggesting that this is their most probable natural egress route, whereas bisoprolol dissociates via both the ECV and the keyhole route. Their difference in the dissociation pathway selection derives from differences in the extended side chain. The long extended side chain of bisoprolol, terminating in a hydrophobic moiety, promotes engagement of the keyhole pathway through both enthalpic stabilization and entropic contributions. Within an appropriate force regime, the framework reproduces the experimental rank order of k_{off} values. Although extrapolated absolute k_{off}^0 values are systematically overestimated by approximately five orders of magnitude relative to experiment, the consistent calibration factor observed across all three ligands underscores the practical utility of this analytical approach for estimating k_{off}^0. Committor analysis further highlights two cyanopindolol configurations with mixed outcomes, identifying them as candidate TSE structures that define the kinetic bottlenecks of dissociation.

Despite these advances, several limitations remain. First, only five committor trajectories are launched per configuration, which is insufficient for robust probability estimation; at least one hundred trajectories will be needed to obtain converged results. Second, the nature of the mixed-committor configurations—whether local or global

energy peaks—could not be resolved. Umbrella sampling along the dissociation pathway, coupled with MBAR analysis, offers a promising strategy to reconstruct the full free-energy landscape and answer this question. To further illuminate the pharmacological basis of β1-selectivity, future studies should compare atenolol and bisoprolol dissociation from both β1- and β2-AR. Finally, a systematic comparison of transient transition states between cyanopindolol and the β1-selective blockers is necessary to explain the striking five-order difference in their dissociation rates.

In summary, this chapter establishes τRAMD as a powerful approach for probing GPCR ligand dissociation and positions the new analysis framework as a rigorous and practical tool for estimating residence times. In combination with committor validation, it offers both quantitative predictions and mechanistic interpretation of transition states.

4. Effect of S-Nitrosylation on Prdx2 Conformational Equilibrium

4.1 Standard MD simulations

4.1.1 Parameterization Validation

SNO-Cys and SO_2^--Cys are non-standard residues and thus require additional force-field parameters. We validated our parameterization by comparing 40-ns MD trajectories with a small-molecule S-nitrosocysteine which is an experimentally determined X-ray crystal structure reported by Yi et al. (Fig 4.1A)[195]. Because the modifications are located in the cysteine side chain, we focused on metrices that reflect on local geometry and conformational preference: the O2–N2–SG–CB dihedral angle and the CB–SG, SG–N2, and N2–O2 bond lengths (illustrated for SNO-Cys).

Among these metrices, the side-chain dihedral is the most significant metric defining protein secondary structure, so accurate dihedral sampling is essential for validating new parameters[196]. In S-nitrosothiols, the C–S–N–O torsion defines the syn/anti conformers. Our O2–N2–SG–CB dihedral distribution shows that the crystallographic value lies within the high-probability region of the MD ensemble (Fig 4.1B).

Bond-length comparisons also support the parameterization. The crystallographic CB–SG and N2–O2 distances fall within the simulation distributions (Fig 4.1C, E). The SG–N2 bond is ~0.1 Å longer in the crystal than the MD average (Fig 4.1D). This minor difference may derive from their variance in the chemical environments. The S-nitrothiol in the dynamics is surrounded by other residues but which is not present in the crystal structure.

Overall, the agreement across dihedral angle and bond-length metrics indicates that the parameters reproduce the expected geometry and conformational preferences of SNO-Cys, and the small SG–N2 deviation is unlikely to have significant influence on the accuracy of the simulation.

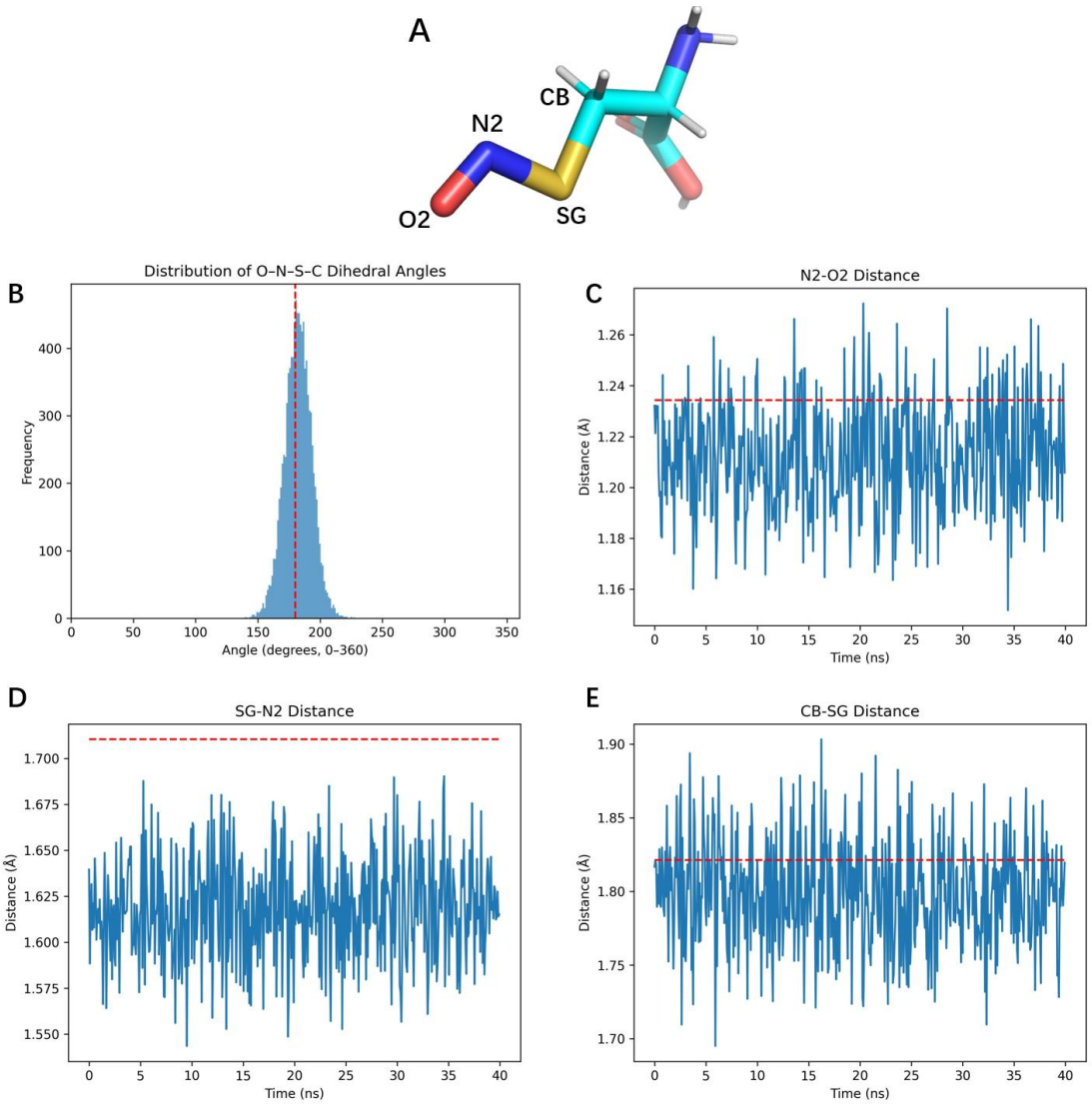


Figure 4.1: Comparison of SNO-Prdx2 dynamics with the crystalized S-nitrosocysteine. Blue traces represent values obtained from dynamics, while red dashed lines indicate corresponding values from the X-ray crystal structure. (A) Structure of S-nitrosocysteine with atom names indicated. (B) Distribution of O2-N2-SG-CB dihedral angle. (C) N2-O2 bond length. (D) SG-N2 bond length. (E) CB-SG bond length.

4.2.2 Principal Component Analysis

To evaluate the impact of different PTMs on the protein dynamics, 40 ns long standard MD simulations were run for five systems: SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P. Strictly speaking, the cysteine thiol is not a PTM; however, for brevity we refer here to each *state/site combination* (including thiol) as a "PTM."

Given the high dimensionality of protein dynamics, which involves thousands of atomic coordinates changing over time, PCA was employed to reduce complexity and capture the most significant motions. Both the FF and LU states were analyzed. For each conformational state, the trajectories of all five PTMs were projected onto the same subspace defined by PC1 and PC2, enabling identification of their conformational sampling differences. The conformational landscapes explored by each PTM are illustrated in Fig 4.2 (FF state) and Fig 4.4 (LU state), where the grey background represents the aggregated explored space from all PTMs, while red marks the sampling of the PTM highlighted in that panel. To further characterize the underlying motions, the structural displacements along PC1 and PC2 were visualized, as presented in Fig 4.3 (FF state) and Fig 4.5 (LU state).

For the FF state, the first two principal components accounted for 14.1% (PC1) and 7.2% (PC2) of the total variance. Among the five PTMs, SH_C_{PR} explored a distinct region of conformational space along PC1 compared with the other four PTMs (SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P) (Fig 4.2). Specifically, the SH_C_{PR} projections on PC1 spanned a wider range (approximately –2.0 to 0.25), whereas the other PTMs remained confined between –1.25 and 1.25. Although this separation suggests some conformational divergence, the variance captured by PC1 is limited, indicating that the effects of different PTMs on the FF-state dynamics cannot be clearly distinguished within 40 ns of standard MD simulation. Along PC2, the separation among PTMs was even less pronounced. Structural representation of PC1 and PC2 is shown in Fig 4.3. Higher-order components (PC3–PC10), which together accounted for

an additional 26.4% of the variance, were also analyzed but did not reveal significant differences in conformational sampling among the systems.

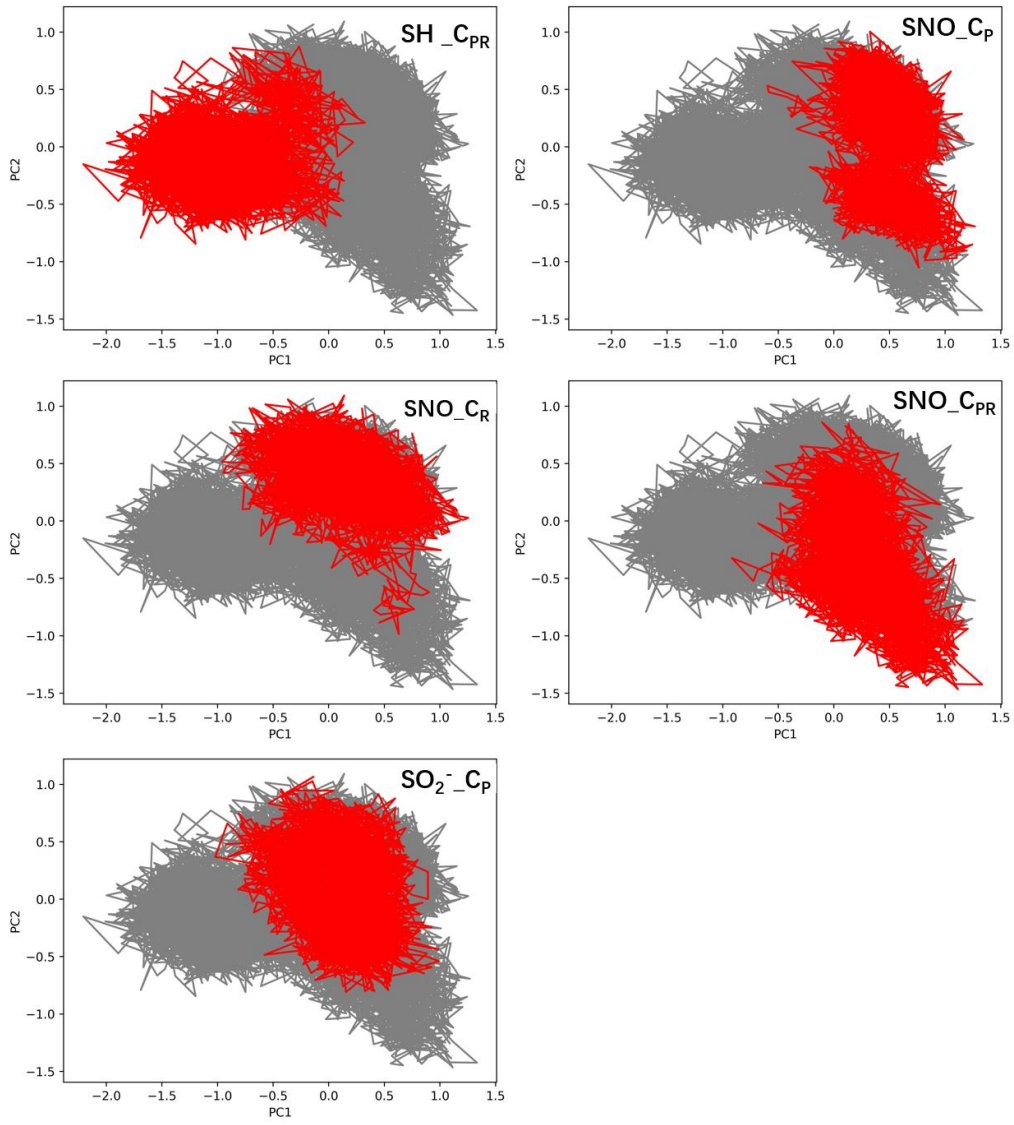


Figure 4.2: FF-state conformational sampling. Trajectories for SH_C$_{PR}$, SNO_C$_{P}$, SNO_C$_{R}$, SNO_C$_{PR}$, and SO_2^-_C$_{P}$ are projected onto a common PC1–PC2 subspace derived from all Cα coordinates. The grey background represents the aggregated explored space from all PTMs, while red marks the sampling of the PTM highlighted in that panel.

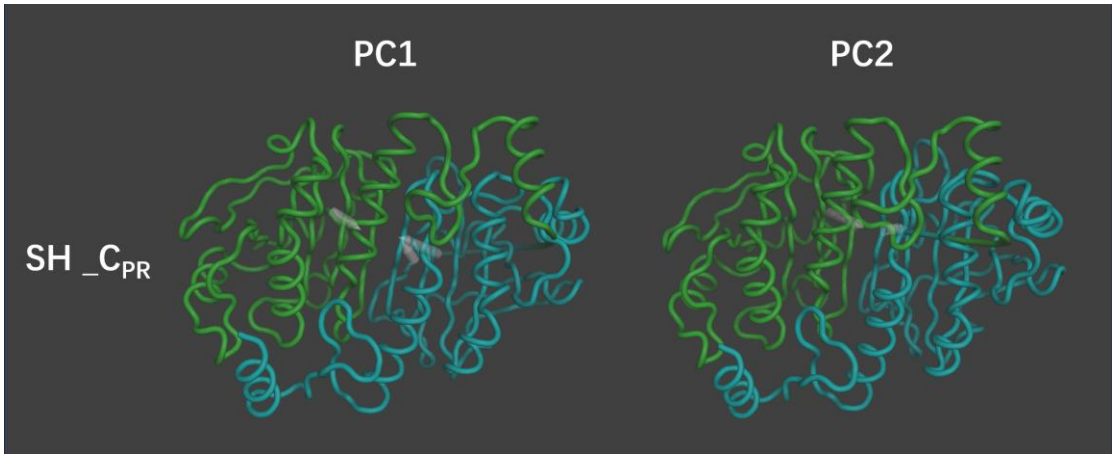


Figure 4.3: Visualization of FF-state motions along PC1 and PC2 exemplified by SH_C_{PR}. The two monomers are colored in green and cyan. The maximal displacements along the principal components are represented by white arrow.

Unlike the FF state, in which the C-terminus forms a well-ordered α-helix that packs against the protein core, the LU state contains a disordered, highly mobile C-terminal tail with minimal core contacts. Including this tail in the PCA can dominate the variance and obscure motions of the protein core; the latter are most informative for discriminating PTM-dependent dynamic differences. Accordingly, we excluded the C-terminus and performed PCA on residues 1–168 and 197–364.

For the LU state, PC1 and PC2 accounted for 29.9% and 14.2% of the total variance, respectively, together capturing 44.1% of the conformational change (Fig 4.4). The resulting conformational landscape displays a characteristic V-shape across all systems. SH_C_{PR} explores both branches of this V-shaped distribution, spanning a wide range along PC1 (approximately –2.5 to 4.5) and PC2 (approximately –1.0 to 3.0). In contrast, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P exhibit more restricted sampling.

PC1 and PC2 reveal that the dominant motions are localized to the dimer interface helices (Fig 4.5). Along PC1, SH_C_{PR} exhibits the largest inter-subunit displacement, with the two monomers undergoing opposite directional shifts at the interface. SNO_C_R and SO_2^-_C_P also display this interfacial rearrangement, albeit with a reduced

magnitude, while SNO_C_P and SNO_C_{PR} remain largely rigid. A similar pattern is also found in PC2, where the motions are involved subtler twisting of the interfacial helices but more localized. Again, SH_C_{PR} retained the highest degree of flexibility, followed by SNO_C_R and SO_2^-_C_P, whereas SNO_C_P and SNO_C_{PR} display minimal displacement. PC3-10, together contribute to the protein motion 23.8%, are also compared but exhibit little difference in conformational sampling among the systems. Collectively, these results suggest that SNO at C_P restricts LU-state flexibility within the 40-ns simulation window.

Overall, the FF state exhibits less conformational flexibility, with PC1 values spanning approximately –2 to 1.5, whereas the LU state explored a broader conformational range, with PC1 values extending from –3 to 5. Across 40-ns unbiased trajectories, PTM-dependent differences were not identified in the FF ensemble and, although detectable in the LU ensemble, they were modest and spatially localized. Because the key functional behaviour of Prdx2 arises from global FF↔LU interconversion and PTM-dependent shifts in their equilibrium, enhanced sampling is required. Accordingly, we employ enhanced-sampling approaches to drive FF↔LU transitions and quantify the corresponding free-energy differences.

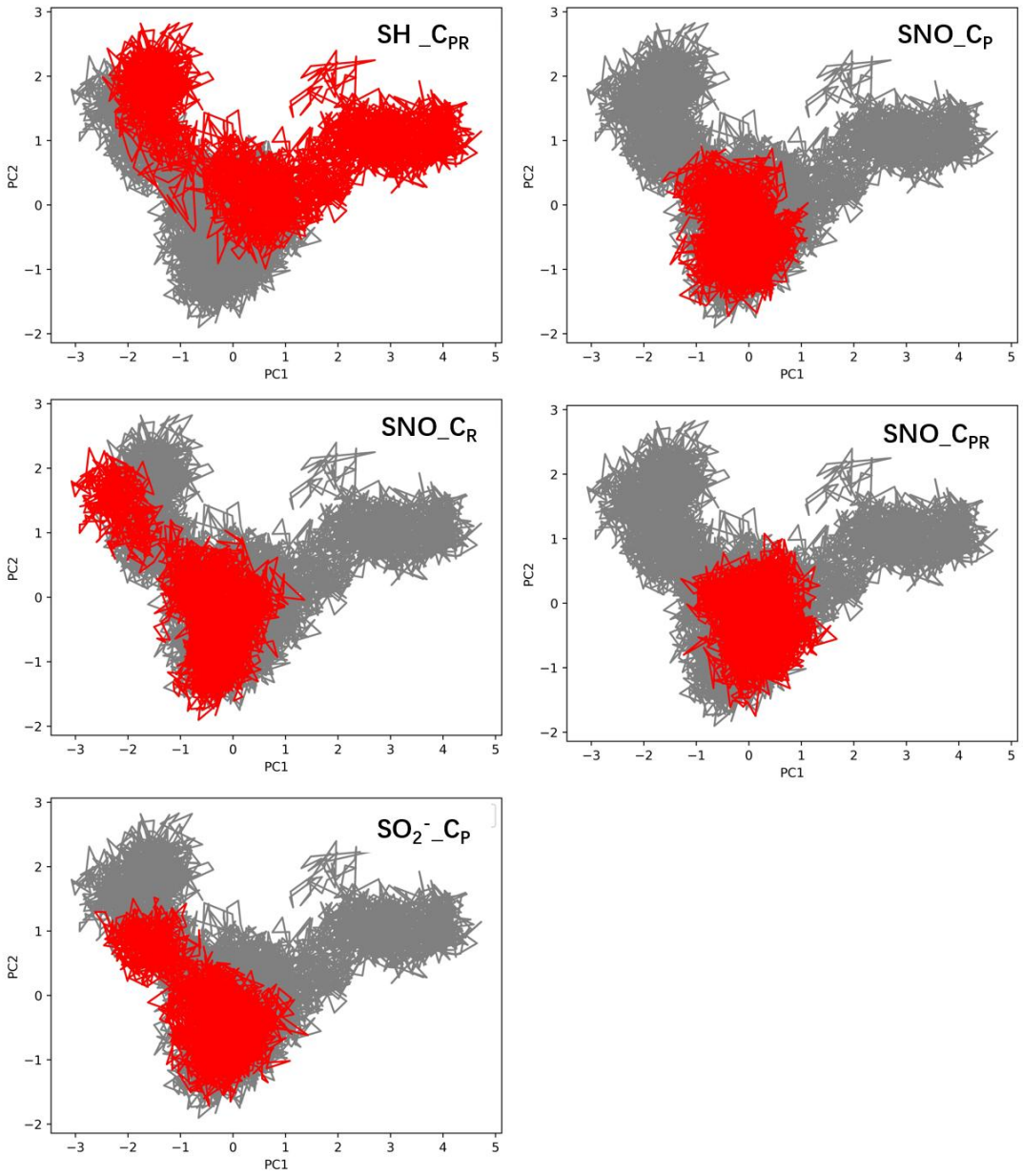


Figure 4.4: LU-state conformational sampling. Trajectories for SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P are projected onto a common PC1–PC2 subspace derived from the Cα coordinates of residues 1–168 and 197–364 (C-termini excluded).

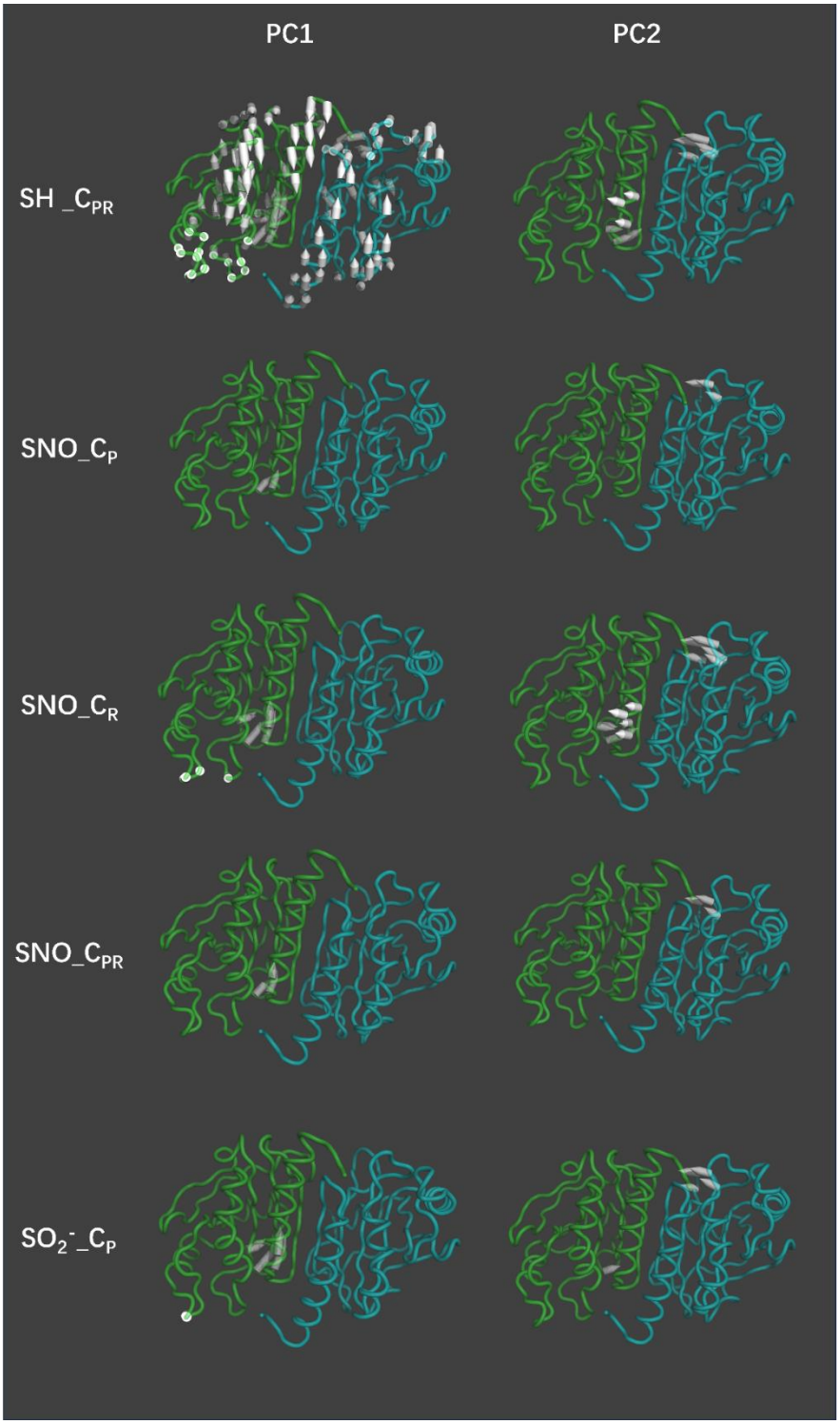


Figure 4.5: Visualization of LU-state motions along PC1 (left panel) and PC2 (right panel) from SH_C_{PR}, SNO_C_P, SNO_C_R, SNO_C_{PR}, and SO_2^-_C_P simulations. The two monomers are colored in green and cyan. The maximal displacements along the principal components are represented by white arrow.

4.2 Free Energy Perturbation

4.2.1 Independence of Alchemical Transitions

In standard MD-based FEP simulations, the limited sampling timescales typically restrict the observed conformational changes to local rearrangements, especially when the perturbation involves only a small number of atoms[197, 198]. In our simulations, SH-Cys was converted to SNO-Cys. Consequently, we assume that such a minor perturbation can only induce local conformational rearrangement within 10 ns simulation window. This assumption further implies that the perturbation exerts negligible influence on distal conformational regions and the overall protein structure.

Prdx2, a dimer, contains two C_P and two C_R residues, and each subunit has one C_P and C_R residue. In the FF state, all four cysteines are spatially distant from one another. In the LU state, two C_P residues remain separated in the space, so as do the two C_R residues. Given these spatial separations and the assumption that the perturbations primarily affect only local environments, modifications at individual cysteine residues can be regarded as independent events. Thus, SNO at one site is unlikely to influence the conformation or reactivity of other sites.

This principle is illustrated in Fig 2.4A, where the 1→2 and 2→3 transitions are compared. Both transitions involve modification of a C_P residue, but in different monomers. Their local chemical and conformational environments are essentially identical; the only distinction lies in the distal redox state of the other C_P. In the 1→2 transition, all three remaining cysteines are reduced, whereas in the 2→3 transition, the second C_P has already been S-nitrosylated. Because of the large spatial separation, this difference can be neglected, and thus the free energies of the two transitions should in principle be equivalent. However, as shown in Table 4.1, the observed ΔG values differ between the 1→2 (ΔG_{FF} = -8.61 kcal/mol; ΔG_{LU} = -9.47 kcal/mol) and 2→3 (ΔG_{FF} = -9.9 kcal/mol; ΔG_{LU} = -10.57 kcal/mol) transitions. These differences arise primarily

from variations in sampling space rather than the effects of the alternative redox states at distal cysteine residues.

In the LU state, a potential disulfide bond can form between a C_P and C_R residue from different monomers, indicating that these two cysteines are in close spatial proximity. This raises the possibility that perturbation of one residue may influence the other, and therefore SNO of C_P and C_R in this pair may not represent independent events. To examine this dependency, the alchemical transitions of SH-C_R to SNO-C_R were divided into two groups according to the redox state of the adjacent C_P: (i) SNO-C_R in the presence of SH-C_P (3→4 and 4→5 transitions), and (ii) SNO-C_R in the presence of SNO-C_P (1→6 and 6→7 transitions). If SNO of C_R is independent of the redox state of the neighboring C_P, then—starting from the same initial conformation—the FEP simulations of both groups should yield comparable ΔG_{LU} values. The degree of similarity between their ΔG_{LU} values thus reflects the extent of coupling: smaller differences indicate weaker dependency. The minimum ΔG_{LU} value for transitions involving SH-C_P is –12.89 kcal/mol, while that for transitions involving SNO-C_P is –13.04 kcal/mol. The minor difference between the two values indicates the low dependency of SNO in C_R to the adjacent C_P redox state, and the SNO in C_R can therefore be considered an essentially independent event.

The same protocol is also applied to assess whether the redox state of C_R influences the SNO of C_P. When the adjacent C_R residue is in the -SH form, the minimum ΔG_{LU} value for SNO modification of C_P is –10.57 kcal/mol (2→3 transition). In comparison, the minimum ΔG_{LU} in the presence of a neighboring SNO-C_R is –10.91 kcal/mol (8→5 transition). The tiny difference between two ΔG_{LU} values supports that the SNO of C_P is essentially independent of the redox state of adjacent C_R. Therefore, even in the LU state—where C_P and C_R residues are in close spatial proximity—their SNO events can be reasonably approximated as independent.

Nevertheless, an important concern remains: the low-energy conformational ensemble of a system in which both C_P and C_R are S-nitrosylated may differ from that of the S–S oxidized, disulfide-bridged reference model on which the S-nitrosylated model was based. Such structural inconsistency could lead to an insufficient sampling of the low-energy conformations, thereby affecting the accuracy of free-energy estimates for the doubly S-nitrosylated system.

Based on the assumption that the perturbation in FEP can only cause tiny impact in the local conformational change in limited sampling time, and the assessment described above, we can conclude that in the FF→LU conformational transition, all the performed perturbations can be treated as independent events. Accordingly, in the alchemical pathway, each perturbation step is independent of the others, and the sequence in which these perturbations are applied is therefore unimportant; only the specific perturbation site is important. As shown in Fig 2.4A, the transitions within the alchemical pathway are categorized to four groups based on their conformational state and perturbation sites: $\Delta G_{FF}_C_P$, $\Delta G_{FF}_C_R$, $\Delta G_{LU}_C_P$, and $\Delta G_{LU}_C_R$. For each group, their transitions and corresponding calculated ΔG values are listed in Table 4.1.

Specially, the FU structure consists of one side in the sub-FF state and the other side in the sub-LU state (Fig 2.3). By approximation, the FF state can be considered as a combination of two sub-FF states, and similarly, the LU state is treated as equivalent to the combination of two sub-LU states. Under this framework, FEP simulations transferring C_P-SH to C_P-SNO in the FU state (9→11 and 9→10 transitions) can be interpreted as sampling more configurational space for FF and LU state. In detail, along FF→FU alchemical pathway, the two C_P residues are modified independently: one located in the sub-FF half and the other in the sub-LU half. The resulting ΔG values are therefore assigned to ΔG_{FF} and ΔG_{LU}, respectively (Table 4.1).

		Transition	ΔG_{FF}	ΔG_{LU} (kcal/mol)
C_P	FF→LU	1→2	-8.61±0.054	-9.47±0.16
		2→3	-9.9±0.17	-10.57±0.085
		7→8	-10.28±0.16	-9.56±0.066
		8→5	-9.02±0.27	-10.91±0.054
	FF→FU	9→11 (pose1)	-9.11±0.27 -8.98±0.18	-
		9→11 (pose2)	-8.86±0.34 -10.75±0.10	-
		9→10 (pose1)	-8.85±0.19	-9.78±0.14
		9→10 (pose2)	-10.68±0.23	-9.61±0.072
C_R	FF→LU	3→4	-11.96±0.061	-10.6±0.057
		4→5	-11.89±0.034	-12.89±0.027
		1→6	-11.98±0.059	-11.33±0.11
		6→7	-12.05±0.047	-13.04±0.091

Table 4.1: Computed free energy changes in selected alchemical transition. ΔG_{FF} and ΔG_{LU} represent the free energy changes associated with given alchemical transition, in FF and LU conformational state, respectively. These alchemical transitions are categorized into two groups based on their perturbation sites: CP and CR. In each group, the variation in calculated ΔG values derives from different sampling configurational spaces. For 9→11 transition, the sub-FF side CP is mutated in the FU conformational state; thus, its calculated ΔG values correspond to the FF state.

4.2.2 Equilibration Time Determination

After assigning all the alchemical transitions and their corresponding ΔG values to either C_P or C_R groups, we focus on the specific FEP simulations themselves and the implication of their results. The free energy differences between redox states in the FF (ΔG_{FF}) and LU (ΔG_{LU}) states are presented in Table 4.1. These values were estimated by using analysis method MBAR as detailed in the Section 2.2.5. The uncertainties reported in Table 4.1 represent the standard deviations obtained from three replicates, reflecting the statistical uncertainties of the free energy estimations. For individual alchemical transitions, these uncertainties range from 0.02 to 0.27 kcal/mol, which falls within an acceptable range. It is important to note that these statistical uncertainties do

not include potential systematic errors in the force field model or other uncertainties arising from incomplete sampling of the system's configurations.

Accurate estimation of ΔG in FEP requires that the system be equilibrated. Since ΔG is the value of interest, it is used as the metric to evaluate whether the system has equilibrated. Specifically, the time that a converged free energy value is shown can be determined as a starting of the equilibrated system, and the time before this point is the equilibrium phase. The protocol for the determination of equilibrium time for a single FEP simulation is detailed in the Section 2.2.5.3.

Specifically, we plotted ΔG difference ($\Delta\Delta G^{equil}$) versus discarded time. As shown in Fig 4.6A, at the early stage of the time, the $\Delta\Delta G^{equil}$ changes dramatically, suggesting a sampling difference between two halves and an un-equilibrated system. As the discarded time increases, the $\Delta\Delta G^{equil}$ decreases and eventually approaches zero, it simply suggests a relative convergence of the ΔG estimates from both halves and, consequently, equilibration of the system in the aspect of free energy.

However, as discarded time increases more, the fluctuation of $\Delta\Delta G^{equil}$ also increases. This undesirably increased fluctuation derives from that fewer samplings are used for free energy calculation. Therefore, an accurate estimation of free energy in FEP requires not only an equilibrated system but also an enough sampling. This presents a trade-off: more discarded time improves the likelihood of equilibration but reduces sampling. To balance the equilibration time and sampling quantity, we defined the start of production sampling as the earliest time point where $\Delta\Delta G^{equil}$ first drops below 0.1 kcal/mol.

As an illustrative example, a FEP simulation along 1→2 transition in the FF state, the system approaches equilibrium at a discarded time of 2.08 ns, at which point the $\Delta\Delta G_{FF}{}^{equil}$ approaches 0.1 kcal/mol (Fig 4.6A). The same protocol was applied to all FEP simulations, revealing varying equilibration times across different cases. The

distribution of these equilibration times along with a fitted Gaussian curve is shown in Fig 4.6B. To ensure consistency in sampling duration and to maximize extent of equilibrium for all systems, the mean equilibration time of 3.58 ns was adopted as a uniform equilibration time cutoff for all simulations.

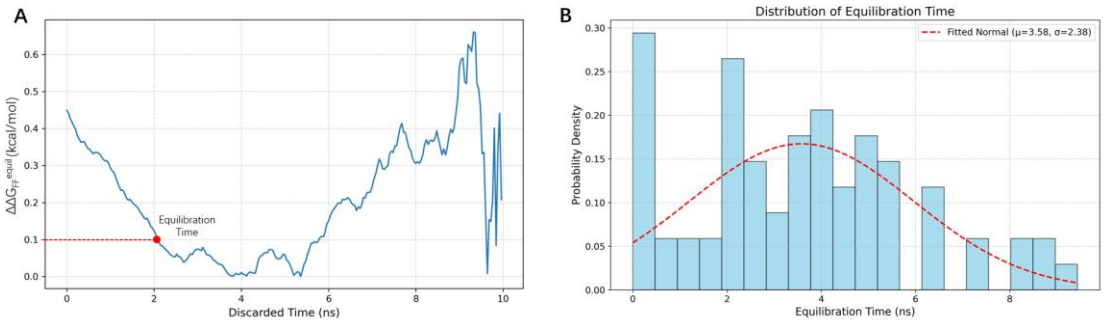


Figure 4.6: Determination of equilibration time. (A) Stability of the calculated ΔG_{FF} for 1→2 transition was assessed by changing discarded simulation time. Data prior to the discarded time point are excluded. The remaining trajectory is split into two halves, and ΔG_{FF} is calculated independently for each half; the $\Delta\Delta G_{FF}^{equil}$ reflects the variation between the two halves. At a discarded time of 2.08 ns, $\Delta\Delta G_{FF}^{equil}$ falls below 0.1 kcal/mol, suggesting that the system has reached relative equilibrium. The obvious severe fluctuation of $\Delta\Delta G_{FF}^{equil}$ after 8 ns derives from an insufficient sampling. (B) Distribution of equilibration times across all performed FEP simulations. The red dashed line refers to a fitted Gaussian curve. The mean equilibration time (μ) is 3.58 ns, indicating that approximately half of the simulations reach equilibrium by this point. To ensure consistency in sampling duration and to maximize the extent of equilibrium for all systems, 3.58 ns is adopted as a uniform equilibration time cutoff for all FEP simulations. All the calculated ΔG values are listed in the Table 4.1.

4.2.3 Error Assessment

Two analysis methods: TI and MBAR were employed independently to analyze the same trajectory and estimate the corresponding ΔG value. Applying both analysis methods enables cross-validation, serving as mutual confirmation of result reliability. Although the potential of mean force (PMF) derived from MBAR showed a slightly

larger magnitude compared to TI, both methods yielded consistent ΔG values for the 1→2 transition, with TI predicting -34.26 kJ/mol and MBAR -37.51 kJ/mol (Fig 4.7). This slight difference between two methods is in expectation, since MBAR generally provides a greater accuracy due to its efficient use of all collected data. For this reason, MBAR was selected as the primary estimator for subsequent analyses.

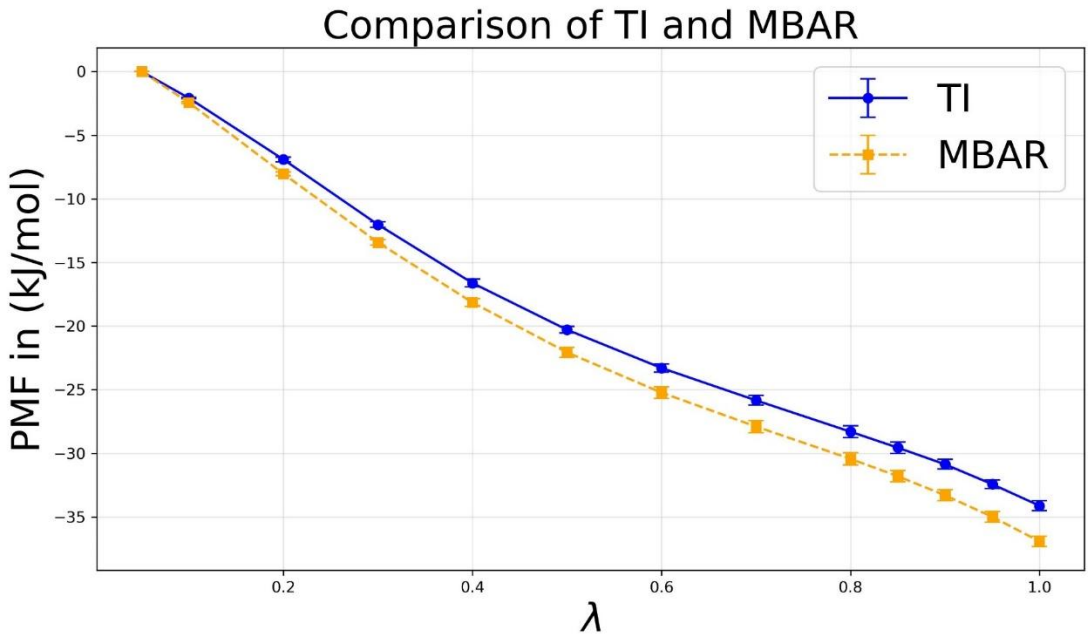


Figure 4.7: Potential of mean force with respect to λ for TI and MBAR. The ΔG_{FF} along 1→2 transition was estimated by using TI and MBAR, respectively.

The accuracy of the free energy estimation can be assessed using a closed thermodynamic cycle, where the starting and ending states are consistent. Since free energy is a state function, the net free energy change along the closed thermodynamic cycle, (1→2→3→4→5) - (1→6→7→8→5), is theoretically expected to be zero (Fig 2.4A). The estimated cycle-closure errors were 0.94 kcal/mol for FF state and 1.31 kcal/mol for the LU state. Nonetheless, considering that the final value results from eight-step alchemical transition path involving a large energy change (~90 kcal/mol), the accuracy of the calculation is quite satisfactory.

4.2.4 Substate Sampling and Approximation of State Free Energy

After confirming the accuracy of the analysis method and verifying that each system reached equilibrium, the calculated ΔG values (Table 4.1) can be considered reliable. However, the interpretation of these values requires careful consideration. In Table 4.1, the transitions within the alchemical pathway are categorized to four groups according to conformational state and perturbation sites: $\Delta G_{FF_C_P}$, $\Delta G_{FF_C_R}$, $\Delta G_{LU_C_P}$, and $\Delta G_{LU_C_R}$. A notable feature of the data is the variability observed within the same group. For example, in the C_P group for ΔG_{FF}, the calculated values range from -8.61 to -10.75 kcal/mol, exhibiting a fluctuation of approximately 2 kcal/mol—an extent that cannot be disregarded.

In principle, transitions belonging to the same group share the same conformational state and identical perturbation sites. They should therefore sample from the same conformational ensemble and result in similar ΔG values. However, in practice, the standard MD-based FEP algorithm can only sample a limited configurational space but not all possible configurations in the state.

Within each conformational state, multiple substates can exist, separated by energy barriers. On the nanosecond timescale, transitions between substates are unlikely to occur, meaning that the sampled configurations typically reside within a single low-energy basin rather than spanning the entire state. Therefore, the sampled configurations highly depend on the initial conformation of the system, here, the pose of the mutated cysteine. The variability of ΔG values in the same group mainly derives from uneven sampling of substates, driven by differences in the initial pose of the perturbation site.

For example, consider State 1, which encompasses all possible poses of SNO-Cys in the FF state (Fig 4.8). This state can be partitioned into three substates (Substate 1–3). In principle, microsecond-scale simulations would allow transitions between these

substates, thereby providing comprehensive sampling of the entire state and yielding a converged ΔG value regardless of the initial pose. However, our simulations are limited to 10 ns, a timescale insufficient to overcome the barriers separating substates. As a consequence, the calculated ΔG values are strongly influenced by the starting conformation: a simulation initiated in Substate 1 yields the ΔG value characteristic of Substate 1, while one initiated in Substate 2 produces the value specific to Substate 2. Thus, the ΔG values obtained in this work should be interpreted as substate-specific rather than representative of the full state.

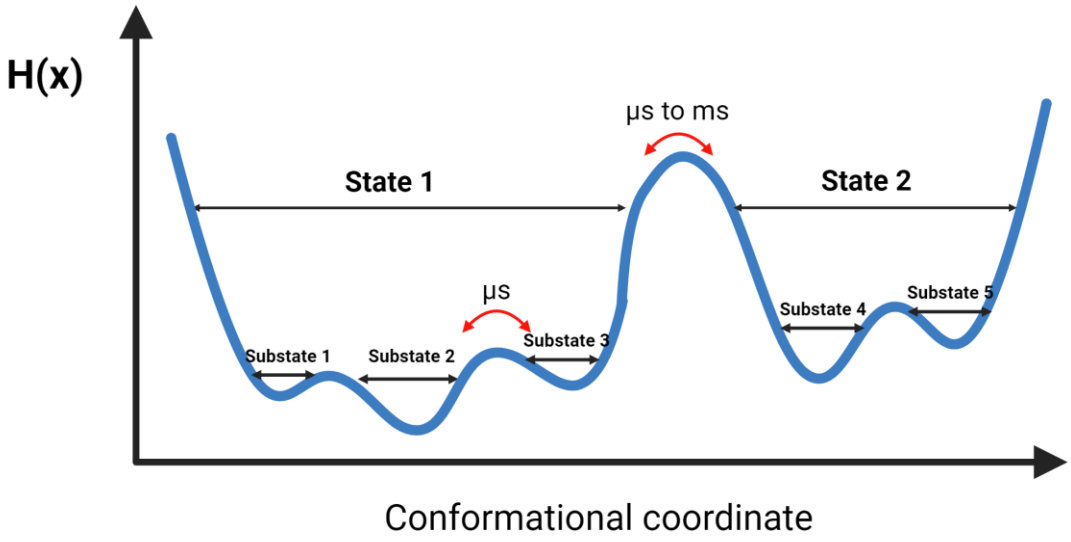


Figure 4.8: Diagram of a conformational energy landscape explaining the relation and difference between substate and state. The horizontal axis denotes the conformational coordinate, while the vertical axis H(x) represents the free energy. Two major conformational states (State 1 and State 2) are separated by a high-energy barrier, with interconversion occurring on the microsecond-to-millisecond timescale. Each state comprises several local minima (Substates 1–5), which are separated by lower barriers, allowing transitions on the microsecond timescale. This hierarchical landscape illustrates the coexistence of fast local motions between sub-state and slower global transitions between states.

This also suggests a strategy to improve sampling: initiating simulations from multiple initial poses. As detailed in the Section 2.2.5.2, during the system buildup process, the manual behavior bias together with nondeterministic protocols of atom addition and energy minimization collectively cause different poses of SNO-Cys. All the poses of mutated C_P in the FF state (including the sub-FF state) are shown in Fig 4.9A. A difference in ΔG_{FF} values can be found in 9→10 between pose1 (-8.85 kcal/mol) and pose2 (-10.68 kcal/mol) (Fig 4.9B). Their difference in the pose mainly lies in the N-CA-CB-S dihedral angle. The ~1.85 kcal/mol difference in energy suggests that the rotation of this dihedral angle is energetically unfavorable, likely requiring the system to overcome an energy barrier inaccessible on the nanosecond timescale of the simulation.

Because ΔG calculations are determined by the initial pose, each simulation explores only the configurational space of a single substate and thus provides a substate-specific rather than state-wide free energy. This raises the important question: which substate is more representative and contributes more to the whole state? Since each simulation reaches thermodynamic equilibrium within its sampled substate, an application of Bolzman factor $P \propto \exp[-G/k_BT]$ allows the calculation of the relative probability of each substate. Accordingly, lower free energy substates are exponentially more probable and contribute more substantially to the overall thermodynamic state. Quantitatively, a difference of 1 kcal/mol corresponds to a ~5.4-fold change in relative probability, while a 2 kcal/mol difference implies a ~29.7-fold difference. For example, in the C_P mutation, the substate $\Delta G_{FF}^{1\rightarrow2}$ (-8.61 kcal/mol) is far less populated and thus contributes less to the whole state ΔG_{FF} than that of $\Delta G_{FF}^{9\rightarrow11\ (pose2)}$ (-10.75 kcal/mol). Therefore, we utilize the lowest free energy value within each group to represent the state's free energy and neglect the others with higher free energy values.

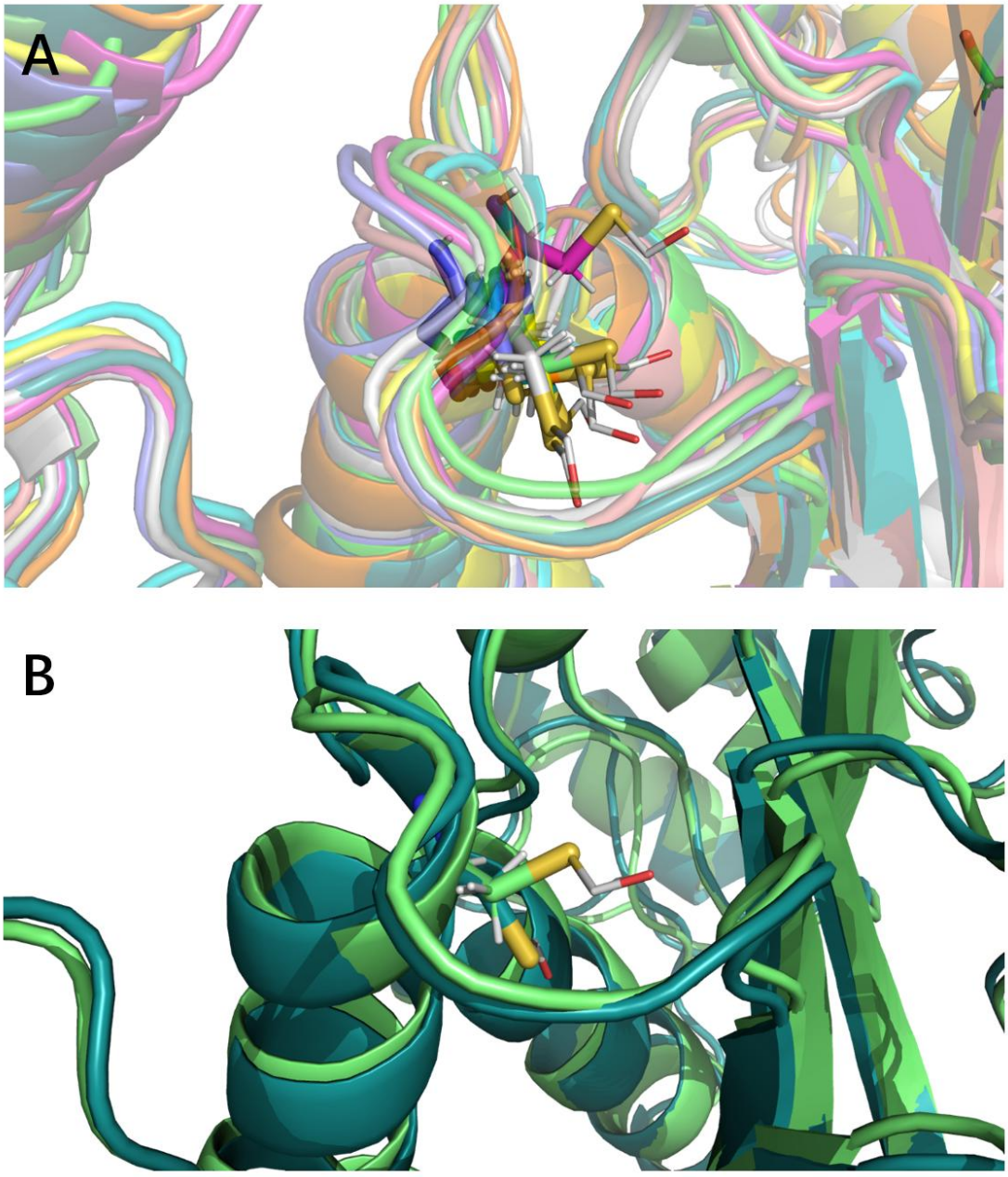


Figure 4.9: Comparison of initial SNO-C_P poses in the FF state. (A) Alignment of all initial SNO-C_P poses. (B) Comparison of the two initial poses: $9 \rightarrow 10^{(pose1)}$ shown in green and $9 \rightarrow 10^{(pose2)}$ shown in dark green.

4.2.5 Effect of SNO on Prdx2 Conformational Equilibrium

Using the approximation that the lowest-energy substate represents the entire state, the effective free energy differences were obtained (Table 4.2). The right-most column expresses the relative free energy ($\Delta\Delta G$) between -SH and -SNO form in the FF→LU

transition process, calculated according to Eq. (1.19). For a mutation from Cys-SH to Cys-SNO, a positive ΔΔG indicates that SNO promotes the conformational transition from FF to LU, thereby favoring the LU state, whereas a negative ΔΔG suggests stabilization of the FF state.

Notably, SNO in C_P and C_R exhibits distinct behaviors. In the C_P group, SNO slightly promotes unfolding, with ΔΔG = 0.23 kcal/mol. In contrast, SNO at C_R exerts a much stronger effect, with ΔΔG = 0.99 kcal/mol, corresponding to an approximately 5.4-fold higher probability of the LU state. Given that the magnitude of the ΔΔG in the C_R (0.99 kcal/mol) is much larger than that the C_P (0.23 kcal/mol), we conclude that SNO at C_R significantly facilitates the FF→LU transition, while SNO at C_P has a relatively minimal impact. Overall, when all four cysteines undergo SNO, the cumulative ΔΔG is 2.44 kcal/mol, predominantly driven by the contribution from C_R. This corresponds to a ~60-fold increase in the population of the LU state compared to the thiol form, suggesting a substantial shift in conformational equilibrium induced by SNO.

Despite its minor role in promoting conformational transitions, SNO at the C_P site directly impairs Prdx2 function. The antioxidant activity of Prdx2 relies on the peroxidatic C_P reducing H_2O_2. Upon SNO, the C_P is no longer capable of reducing H_2O_2. By contrast, the C_R plays a key role in the function regeneration of Prdx2 activity through disulfide bond formation with C_P during the catalytic cycle. The SNO-C_R does not directly influence the Prdx2's function but change the conformational equilibrium between FF and LU. Our FEP results demonstrate that SNO-C_R promotes protein unfolding, thereby increasing the likelihood of C_P–C_R interaction and potentially accelerating disulfide bond formation and Prdx2 regeneration.

Given that no available evidence supports that the two sides of the Prdx2 dimer unfold simultaneously, an intermediate state—designated FU—is postulated, in which one side of the protein adopts a sub-FF conformation while the other is in a sub-LU

conformation (Fig 2.3). This implies a two-step FF→FU→LU transition, where unfolding occurs sequentially. In the functional activated FF state, the nucleophilic C_P can easily react with electrophilic molecules such as NO^+ donors and H_2O_2, while the C_R remains buried. This structural arrangement suggests that C_P may be more accessible to SNO than C_R. Consequently, we specifically evaluated the impact of SNO at the C_P site during the FF→FU transition. As shown in Table 4.2, SNO of C_P, whether located on the sub-FF or sub-LU side of the dimer, exerts minimal influence on the equilibrium between FF and FU states. These findings underscore the relatively weak conformational effect of C_P SNO during unfolding, despite its significant role Prdx2 function.

	ΔG_{FF}	ΔG_{LU}	$\Delta\Delta G$ (kcal/mol)
C_P	-10.75±0.10	-10.91±0.054	0.23±0.11
C_R	-12.05±0.047	-13.04±0.091	0.99±0.10

Table 4.2: Effective free energy differences and relative free energy difference ($\Delta\Delta G$).

4.2.6 How C_R is S-nitrosylated?

Fang and his colleagues demonstrated that both C_P and C_R are the target of SNO, using S-nitrosocysteine (SNOC) as an NO^+ donor[156]. SNOC, a low-molecular-weight nitrosothiol, mediates transnitrosylation, in which the nucleophilic sulfur atom of a thiol attacks the electrophilic nitrogen of SNOC, thereby transferring the nitroso group to the cysteine. Consequently, both the rate of NO^+ donor delivery to the target cysteine and of the transnitrosylation reaction determine the overall kinetics and selectivity of SNO[199, 200].

In the FF state of Prdx2, the C_P site presents a positively charged Arg side chain that lowers the p*K*a of the C_P thiol, resulting in a thiolate form C_P (S^--C_P), which is more nucleophilic. Additionally, a nearby negatively charged Thr side chain contributes to destabilizing NO^+ donors and thereby promoting the NO^+ donation. This acid–base motif cooperatively catalyzes the SNO at the C_P site.

By contrast, in the FF state, the C_R is buried within the protein and inaccessible to solvent, making direct contact with NO^+ donors unlikely. A fundamental requirement for the SNO of C_R is a physical contact of C_R to the NO^+ donor. Two potential SNO mechanisms for C_R have been proposed here.

Direct SNO of C_R by Small-Molecule NO^+ Donors

The first mechanism proposes that C_R can be directly S-nitrosylated by small-molecule NO^+ donors when it becomes transiently solvent-exposed in an intermediate state. Although C_R is buried in the FF conformation, this state may transition to a relatively stable intermediate in which C_R is accessible to RNS, enabling direct SNO (Fig 4.10).

Several lines of evidence support this mechanism. First, in the FF state, the C_R is in a relative hydrophobic environment formed by its surrounding residues. Given that NO and O_2 preferentially partition into the hydrophobic phase and that hydrolysis of NO_X species is suppressed in such media, a hydrophobic site favours the formation of reactive NO^+ donors such as NO_2, N_2O_3, and N_2O_4, thereby facilitating the SNO[200, 201]. Consequently, the hydrophobic microenvironment surrounding C_R in the FF state suggests the potential for a hydrophobic cavity in the intermediate state that could facilitate SNO, where C_R becomes accessible to both solvent and NO^+ donors. Second, mutational studies have demonstrated that C_R can still undergo SNO even when C_P is substituted. This finding indicates that the SNO of C_R may occur independently of C_P and is likely mediated directly by small reactive species rather than through trans-nitrosylation from C_P.

For this mechanism to be feasible, the intermediate state must be relatively stable, and the transition from the FF state must be kinetically favorable. Otherwise, the intermediate would either rapidly convert into alternative conformations or fail to form efficiently, thereby preventing SNO of C_R. Identification of such an intermediate state would therefore provide strong support for this hypothesis. In our study, although

intermediate conformations were observed in string method simulations, the kinetics of their interconversion with the FF state were not determined. As a result, the stability of these intermediates remains unresolved.

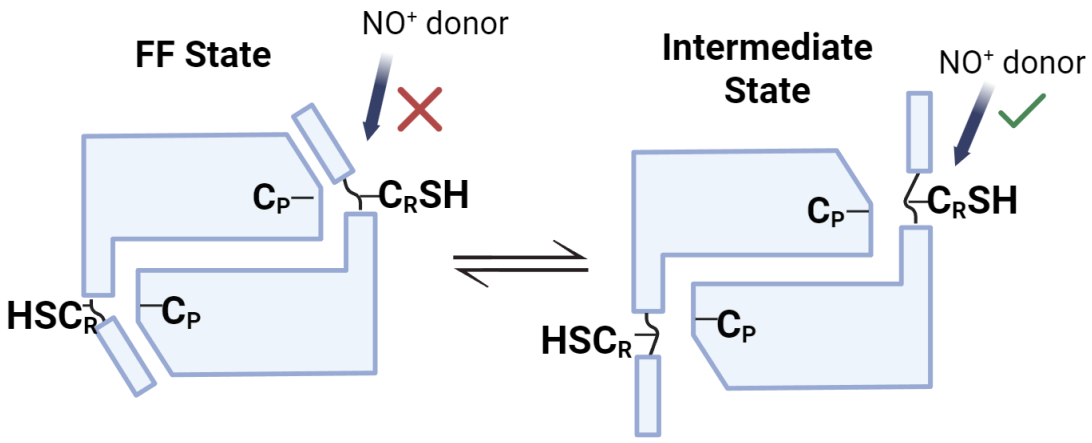


Figure 4.10: Proposed mechanism of C_R SNO by small-molecule NO^+ donors. In an intermediate state where C_R becomes solvent-exposed, RNS can directly modify C_R.

Alternative Mechanism of C_R SNO: Transnitrosylation from C_P

An alternative mechanism for C_R modification involves transnitrosylation from SNO-C_P (Fig 4.11). In this model, once C_P is S-nitrosylated, the resulting SNO-C_P promotes the FF → LU conformational transition. In the LU state, the SNO-C_P is in close spatial proximity to SH-C_R, thereby enabling transfer of the NO group to C_R to form SNO-C_R. After this transfer, the protein is proposed to refold into the FF state, with SNO-C_R and SH-C_P available for subsequent round SNO. This model would thus provide a means by which both cysteines are modified.

However, our FEP simulations suggest that the SNO at C_P has minor influence on the conformational equilibrium between the FF and LU states, thereby reducing the likelihood that this pathway is the dominant mechanism of C_R modification.

When the SNO-C_P and SH-C_R are in close spatial proximity, two competing reaction pathways are possible. The nucleophilic sulfur atom of SH-C_R can attack either

the electrophilic sulfur or nitrogen of the SNO-C_P. Nucleophilic attack at the sulfur leads to the formation of a disulfide bond, whereas attack at the nitrogen results in transnitrosylation, transferring the NO group to the thiol-containing cysteine. The former reaction yields a stable disulfide linkage and is considered irreversible in the absence of external reductants, whereas the latter is a reversible event that simply regenerates SNO-Cys on a different cysteine.

In SNO-Cys, the nitrogen atom is more electrophilic than the sulfur, making it a preferred target for nucleophilic attack by thiol groups. Consequently, transnitrosylation is more likely to proceed. If the LU state structure is preserved, this transnitrosylation cycle can continue until a disulfide bond is eventually formed. This implies that a kinetic competition between disulfide bond formation and disruption of the LU state through global conformational changes that spatially separate the C_P and C_R residues. If disulfide formation proceeds faster than conformational rearrangement, the covalent disulfide bond forms. Conversely, if the LU state collapses before disulfide formation, the NO moiety may be transferred instead. Validation of this mechanism would require a direct comparison of the reaction rates between LU structure destruction and disulfide formation.

Following NO transfer from C_P to C_R, refolding into the FF state would allow C_P to undergo further SNO in its favorable microenvironment. However, our FEP results suggest that SNO-C_R biases the equilibrium toward the LU state rather than the FF state, which complicates this mechanistic model.

Finally, mutational assays demonstrate that C_R can still be S-nitrosylated in the absence of C_P[156]. This independence implies that C_R modification does not strictly rely on transnitrosylation from C_P, which is inconsistent with this proposed mechanism. In these assays, C_P is replaced by histidine, raising the possibility that this substitution itself alters structure and promotes exposure of C_R. Therefore, identification of

structural evidence showing that substitution of cysteine with histidine directly promotes C_R exposure would be necessary to substantiate this pathway.

Despite that no direct evidence can prove either of the mechanisms, our FEP simulation results and mutation experiments are more consistent with the first mechanism. Except for C_P and C_R, Prdx2 also possesses a third cysteine, Cys69, which remains buried within the protein in both FF and LU states. This persistent isolation from solvent prohibits interactions between Cys69 and NO^+ donor, and thereby the SNO of Cys69[156].

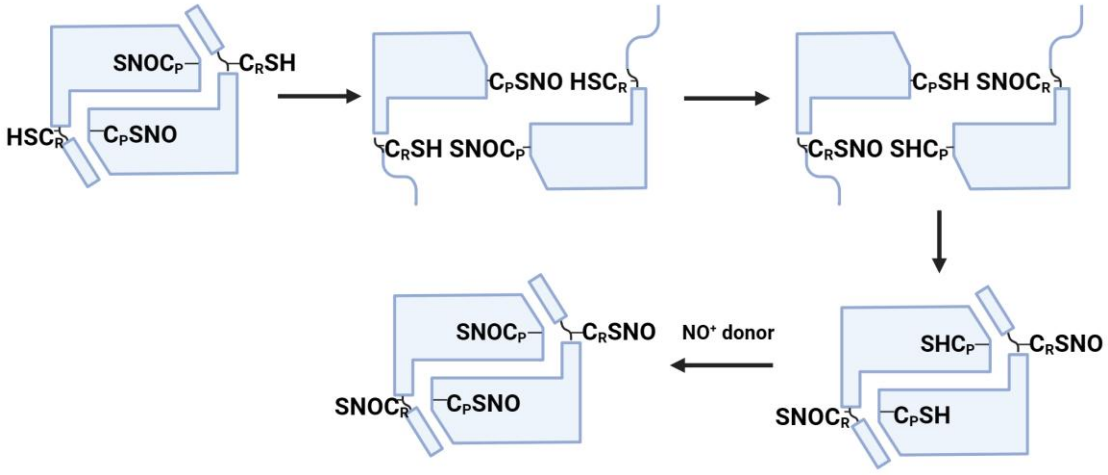


Figure 4.11: Proposed mechanism of C_R SNO by SNO-C_P.

4.3 Transition Pathway Identification

FEP analysis provides quantitative insight into how PTMs modulate the equilibrium between the FF and FU conformations of Prdx2. However, FEP alone does not reveal the transition pathways underlying this conformational switch. To complement this, we employed the FTS method to identify the most probable transition route connecting the FF and FU states under different PTMs.

A prerequisite for FTS is the specification of a CV that can describe progression along the transition. Despite the simplicity of one-dimensional RMSD, it well reflects the conformational fluctuations. Here, as in committor analysis, we used the cumulative RMSD between successive intermediate structures, computed over all protein atoms, to define the string coordinate. This choice provides a straightforward yet robust representation of conformational progression, though direct visualization in this one-dimensional space is limited. To aid interpretation, the resulting transition pathways were projected onto the PC subspace.

The optimization process of the FF→FU transition pathway is exemplified by the system of SNO_Cp. Successive iterations of the string (colored from purple to bright red) progressively converge, ultimately yielding the most probable pathway (Fig 4.12, left). This convergence demonstrates the robustness of the FTS approach in refining an initially interpolated path into a physically meaningful transition route. The structural overlap of the FF and FU endpoints with their intermediate states is visualized in Fig 4.12, right.

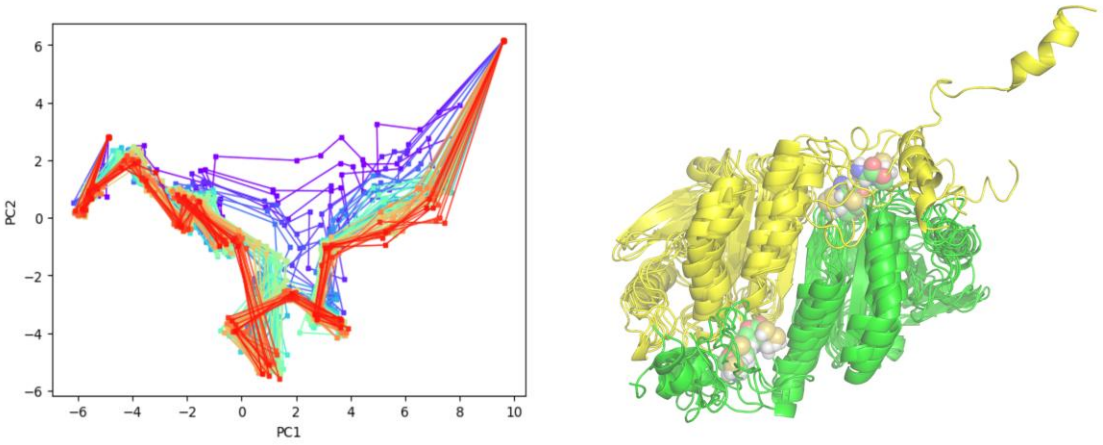


Figure 4.12: Optimization process of the FF→FU transition pathway for the SNO_C_P system. LEFT: Successive iterations of the string, projected onto the PC1–PC2 subspace. Strings are colored from purple (initial) to bright red (final), showing convergence toward the most probable pathway. PC1 accounts for 43% of the variance and PC2 for 6%. RIGHT: Structural overlap of the intermediate states, with the two subunits colored yellow and green. C_P and C_R are highlighted in sphere style.

Finally, to directly compare the influence of different PTMs on the conformational transition, the final optimized strings for all systems were projected onto the same PC1–PC2 subspace (Fig 4.13). PC1 alone accounts for 45% of the total variance, while PC2 contributes 8.0%. Despite PTM-dependent variations in detail, all pathways share a consistent global trend: an initial progression from left to right along PC1, accompanied by a modest downward shift in PC2, followed by an upward deflection toward the FU state. This indicates that, although the PTMs modulate local structural preferences and transition barriers, they generally preserve a similar FF→FU pathway.

However, the FTS analysis itself only yields the most probable transition pathway and does not provide the corresponding free-energy profile. As such, it cannot directly quantify barrier heights or relative basin stabilities along the route. To complement this limitation, enhanced sampling strategies such as weighted ensemble simulations or path-based umbrella sampling can be employed. These methods can project the dynamics onto the FTS-derived pathways (Fig 4.13), thereby generating the potential

of mean force and offering quantitative insight into both the thermodynamics and kinetics of the FF→FU transition.

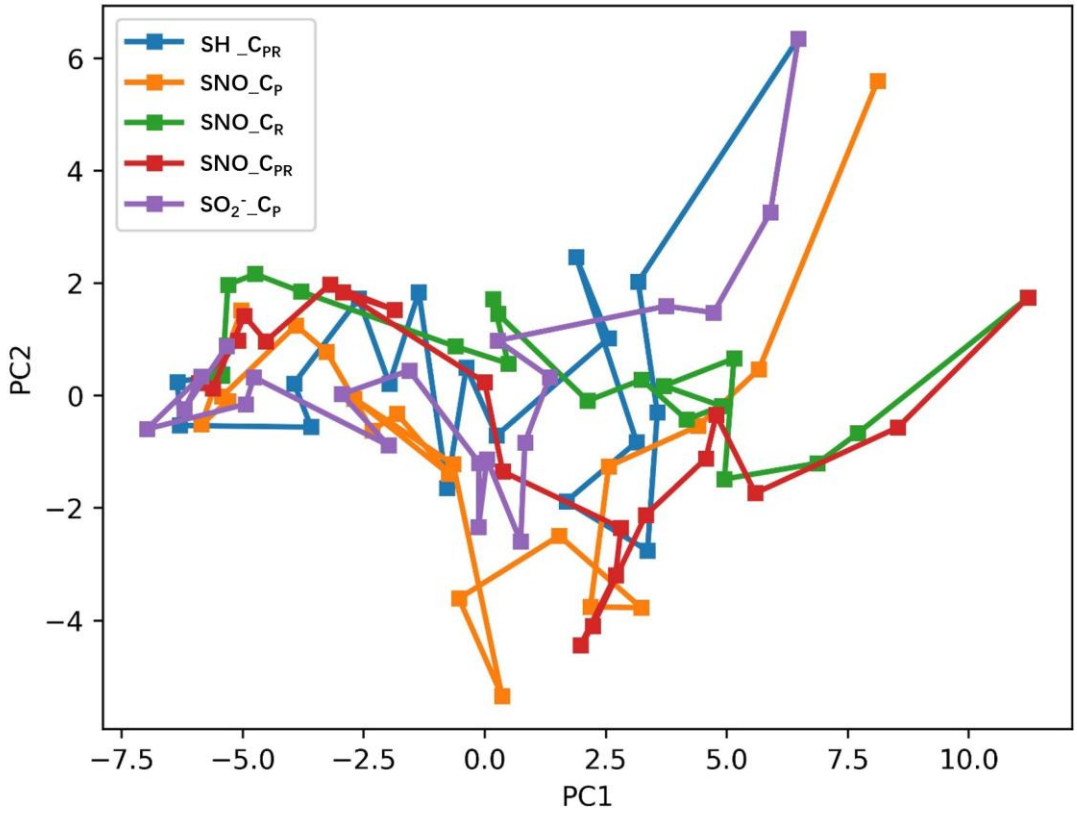


Figure 4.13: Comparison of optimized FF→FU transition pathways under different PTMs. Final strings for SH_C$_{PR}$, SNO_C$_{P}$, SNO_C$_{R}$, SNO_C$_{PR}$, and SO_2^-_C$_{P}$ systems are projected onto the PC1–PC2 subspace. PC1 and PC2 account for 45% and 8% of the variance, respectively.

4.4 Conclusion

This chapter establishes an integrated framework—PCA for mode characterization, FEP for thermodynamics, and FTS for pathway discovery—to quantify how cysteine nitrosylation reshapes the FF↔LU equilibrium in Prdx2. PCA shows that the FF state remains comparatively rigid while the LU state displays more flexibility, with motion concentrated at the dimer interface helices (the C-terminal is excluded); nitrosylation at C_P restricts LU-state flexibility, highlighting a PTM-dependent constraint on interfacial motions within the 40 ns timescale.

A rigorously designed FEP workflow is introduced to quantify how cysteine nitrosylation shifts the FF↔LU equilibrium. This protocol incorporates equilibration-time analysis, cross-validation with TI and MBAR, and thermodynamic cycle checks to ensure accuracy and convergence. Application of this workflow shows that nitrosylation at C_P slightly stabilizes the LU state, whereas nitrosylation at C_R, it stabilizes the LU state to a greater extent. FTS supplies the most probable FF→FU routes under each PTM and, when projected into a common PC subspace, reveals a conserved global path.

At the same time, the work clarifies important limitations that guide future improvements. First, in ligand-optimization contexts, a substituent can shift ΔΔG by >10 kcal/mol, so 1–2 kcal/mol fluctuations from incomplete sampling are often negligible; by contrast, protein-mutation or PTM studies typically seek ΔΔG changes on the order of −2 to 6 kcal/mol, where a 1–2 kcal/mol spread can be decisive. In our study, ΔG uncertainties approach ~2 kcal/mol while the ΔΔG shifts lie near 0–1 kcal/mol, underscoring that the limited sampling cannot be neglected. To broaden coverage, simulations are initialized from multiple side-chain poses and the lowest-free-energy pose is used as the state representative Although this strategy reduces substate bias, it does not fully resolve the problem of limited sampling.

To achieve complete sampling, the entire conformational landscape of each state must be explored. This can be accomplished by coupling FEP with enhanced-sampling approaches such as umbrella sampling or replica exchange. These methods enable a more reliable free-energy calculation. Similarly, while FTS provides the most likely structural pathway, the absence of free-energy data limits quantitative interpretation. Combining FTS-derived pathways with umbrella sampling or weighted ensemble simulations would allow barrier heights and kinetics to be computed directly, completing the link between mechanism, thermodynamics, and rates.

Another key future direction is the identification of kinetically stable intermediate states along the FF→LU transition. Detecting and characterizing such intermediates—either experimentally or through long-timescale simulations—would be critical to validating the first proposed nitrosylation mechanism, in which the C_R residue is S-nitrosylated by small molecular donors during the early stages of unfolding.

Finally, PropKa analysis predicts that the C_P sulfur exists primarily in a deprotonated state, but alchemical FEP is intrinsically challenged by electronic changes. As a compromise, this study quantifies ΔG between the –SH and –SNO instead of the thiolate form. A more rigorous treatment that could model protonation and electronic effects could offer more reliable insights.

Reference

1. Anjo, S. I.; He, Z.; Hussain, Z.; Farooq, A.; McIntyre, A.; Laughton, C. A.; Carvalho, A. N.; Finelli, M. J., Protein oxidative modifications in neurodegenerative diseases: from advances in detection and modelling to their use as disease biomarkers. *Antioxidants* **2024,** *13* (6), 681.

2. Hollingsworth, S. A.; Dror, R. O., Molecular Dynamics Simulation for All. *Neuron* **2018,** *99* (6), 1129-1143.

3. Karplus, M.; McCammon, J. A., Molecular dynamics simulations of biomolecules. *Nature structural biology* **2002,** *9* (9), 646-652.

4. Dill, K.; Bromberg, S., *Molecular driving forces: statistical thermodynamics in biology, chemistry, physics, and nanoscience*. Garland Science: 2010.

5. Karplus, M.; Petsko, G. A., Molecular dynamics simulations in biology. *Nature* **1990,** *347* (6294), 631-9.

6. Sinha, S.; Tam, B.; Wang, S. M., Applications of molecular dynamics simulation in protein study. *Membranes* **2022,** *12* (9), 844.

7. Warshel, A., Molecular dynamics simulations of biological reactions. *Accounts of chemical research* **2002,** *35* (6), 385-395.

8. Frauenfelder, H.; Sligar, S. G.; Wolynes, P. G., The energy landscapes and motions of proteins. *Science* **1991,** *254* (5038), 1598-1603.

9. Shaw, D. E.; Maragakis, P.; Lindorff-Larsen, K.; Piana, S.; Dror, R. O.; Eastwood, M. P.; Bank, J. A.; Jumper, J. M.; Salmon, J. K.; Shan, Y., Atomic-level characterization of the structural dynamics of proteins. *Science* **2010,** *330* (6002), 341-346.

10. Copeland, R. A.; Pompliano, D. L.; Meek, T. D., Drug–target residence time and its implications for lead optimization. *Nature reviews Drug discovery* **2006,** *5* (9), 730-739.

11. Pan, A. C.; Borhani, D. W.; Dror, R. O.; Shaw, D. E., Molecular determinants of drug–receptor binding kinetics. *Drug discovery today* **2013,** *18* (13-14), 667-673.

12. Bernetti, M.; Masetti, M.; Rocchia, W.; Cavalli, A., Kinetics of drug binding and residence time. *Annual review of physical chemistry* **2019,** *70* (1), 143-171.

13. Wang, A.-h.; Zhang, Z.-c.; Li, G.-h., Advances in enhanced sampling molecular dynamics simulations for biomolecules. *Chinese Journal of Chemical Physics* **2019,** *32* (3), 277-286.

14. Mey, A.; Allen, B. K.; Macdonald, H. E. B.; Chodera, J. D.; Hahn, D. F.; Kuhn, M.; Michel, J.; Mobley, D. L.; Naden, L. N.; Prasad, S.; Rizzi, A.; Scheen, J.; Shirts, M. R.; Tresadern, G.; Xu, H., Best Practices for Alchemical Free Energy Calculations [Article v1.0]. *Living J Comput Mol Sci* **2020,** *2* (1).

15. Badaoui, M.; Buigues, P. J.; Berta, D.; Mandana, G. M.; Gu, H.; Foldes, T.; Dickson, C. J.; Hornak, V.; Kato, M.; Molteni, C.; Parsons, S.; Rosta, E., Combined Free-Energy Calculation and Machine Learning Methods for Understanding Ligand Unbinding Kinetics. *J Chem Theory Comput* **2022,** *18* (4), 2543-2555.

16. Bernardi, R. C.; Melo, M. C. R.; Schulten, K., Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim Biophys Acta* **2015,** *1850* (5), 872-877.

17. Williams-Noonan, B. J.; Yuriev, E.; Chalmers, D. K., Free Energy Methods in Drug Design: Prospects of "Alchemical Perturbation" in Medicinal Chemistry. *J Med Chem* **2018,** *61* (3), 638-649.

18. Hänggi, P.; Talkner, P.; Borkovec, M., Reaction-rate theory: fifty years after Kramers. *Reviews of modern physics* **1990,** *62* (2), 251.

19. Zuckerman, D. M., *Statistical physics of biomolecules: an introduction*. CRC press: 2010.

20. Onuchic, J. N.; Luthey-Schulten, Z.; Wolynes, P. G., Theory of protein folding: the energy landscape perspective. *Annual review of physical chemistry* **1997,** *48* (1), 545-600.

21. Copeland, R. A., The drug–target residence time model: a 10-year retrospective. *Nature Reviews Drug Discovery* **2016,** *15* (2), 87-95.

22. Held, C.; Sadowski, G., Thermodynamics of bioreactions. *Annual review of chemical and biomolecular engineering* **2016,** *7* (1), 395-414.

23. Frenkel, D.; Smit, B., *Understanding molecular simulation: from algorithms to applications*. Elsevier: 2023.

24. Chodera, J. D.; Noé, F., Markov state models of biomolecular conformational dynamics. *Current opinion in structural biology* **2014,** *25*, 135-144.

25. McQuarrie, D., Statistical mechanics University Science Books. New York: Harper and Row: 1997.

26. Gibbs, J. W., *Elementary principles in statistical mechanics: developed with especial reference to the rational foundations of thermodynamics*. C. Scribner's sons: 1902.

27. Hill, T. L., *An introduction to statistical thermodynamics*. Courier Corporation: 1986.

28. Chandler, D., Introduction to modern statistical. *Mechanics. Oxford University Press, Oxford, UK* **1987,** *5* (449), 11.

29. Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J., Current status of transition-state theory. *The Journal of physical chemistry* **1996,** *100* (31), 12771-12800.

30. Copeland, R. A., The dynamics of drug-target interactions: drug-target residence time and its impact on efficacy and safety. *Expert opinion on drug discovery* **2010,** *5* (4), 305-310.

31. Tummino, P. J.; Copeland, R. A., Residence time of receptor− ligand complexes and its effect on biological function. *Biochemistry* **2008,** *47* (20), 5481-5492.

32. Vauquelin, G.; Charlton, S. J., Long-lasting target binding and rebinding as mechanisms to prolong in vivo drug action. *British journal of pharmacology* **2010,** *161* (3), 488-508.

33. Swinney, D. C., Biochemical mechanisms of drug action: what does it take for success? *Nature reviews Drug discovery* **2004,** *3* (9), 801-808.

34. Hyre, D. E.; LE TRONG, I.; Freitag, S.; Stenkamp, R. E.; Stayton, P. S., Ser45 plays an important role in managing both the equilibrium and transition state energetics of the streptavidin–biotin system. *Protein Science* **2000,** *9* (5), 878-885.

35. Lu, H.; England, K.; am Ende, C.; Truglio, J. J.; Luckner, S.; Reddy, B. G.; Marlenee, N. L.; Knudson, S. E.; Knudson, D. L.; Bowen, R. A., Slow-onset inhibition of the FabI enoyl reductase from Francisella tularensis: residence time and in vivo activity. *ACS chemical biology* **2009,** *4* (3), 221-231.

36. Guo, D.; Mulder-Krieger, T.; IJzerman, A. P.; Heitman, L. H., Functional efficacy of adenosine A2A receptor agonists is positively correlated to their receptor residence time. *British journal of pharmacology* **2012,** *166* (6), 1846-1859.

37. de Witte, W. E.; Danhof, M.; van der Graaf, P. H.; de Lange, E. C., In vivo target residence time and kinetic selectivity: the association rate constant as determinant. *Trends in pharmacological sciences* **2016,** *37* (10), 831-842.

38. Guo, D.; Dijksteel, G. S.; van Duijl, T.; Heezen, M.; Heitman, L. H.; AP, I. J., Equilibrium and kinetic selectivity profiling on the human adenosine receptors. *Biochem Pharmacol* **2016,** *105*, 34-41.

39. Sykes, D. A.; Stoddart, L. A.; Kilpatrick, L. E.; Hill, S. J., Binding kinetics of ligands acting at GPCRs. *Molecular and cellular endocrinology* **2019,** *485*, 9-19.

40. Baker, J. G., The selectivity of β-adrenoceptor agonists at human β1-, β2-and β3-adrenoceptors. *British journal of pharmacology* **2010,** *160* (5), 1048-1061.

41. Moore, R. H.; Millman, E. E.; Godines, V.; Hanania, N. A.; Tran, T. M.; Peng, H.; Dickey, B. F.; Knoll, B. J.; Clark, R. B., Salmeterol stimulation dissociates beta2-adrenergic receptor phosphorylation and internalization. *Am J Respir Cell Mol Biol* **2007,** *36* (2), 254-61.

42. Masureel, M.; Zou, Y.; Picard, L. P.; van der Westhuizen, E.; Mahoney, J. P.; Rodrigues, J.; Mildorf, T. J.; Dror, R. O.; Shaw, D. E.; Bouvier, M.; Pardon, E.; Steyaert, J.; Sunahara, R. K.; Weis, W. I.; Zhang, C.; Kobilka, B. K., Structural insights into binding specificity, efficacy and bias of a beta(2)AR partial agonist. *Nat Chem Biol* **2018,** *14* (11), 1059-1066.

43. Sykes, D. A.; Charlton, S. J., Slow receptor dissociation is not a key factor in the duration of action of inhaled long-acting β2-adrenoceptor agonists. *British journal of pharmacology* **2012,** *165* (8), 2672-2683.

44. Stoddart, L. A.; Johnstone, E. K.; Wheal, A. J.; Goulding, J.; Robers, M. B.; Machleidt, T.; Wood, K. V.; Hill, S. J.; Pfleger, K. D., Application of BRET to monitor ligand binding to GPCRs. *Nature methods* **2015,** *12* (7), 661-663.

45. Bouzo-Lorenzo, M.; Stoddart, L. A.; Xia, L.; IJzerman, A. P.; Heitman, L. H.; Briddon, S. J.; Hill, S. J., A live cell NanoBRET binding assay allows the study of ligand-binding kinetics to the adenosine A3 receptor. *Purinergic Signalling* **2019,** *15* (2), 139-153.

46. Motulsky, H. J.; Mahan, L., The kinetics of competitive radioligand binding predicted by the law of mass action. *Molecular pharmacology* **1984,** *25* (1), 1-9.

47. Patching, S. G., Surface plasmon resonance spectroscopy for characterisation of membrane protein–ligand interactions and its potential for drug discovery. *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2014,** *1838* (1), 43-55.

48. Capelli, D.; Parravicini, C.; Pochetti, G.; Montanari, R.; Temporini, C.; Rabuffetti, M.; Trincavelli, M. L.; Daniele, S.; Fumagalli, M.; Saporiti, S., Surface plasmon resonance as a tool for ligand binding investigation of engineered GPR17 receptor, a G protein coupled receptor involved in myelination. *Frontiers in chemistry* **2020,** *7*, 910.

49. Abdiche, Y.; Malashock, D.; Pinkerton, A.; Pons, J., Determining kinetics and affinities of protein interactions using a parallel real-time label-free biosensor, the Octet. *Analytical biochemistry* **2008,** *377* (2), 209-217.

50. Walkup, G. K.; You, Z.; Ross, P. L.; Allen, E. K.; Daryaee, F.; Hale, M. R.; O'donnell, J.; Ehmann, D. E.; Schuck, V. J.; Buurman, E. T., Translating slow-binding inhibition kinetics into cellular and in vivo effects. *Nature chemical biology* **2015,** *11* (6), 416-423.

51. Hulme, E. C.; Trevethick, M. A., Ligand binding assays at equilibrium: validation and interpretation. *British journal of pharmacology* **2010,** *161* (6), 1219-1237.

52. Jameson, D. M.; Sawyer, W. H., [12] Fluorescence anisotropy applied to biomolecular interactions. In *Methods in enzymology*, Elsevier: 1995; Vol. 246, pp 283-300.

53. Selvin, P. R., The renaissance of fluorescence resonance energy transfer. *Nature structural biology* **2000,** *7* (9), 730-734.

54. Velazquez-Campoy, A.; Freire, E., Isothermal titration calorimetry to determine association constants for high-affinity ligands. *Nature protocols* **2006,** *1* (1), 186-191.

55. Hinterdorfer, P.; Dufrêne, Y. F., Detection and localization of single molecular recognition events using atomic force microscopy. *Nature methods* **2006,** *3* (5), 347-355.

56. Merkel, R.; Nassoy, P.; Leung, A.; Ritchie, K.; Evans, E., Energy landscapes of receptor–ligand bonds explored with dynamic force spectroscopy. *Nature* **1999,** *397* (6714), 50-53.

57. Berman, H. M.; Battistuz, T.; Bhat, T. N.; Bluhm, W. F.; Bourne, P. E.; Burkhardt, K.; Feng, Z.; Gilliland, G. L.; Iype, L.; Jain, S., The protein data bank. *Biological Crystallography* **2002,** *58* (6), 899-907.

58. Kubinyi, H., QSAR and 3D QSAR in drug design Part 1: methodology. *Drug discovery today* **1997,** *2* (11), 457-467.

59. Hall, A.; Nelson, K.; Poole, L. B.; Karplus, P. A., Structure-based insights into the catalytic power and conformational dexterity of peroxiredoxins. *Antioxidants & redox signaling* **2011,** *15* (3), 795-815.

60. Henzler-Wildman, K.; Kern, D., Dynamic personalities of proteins. *Nature* **2007,** *450* (7172), 964-72.

61. Bolduc, J. A.; Nelson, K. J.; Haynes, A. C.; Lee, J.; Reisz, J. A.; Graff, A. H.; Clodfelter, J. E.; Parsonage, D.; Poole, L. B.; Furdui, C. M., Novel hyperoxidation resistance motifs in 2-Cys peroxiredoxins. *Journal of Biological Chemistry* **2018,** *293* (30), 11901-11912.

62. Peskin, A. V.; Meotti, F. C.; Kean, K. M.; Göbl, C.; Peixoto, A. S.; Pace, P. E.; Horne, C. R.; Heath, S. G.; Crowther, J. M.; Dobson, R. C., Modifying the resolving cysteine affects the structure and hydrogen peroxide reactivity of peroxiredoxin 2. *Journal of Biological Chemistry* **2021,** *296*.

63. Rhee, S. G.; Woo, H. A., Multiple functions of 2-Cys peroxiredoxins, I and II, and their regulations via post-translational modifications. *Free Radic Biol Med* **2020,** *152*, 107-115.

64. Wales, D. J.; Doye, J. P.; Miller, M. A.; Mortenson, P. N.; Walsh, T. R., Energy landscapes: from clusters to biomolecules. *Advances in Chemical Physics* **2000,** *115*, 1-111.

65. Kokh, D. B.; Amaral, M.; Bomke, J.; Grädler, U.; Musil, D.; Buchstaller, H.-P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F., Estimation of drug-target residence times by τ-random acceleration molecular dynamics simulations. *Journal of chemical theory and computation* **2018,** *14* (7), 3859-3869.

66. Bruce, N. J.; Ganotra, G. K.; Kokh, D. B.; Sadiq, S. K.; Wade, R. C., New approaches for computing ligand–receptor binding kinetics. *Current opinion in structural biology* **2018,** *49*, 1-10.

67. Buch, I.; Giorgino, T.; De Fabritiis, G., Complete reconstruction of an enzyme-inhibitor binding process by molecular dynamics simulations. *Proceedings of the National Academy of Sciences* **2011,** *108* (25), 10184-10189.

68. Zuckerman, D. M.; Chong, L. T., Weighted ensemble simulation: review of methodology, applications, and software. *Annual review of biophysics* **2017,** *46*, 43-57.

69. Barducci, A.; Bonomi, M.; Parrinello, M., Metadynamics. *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2011,** *1* (5), 826-843.

70. Vanden-Eijnden, E.; Venturoli, M.; Ciccotti, G.; Elber, R., On the assumptions underlying milestoning. *The Journal of chemical physics* **2008,** *129* (17).

71. Isralewitz, B.; Gao, M.; Schulten, K., Steered molecular dynamics and mechanical functions of proteins. *Current opinion in structural biology* **2001,** *11* (2), 224-230.

72. Kokh, D. B.; Wade, R. C., G Protein-Coupled Receptor-Ligand Dissociation Rates and Mechanisms from tauRAMD Simulations. *J Chem Theory Comput* **2021,** *17* (10), 6610-6623.

73. Nunes-Alves, A.; Kokh, D. B.; Wade, R. C., Ligand unbinding mechanisms and kinetics for T4 lysozyme mutants from tauRAMD simulations. *Curr Res Struct Biol* **2021,** *3*, 106-111.

74. Kokh, D. B.; Doser, B.; Richter, S.; Ormersbach, F.; Cheng, X.; Wade, R. C., A workflow for exploring ligand dissociation from a macromolecule: Efficient random acceleration molecular dynamics simulation and interaction fingerprint analysis of ligand trajectories. *J Chem Phys* **2020,** *153* (12), 125102.

75. Berger, B. T.; Amaral, M.; Kokh, D. B.; Nunes-Alves, A.; Musil, D.; Heinrich, T.; Schroder, M.; Neil, R.; Wang, J.; Navratilova, I.; Bomke, J.; Elkins, J. M.; Muller, S.; Frech, M.; Wade, R. C.; Knapp, S., Structure-kinetic relationship reveals the mechanism of selectivity of FAK inhibitors over PYK2. *Cell Chem Biol* **2021,** *28* (5), 686-698 e7.

76. D'Arrigo, G.; Kokh, D. B.; Nunes-Alves, A.; Wade, R. C., Computational screening of the effects of mutations on protein-protein off-rates and dissociation mechanisms by τRAMD. *Communications Biology* **2024,** *7* (1), 1159.

77. Kokh, D. B.; Amaral, M.; Bomke, J.; Gradler, U.; Musil, D.; Buchstaller, H. P.; Dreyer, M. K.; Frech, M.; Lowinski, M.; Vallee, F.; Bianciotto, M.; Rak, A.; Wade, R. C., Estimation of Drug-Target Residence Times by tau-Random Acceleration Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation* **2018,** *14* (7), 3859-3869.

78. Kokh, D. B.; Wade, R. C., G Protein-Coupled Receptor-Ligand Dissociation Rates and Mechanisms from tau RAMD Simulations. *Journal of Chemical Theory and Computation* **2021,** *17* (10), 6610-6623.

79. Zhang, Q. Q.; Han, J. T.; Zhu, Y. C.; Tan, S. Y.; Liu, H. X., Binding Thermodynamics and Dissociation Kinetics Analysis Uncover the Key Structural Motifs of Phenoxyphenol Derivatives as the Direct InhA Inhibitors and the Hotspot Residues of InhA. *International Journal of Molecular Sciences* **2022,** *23* (17).

80. Nunes-Alves, A.; Kokh, D. B.; Wade, R. C., Ligand unbinding mechanisms and kinetics for T4 lysozyme mutants from tau RAMD simulations. *Current Research in Structural Biology* **2021,** *3*, 106-111.

81. Wang, Q. Q.; Zhang, Q. G.; Leung, E. L. H.; Chen, Y. Q.; Yao, X. J., Exploring the thermodynamic, kinetic and inhibitory mechanisms of 5-iTU targeting mitotic kinase haspin by integrated molecular dynamics. *Physical Chemistry Chemical Physics* **2021,** *23* (34), 18404-18413.

82. Berger, B. T.; Amaral, M.; Kokh, D. B.; Nunes-Alves, A.; Musil, D.; Heinrich, T.; Schroder, M.; Neil, R.; Wang, J.; Navratilova, I.; Bomke, J.; Elkins, J. M.; Muller, S.; Frech,

M.; Wade, R. C.; Knapp, S., Structure-kinetic relationship reveals the mechanism of selectivity of FAK inhibitors over PYK2. *Cell Chemical Biology* **2021,** *28* (5), 686-+.

83. Pietrucci, F., Strategies for the exploration of free energy landscapes: Unity in diversity and challenges ahead. *Reviews in Physics* **2017,** *2*, 32-45.

84. Wales, D. J.; Bogdan, T. V., Potential energy and free energy landscapes. ACS Publications: 2006; Vol. 110, pp 20765-20776.

85. Noé, F.; Clementi, C., Collective variables for the study of long-time kinetics from molecular trajectories: theory and methods. *Current opinion in structural biology* **2017,** *43*, 141-147.

86. Peters, B., Reaction Coordinates and Mechanistic Hypothesis Tests. *Annu Rev Phys Chem* **2016,** *67*, 669-90.

87. Bolhuis, P. G.; Chandler, D.; Dellago, C.; Geissler, P. L., Transition path sampling: Throwing ropes over rough mountain passes, in the dark. *Annual review of physical chemistry* **2002,** *53* (1), 291-318.

88. Best, R. B.; Hummer, G., Reaction coordinates and rates from transition paths. *Proceedings of the National Academy of Sciences* **2005,** *102* (19), 6732-6737.

89. Zwanzig, R. W., High-temperature equation of state by a perturbation method. I. Nonpolar gases. *The Journal of Chemical Physics* **1954,** *22* (8), 1420-1426.

90. Tuckerman, M. E., Free Energy Calculations: Theory and Applications in Chemistry and Biology Springer Series in Chemical Physics, 86 Edited by Christophe Chipot (Université Henri Poincaré Vandoeuvre-lès-Nancy, France) and Andrew Pohorille (University of California, San Francisco, USA). Springer: Berlin, Heidelberg, New York. 2007. xviii+ 518 pp. $199.00. ISBN 978-3-540-38447-2. ACS Publications: 2007.

91. Zara, L.; Efrem, N. L.; van Muijlwijk-Koezen, J. E.; de Esch, I. J. P.; Zarzycka, B., Progress in Free Energy Perturbation: Options for Evolving Fragments. *Drug Discov Today Technol* **2021,** *40*, 36-42.

92. Palma, P. N.; Bonifacio, M. J.; Loureiro, A. I.; Soares-da-Silva, P., Computation of the binding affinities of catechol-O-methyltransferase inhibitors: multisubstate relative free energy calculations. *J Comput Chem* **2012,** *33* (9), 970-86.

93. Aldeghi, M.; Heifetz, A.; Bodkin, M. J.; Knapp, S.; Biggin, P. C., Accurate calculation of the absolute free energy of binding for drug molecules. *Chem Sci* **2016,** *7* (1), 207-218.

94. Zhu, F.; Bourguet, F. A.; Bennett, W. F. D.; Lau, E. Y.; Arrildt, K. T.; Segelke, B. W.; Zemla, A. T.; Desautels, T. A.; Faissol, D. M., Large-scale application of free energy perturbation calculations for antibody design. *Sci Rep* **2022,** *12* (1), 12489.

95. Steinbrecher, T.; Zhu, C.; Wang, L.; Abel, R.; Negron, C.; Pearlman, D.; Feyfant, E.; Duan, J.; Sherman, W., Predicting the Effect of Amino Acid Single-Point Mutations on Protein Stability-Large-Scale Validation of MD-Based Relative Free Energy Calculations. *J Mol Biol* **2017,** *429* (7), 948-963.

96. Scarabelli, G.; Oloo, E. O.; Maier, J. K.; Rodriguez-Granillo, A., Accurate prediction of protein thermodynamic stability changes upon residue mutation using free energy perturbation. *Journal of Molecular Biology* **2022,** *434* (2), 167375.

97. Case, D. A.; Aktulga, H. M.; Belfon, K.; Ben-Shalom, I. Y.; Berryman, J. T.; Brozell, S. R.; Cerutti, D. S.; Cheatham III, T. E.; Cisneros, G. A.; Cruzeiro, V. W. D., *Amber 2023*. University of California, San Francisco: 2023.

98. Kulshrestha, A.; Punnathanam, S. N.; Ayappa, K. G., Finite temperature string method with umbrella sampling using path collective variables: application to secondary structure change in a protein. *Soft Matter* **2022,** *18* (39), 7593-7603.

99. Chen, H.; Ogden, D.; Pant, S.; Cai, W.; Tajkhorshid, E.; Moradi, M.; Roux, B.; Chipot, C., A Companion Guide to the String Method with Swarms of Trajectories: Characterization, Performance, and Pitfalls. *J Chem Theory Comput* **2022,** *18* (3), 1406-1422.

100. Hasyim, M. R.; Batton, C. H.; Mandadapu, K. K., Supervised learning and the finite-temperature string method for computing committor functions and reaction rates. *The Journal of chemical physics* **2022,** *157* (18).

101. Song, H. D.; Zhu, F., Finite Temperature String Method with Umbrella Sampling: Application on a Side Chain Flipping in Mhp1 Transporter. *J Phys Chem B* **2017,** *121* (15), 3376-3386.

102. Rosta, E.; Hummer, G., Free Energies from Dynamic Weighted Histogram Analysis Using Unbiased Markov State Model. *Journal of Chemical Theory and Computation* **2015,** *11* (1), 276-285.

103. Giese, T. J.; Ekesan, S. o. l.; York, D. M., Extension of the variational free energy profile and multistate Bennett acceptance ratio methods for high-dimensional potential of mean force profile analysis. *The Journal of Physical Chemistry A* **2021,** *125* (19), 4216-4232.

104. Zuckerman, D. M.; Chong, L. T., Weighted Ensemble Simulation: Review of Methodology, Applications, and Software. *Annu Rev Biophys* **2017,** *46*, 43-57.

105. Latorraca, N. R.; Venkatakrishnan, A.; Dror, R. O., GPCR dynamics: structures in motion. *Chemical reviews* **2017,** *117* (1), 139-155.

106. Hilger, D.; Masureel, M.; Kobilka, B. K., Structure and dynamics of GPCR signaling complexes. *Nature structural & molecular biology* **2018,** *25* (1), 4-12.

107. Congreve, M.; De Graaf, C.; Swain, N. A.; Tate, C. G., Impact of GPCR structures on drug discovery. *Cell* **2020,** *181* (1), 81-91.

108. Proft, J.; Weiss, N., G protein regulation of neuronal calcium channels: back to the future. *Molecular pharmacology* **2015,** *87* (6), 890-906.

109. Castellano, M.; Böhm, M., The cardiac β-adrenoceptor–mediated signaling pathway and its alterations in hypertensive heart disease. *Hypertension* **1997,** *29* (3), 715-722.

110. Najafi, A.; Sequeira, V.; Kuster, D. W.; van der Velden, J., β-adrenergic receptor signalling and its functional consequences in the diseased heart. *European journal of clinical investigation* **2016,** *46* (4), 362-374.

111. Molinoff, P. B., α-and β-adrenergic receptor subtypes: properties, distribution and regulation. *Drugs* **1984,** *28* (Suppl 2), 1-15.

112. Brodde, O.-E.; Bruck, H.; Leineweber, K.; Seyfarth, T., Presence, distribution and physiological function of adrenergic and muscarinic receptor subtypes in the human heart. *Basic research in cardiology* **2001,** *96* (6), 528-538.

113. Frishman, W. H., β-Adrenergic blockers: a 50-year historical perspective. *American Journal of Therapeutics* **2008,** *15* (6), 565-576.

114. Salpeter, S. R.; Ormiston, T. M.; Salpeter, E. E., Cardioselective beta-blockers for chronic obstructive pulmonary disease. *Cochrane database of systematic reviews* **2005,** (4).

115. Andrus, M. R.; Holloway, K. P.; Clark, D. B., Use of β-blockers in patients with COPD. *Annals of Pharmacotherapy* **2004,** *38* (1), 142-145.

116. Baker, J. G.; Wilcox, R. G., β-Blockers, heart disease and COPD: current controversies and uncertainties. *Thorax* **2017,** *72* (3), 271-276.

117. Bhatt, S. P.; Wells, J. M.; Kinney, G. L.; Washko, G. R.; Budoff, M.; Kim, Y.-i.; Bailey, W. C.; Nath, H.; Hokanson, J. E.; Silverman, E. K., β-Blockers are associated with a reduction in COPD exacerbations. *Thorax* **2016,** *71* (1), 8-14.

118. do Vale, G. T.; Ceron, C. S.; Gonzaga, N. A.; Simplicio, J. A.; Padovan, J. C., Three generations of β-blockers: history, class differences and clinical applicability. *Current hypertension reviews* **2019,** *15* (1), 22-31.

119. Cherezov, V.; Rosenbaum, D. M.; Hanson, M. A.; Rasmussen, S. G.; Thian, F. S.; Kobilka, T. S.; Choi, H.-J.; Kuhn, P.; Weis, W. I.; Kobilka, B. K., High-resolution crystal structure of an engineered human β2-adrenergic G protein–coupled receptor. *science* **2007,** *318* (5854), 1258-1265.

120. Warne, T.; Moukhametzianov, R.; Baker, J. G.; Nehmé, R.; Edwards, P. C.; Leslie, A. G.; Schertler, G. F.; Tate, C. G., The structural basis for agonist and partial agonist action on a β1-adrenergic receptor. *Nature* **2011,** *469* (7329), 241-244.

121. Dixon, R.; Sigal, I.; Candelore, M.; Register, R.; Scattergood, W.; Rands, E.; Strader, C., Structural features required for ligand binding to the beta-adrenergic receptor. *The EMBO journal* **1987,** *6* (11), 3269-3275.

122. Rasmussen, S. G.; Choi, H.-J.; Fung, J. J.; Pardon, E.; Casarosa, P.; Chae, P. S.; DeVree, B. T.; Rosenbaum, D. M.; Thian, F. S.; Kobilka, T. S., Structure of a nanobody-stabilized active state of the β2 adrenoceptor. *Nature* **2011,** *469* (7329), 175-180.

123. Rasmussen, S. G.; DeVree, B. T.; Zou, Y.; Kruse, A. C.; Chung, K. Y.; Kobilka, T. S.; Thian, F. S.; Chae, P. S.; Pardon, E.; Calinski, D., Crystal structure of the β2 adrenergic receptor–Gs protein complex. *Nature* **2011,** *477* (7366), 549-555.

124. Ishchenko, A.; Stauch, B.; Han, G. W.; Batyuk, A.; Shiriaeva, A.; Li, C.; Zatsepin, N.; Weierstall, U.; Liu, W.; Nango, E., Toward G protein-coupled receptor structure-based drug design using X-ray lasers. *IUCrJ* **2019,** *6* (6), 1106-1119.

125. Warne, T.; Serrano-Vega, M. J.; Baker, J. G.; Moukhametzianov, R.; Edwards, P. C.; Henderson, R.; Leslie, A. G.; Tate, C. G.; Schertler, G. F., Structure of a β1-adrenergic G-protein-coupled receptor. *Nature* **2008,** *454* (7203), 486-491.

126. Smith, C.; Teitler, M., Beta-blocker selectivity at cloned human beta1-and beta2-adrenergic receptors. *Cardiovascular drugs and therapy* **1999,** *13* (2), 123-126.

127. do Vale, G. T.; Ceron, C. S.; Gonzaga, N. A.; Simplicio, J. A.; Padovan, J. C., Three Generations of beta-blockers: History, Class Differences and Clinical Applicability. *Curr Hypertens Rev* **2019,** *15* (1), 22-31.

128. Emtage, A. L.; Mistry, S. N.; Fischer, P. M.; Kellam, B.; Laughton, C. A., GPCRs through the keyhole: the role of protein flexibility in ligand binding to β-adrenoceptors. *Journal of Biomolecular Structure and Dynamics* **2017,** *35* (12), 2604-2619.

129. Ian M. Withers, M. P. M., Hao Wang, Peter M. Fischer, and Charles A. Laughton, Active Site Pressurization: A New Tool for Structure-Guided Drug Design and Other Studies of Protein Flexibility. *J. Chem. Inf. Model 48*, 1448-1454.

130. Mazanetz, M. P.; Laughton, C. A.; Fischer, P. M., Investigation of the flexibility of protein kinases implicated in the pathology of Alzheimer's disease. *Molecules* **2014,** *19* (7), 9134-59.

131. Emtage, A. L.; Mistry, S. N.; Fischer, P. M.; Kellam, B.; Laughton, C. A., GPCRs through the keyhole: the role of protein flexibility in ligand binding to beta-adrenoceptors. *J Biomol Struct Dyn* **2017,** *35* (12), 2604-2619.

132. Sykes, D. A.; Parry, C.; Reilly, J.; Wright, P.; Fairhurst, R. A.; Charlton, S. J., Observed drug-receptor association rates are governed by membrane affinity: the importance of establishing "micro-pharmacokinetic/pharmacodynamic relationships" at the beta2-adrenoceptor. *Mol Pharmacol* **2014,** *85* (4), 608-17.

133. Rhee, S. G.; Woo, H. A.; Kil, I. S.; Bae, S. H., Peroxiredoxin functions as a peroxidase and a regulator and sensor of local peroxides. *J Biol Chem* **2012,** *287* (7), 4403-10.

134. Bolduc, J.; Koruza, K.; Luo, T.; Malo Pueyo, J.; Vo, T. N.; Ezerina, D.; Messens, J., Peroxiredoxins wear many hats: Factors that fashion their peroxide sensing personalities. *Redox Biol* **2021,** *42*, 101959.

135. Rhee, S. G., Overview on Peroxiredoxin. *Mol Cells* **2016,** *39* (1), 1-5.

136. Wood, Z. A.; Schroder, E.; Robin Harris, J.; Poole, L. B., Structure, mechanism and regulation of peroxiredoxins. *Trends Biochem Sci* **2003,** *28* (1), 32-40.

137. Gomes, F.; Turano, H.; Haddad, L. A.; Netto, L. E., Human mitochondrial peroxiredoxin Prdx3 is dually localized in the intermembrane space and matrix subcompartments. *Redox Biology* **2024,** *78*, 103436.

138. Dip, P. V.; Kamariah, N.; Nartey, W.; Beushausen, C.; Kostyuchenko, V. A.; Ng, T.-S.; Lok, S.-M.; Saw, W. G.; Eisenhaber, F.; Eisenhaber, B., Key roles of the Escherichia coli AhpC C-terminus in assembly and catalysis of alkylhydroperoxide reductase, an enzyme essential for the alleviation of oxidative stress. *Biochimica et Biophysica Acta (BBA)-Bioenergetics* **2014,** *1837* (12), 1932-1943.

139. Perkins, A.; Nelson, K. J.; Parsonage, D.; Poole, L. B.; Karplus, P. A., Peroxiredoxins: guardians against oxidative stress and modulators of peroxide signaling. *Trends Biochem Sci* **2015,** *40* (8), 435-45.

140. Stancill, J. S.; Corbett, J. A., The role of thioredoxin/peroxiredoxin in the β-cell defense against oxidative damage. *Frontiers in endocrinology* **2021,** *12*, 718235.

141. Rhee, S. G.; Jeong, W.; Chang, T. S.; Woo, H. A., Sulfiredoxin, the cysteine sulfinic acid reductase specific to 2-Cys peroxiredoxin: its discovery, mechanism of action, and biological significance. *Kidney Int Suppl* **2007,** (106), S3-8.

142. Wood, Z. A.; Schröder, E.; Harris, J. R.; Poole, L. B., Structure, mechanism and regulation of peroxiredoxins. *Trends in biochemical sciences* **2003,** *28* (1), 32-40.

143. Jang, H. H.; Lee, K. O.; Chi, Y. H.; Jung, B. G.; Park, S. K.; Park, J. H.; Lee, J. R.; Lee, S. S.; Moon, J. C.; Yun, J. W., Two enzymes in one: two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* **2004,** *117* (5), 625-635.

144. Moon, J. C.; Hah, Y.-S.; Kim, W. Y.; Jung, B. G.; Jang, H. H.; Lee, J. R.; Kim, S. Y.; Lee, Y. M.; Jeon, M. G.; Kim, C. W., Oxidative stress-dependent structural and functional switching of a human 2-Cys peroxiredoxin isotype II that enhances HeLa cell resistance to H2O2-induced cell death. *Journal of Biological Chemistry* **2005,** *280* (31), 28775-28784.

145. Schroder, E.; Littlechild, J. A.; Lebedev, A. A.; Errington, N.; Vagin, A. A.; Isupov, M. N., Crystal structure of decameric 2-Cys peroxiredoxin from human erythrocytes at 1.7 A resolution. *Structure* **2000,** *8* (6), 605-15.

146. Kamariah, N.; Sek, M. F.; Eisenhaber, B.; Eisenhaber, F.; Gruber, G., Transition steps in peroxide reduction and a molecular switch for peroxide robustness of prokaryotic peroxiredoxins. *Sci Rep* **2016,** *6*, 37610.

147. Bolduc, J. A.; Nelson, K. J.; Haynes, A. C.; Lee, J.; Reisz, J. A.; Graff, A. H.; Clodfelter, J. E.; Parsonage, D.; Poole, L. B.; Furdui, C. M.; Lowther, W. T., Novel hyperoxidation resistance motifs in 2-Cys peroxiredoxins. *J Biol Chem* **2018,** *293* (30), 11901-11912.

148. Wood, Z. A.; Poole, L. B.; Hantgan, R. R.; Karplus, P. A., Dimers to doughnuts: redox-sensitive oligomerization of 2-cysteine peroxiredoxins. *Biochemistry* **2002,** *41* (17), 5493-504.

149. Finelli, M. J., Redox Post-translational Modifications of Protein Thiols in Brain Aging and Neurodegenerative Conditions-Focus on S-Nitrosation. *Front Aging Neurosci* **2020,** *12*, 254.

150. Hall, A.; Parsonage, D.; Poole, L. B.; Karplus, P. A., Structural evidence that peroxiredoxin catalytic power is based on transition-state stabilization. *J Mol Biol* **2010,** *402* (1), 194-209.

151. Perkins, A.; Parsonage, D.; Nelson, K. J.; Ogba, O. M.; Cheong, P. H.-Y.; Poole, L. B.; Karplus, P. A., Peroxiredoxin catalysis at atomic resolution. *Structure* **2016,** *24* (10), 1668-1678.

152. Jonsson, T. J.; Johnson, L. C.; Lowther, W. T., Structure of the sulphiredoxin-peroxiredoxin complex reveals an essential repair embrace. *Nature* **2008,** *451* (7174), 98-101.

153. Teixeira, F.; Tse, E.; Castro, H.; Makepeace, K. A.; Meinen, B. A.; Borchers, C. H.; Poole, L. B.; Bardwell, J. C.; Tomás, A. M.; Southworth, D. R., Chaperone activation and client binding of a 2-cysteine peroxiredoxin. *Nature communications* **2019,** *10* (1), 659.

154. Rhee, S. G.; Woo, H. A., Multiple functions of peroxiredoxins: peroxidases, sensors and regulators of the intracellular messenger H2O2, and protein chaperones. *Antioxidants & redox signaling* **2011,** *15* (3), 781-794.

155. Jang, H. H.; Lee, K. O.; Chi, Y. H.; Jung, B. G.; Park, S. K.; Park, J. H.; Lee, J. R.; Lee, S. S.; Moon, J. C.; Yun, J. W.; Choi, Y. O.; Kim, W. Y.; Kang, J. S.; Cheong, G. W.; Yun, D. J.; Rhee, S. G.; Cho, M. J.; Lee, S. Y., Two enzymes in one; two yeast peroxiredoxins display oxidative stress-dependent switching from a peroxidase to a molecular chaperone function. *Cell* **2004,** *117* (5), 625-35.

156. Fang, J.; Nakamura, T.; Cho, D.-H.; Gu, Z.; Lipton, S. A., S-nitrosylation of peroxiredoxin 2 promotes oxidative stress-induced neuronal cell death in Parkinson's disease. *Proceedings of the National Academy of Sciences* **2007,** *104* (47), 18742-18747.

157. Engelman, R.; Weisman-Shomer, P.; Ziv, T.; Xu, J.; Arner, E. S.; Benhar, M., Multilevel regulation of 2-Cys peroxiredoxin reaction cycle by S-nitrosylation. *J Biol Chem* **2013,** *288* (16), 11312-24.

158. Phaniendra, A.; Jestadi, D. B.; Periyasamy, L., Free radicals: properties, sources, targets, and their implication in various diseases. *Indian J Clin Biochem* **2015,** *30* (1), 11-26.

159. Lushchak, V. I., Free radicals, reactive oxygen species, oxidative stress and its classification. *Chemico-biological interactions* **2014,** *224*, 164-175.

160. Halliwell, B.; Gutteridge, J. M., *Free radicals in biology and medicine*. Oxford university press: 2015.

161. Imlay, J. A., Pathways of oxidative damage. *Annual Reviews in Microbiology* **2003,** *57* (1), 395-418.

162. Bedard, K.; Krause, K.-H., The NOX family of ROS-generating NADPH oxidases: physiology and pathophysiology. *Physiological reviews* **2007,** *87* (1), 245-313.

163. Förstermann, U.; Sessa, W. C., Nitric oxide synthases: regulation and function. *European heart journal* **2012,** *33* (7), 829-837.

164. Sies, H., Hydrogen peroxide as a central redox signaling molecule in physiological oxidative stress: Oxidative eustress. *Redox biology* **2017,** *11*, 613-619.

165. Hussain, S. P.; Hofseth, L. J.; Harris, C. C., Radical causes of cancer. *Nat Rev Cancer* **2003,** *3* (4), 276-85.

166. Shah, A. M.; Channon, K. M., Free radicals and redox signalling in cardiovascular disease. *Heart* **2004,** *90* (5), 486-7.

167. Raza, C.; Anjum, R.; Shakeel, N. U. A., Parkinson's disease: Mechanisms, translational models and management strategies. *Life Sci* **2019,** *226*, 77-90.

168. Chinta, S. J.; Andersen, J. K., Dopaminergic neurons. *Int J Biochem Cell Biol* **2005,** *37* (5), 942-6.

169. Sonnen, J. A.; Breitner, J. C.; Lovell, M. A.; Markesbery, W. R.; Quinn, J. F.; Montine, T. J., Free radical-mediated damage to brain in Alzheimer's disease and its transgenic mouse models. *Free Radic Biol Med* **2008,** *45* (3), 219-30.

170. Lopez-Otin, C.; Blasco, M. A.; Partridge, L.; Serrano, M.; Kroemer, G., The hallmarks of aging. *Cell* **2013,** *153* (6), 1194-217.

171. Hardy, J., Genetic analysis of pathways to Parkinson disease. *Neuron* **2010,** *68* (2), 201-206.

172. Michel, P. P.; Hirsch, E. C.; Hunot, S., Understanding dopaminergic cell death pathways in Parkinson disease. *Neuron* **2016,** *90* (4), 675-691.

173. Miller, R. L.; James-Kracke, M.; Sun, G. Y.; Sun, A. Y., Oxidative and inflammatory pathways in Parkinson's disease. *Neurochemical research* **2009,** *34* (1), 55-65.

174. Papasergi-Scott, M. M.; Pérez-Hernández, G.; Batebi, H.; Gao, Y.; Eskici, G.; Seven, A. B.; Panova, O.; Hilger, D.; Casiraghi, M.; He, F., Time-resolved cryo-EM of G-protein activation by a GPCR. *Nature* **2024,** *629* (8014), 1182-1191.

175. Xu, J.; Chen, G.; Wang, H.; Cao, S.; Heng, J.; Deupi, X.; Du, Y.; Kobilka, B. K., Calcineurin-fusion facilitates cryo-EM structure determination of a Family A GPCR. *Proceedings of the National Academy of Sciences* **2024,** *121* (48), e2414544121.

176. Yang, F.; Ling, S.; Zhou, Y.; Zhang, Y.; Lv, P.; Liu, S.; Fang, W.; Sun, W.; Hu, L. A.; Zhang, L., Different conformational responses of the β2-adrenergic receptor-Gs complex upon binding of the partial agonist salbutamol or the full agonist isoprenaline. *National Science Review* **2021,** *8* (9), nwaa284.

177. Abraham, M. J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J. C.; Hess, B.; Lindahl, E., GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **2015,** *1*, 19-25.

178. McGibbon, R. T.; Beauchamp, K. A.; Harrigan, M. P.; Klein, C.; Swails, J. M.; Hernández, C. X.; Schwantes, C. R.; Wang, L.-P.; Lane, T. J.; Pande, V. S., MDTraj: a modern open library for the analysis of molecular dynamics trajectories. *Biophysical journal* **2015,** *109* (8), 1528-1532.

179. Warburton, D. E.; Jamnik, V. K.; Bredin, S. S.; Gledhill, N., The physical activity readiness questionnaire for everyone (PAR-Q+) and electronic physical activity readiness medical examination (ePARmed-X+). *The Health & Fitness Journal of Canada* **2011,** *4* (2), 3-17.

180. Ryckaert, J.-P.; Ciccotti, G.; Berendsen, H. J., Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. *Journal of computational physics* **1977,** *23* (3), 327-341.

181. Pronk, S.; Páll, S.; Schulz, R.; Larsson, P.; Bjelkmar, P.; Apostolov, R.; Shirts, M. R.; Smith, J. C.; Kasson, P. M.; Van Der Spoel, D., GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **2013,** *29* (7), 845-854.

182. Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J., SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods* **2020,** *17* (3), 261-272.

183. Peskin, A. V.; Meotti, F. C.; Kean, K. M.; Gobl, C.; Peixoto, A. S.; Pace, P. E.; Horne, C. R.; Heath, S. G.; Crowther, J. M.; Dobson, R. C. J.; Karplus, P. A.; Winterbourn, C. C., Modifying the resolving cysteine affects the structure and hydrogen peroxide reactivity of peroxiredoxin 2. *J Biol Chem* **2021,** *296*, 100494.

184. Jo, S.; Cheng, X.; Lee, J.; Kim, S.; Park, S. J.; Patel, D. S.; Beaven, A. H.; Lee, K. I.; Rui, H.; Park, S., CHARMM-GUI 10 years for biomolecular modeling and simulation. *Journal of computational chemistry* **2017,** *38* (15), 1114-1124.

185. Sgro, J.-Y. In *UCSF ChimeraX-I-Overview*, Workshop goal, 2018.

186. Maestro, S., Maestro. *Schrödinger, LLC, New York, NY* **2020,** *2020*, 682.

187. Case, D.; Babin, V.; Berryman, J.; Betz, R.; Cai, Q.; Cerutti, D.; Cheatham III, T.; Darden, T., RE, Duke, H. *Gohlke, AW Goetz, S. Gusarov, N. Homeyer, P. Janowski, J. Kaus, I. Kolossváry, AK* **2014**.

188. O'Boyle, N. M.; Banck, M.; James, C. A.; Morley, C.; Vandermeersch, T.; Hutchison, G. R., Open Babel: An open chemical toolbox. *Journal of cheminformatics* **2011,** *3* (1), 33.

189. Wang, J.; Wang, W.; Kollman, P. A.; Case, D. A., Antechamber: an accessory software package for molecular mechanical calculations. *J. Am. Chem. Soc* **2001,** *222* (1), 2001.

190. Kim, S.; Chen, J.; Cheng, T.; Gindulyte, A.; He, J.; He, S.; Li, Q.; Shoemaker, B. A.; Thiessen, P. A.; Yu, B., PubChem 2019 update: improved access to chemical data. *Nucleic acids research* **2019,** *47* (D1), D1102-D1109.

191. Shkurti, A.; Goni, R.; Andrio, P.; Breitmoser, E.; Bethune, I.; Orozco, M.; Laughton, C. A., pyPcazip: A PCA-based toolkit for compression and analysis of molecular simulation data. *SoftwareX* **2016,** *5*, 44-50.

192. Hedges, L. O.; Mey, A. S.; Laughton, C.; Gervasio, F. L.; Mulholland, A. J.; Woods, C. J.; Michel, J., BioSimSpace: An interoperable Python framework for biomolecular simulation. *Journal of Open Source Software* **2019,** *4*, 1831.

193. Jorge, M.; Garrido, N. M.; Queimada, A. J.; Economou, I. G.; Macedo, E. A., Effect of the integration method on the accuracy and computational efficiency of free energy calculations using thermodynamic integration. *Journal of Chemical Theory and Computation* **2010,** *6* (4), 1018-1027.

194. Peters, B., Reaction coordinates and mechanistic hypothesis tests. *Annual review of physical chemistry* **2016,** *67* (1), 669-690.

195. Yi, J.; Khan, M. A.; Lee, J.; Richter-Addo, G. B., The solid-state molecular structure of the S-nitroso derivative of L-cysteine ethyl ester hydrochloride. *Nitric Oxide* **2005,** *12* (4), 261-6.

196. Shapovalov, M. V.; Dunbrack, R. L., A smoothed backbone-dependent rotamer library for proteins derived from adaptive kernel density estimates and regressions. *Structure* **2011,** *19* (6), 844-858.

197. Ikeguchi, M.; Ueno, J.; Sato, M.; Kidera, A., Protein structural change upon ligand binding: linear response theory. *Physical review letters* **2005,** *94* (7), 078102.

198. Unciuleac, M.-C.; Goldgur, Y.; Shuman, S., Caveat mutator: alanine substitutions for conserved amino acids in RNA ligase elicit unexpected rearrangements of the active site for lysine adenylylation. *Nucleic Acids Research* **2020,** *48* (10), 5603-5615.

199. Gusarov, I.; Nudler, E., Protein S-Nitrosylation: Enzymatically Controlled, but Intrinsically Unstable, Post-translational Modification. *Mol Cell* **2018,** *69* (3), 351-353.

200. Hess, D. T.; Matsumoto, A.; Kim, S. O.; Marshall, H. E.; Stamler, J. S., Protein S-nitrosylation: purview and parameters. *Nat Rev Mol Cell Biol* **2005,** *6* (2), 150-66.

201. Hess, D. T.; Matsumoto, A.; Nudelman, R.; Stamler, J. S., S-nitrosylation: spectrum and specificity. *Nature Cell Biology* **2001,** *3* (2), E46-E48.