

Understanding the mechanism of NO/CO regulation of human Cystathionine β -Synthase: A Molecular Dynamics Simulations study

Jiangli Gui

Promotor: Prof. Jeremy Harvey

Department: Department of Chemistry

Division: Quantum Chemistry and Physical Chemistry

Dissertation presented in fulfillment of the requirements for the degree of Master of Science:
Bioinformatics

Tuesday 8th August, 2023

Declaration

This dissertation is part of the examination and has not been corrected after defense for eventual errors.
Use as a reference is permitted subject to written approval of the promotor stated on the front page.

Jiangli Gui, Tuesday 8th August, 2023

Acknowledgements

First of all, I would like to express my heartfelt gratitude to professor Harvey, who supported me when I was feeling low last summer. I can still vividly remember his encouraging words: 'I am optimistic about you,' which truly lifted my spirits. Apart from providing invaluable mental support, Prof. Jeremy offered me priceless help throughout my thesis project. Whenever I encountered obstacles, he patiently shared his insights and knowledge, guiding me towards new solutions. Additionally, Prof. Jeremy's sense of humor brightened our interactions, and I made sure to jot down some of his jokes and stories.

I extend my sincere thanks to Neil, of course, my daily supervisor. He exemplifies patience and kindness, always addressing my inquiries attentively, even when I had asked the same questions multiple times before. His unwavering support proved indispensable whenever I faced challenges. Frankly, I would not have been able to complete this project without his guidance and encouragement.

I am also deeply grateful to the other members of Prof. Jeremy's group. Our weekly group meetings were not only enjoyable but also provided me with the opportunity to expand my horizons in various computational chemistry areas. Moreover, I learned valuable presentation skills and plotting tricks, all while relishing the delightful jokes and stories shared by Prof. Jeremy.

I am indebted to the two GPU nodes in our group. Their dedicated work for three months on the calculations played a crucial role in my ability to conduct molecular dynamics simulations for a 130,000 atoms system, a task I could not have accomplished by hand.

I am deeply grateful to Agora and the Law and Criminology library, where I dedicated the majority of my study hours in Leuven.

My appreciation extends to all my friends, both in Belgium and China. Our shared experiences of dining, chatting, traveling, and playing video games have added vibrant hues to my life.

Lastly, I want to express my profound gratitude to my parents for their unwavering support. I regret that my grandmother did not witness my graduation; I wish I could have completed my studies a bit earlier to grant her that opportunity. May she rest in peace and find happiness in the heavens. To myself, I acknowledge the strength and determination that carried me through to the finish line.

Abstract

Human Cystathionine beta synthase (hCBS) plays a critical role in the reverse transsulfuration pathway in the body, and its dysfunction has been associated with various diseases, notably Hyperhomocysteinaemia. Previous studies have shown that Nitric Oxide (NO) and carbon monoxide (CO) can deactivate CBS activity, yet the underlying mechanism remains elusive. In this study, we explored the inhibition mechanism of NO/CO on hCBS using a combination of multiple sequence alignment (MSA) and classical molecular dynamics (MD) simulations.

MSA analysis revealed that residues implicated in the hypothesized regulation mechanism were highly conserved, supporting their potential importance in CBS regulation. Subsequently, MD simulations provided deeper insights into the process. Upon the disruption of the salt bridge between Cys52-SG and Arg266-NH1, Arg266 underwent rotational motion, propagating to the end of α helix 8 (Thr260). This helix is part of a phosphate-binding loop crucially involved in strong hydrogen bonding with pyridoxal 5'-phosphate (PLP), the key catalytic region of CBS. The rotation of PLP led to the breakage of two hydrogen bonds, namely, Asn149-ND2 to PLP119-O and Ser349-OG to PLP119-N1. As a result, PLP119-O became free to receive an internal hydrogen from PLP119-NZ, transitioning into the enolimine tautomer (deactive form) instead of the active ketoenamine tautomer.

To the best of our knowledge, this study is the first to elucidate the inhibition mechanism of NO/CO on hCBS using MD simulations, providing valuable insights into the elusive motion of α helix 8. However, further replicates and sample enhancement methods are needed to increase the efficiency of MD simulations. Additionally, the function of decoordination of His65 from Heme, as uncovered by our current findings, remains unknown and warrants further investigation.

Keywords: Cystathionine beta synthase, Molecular dynamics simulations, Inhibition mechanism, Nitric Oxide, Carbon monoxide

Contents

1	Introduction	1
1.1	Physiological role of human cystathionine beta-synthase	1
1.2	5 types of Cystathionine beta-synthase	3
1.3	Structure of human cystathionine beta-synthase	4
1.4	NO/CO heme binding regulates hCBS activity	11
1.5	Previous Molecular dynamics(MD) simulations study on hCBS	11
1.6	Hypothesis	12
1.7	Aims and Objectives	13
2	Theoretical background	15
2.1	Molecular dynamics simulations	15
2.2	Analysis Theory	19
3	Computational methods	21
3.1	Multiple sequence alignment	21
3.2	Molecular dynamic simulations system preparation	21
3.3	Trajectory analysis	22
4	Results and discussion	24
4.1	Evolution conservation study	24
4.2	Trajectory analysis	25
4.3	Dynamical cross correlation analysis	43
5	Conclusion	47
6	Appendix	49

6.1	Abbreviation	49
6.2	Supplementary Figures	49

Chapter 1

Introduction

Cystathionine beta-synthase (CBS) plays a crucial role in the metabolism of methionine, sulfur, and cellular redox balance across all kingdoms. It is involved in the transsulfuration pathway (it is also called the reverse transsulfuration pathway in mammals), where CBS catalyzes the condensation of homocysteine and serine to form cystathionine, which is further metabolized into cysteine. Deficiencies in CBS activity disrupt the reverse transsulfuration pathway metabolism, leading to elevated levels of homocysteine in the blood and urine, a condition known as hyperhomocysteinemia, which can cause various complications.

1.1 Physiological role of human cystathionine beta-synthase

1.1.1 Reverse transsulfuration pathway

The reverse transsulfuration pathway plays a pivotal role in various fundamental physiological processes in mammalian animals. Dysfunctions within this pathway have been implicated in several health conditions. The pathway initiates with the production of homocysteine (Hcy), a byproduct of the methionine cycle. Hcy, along with serine or cysteine, acts as a substrate and undergoes catalysis by the enzyme Cystathionine-beta-synthase (CBS). The resulting product, cystathionine, is further metabolized by another enzyme, Cystathionase (CSE), to produce hydrogen sulfide (H_2S) (in the case of cysteine catalysis by CBS) and cysteine. The generated cysteine then participates in sulfur catabolism, ultimately leading to the generation of H_2S (shown in Figure 1.1).

The reverse transsulfuration pathway is intricately interconnected with multiple metabolic pathways and has been implicated in various diseases. Extensive evidence from experiments conducted on both mice and humans has demonstrated its involvement in conditions such as hyperhomocysteinemia, homocystinuria, cardiovascular diseases, central nervous system disorders, liver dysfunction, and more.

1.1.2 Hyperhomocysteinemia

The most prominent disease observed in these mice is hyperhomocysteinemia, a condition that results from the high retention of homocysteine in the plasma. The normal concentration of homocysteine in serum ranges from 5 to 15 μM , while deficient mice have been observed to have 150 μM of homocysteine [2]. The extreme effects of CBS gene loss lead to neonatal lethality, which is not solely due to high homocysteine levels, suggesting that CBS proteins may play additional roles[3]. Hyperhomocysteinemia can result in various complications, such as oxidative damage in the liver (lipid peroxidation, carbonyl residues in proteins)[4], facial alopecia[5], osteoporosis[5], and loss of fat mass[6]. The correlations between CBS deficiency and dysfunctions were not entirely understood and may not be solely attributed to high homocysteine levels. The other critical product in the reverse transsulfuration pathway is H_2S , which

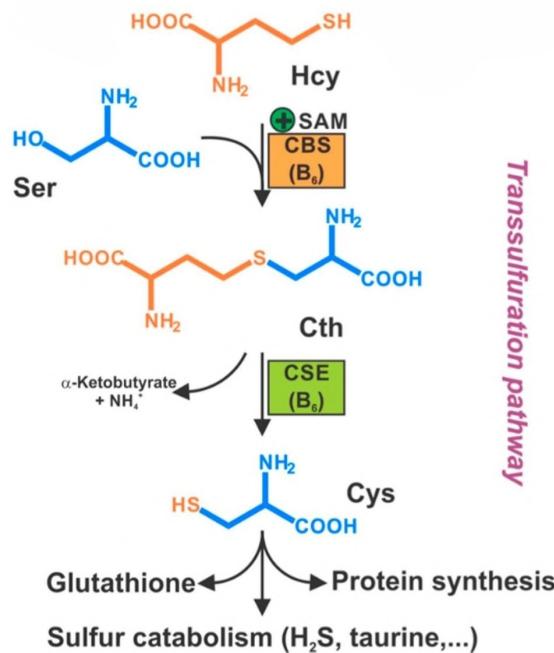


Figure 1.1: Reverse transsulfuration pathway figure modified from [1]

may have interaction effects with homocysteine. For example, it was suggested that both homocysteine and CBS-derived H₂S modulate adipogenesis and adipocyte maturation[7].

1.1.3 Homocystinuria

Homocystinuria is an additional disease that arises from the excessive accumulation of homocysteine, but it is specifically attributed to genetic disorders caused by mutations in the CBS gene. Homocystinuria manifests with various complications, such as abnormal bone development characterized by elongation of long bones or loss of bone density and deterioration of bone tissue[8]. Other notable complications include misalignment of the lens within the eye, thromboembolism (formation of blood clots that obstruct blood vessels), and intellectual disability [8].

1.1.4 Cardiovascular disease

In addition to hyperhomocysteinemia and homocystinuria, elevated concentrations of homocysteine are widely recognized as a significant risk factor for cardiovascular disease. Experimental studies using CBS-deficient mice have provided compelling evidence that the absence of CBS activity can give rise to detrimental effects on cardiovascular health: Specifically, these mice exhibit endothelial dysfunction [9], cerebral vessel hypertrophy [10], and accelerated atherosclerosis [11]. Notably, a similar risk of cardiovascular dysfunction arises from reduced biogenerated H₂S levels. Remarkably, compensatory experiments involving the administration of supplemental H₂S have demonstrated a significant attenuation of cardiovascular alterations in CBS-deficient mice [12, 13]. Given that CBS is one of the three principal mammalian enzymes responsible for H₂S production and was found highly correlated with the production of H₂S in gut[14], liver[15], and plasma[16], the CBS-dependent H₂S pathway serves as an essential endogenous vasodilator that regulates blood pressure[17], and angiogenesis[18] in mammals.

1.1.5 Central nervous system

It has been observed that H₂S generated from CBS in brain homogenates plays a crucial role in various physiological functions, including cognition, memory, regulation of heart and lung activities, and neu-

protection against damage and degeneration [19, 20, 21]. However, it is important to note that these conclusions were primarily drawn from studies involving chemically generated H₂S.

1.1.6 Liver dysfunction

The liver is another organ with high CBS expression and it is seemingly that the dysfunction of the liver is due to the combination of accumulation of hcy and the low production of H₂S, both of which were proven to result in the development of liver fibrosis and cirrhosis[22, 23]. The underlying reason is not clear yet, but it might be explained by CBS/H₂S homeostasis can regulate glucose metabolism, insulin responsiveness, and lipoprotein synthesis[24]. Or it was found that the dysregulation of arginine methylation[25] and reduced expression of DYRK1A[26] in CBS-deficient mice.

1.1.7 Others

In addition to the aforementioned diseases and conditions, down-regulation of CBS has been associated with lower ATP levels in skeletal muscle[27], cognitive impairment[28], and an increased risk of autoimmune disease[29]. Conversely, overexpression of CBS has been linked to Down syndrome [30] as well as reduced food consumption and body weight[31].

It is important to note that drawing indisputable conclusions from these experiments is currently challenging for several reasons [1]:

1. There are some severe phenotypes were observed in CBS-deficient mice during the early embryonic development stage, rather than being solely associated with the physiological roles of the CBS enzyme in a fully developed, adult organism.
2. Inhibitors of CBS do not exhibit specific and selective action on CBS in different tissues and cells, necessitating caution when interpreting the effects of these inhibitors.
3. Even when considering the aforementioned challenges, it remains challenging to discern the specific phenotypes resulting from either deficient or inhibited experiments and determine which aspects of the transsulfuration pathway are primarily responsible for the correlation with diseases or illnesses. This ambiguity arises from the potential dysfunctions occurring upstream (e.g., accumulation of Hcy in cells/tissues due to CBS dysfunction) and downstream (e.g., impaired production of H₂S).

1.2 5 types of Cystathionine beta-synthase

CBS exhibits diverse architectural variations, which can be categorized into five known classes [32] (shown in Figure 1.2):

(A) In bacteria and protozoa, CBS adopts the simplest form, existing as a dimeric enzyme that relies on pyridoxal 5' -phosphate (PLP) as a cofactor.

(B) A rare variant of CBS identified in roundworms features a tandem arrangement of two catalytic domains, with only one domain containing PLP.

(C) Class C architecture represents a higher level of complexity, wherein CBS incorporates a novel module known as the Bateman module. This module consists of two CBS motifs and is present in certain eukaryotes such as yeasts (in tetrameric form) and Apicomplexa, as well as in certain Gram-negative bacteria. The precise functions of the Bateman module in this particular class remain largely unknown[33].

(D) In insects, including fruit flies and honey bees, CBS architecture builds upon class C by incorporating a heme group. Notably, the loop length in CBS2 of this architecture is shorter compared to its counterpart in class C.

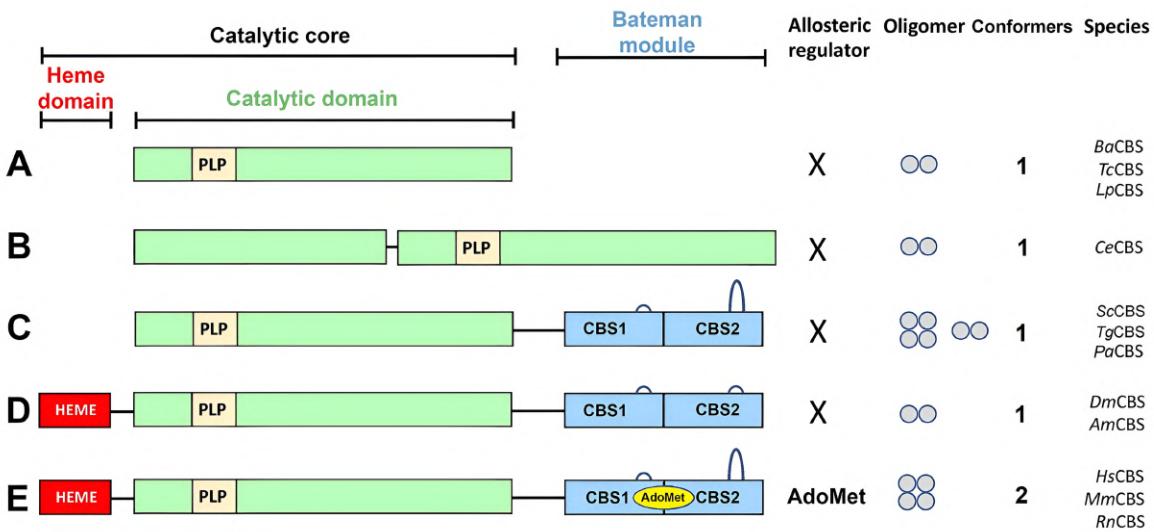


Figure 1.2: Five architectures of CBS protein, figure taken from [32]

(E) Among mammals, CBS exhibits the most sophisticated architecture. It is allosterically regulated by S-adenosyl methionine (AdoMet)[34]. The mammalian version of CBS consists of a catalytic core (comprising the heme-binding domain and catalytic domain) and the Bateman module (CBS1 and CBS2), connected by a peptide linker. AdoMet activation induces a conformational change in the Bateman module, transitioning from a basal conformation (e.g., PDB 4L0D) to an activated conformation (e.g., PDB 4PCU and 4COO). This activation enhances substrate accessibility to the catalytic cores[35, 36].

The species possessing CBS in the five categories mentioned above have been compiled and listed in Table 1.1, along with their corresponding PDB codes (where available).

As previously mentioned, human CBS (hCBS) possesses the most complex architecture among the categories discussed. It plays a crucial role in the reverse transsulfuration pathway in humans, being responsible for metabolizing homocysteine and generating H₂S. Dysfunctions of CBS have been associated with a range of diseases and disorders, which has been discussed in the previous section in detail. Gaining a thorough understanding of the functioning and regulatory mechanisms of hCBS is essential for establishing the groundwork for future drug development endeavors.

Table 1.1: Summary of CBS Protein Types and Corresponding PDB Codes (if available)

Types	PDB code available
A: Bacteria and protozoa	<i>Helicobacter pylori</i> (6AHI, 5HBG)[37], <i>Bacillus anthracis</i> (5XW3)[38], <i>Coxiella burnetii</i> (4OFX), <i>Lactobacillus plantarum</i> (5B1H, 5B1I) [39], <i>Trypanosoma cruzi</i> [40]
B: Trypanosomatids	<i>Caenorhabditis elegans</i> [41]
C: Yeast, Apicomplexa and some Gram-negative bacteria	<i>Saccharomyces cerevisiae</i> (PDB 6C2Z, 6C4P, 6C2Q, 6C2H)[42], <i>Toxoplasma gondii</i> [32], <i>Pseudomonas aeruginosa</i>
D: Insects	<i>Drosophila melanogaster</i> (PDB IDs 3PC2, 3PC3, 3PC4)[43], <i>Apis mellifera</i> (5OHX) [44]
E: Mammals	<i>Homo sapiens</i> (1JBQ, 1M54, 4COO, 4L0D, 4L27, 4L28, 4L3V, 4PCU, and 5MMS), <i>Mus musculus</i> , <i>Rattus norvegicus</i>

1.3 Structure of human cystathione beta-synthase

Human CBS is composed of 551 amino acids, with a molecular weight of approximately 63 kDa. It naturally exists in the form of homotetramers[45]. It is a very modulated protein with an N-terminal

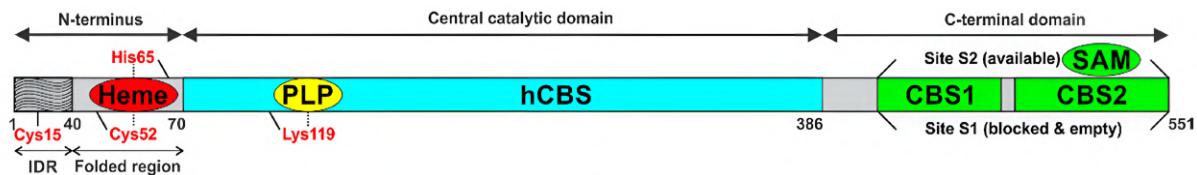


Figure 1.3: Three modules of human CBS: N-terminus, central catalytic domain and C-terminal domain figure taken from [1]

region with heme, a catalytic domain with PLP, and a Bateman module at the C-terminal region (shown in Figure 1.3).

1.3.1 N-terminal domain

The heme domain of human CBS encompasses the first 70 residues and can be further divided into two subregions.

Intrinsically disordered region

The initial 40 amino acids form an intrinsically disordered region. Crystal structures have not resolved this region extensively, likely due to its limited interaction with the rest of the protein. Notably, Kumar et al. proposed that this region serves as a noncanonical heme-binding region based on the presence of two conserved cysteine-proline (CP) motifs at positions 15 and 16, along with two histidines at positions 17 and 22[46]. These motifs exhibit structural similarity to other intrinsically disordered regions with heme-binding properties[47]. NMR studies conducted by Kumar et al. confirmed the involvement of the CP motifs in interactions with the heme group[46]. Additionally, their mutation study, specifically substituting cysteine 15 with serine, demonstrated a significant decrease in protein activity by approximately 30 %[46]. However, it is worth noting that other studies did not observe a reduction in protein activity upon the same C15S mutation[48, 49].

Canonical heme-binding region

Residues 40 to 70 within human CBS have been identified as the canonical heme-binding region. In this region, the ferric iron in the heme group is coordinated by the distal $N_{\epsilon 2}$ atom of His65 and the proximal sulfur atom of the thiolate of Cys52 (shown in Figure 1.4a). In the heme-binding region the $N_{\delta 1}$ atom of His65 is protonated, which does not engage in hydrogen bonding with other residues. On the other hand, the sulfur atom of Cys52 is deprotonated, allowing it to form hydrogen bonds with the nitrogen atom of the main chain of Trp54 and the N_{H1} atom of the guanidine group of Arg266, which is located within the α -helix 8 (Note the nomenclature of helices and sheets in this study is fully based on annotation from Figure 4.2.), where two hydrogen bonds are formed between the phosphate group of PLP and the Thr257 and Thr260 residues (shown in Figure 1.4b).

The salt bridge formed between Cys52-SG and Arg266-NH1 is hypothesized to play a significant role in the communication between the heme-binding and PLP domains, potentially influencing the activity of human CBS[50]. However, the precise underlying mechanism of this interaction remains unclear.

1.3.2 Catalytic domain

The catalytic domain of human CBS spans residues 70 to 386 and contains a catalytic core with pyridoxal 5'-phosphate (PLP) that forms a Schiff base with Lys119, also known as the internal aldimine [51].

The catalytic domain is composed of 10 α -helices and two β -sheets. The β -sheets consist of four strands ($\beta 3-\beta 6$) and six strands ($\beta 1-\beta 2$ and $\beta 7-\beta 10$), following the nomenclature defined in Figure 4.2. Based on their dynamic properties, these α -helices and β -sheets can be grouped into two blocks (shown in Figure 1.5. The first block, known as the static subdomain, includes $\alpha 1-\alpha 2$, $\alpha 6-\alpha 10$, $\beta 1-\beta 2$, and $\beta 7-\beta 10$, and exhibits a relatively rigid structure [44]. This subdomain contains two loops (B3 and B5) that are responsible for substrate accessibility. Five consensus sequences within this region have been identified,

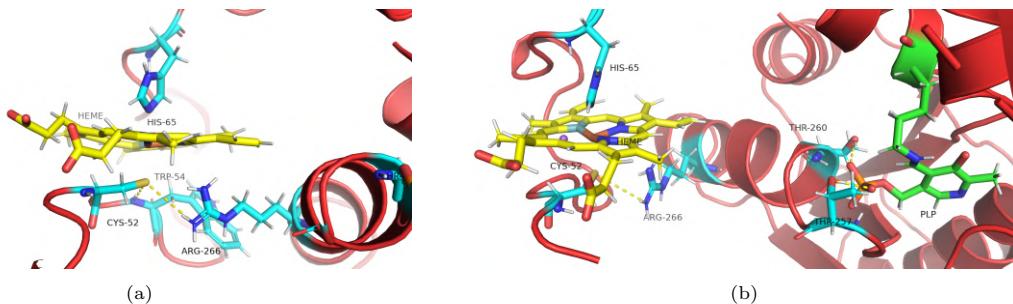


Figure 1.4: Canonical Heme-Binding Region: (a) Two hydrogen bonds between Cys52-SG and Trp54-N, and Cys52-SG and Arg266-NH1; (b) Bidirectional communication between Heme group and PLP via the α -helix 8. The protein is represented in red, while the heme group is depicted in yellow. The important residues Tpr54, Cys52, His65, Arg266, Thr257, and Thr260 are highlighted in cyan. The pyridoxal-5'-phosphate (PLP) group is visualized in green. The hydrogen bonds are illustrated by yellow dotted lines.

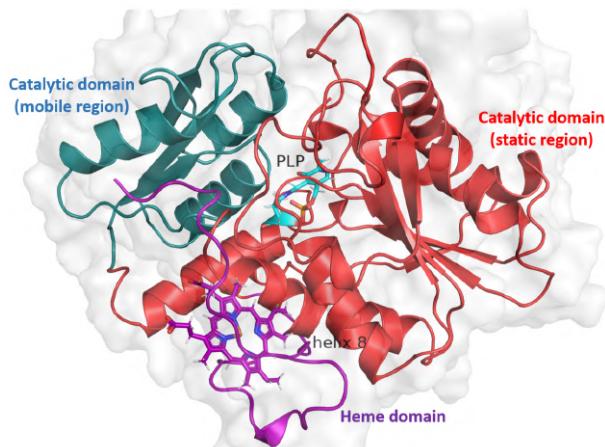


Figure 1.5: Heme domain and catalytic domain: The protein is divided into distinct domains, with the heme domain in purple and the catalytic domain further divided into a static domain (red) and a mobile domain (blue)

and they play important roles in catalytic reactions and substrate interactions [37] (shown in Figure 1.6). Block 1, highlighted in red and characterized by the conserved sequence TSGNTG_—, encompasses the "asparagine loop." This pivotal region assumes the responsibility of interacting with the substrate serine, thereby contributing to the stabilization of the aminoacrylate intermediate. Additionally, it plays a crucial role in the closure of the enzyme's active site. Block 2, distinguished in yellow and marked by the conserved sequence GTGGT, corresponds to the "phosphate loop." This loop functions as an anchor, ensuring the constant positioning of the pyridoxal 5'-phosphate (PLP) cofactor within the active site. This structural role is also depicted in Figure 1.8. Block 3, depicted in green and characterized by the conserved sequence VEGIGYD, engages with the substrate homocysteine (Hcy). Its involvement facilitates the progression of the intermediate reaction from aminoacrylate to external aldimine and aids in the reformation of the internal aldimine (as illustrated in Figure 1.10). Notably, upon substrate binding to PLP, this region contributes to the closure of the active site. Block 4, rendered in orange and defined by the conserved sequence GSSaG, assumes the role of stabilizing the pyridine ring of the PLP cofactor. Block 5, distinguished in purple and characterized by the conserved sequence NYMITSK, interacts with the active site's lip region, contributing to the overall structural integrity and functionality of the enzyme.

The remaining residues form the mobile subdomain (α 3- α 5 and β 3- β 6). Although not physically separated from the static subdomain, the mobile subdomain is connected to it by two loops: one connecting helix α 2 with strand β 3 and the other connecting helix α 6 with strand β 6. The main function of the mobile subdomain is to act as a "lid" that controls the accessibility of substrates to the catalytic core. The allosteric regulation of CBS will be further discussed in the Bateman module subsection.

The combined effects of the static and mobile subdomains result in the formation of a narrow entrance that exposes the PLP to potential substrates.

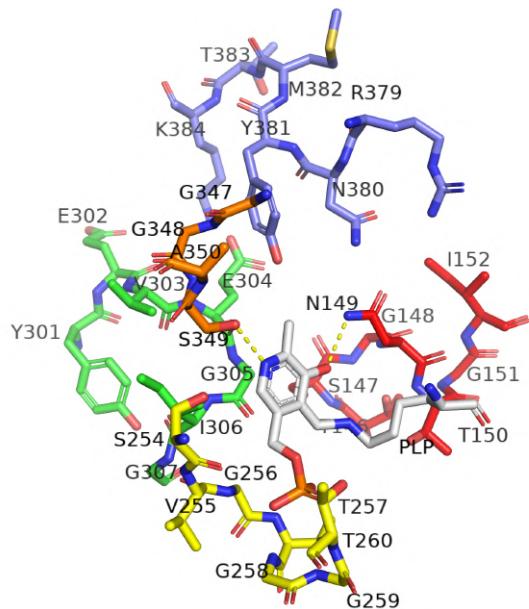


Figure 1.6: Catalytic Domain of CBS: Unveiling five distinct consensus sequences within the catalytic domain, delineated by color — Block 1 (red), Block 2 (yellow), Block 3 (green), Block 4 (orange), and Block 5 (purple). These sequences orchestrate specific roles in catalytic reactions, while the essential pyridoxal phosphate (PLP) cofactor resides in understated grey [52]

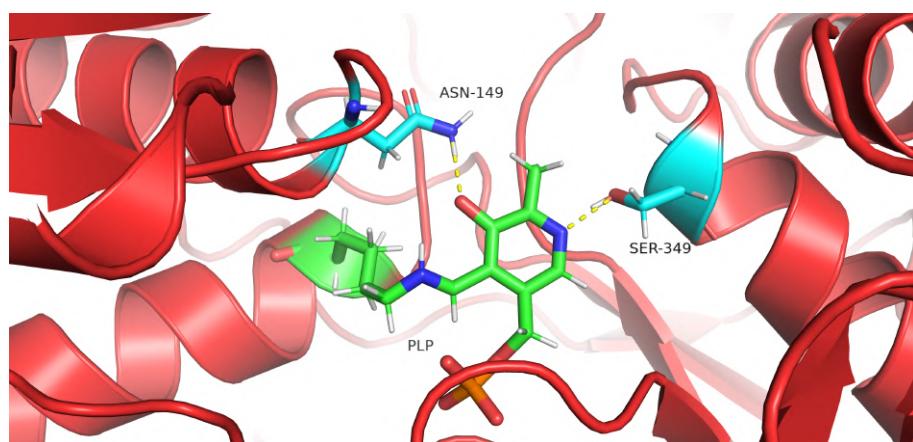


Figure 1.7: Key Hydrogen Bonds Stabilizing PLP in CBS Protein: Vital for PLP (green) stability in CBS, two crucial hydrogen bonds are notable. Asn149 (cyan) connects through Asn149-ND2 to PLP-O, and Ser349 (cyan) forms a bond linking Ser349-OG with PLP-N1. In this molecular interaction, the CBS protein takes on a prominent red depiction.

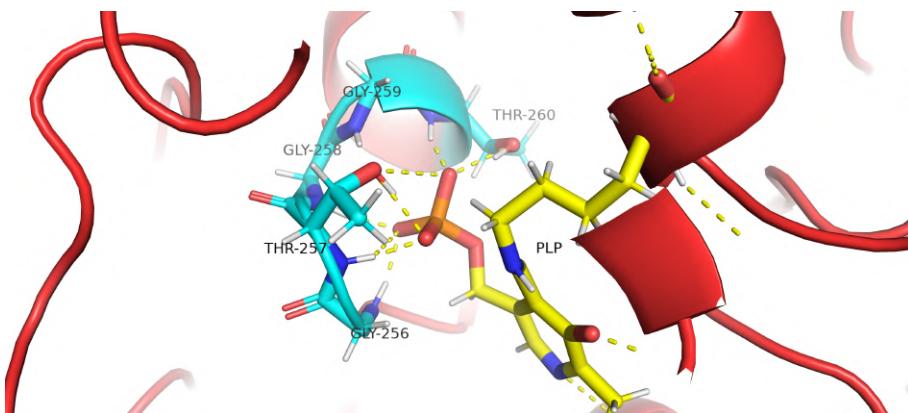


Figure 1.8: Phosphate Binding Loop and Hydrogen Interactions in CBS: The phosphate binding loop, consisting of Gly256, Thr257, Gly258, Gly259, and Thr260, is highlighted in cyan. The loop interacts with the phosphate group from the PLP (pyridoxal 5'-phosphate) cofactor, depicted in yellow, while the protein is shown in red. Notably, there are eight hydrogen interactions between the phosphate binding group and the PLP phosphate group.

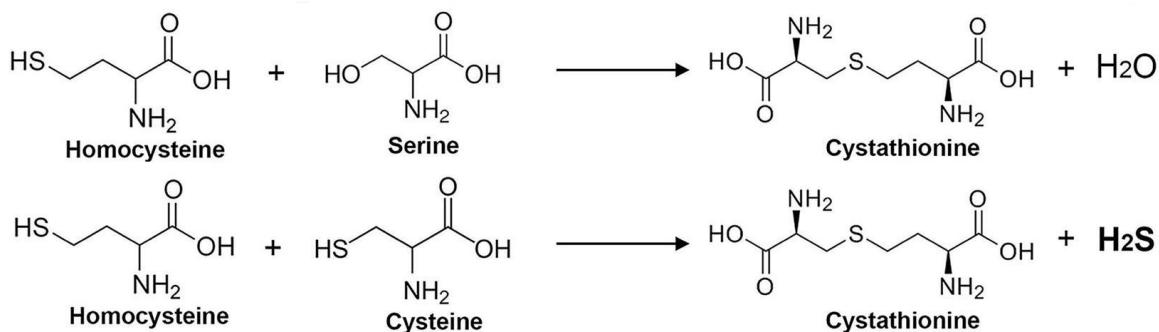


Figure 1.9: CBS catalytic chemical reaction equation, figure taken from [32]

As previously mentioned, PLP forms a Schiff base with Lys119, and additional hydrogen bond interactions contribute to the anchoring of PLP. One hydrogen bond is formed between the nitrogen atom of the pyridine ring of PLP and the O_γ atom of Ser349 [52] (shown in Figure 1.7). Another important hydrogen bond is observed between the O3' atom of the pyridine ring of PLP and the N_{ε2} atom of Asn149 (shown in Figure 1.7). It has been proposed that the stability of this hydrogen bond is crucial for the transfer of PLP from the ketoenamine (active) to the enolimine (deactive) form [53]. Further details regarding this hypothesis will be discussed later. There are two hydrogen bonds between the phosphate group of PLP and the end of α-helix 8 (Thr257 and Thr260). However, it should be noted that three nearby glycine residues, Gly256, Gly258, and Gly259, can also potentially contribute hydrogen bonds to the phosphate group [54]. Collectively, these five residues form a region known as the "phosphate binding loop", which exhibits strong hydrogen bonding and contributes to the rigidity of this region (shown in Figure 1.8).

β-elimination chemistry

The enzymatic reactions catalyzed by CBS involve the transfer of homocysteine together with serine or cysteine to form cystathionine, along with the production of a byproduct, either water (in the case of serine) or hydrogen sulfide (in the case of cysteine) (shown in Figure 1.9). The detailed steps of these reactions have been elucidated through the crystal structure of *Drosophila* CBS and stopped-flow spectroscopy experiments [43].

The catalytic reaction initiated by CBS involves the transfer of the "internal aldimine" to the "external aldimine" as serine or cysteine attacks the C4 position of PLP. This process is followed by a hydrogen abstraction event, leading to the formation of a carbanion intermediate. Subsequently, through β-elimination of water (in the presence of serine) or hydrogen sulfide (in the presence of cysteine), the carbanion is transformed into an aminoacrylate. The aminoacrylate then undergoes a reaction with another substrate

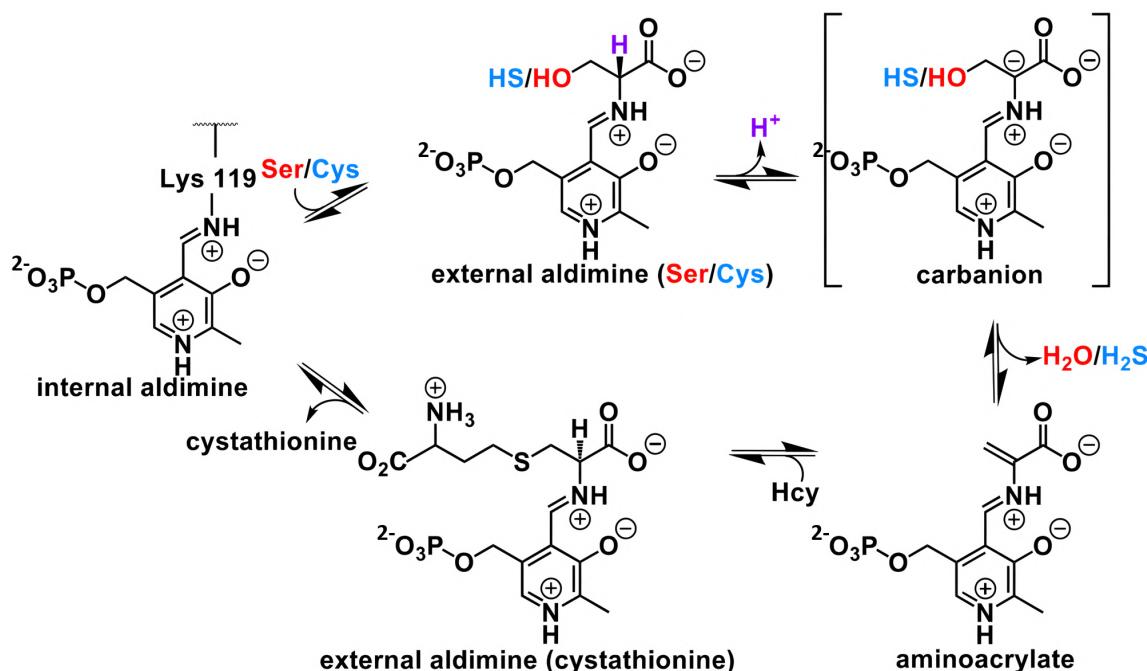


Figure 1.10: Reaction mechanism of PLP (follow the clockwise direction in the figure), figure taken from [55]

molecule, homocysteine, resulting in the formation of a new external aldimine intermediate. Finally, a transchiffization reaction occurs, leading to the release of cystathionine from the external aldimine. This final step restores PLP to its basal state, forming a Schiff base with Lys119 (shown in Figure 1.10).

These sequential reactions are essential for the conversion of homocysteine and serine or cysteine into cystathionine, with PLP serving as a crucial cofactor in mediating these transformations.

1.3.3 Bateman module

The C-terminal region of the CBS protein, spanning residues 387 to 551, consists of two tandem CBS motifs known as the "Bateman module" [56]. The first CBS motif, CBS1, adopts an $\alpha\alpha\beta\beta\alpha$ fold, while the second CBS motif, CBS2, adopts an $\alpha\beta\alpha\beta\beta\alpha$ fold. Despite sharing only 7 % sequence identity, these motifs are structurally conserved and have been found in functionally unrelated regions of other proteins, where they play a regulatory role by binding adenosine analogs [56].

In the case of CBS, the specific ligand that binds to this Bateman module is S-adenosyl-L-methionine (AdoMet), which can increase enzyme activity by 2-3 fold [57]. The regulatory mechanism underlying this modulation remained elusive until the crystal structures of the AdoMet-bound and AdoMet-free conformations were determined [35] (shown in Figure 1.11).

Unexpectedly, in the basal state, one subunit of the Bateman module interacts with the catalytic core of the other subunit to maintain equilibrium. However, upon AdoMet binding to the S2 binding site on the motif, the two dispersed motifs cluster together, adopting a compact conformation. Simultaneously, the two subunits reorient themselves to fit their corresponding regions in the catalytic domain, facilitated by the linker region (residues 386-411) that allows for the necessary rearrangement [58] that in this case, the catalytic cavity is exposed and becomes accessible for the substrates. It is important to note that the crystal structure was obtained by deleting a flexible loop (residues 516-525) in the CBS2 domain, and no significant differences in biochemical activity were observed compared to the wild-type CBS [35]. However, the deletion of the loop resulted in the CBS existing as two dimers instead of a tetrameric state, indicating that the deleted loop plays a crucial role in maintaining the interaction between the two dimers [35].

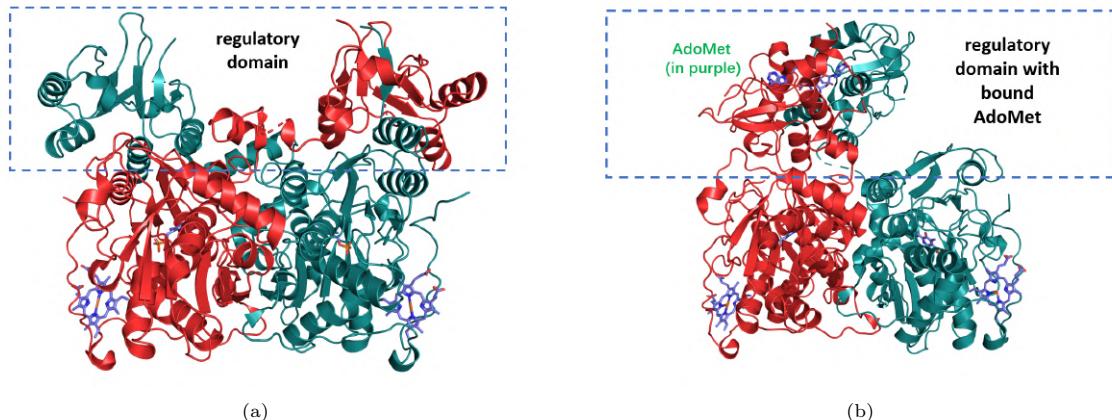


Figure 1.11: Dimeric CBS Crystal Structures: (a) Basal conformation, hCBS Δ 516-525 (PDB: 4L0D); (b) AdoMet-bound activated conformation, hCBS Δ 516-525 E201S (PDB: 4PCU). In the basal state, the regulatory domain (red) of one subunit interacts with the catalytic core (blue) of another subunit, leading to auto-inhibition of enzyme activity. The conformational change induced by AdoMet binding at the S2 site exposes the catalytic core for activation.

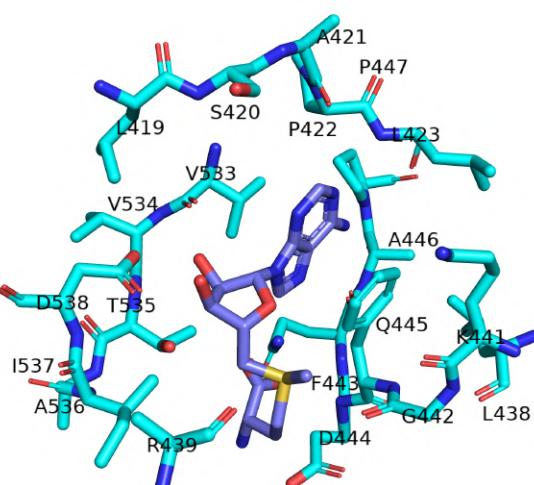


Figure 1.12: S2 AdoMet binding environment in stick: hCBS Δ 516-525 E201S (PDB: 4PCU) (basal conformation)

Isothermal titration calorimetric analyses of tetrameric CBS have identified two sets of AdoMet-binding sites. One set consists of two high-affinity sites (K_d : 10 nM), which are likely responsible for the kinetic stabilization of the regulatory domain. The other set comprises four sites with lower affinity (K_d : 400 nM), and these sites are likely involved in enzyme activation [59].

Within the regulatory region, only two AdoMet-binding sites, S1 and S2, have been found. The S1 site is obstructed by bulky hydrophobic residues, preventing AdoMet binding [60]. On the other hand, the residues surrounding the S2 site provide a favorable environment for AdoMet binding. Specifically, Asp538, Thr535, and Ser420 orient the ribose ring of AdoMet, 537Ile creates a hydrophobic pocket for the methyl moiety of AdoMet, and the strand β 11, helix α 14, and residues between α 12 and α 13 form a suitable hydrophobic environment for the adenine ring of AdoMet [58, 52] (shown in Figure 1.12).

The discrepancy between the isothermal titration calorimetric analyses and the crystal structure of the AdoMet-bound dimeric hCBS regarding the number of AdoMet binding sites can be attributed to the fact that the six identified sites are only present in tetrameric or higher-order oligomers of hCBS [52]. This suggests that the tetrameric or oligomeric state of hCBS is required for the formation of these additional AdoMet binding sites, which might not be observed in the dimeric form captured in the crystal structure.

1.4 NO/CO heme binding regulates hCBS activity

In the basal state of hCBS, the heme iron is in the ferric state with six coordination, where His65 coordinates in the distal pocket and Cys52 coordinates in the proximal pocket. This ferric heme state is characterized by low spin and is inert to exogenous ligands, showing no affinity for common ferric ligands [61, 62].

On the other hand, the ferrous state of CBS is more active and can accept other ligands due to the weakening of the His-Fe-Cys bond [63]. Both NO and CO can bind to the ferrous Fe^{2+} , leading to the decoordination of Cys52 from the iron. In the case of NO binding, His65 also becomes decoordinated, resulting in the formation of a five-coordinate ferrous nitrosyl species [62].

The binding of NO or CO to hCBS results in complete inactivation, although the binding affinities of NO and CO remain a subject of controversy, shown in Table 1.2. The inhibition mechanism is proposed to involve the modulation of the salt bridge between Cys52-SG and Arg266-NH1 [64], which is part of the α -helix 8 connecting the heme and the catalytic core PLP, located 20 Å away. A subsequent study using Fluorescence and Raman Spectroscopy found that CO binding leads to a tautomeric shift of PLP from its active ketoenamine form to the inactive enolimine form [53].

Table 1.2: Binding Constants for Gases

Gas	K_d (μM)	Reference
NO	281 ± 50	Full-length tetramer [62]
NO	350 ± 44	Truncated dimer [62]
NO	30	[65]
NO	≤ 0.23	[66]
CO	$K_{d1} = 1.5 \pm 0.1$, $K_{d2} = 68 \pm 14$	Truncated dimer [61]
CO	$K_{d1} = 3.9 \pm 2.0$, $K_{d2} = 50 \pm 8$	Truncated CBS lacking 143 C-terminal residues [64]
CO	$K_{d1} = 0.8 \pm 0.4$, $K_{d2} = 50 \pm 10$	[66]

Note: K_d represents the dissociation constant, which measures the affinity between a ligand and an enzyme. A lower K_d value indicates a higher affinity.

1.5 Previous Molecular dynamics(MD) simulations study on hCBS

At this stage, there have been limited studies using MD simulations to investigate the disturbance of hCBS activity (summarized in Table 1.3).

Three single nucleotide polymorphisms (C109R, E176K, and D376N) were identified as having significant effects on CBS protein function and stability through the use of various computational tools[67] . To examine the impact of these mutations, the crystal structure of hCBS (1JBQ) was utilized as a starting point, and the mutations were introduced. Subsequently, 6 ns MD simulations were performed for each of the four species. The results indicated that the D376N mutation, located near the catalytic region, formed hydrogen bonds with adjacent residues. As a result, this mutation potentially disrupts the pocket and affects the accessibility of the active site. Moreover, the presence of fewer intermolecular hydrogen bonds in this structure renders the protein less rigid and more flexible, which could decrease its stability by influencing protein folding.

Another commonly observed mutation, G307S, which is frequently associated with CBS deficiency in Ireland, was subjected to 200 ns MD simulations using the PDB code 4COO (without the C-terminal region) [68]. The simulations revealed that the mutation caused the displacement of Tyr308, which is responsible for removing the hydroxyl group from the external aldimine formed by serine combined with PLP. Consequently, the activity of CBS is disrupted. However, no effects on protein folding or stability were observed with the G307S mutation.

In a recent study focusing on the bidirectional communication between the heme group and PLP, the R266K mutation was investigated [69]. Based on the PDB code 4L0D, MD simulations were performed for a total of 500 ns, and the last 250 ns corresponding to the equilibrium state of the protein were analyzed. The results indicated that the salt bridge between Cys52 and 266Lys was weakened, and there was reduced electron density overlap between these residues compared to the wild-type structure. Additionally, the mutation disrupted the salt bridge between Arg266 and 316Asp, which is part of the catalytic loop and likely plays a crucial role in substrate accessibility.

These findings offer significant insights into the structural stability of the examined mutations in hCBS. The investigation of the R266K mutation, in particular, has illuminated the bidirectional communication between the Heme and PLP moieties. This communication is postulated to hold a pivotal role in the NO/CO regulation mechanism, as articulated in the 'hypothesis' section. Drawing inspiration from these illuminating studies, we intend to employ intensive MD simulations to rigorously test and further validate our hypothesis concerning the regulatory mechanism of NO/CO on hCBS.

Table 1.3: Summary of Studies on CBS Protein Mutations

Study	Crystal Structure	Simulation Details	De-	Analysis Techniques	Tech-	Conclusions	Reference
SNPs (C109R, E176K, and D376N) Influence on CBS protein stability	1JBQ	GROMACS 4.5.3, GROMOS9643a1, 6ns simulations	RMSD, Radius of gyration, SASA, PCA, Gibbs free energy landscape, hydrogen bonds	RMSF, RMSF, decreases protein stability	D376N disrupts active site access, decreases protein stability	[67]	
G307S Mutation and CBS protein activity	4COO (resid 43-401)	GROMACS 5.1.2, CGenFF (PLP-Lys), CHARMM36, 200ns simulations	RMSD, dihedrals, distance	RMSF, distance	G307S distorts catalytic mechanism, no impact on folding/stability	[68]	
R266K Mutation and bidirectional communication in CBS	4L0D	NAMD-2.13, MCPB (heme with His65, Cys52), Antechamber (PLP-Lys), Amber ff14SB, 500ns simulations	DCCM, electron density, distance, electrostatic potential	RMSD, RMSF, electron density, distance, electrostatic potential	R266K weakens C52-K266 hydrogen bond, modulates electronic communications	[69]	

1.6 Hypothesis

Indeed, the distance between the heme group and PLP in CBS is approximately 20 Å[54] , with an intermediate α -helix 8 connecting them. The observed tautomeric shift of PLP and the fluorescence spectra support the existence of communication between the heme and PLP, indicating that the heme group plays a regulatory role in CBS activity[53].

In the ketoenamine form of PLP, the imine C=N double bond is protonated, making the C4A more susceptible to nucleophilic attack by substrates. However, in the enolimine form, an internal hydrogen atom transfer occurs, where the hydrogen atom that was originally attached to the NZ nitrogen atom jumps to the nearby O3 oxygen atom. As a result, the C4A carbon becomes neutral and less prone to nucleophilic attack by substrates.

In the ketoenamine form, there is a hydrogen bond between PLP-O3 and the side chain of Asn149 (ND2)[54, 53, 70], which prevents the tautomerization from occurring. It has been proposed that disrupting the salt bridge between Cys52-SG and Arg266-NH1 could lead to a displacement of α -helix 8, which in

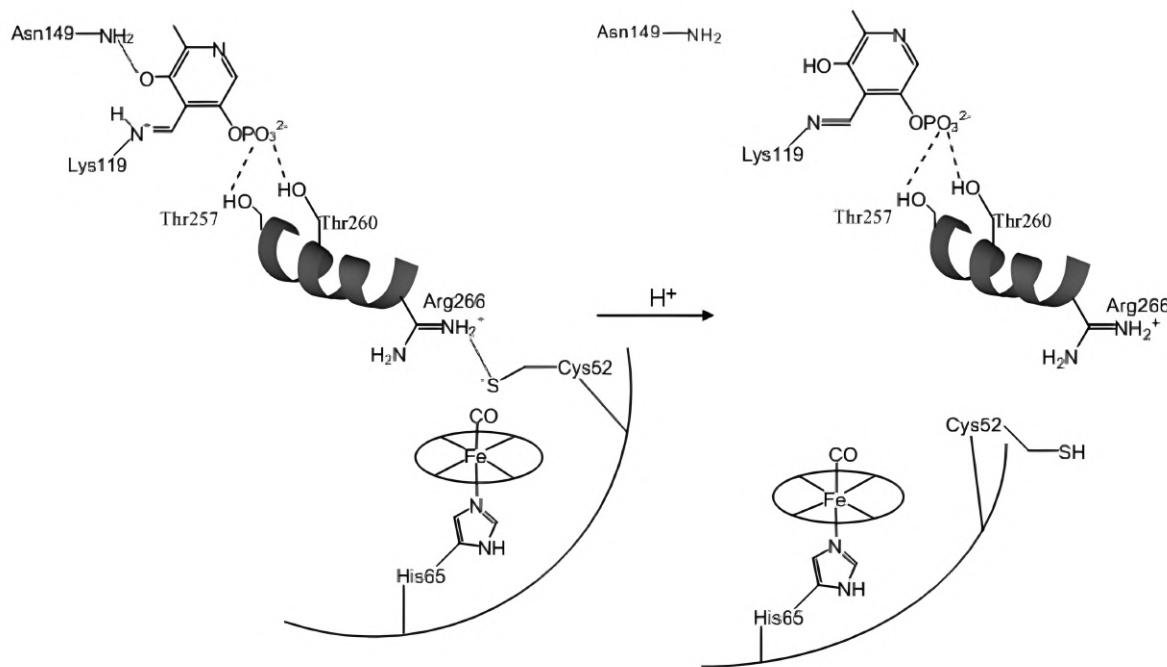


Figure 1.13: Proposed mechanistic insights into hCBS deactivation process, figure taken from [53]

turn affects the positioning of PLP through strong hydrogen bonds between α -helix 8 and PLP. This disruption of the hydrogen bond between PLP-O3 and Asn149-ND2 allows for the hydrogen transfer event, leading to the conversion of PLP from the ketoenamine form to the enolimine form [53] (shown in Figure 1.13).

1.7 Aims and Objectives

1.7.1 Aims

In this study, we aim to investigate the mechanism by which the binding of NO/CO to the heme group in CBS induces a tautomeric shift of PLP. This phenomenon has been observed, but the underlying mechanism remains unclear. By conducting detailed investigations, we seek to uncover the precise molecular events that lead to this tautomeric shift and elucidate the regulatory role of the heme group in PLP dynamics.

1.7.2 Objectives

To achieve these aims, we have outlined the following objectives for our research. Firstly, we will conduct a rigorous multi-sequence alignment of CBS protein sequences from insects and mammals that contain the heme group. This analysis will allow us to identify conserved regions that are likely to play a crucial role in the regulation of CBS activity.

Next, we will employ molecular dynamics (MD) simulations to investigate the dynamics of CBS in different states (shown in Figure 1.14). Specifically, we will simulate the initial state of CBS without NO/CO binding, as well as three proposed states after NO/CO binding. Through long-time MD simulations using appropriate force fields, we will monitor the conformational changes and examine the distortion of the active site in each state.

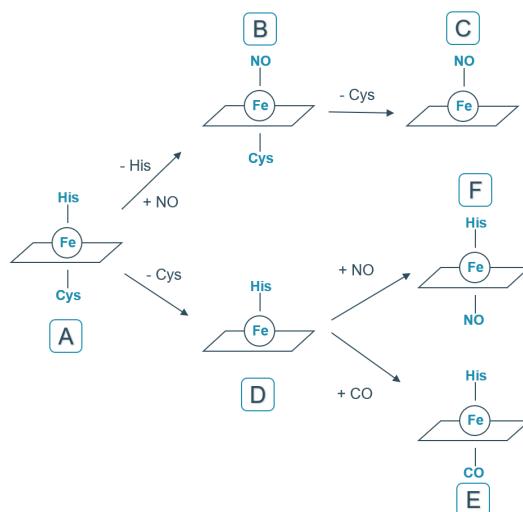


Figure 1.14: Evolutionary States of hCBS: NO/CO Binding and Structural Transitions: started from state A (basal), after NO/CO binding to heme, both histidine and cysteine could decoordinate from the heme and evolve into NO bound (C) or His, CO bound (E) or His, NO bound (F)

Chapter 2

Theoretical background

2.1 Molecular dynamics simulations

Molecular dynamics (MD) simulations are computational techniques used to simulate the motion and behavior of particles in a system over time [71]. These simulations are valuable tools in various scientific disciplines, including chemistry, biology, and materials science.

The MD simulation process begins with an initial configuration of the system, including the positions and velocities of all atoms or particles. The initial velocities are typically assigned based on the Maxwell-Boltzmann distribution, which describes the distribution of velocities at a given temperature.

The equation of motion (see the next section for details), usually derived from Newton's laws, is then solved numerically to update the positions and velocities of the particles at each time step. By iteratively updating the positions and velocities, the simulation progresses and provides information on the dynamic behavior of the system over time.

The accuracy and reliability of MD simulations depend on several factors, such as the force field used to describe interatomic interactions, the time step used in the simulation, and the treatment of long-range interactions. These considerations must be carefully addressed to ensure meaningful and accurate results from MD simulations.

2.1.1 Principles of Mathematics and Physics

In order to predict the positions and velocities of atoms in a system at a given time point t , it is necessary to calculate the new positions $q(t + \Delta t)$ and velocities $v(t + \Delta t)$, where Δt represents the time step.

To approximate the new positions and velocities based on the known information of $q(t)$ and $v(t)$, a Taylor expansion with two derivatives is employed. This expansion allows us to express the new positions and velocities as follows (Eq. 2.1, Eq. 2.2):

$$q(t + \Delta t) = q(t) + \Delta t \frac{dq(t)}{dt} + \frac{\Delta t^2}{2!} \frac{d^2q(t)}{dt^2} \quad (2.1)$$

$$v(t + \Delta t) = v(t) + \Delta t \frac{dv(t)}{dt} + \frac{\Delta t^2}{2!} \frac{d^2v(t)}{dt^2} \quad (2.2)$$

Here, $\frac{dq(t)}{dt}$ and $\frac{dv(t)}{dt}$ represent the first derivatives of position and velocity with respect to time, respectively. The terms $\frac{d^2q(t)}{dt^2}$ and $\frac{d^2v(t)}{dt^2}$ correspond to the second derivatives of the position and velocity, which

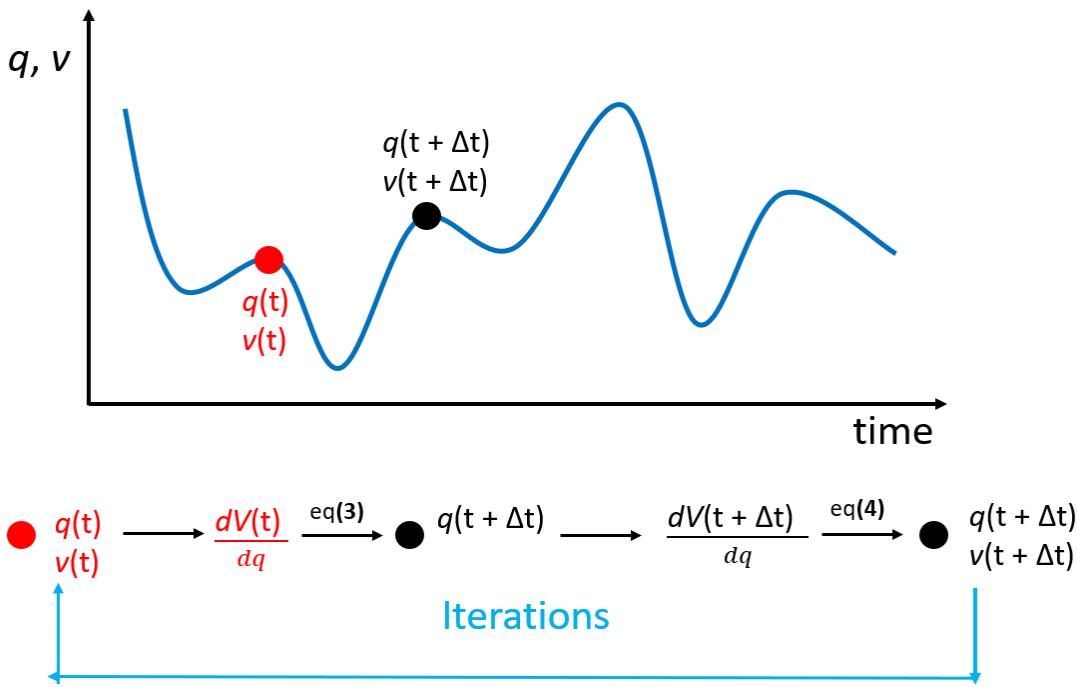


Figure 2.1: simple representation of two states of a system at time t and time $t + \Delta t$ and iterative way of predicting the $q(t + \Delta t)$, and $v(t + \Delta t)$

account for the acceleration of the atoms.

To solve the equations, we can start by applying Newton's second law:

$$F = m \cdot a = m \frac{dv}{dt} \quad (a)$$

where $\frac{dv}{dt}$ represents the second derivative of the position q with respect to time. And we can get $\frac{dv}{dt} = \frac{d^2q}{dt^2}$.

Furthermore, considering the equation of motion:

$$F = -\frac{dV}{dq} \quad (b)$$

where V represents the potential energy.

By combining equations (a) and (b), the new position $q(t + \Delta t)$ can be represented as:

$$q(t + \Delta t) = q(t) + \Delta t \cdot v(t) - \frac{\Delta t^2}{2m} \cdot \frac{dV(t)}{dq} \quad (2.3)$$

In this equation, if we know the coordinates of the previous time point, $q(t)$, the velocity of the previous time point, $v(t)$, and the potential energy gradient from the previous time point, $\frac{dV(t)}{dq}$, we can calculate the new position $q(t + \Delta t)$. This equation takes into account the effect of the potential energy on the motion and allows us to predict the updated position at the next time step Δt .

Solving equation (2.2) can be achieved by applying Taylor expansion with two derivatives around $\frac{dv(t)}{dt}$. By rearranging the terms obtained from the Taylor expansion and multiplying both sides by $\frac{\Delta t}{2}$, the second derivative $\frac{\Delta t^2}{2} \frac{d^2v(t)}{dt^2}$ can be expressed as $\frac{\Delta t}{2} \frac{dv(t+\Delta t)}{dt} - \frac{\Delta t}{2} \frac{dv(t)}{dt}$. Utilizing Newton's second law (equation (a)) and the equation of motion (equation (b)), equation (2.2) can be written as:

$$v(t + \Delta t) = v(t) - \frac{\Delta t}{2m} \left(\frac{dV(t + \Delta t)}{dq} + \frac{dV(t)}{dq} \right) \quad (2.4)$$

Equations (2.3) and (2.4) form the basis of the Velocity Verlet algorithm. However, it should be noted that equation (2.4) requires knowledge of the current potential energy, $\frac{dV(t+\Delta t)}{dq}$.

To perform molecular dynamics simulations, an iterative process is followed. Starting from a state with coordinates (q) and velocities (v) at time point t , the force exerted on each atom, $\frac{dV(t)}{dq}$, is calculated. Using equation (2.3), the coordinates of all atoms at time point $t + \Delta t$ are updated. This allows the calculation of the potential energy gradient, $\frac{dV(t+\Delta t)}{dq}$. By applying equation (2.4), the velocities (v) at time point $t + \Delta t$ are also updated. The state at $t + \Delta t$ becomes the new state, and the process is iterated until the desired end of the simulation is reached.

2.1.2 Potential energy

To update the positions $q(t + \Delta t)$ and velocities $v(t + \Delta t)$, the calculation of the potential energy gradient is necessary. There are various approaches to obtain the potential energy gradient, such as using force fields, Schrödinger equation, or a combination of Quantum Mechanics/Molecular Mechanics (QM/MM). For classical molecular dynamics simulations, we employ force fields.

The potential energy is composed of two main types of interactions based on the force field: bond interactions (including bond, angle, and dihedral potential energy) and nonbonded interactions (nonbonded potential energy) (equation (2.5)).

$$\mathbf{V}_{\text{total}} = \sum_{i=1}^{N_{\text{bond}}} V_{\text{bond}} + \sum_{i=1}^{N_{\text{angle}}} V_{\text{angle}} + \sum_{i=1}^{N_{\text{dihed}}} V_{\text{dihed}} + \sum_{i=1}^{N_{\text{nb}}} V_{\text{nb}} \quad (2.5)$$

Here, N_{bond} represents the number of atom pairs that are directly bonded, N_{angle} represents the number of atom triplets separated by two bonds, N_{dihed} represents the number of atom pairs separated by three bonds, and N_{nb} represents the number of atom pairs separated by more than three bonds.

Atom types

In the context of force fields, atom types are used to distinguish different chemical environments or bonding conditions of atoms within a molecule. Each atom type is assigned specific force field parameters, such as partial charges, bond lengths, bond angles, and dihedral angles, which govern the interatomic interactions during a molecular dynamics simulation.

For example, in the Charmm36 force field[72], different types of carbon atoms are defined to capture their specific characteristics and behaviors in different bonding environments. These atom types are named based on their functional roles or structural features. For instance:

"C" represents carbonyl carbon, typically found in peptide backbone structures. "CA" refers to aromatic carbon atoms, commonly encountered in aromatic rings. "CC" represents carbonyl carbon in specific contexts, such as those found in asparagine, aspartate, glutamine, glutamate, C-terminal ends, and some other cases. "CC3151" is a specific atom type for a carbon atom in a furan ring. These different atom types allow for more accurate descriptions of the chemical and physical properties of atoms in diverse bonding environments, enhancing the realism and accuracy of molecular dynamics simulations.

Bonded interactions

The bonded interactions in the potential energy are represented by the harmonic potentials for bond (V_{bond}) and angle (V_{angle}) interactions (equations (2.6) and (2.7)). In the case of bond potential, the energy is described by a harmonic function with force constant k_b , where r is the current separation between two atoms and r_0 is the equilibrium distance between them. For angle potential, the energy is also described by a harmonic function with force constant k_a , where j represents the angle between three atoms and j_0 is the equilibrium angle.

$$V_{\text{bond}} = \frac{1}{2}k_b(r - r_0)^2 \quad (2.6)$$

$$V_{\text{angle}} = \frac{1}{2}k_a(\theta - \theta_0)^2 \quad (2.7)$$

It is important to note that these harmonic potentials provide a good approximation only around the equilibrium geometry and are unable to describe bond breaking or formation events, as they assume a harmonic shape for the energy curve and do not allow for molecular dissociation.

The dihedral (or torsion, or 1-4) interactions are described by Fourier potentials (equation (2.8)). In this case, V_{dihed} depends on the dihedral angle ω formed by two planes: one formed by atoms 1-2-3 and the other formed by atoms 2-3-4. The parameters k_t , n , and γ are defined by the force field and determine the shape of the potential. The parameter n corresponds to the number of minima in the potential energy, k_t represents the energy barriers between the minima, and γ determines the position of a minimum.

$$V_{\text{dihed}} = k_t(1 + \cos(n\omega - \gamma)) \quad (2.8)$$

Non-bonded interactions

Non-bonded interactions in the potential energy consist of Coulomb interactions (2.9), van der Waals interactions (2.10), and repulsion interactions (2.11). In fixed charge force fields, these interactions are described as follows:

For two atoms, atom i and atom j , with charges q_i and q_j , and a distance r_{ij} between them, the Coulomb interaction is given by equation (2.9). The van der Waals interactions are considered through equation (2.10), which describes the dispersion forces arising due to transient dipole-dipole interactions. When the separation r_{ij} between atoms is very small, there is a repulsion interaction described by equation (2.11).

$$V_C = \frac{q_i q_j}{4\pi \varepsilon r_{ij}} \quad (2.9)$$

$$V_{vdW} \propto -\frac{1}{r_{ij}^6} \quad (2.10)$$

$$V_{rep} \propto e^{-r_{ij}} \approx \frac{1}{r_{ij}^{12}} \quad (2.11)$$

The Lennard-Jones potential (2.12) is used to model the van der Waals and repulsion interactions. Here, σ represents the distance at which the intermolecular potential energy between particles is zero, and ε represents the depth of the potential energy well between particles. Each atom type has its own σ and ε . The combined parameters for paired atom types, σ_{ij} and ε_{ij} , can be calculated as follows: $\varepsilon_{ij} = \sqrt{\varepsilon_i \varepsilon_j}$ and $\sigma_{ij} = \frac{\sigma_i + \sigma_j}{2}$.

The Lennard-Jones potential is then expressed as:

$$V_{LJ} = 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \quad (2.12)$$

Finally, the total potential energy (2.5) can be written as:

$$\begin{aligned} V_{total} = & \sum_{i=1}^{N_{bond}} \frac{1}{2} k_b (r - r_0)^2 + \sum_{i=1}^{N_{angle}} \frac{1}{2} k_a (\theta - \theta_0)^2 + \sum_{i=1}^{N_{dihed}} k_t (1 + \cos(n\omega - \gamma)) + \\ & \sum_{i>j} \frac{q_i q_j}{4\pi\epsilon r_{ij}} + 4\epsilon_{ij} \left[\left(\frac{\sigma_{ij}}{r_{ij}} \right)^{12} - \left(\frac{\sigma_{ij}}{r_{ij}} \right)^6 \right] \end{aligned} \quad (2.13)$$

2.1.3 Periodic Boundary Conditions

The current computational power allows for the simulation of approximately 10^6 atoms in a cubic environment. However, it is important to note that around 10% of these molecules are located at the surface, and their behavior can differ from those in the bulk. This can introduce artifacts in simulations. In real systems, such as a 1 mol solution with about 10^{23} molecules, only an extremely small fraction of approximately $10^{-6}\%$ of the molecules are exposed to the vacuum, making their effects negligible. Unfortunately, current computational capabilities are not sufficient to directly simulate such large-scale systems.

To address the issue of surface-exposed molecules and simulate larger systems, the concept of Periodic Boundary Conditions (PBC) is employed. PBC involves creating a primary cell that contains the atoms of interest, and this cell is then replicated in all three dimensions to represent the entire system. The replicated cells maintain the same relative positions of atoms, and any motion of atoms within the primary cell is mirrored in the corresponding atoms of other cells. This ensures that the total number of atoms in each cell remains constant throughout the simulation, effectively mitigating the effects of surface exposure and allowing for more realistic simulations.

MD simulations can then be carried out in these replicated cells, but the Hamiltonian equations are only calculated for the primary cell. However, to account for the interactions between atoms in neighboring cells, it is necessary to consider neighboring cells within a certain cutoff distance. This introduces computational limitations and errors in the simulation.

Specifically, the non-bonded interactions between atoms are typically described by Coulomb interactions (Equation 2.9) and the Lennard-Jones potential (Equation 2.12). The Lennard-Jones potential rapidly approaches zero as the distance between atoms increases, making it negligible at long distances. However, Coulomb interactions persist even at long distances and cannot be ignored, leading to inaccuracies in the simulation.

To overcome this issue and obtain more accurate and physically realistic results, the Ewald summation [73] method was introduced. Ewald summation allows for the separation of electrostatic interactions into two components: short-range interactions and long-range interactions. The short-range interactions are typically handled using a cutoff distance, while the long-range interactions are computed using Fourier transforms. This approach improves the accuracy of the simulation by properly accounting for the long-range electrostatic interactions, reducing errors associated with the use of a simple cutoff distance.

2.2 Analysis Theory

After conducting molecular dynamics simulations, several analysis methods were employed to gain deeper insights into our system. Here, we present the mathematical foundations of these methods:

2.2.1 Root Mean Square Deviation

Root Mean Square Deviation(RMSD) quantifies the extent to which the system changes over time compared to a reference structure. It is one of the very important indicator that can make sure whether a system reach equilibrium.

$$\begin{aligned} \text{RMSD}(\mathbf{q}, \mathbf{q}') &= \sqrt{\frac{1}{n} \sum_{i=1}^n \|q_i - q'_i\|^2} \\ &= \sqrt{\frac{1}{n} \sum_{i=1}^n ((q_{ix} - q'_{ix})^2 + (q_{iy} - q'_{iy})^2 + (q_{iz} - q'_{iz})^2)} \end{aligned} \quad (2.14)$$

Where q and q' are two sets of systems, q as the reference structure of the system (protein in our case), and q' the later-on configurations during the simulations. The RMSD value is expressed in length unit, which is Ångström (Å).

2.2.2 Root Mean Square Fluctuation

Root Mean Square Fluctuation(RMSF) measures the fluctuation of individual atoms or residues throughout the entire simulation period relative to their average positions. It helps to identify the most flexible regions within the protein.

$$\rho_i = \sqrt{\langle (\mathbf{x}_i - \langle \mathbf{x}_i \rangle)^2 \rangle} \quad (2.15)$$

Where x_i represents the coordinates of residues i and $\langle \mathbf{x}_i \rangle$ represents the average position of residues i . The unit is expressed the same as the RMSD.

2.2.3 Dynamical Cross Correlation Matrix

Dynamical Cross Correlation Matrix(DCCM) is used to identify atoms or residues that exhibit correlated motions during the simulation. It allows us to detect regions within the protein that move together, providing insights into functional dynamics and communication pathways.

$$C(i, j) = \frac{1}{N} \sum \left[\frac{(r_i(t) - \langle r_i \rangle)(r_j(t) - \langle r_j \rangle)}{\sqrt{\sigma_i^2} \cdot \sqrt{\sigma_j^2}} \right] \quad (2.16)$$

Here, $C(i, j)$ represents the element at the i^{th} row and j^{th} column, N is the total number of frame, $r_i(t)$ and $r_j(t)$ store the coordinations of the i^{th} and j^{th} atoms/residues at the time step t , $\langle r_i \rangle$ and $\langle r_j \rangle$ are the average positions of the i^{th} and j^{th} atoms/residues over all the frames, σ_i^2 and σ_j^2 are the standard deviations of the i^{th} and j^{th} atoms/residues.

Chapter 3

Computational methods

3.1 Multiple sequence alignment

In this study, we focused on class E (insects) and F (mammals) of CBS, as they are the only classes known to incorporate a Heme group (as shown in Table 1.1). Consequently, we selected the CBS protein sequences from *Drosophila melanogaster* (PDB ID 3PC2), *Apis mellifera* (PDB ID 5OHX), and *Homo sapiens* (PDB ID 4COO) for multiple sequence alignment analysis. To ensure consistency, the initial "MHHHHHHSSGVDLGTENLYFQS" amino acid sequence from 4COO was removed to match the sequence ID used in our study. Additionally, the protein sequences of *Mus musculus* and *Rattus norvegicus* were obtained from the [UniProt](#) database.

To assess the diversity of individual residues at each position, we employed the Simpson diversity index.

For multiple sequence alignment, we utilized the [SeaView](#) [74] software (version 5.0.5), followed by the identification of conserved regions (with a cutoff of 70% similarity) among the CBS protein sequences using the online platform [ESPrift 3.0](#) [75].

3.2 Molecular dynamic simulations system preparation

Based on the current knowledge, nine CBS 3D structures are available in the PDB database (accessed via [RSCB PDB](#)). For the purposes of the research, the 4COO [36] structure with a resolution of 2.0 Å was selected as the initial model. As CBS is a biological dimer, the two chains are nearly identical and symmetrical. To focus on the research interests, the C-terminal regulatory domains were removed from both chains since they have no effect on the heme or PLP domain[68]. Missing atoms were added using the [SCWRL](#) [76] program.

The modified structure file underwent initial MD simulations for state A (as depicted in Figure 1.14) using the CHARMM36 force field (Feb 2021 version) [77], with some modifications to incorporate the PLP ligand by CGenFF [78]. To properly define the bonding state of the heme group, a file that defines the bonding information (Specbond.bat) was included. Prior to generating the topology file, the protonation state of aspartate, glutamate, and histidine was carefully determined based on their pKa values, as calculated by [Propka](#), and the surrounding environment.

To better simulate the protein in a biologically relevant environment, 0.15 mol Na⁺ and Cl⁻ ions were added to the water solution(TIP3P) [68], using a rhombic dodecahedron box to ensure that the protein was at least 20 Å away from the box boundary. Energy minimization was then applied to avoid any steric clashes and to relax the protein system properly, with convergence criteria: maximum force of 500.0 kJ/mol/nm. Subsequently, both constant NVT and NPT ensembles were applied sequentially to

the entire protein-solution system for 1 ns each, using Berendsen thermostats and pressure couplings methods, respectively, to equilibrate the system at a temperature of 300K and pressure of 1 bar.

The well-equilibrated system underwent 450 ns of MD simulations using GROMACS (version 2021.4) [79]. After removing the periodic boundary condition (PBC), the trajectory file was obtained for later data analysis and structure modifications.

After the RMSD analysis of the “Backbone,” a representative frame at the convergence state that was as close to the end as possible was selected. Due to an error given by GROMACS, both “Na⁺” and “Cl⁻” were removed from the solution. Additionally, all four-letter residue names indicating the protonation state of the residues were changed back to three-letter names to make the PDB file readable by GROMACS. Furthermore, 52CYF in both chains were renamed as CYS, as they would be protonated later by GROMACS.

To obtain proper structures of states C, F, and E (as depicted in Figure 1.14), “NO” or “CO” molecules were constructed using Avogadro2 and attached to the corresponding positions in both chains of CBS. The bonding between His65-N and Fe²⁺ and Cys52-S and Fe²⁺ were removed based on the specific state(C or F or E), and the atoms N and S were slightly shifted to avoid any steric clashes. After obtaining the new coordinates of the atoms of interest, they were added to the PDB file of the selected frame. Modifications were made, such as renumbering the atoms, to ensure that the file was readable for GROMACS.

The modified PDB files were now ready for MD simulations using the same procedures defined earlier. The protonation states of aspartate, glutamate, and histidine were defined the same as before, and 0.15 mol NaCl was added back to the solvent. A maximum cutoff force of 500.0 kJ/mol/nm was set to make the system as relaxed as possible. The lowest potential energy system underwent 1 ns of NVT and NPT simulations to reach a well-equilibrated state. Subsequently, 1000 ns of MD simulations were conducted after NPT. The trajectory files were used for further analysis to extract essential information that demonstrates the conformation change at the α -helix 8 and PLP region.

3.3 Trajectory analysis

Root Mean Square Deviation

To evaluate the stability of the protein structure, we conducted RMSD analysis based on the backbone atoms (C, N, O, and C_a).

For each state, RMSD analysis was performed on the entire trajectory to capture the overall changes in the protein structure over time. Additionally, to gain a better understanding of the movement of α -helix 8 during the simulations, RMSD analysis was specifically conducted on the two helices in two chains for each state.

Root Mean Square Fluctuation

Furthermore, we utilized RMSF analysis based on the backbone atoms to identify the residues with the highest flexibility following the introduction of NO/CO bonding to the Heme group. The RMSF analysis is a valuable tool for understanding the conformational changes in the protein, particularly the differences observed after NO/CO binding.

Distance/Angle measurements

To validate our hypothesis, we conducted distance calculations between specific atoms and plotted the changes over time. The atom pairs of interest are detailed in Table 3.1. Through these calculations and subsequent analyses, we gained valuable insights into the conformational changes occurring in the α -helix and PLP region. Furthermore, these investigations shed light on the structural alterations induced by the introduction of NO/CO bonding to the Heme group. The results of these analyses contribute significantly to our understanding of the protein’s behavior and its functional implications.

Table 3.1: Distance/Angle of Interest

	chain a	chain b
h1	Cys52-SG to Arg266-NH1	Cys52-SG to Arg266-NH1
h2	Thr257-OG1 to PLP119-O2P	Thr257-OG1 to PLP119-O3P
h3	Thr260-OG1 to PLP119-O3P	Thr260-OG1 to PLP119-O1P
h4	Asn149-ND2 to PLP119-O	Asn149-ND2 to PLP119-O
h5	Ser349-OG to PLP119-N1	Ser349-OG to PLP119-N1
t	COG of Heme to COG of α helix 8 (resid 259-271)	COG of Heme to COG of α helix 8 (resid 259-271)
r1	Angle between the normal of the Heme plane and a vector from Heme-Fe to Arg266-CA	Angle between the normal of the Heme plane and a vector from Heme-Fe to Arg266-CA
r2	Angle between the normal of the Heme plane and a vector from Heme-Fe to Thr260-CA	Angle between the normal of the Heme plane and a vector from Heme-Fe to Thr260-CA

Dynamical cross-correlation matrix

In addition to identifying the most flexible regions resulting from NO/CO binding to the Heme group and the disruption of the salt bridge between Cys52-SG and Arg266-NH1, understanding the interplay between different residues and their correlated motions is of paramount importance. To achieve this, we employed Dynamical Cross Correlation Matrix (DCCM) analysis on each state. The DCCM calculations were performed using a Python package called [correlationplus](#) [80], and α -carbon atoms were used as the default atoms for the analysis.

Modevectors

To visually represent the protein motions in a more intuitive manner, we utilized the PyMol script package called [Modevectors](#) (Author: Sean M. Law). This package allowed us to create arrows that visually depict the direction and magnitude of the motions observed in the protein structure.

3.3.1 Software

In this study, we utilized various software tools to conduct data analysis and visualize the results. PyMol, version 4.6.0, was employed to visualize the protein structures represented in .pdb or .gro format. For visualizing trajectory files in .xtc format obtained from GROMACS, we used VMD version 1.9, 4a53. Python, version 3.11, and RStudio, version 4.1.1, were employed for data analysis and generating plots to support our findings.

Chapter 4

Results and discussion

4.1 Evolution conservation study

As shown in Figure 4.1, the conservation analysis of CBS proteins from five species (*Homo sapiens*, *Mus musculus*, *Rattus norvegicus*, *Drosophila melanogaster*, and *Apis mellifera*) reveals a high degree of conservation in the protein sequences. The conservation is represented by the presence of red and purple dots, indicating sites with a Simpson diversity rate less than or equal to 0.4. In this context, a value of 0.4 implies that each site tolerates only one variant.

Overall, the conservation rate among these CBS sequences, which consist of approximately 550 amino acids, is determined to be 52.77% based on the threshold of a Simpson diversity rate of 0.4. Additionally, 34.26% of the residues are found to be fully conserved across all species.

The distribution of conservation aligns with the known information depicted in Figure 1.3, where residues 70 to 386, corresponding to the catalytic domain responsible for enzyme activity, exhibit the highest degree of conservation. On the other hand, the heme binding region and regulatory domain show more variability, as expected.

These findings highlight the importance of the catalytic domain in maintaining the core functionality of CBS, while allowing for variability in other regions involved in heme binding and regulation. The multiple sequence alignment shown in Figure 4.2 provides valuable insights into protein conservation and

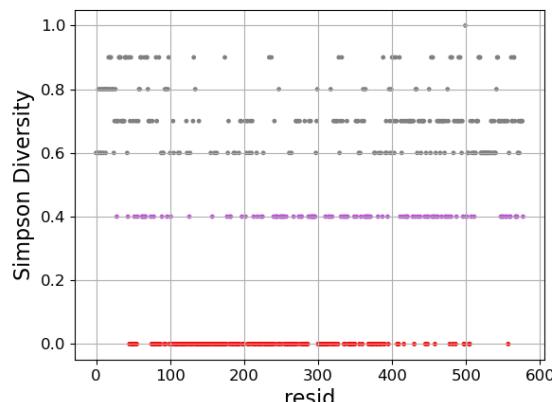


Figure 4.1: Simpson diversity rate plot among the sequence: Red dots represent those sites are 100% conserved, purple dots are those with Simpson diversity = 0.4, while grey dots represent those sites are variable sites

its association with CBS function. A cutoff of 70% similarity (in this case, equivalent to one tolerance) was employed to identify regions where the five species exhibited significant similarity. These regions were outlined with a blue frame. Fully conserved sites, where all five species share identical amino acids, were highlighted in red.

In the alignment, it is evident that the intrinsically disordered region (residues 1-39) exhibits considerable variability. However, it is worth noting that this variability may be attributed to the fact that these regions have not been resolved in fruit flies and honey bees. Notably, there is significant conservation observed in mammals.

The heme binding region (residues 40-70) is of particular interest as it is associated with the regulation of protein activity by heme. It has been established that the heme group is coordinated by residues Cys52 and His65 in human CBS. Interestingly, while residue Cys52 is fully conserved among the five species, a variant is observed in honey bees, where residue 65 is replaced by a threonine. This suggests that heme stabilization may be accomplished by a cysteine residue alone. Furthermore, residue 54Tyr, which forms a hydrogen bonding interaction with Cys52, is conserved across all five species.

Another residue of interest is Arg266, which acts as an intermediate between heme and PLP through α -helix 8. As depicted in Figure 4.2, this residue is fully conserved among the five species, and α -helix 8 (residues 259-271) shows high conservation as well, with the sequence "GTI_GI_RKIKE_". Here, uppercase letters indicate complete conservation, lowercase letters denote a single variant allowed under the 70% similarity cutoff, and "_" represents a variable site. This conservation strongly suggests that the bidirectional communication between heme and α -helix 8 likely plays a significant role, although the precise details are yet to be elucidated.

The "Phosphate Binding Loop" (residues 256-260), as identified by Meier et al. [54], is also conserved and can be represented by the sequence "GTgGT". This loop, located at the other end of α -helix 8, contributes to the firm anchoring of PLP.

At the opposite end of PLP, two additional residues, Asn149 and Ser349, form hydrogen bonds, as shown in Figure 1.7. These sites are highly conserved across species and serve to anchor PLP within the protein structure.

The C-terminal region of CBS (from residue 386 onwards) exhibits variability among the five species, although many sites still fall within the blue frames indicating significant similarity. This can be attributed to the fact that in mammals, this region has evolved to acquire the ability to bind AdoMet, as depicted in Figure 1.2. In contrast, CBS from insects lacks this property. Consequently, some sites within the range of residue 386 to 515 are covered by blue frames. The most flexible region is observed from residue 516 to 519, corresponding to the loop in CBS2. It is worth noting that the length of this loop is longer in mice and rats compared to humans and insects. However, since the primary focus of this project is the heme and PLP domains, we will not delve deeper into this particular region.

4.2 Trajectory analysis

4.2.1 Initial His65 and Cys52 bound Heme state (AA)

Before conducting MD simulations for His, CO bound state, His, NO bound state, and NO bound state, ensuring a well-established initial structure is crucial. After performing 450ns MD simulations, we obtained a reasonably stable His, Cys bound state (Figure 1.14), characterized by an RMSD value of 3 Å at the end of simulations (Figure 6.1a). This result is consistent with the RMSD values reported in other studies of CBS[67, 69].

The RMSF analysis (Figure 6.1b) revealed two regions with high flexibility. The first region comprises residues 52 to 65, exhibiting an RMSF of approximately 4 Å. The second region includes residues 290 to 310, displaying an RMSF of roughly 7 Å. These findings align with our expectations and are consistent

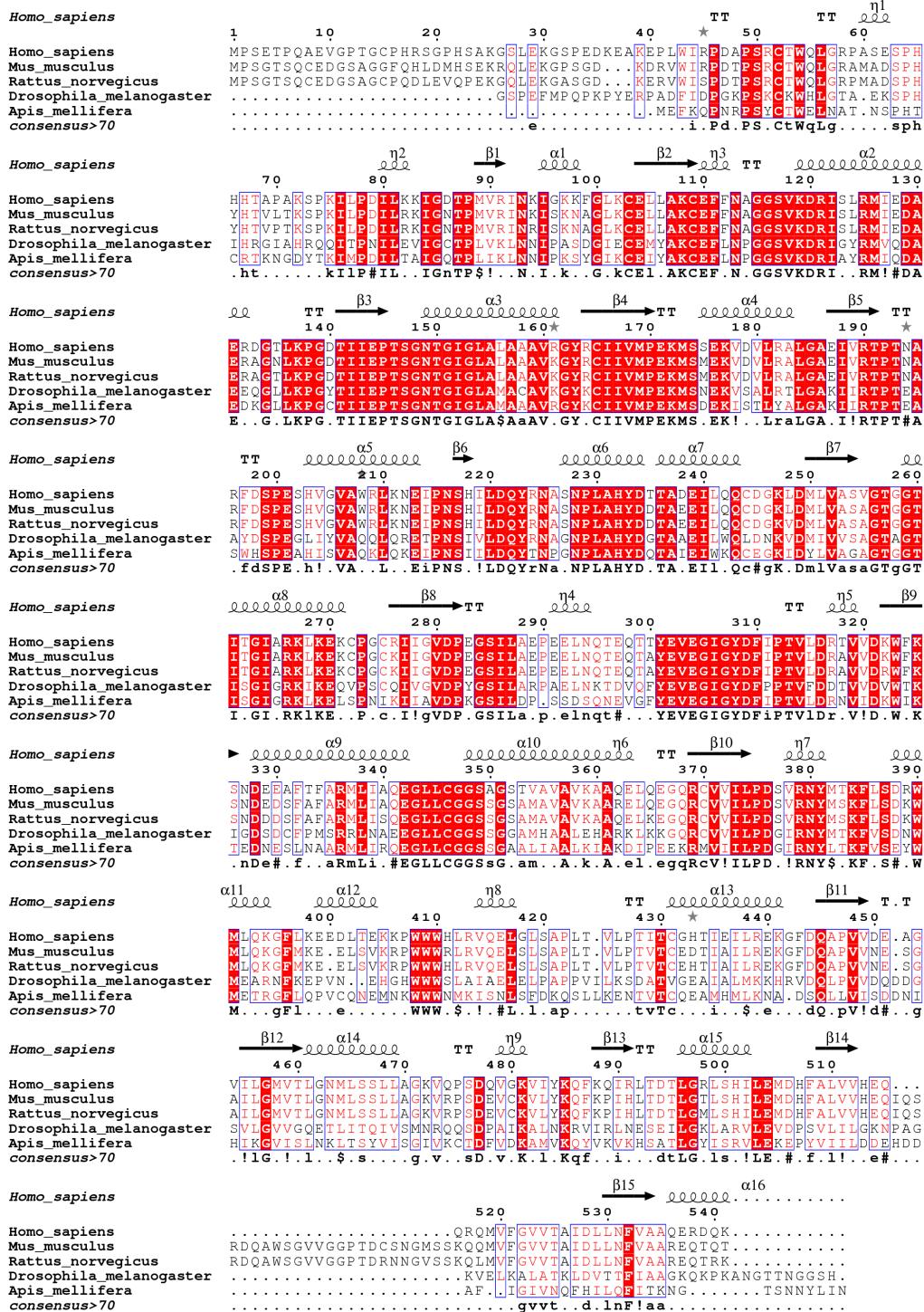


Figure 4.2: Multiple sequence alignment result: At each column, the similarity value is calculated and a cut off with 0.7 is set, when the similarity of a site is larger than the cutoff value, it will be covered by a blue frame, and when the site is 100% conserved, it will colored in red

with the RMSF results reported in other studies of CBS[68, 69]. Both flexible regions correspond to loop regions that are in close proximity to the solvent environment.

Understanding the bidirectional communication between the heme and PLP in a stable state is of paramount importance. The salt bridge interaction between Cys52-SG and Arg266-NH1 has been found to exhibit remarkable stability (Figure 4.3a). Specifically, the average distance of h1 in chain a is approximately 3.45 Å, while in chain b, it is approximately 3.56 Å. Moreover, two hydrogen bonds, one between Thr257-OG1 and the phosphate group of PLP and the other between Thr260-OG1 and the phosphate group of PLP, were also analyzed over time (Figure 4.3b). These hydrogen bonds consistently displayed a typical hydrogen bond distance, hovering around 2.70 Å. This observation aligns with the structural rigidity observed in the region containing eight hydrogen bonds (Figure 1.8). Additionally, for the two hydrogen bonds located at the other end of PLP, namely h4 and h5, they also exhibited remarkable stability (Figure 4.3c and Figure 4.3d). The average distance of h4 was approximately 3.10 Å. These findings provide essential insights into the dynamic interactions between the heme and PLP, revealing their role in maintaining the stability and functionality of the protein. Upon the assumption of breaking

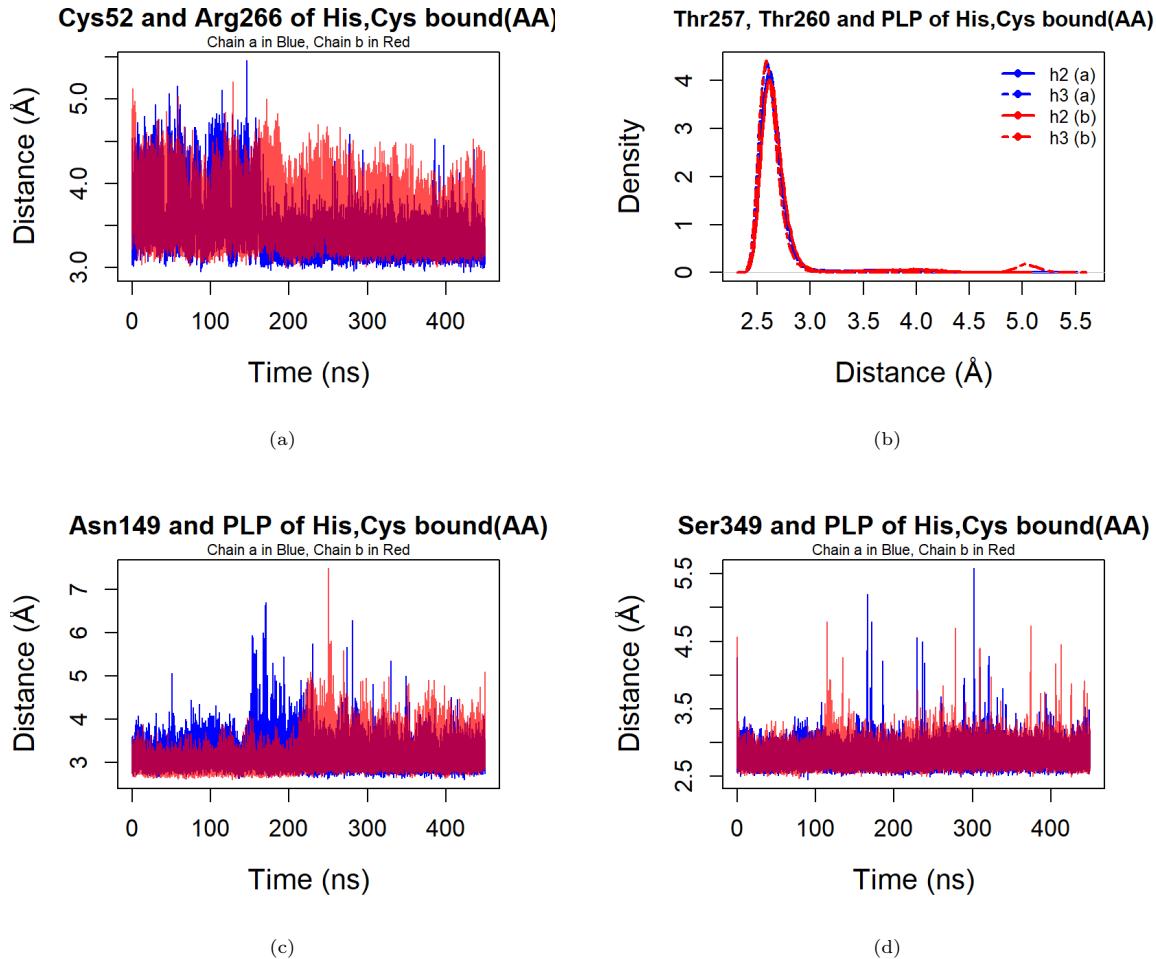


Figure 4.3: Analyzing Distances in His-Cys Bound State (AA) for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1: (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

the salt bridge between Cys52-SG and Arg266-NH1, it was hypothesized that α -helix 8 would undergo slight conformational changes. To evaluate this hypothesis, distances and angles were measured over time. The distance between the center of geometry (COG) of the heme group and the COG of α -helix

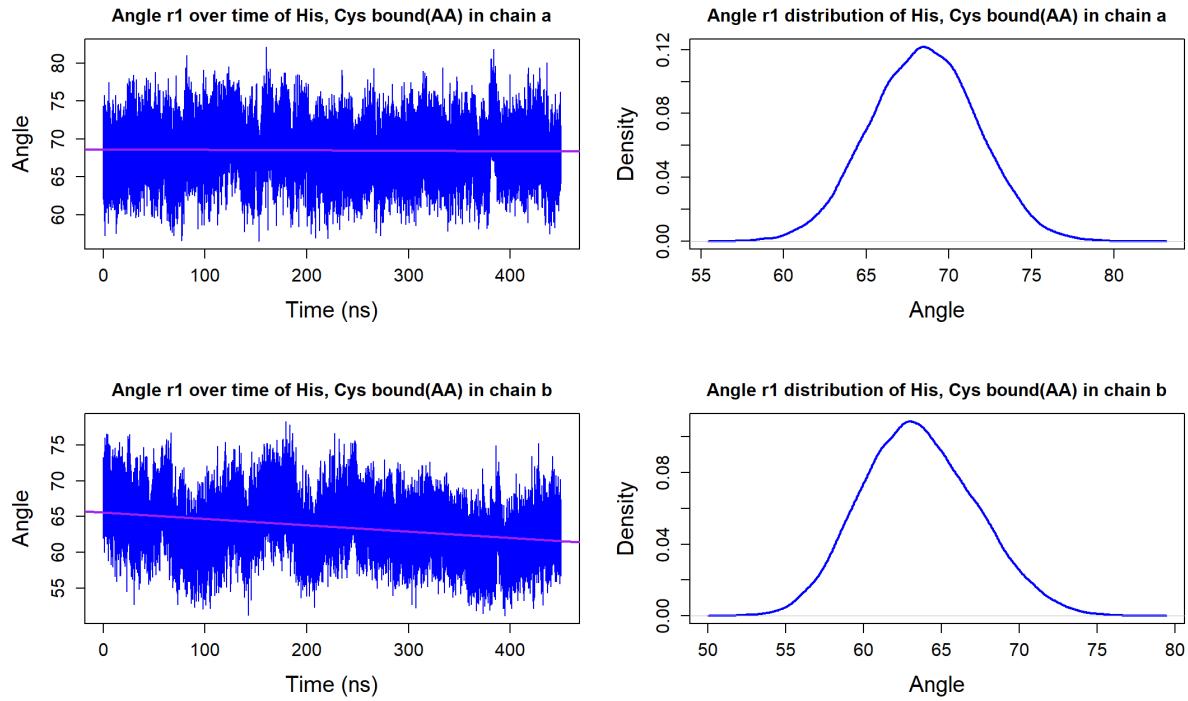


Figure 4.4: Rotational Dynamics of α -Helix 8 in His-Cys Bound State (AA) for Chains a and b: Evolution of angle r_1 (left) and its distribution (right) over time. Angle r_1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

8 (t) (Figure 6.2), the angle between the normal of the heme and the vector from the iron of the heme to Arg266-CA (r_1) (Figure 4.4), and the angle between the normal of the heme and the vector from the iron of the heme to Thr260-CA (r_2) (Figure 4.5) were analyzed.

In state AA, the distance between the heme and α -helix 8 (t) displayed slight variations between chain a and chain b. The average distance of t in chain a was approximately 12.8 Å, while in chain b, it was approximately 13.2 Å.

The average angle r_1 exhibited stability over time in chain a, with an average value of approximately 69°. In chain b, the average angle r_1 was approximately 63°, and it exhibited a small decrease from 66° to 62° over time, which is within an acceptable range.

The average angle r_2 in both chain a and chain b in state AA was approximately 82°. Over time, chain a showed a slight increasing trend, while chain b exhibited a minor decreasing trend, indicating some variability in the conformation of α -helix 8.

These findings provide insights into the stability and dynamics of α -helix 8 before breaking the salt bridge.

Once a stable initial His, Cys bound state (AA) was obtained, modifications were made to create His, CO bound state (EE), His, NO bound state (FF), and NO bound state (CC), as described in Section 3.2. The last frame from the 450ns trajectory file was utilized to construct these states. Additionally, frames at 400ns were used to create His, CO bound state (EE') and His, NO bound state (FF') as replicates. These modifications allowed for the exploration of various states and the investigation of the system's behavior under different conditions.

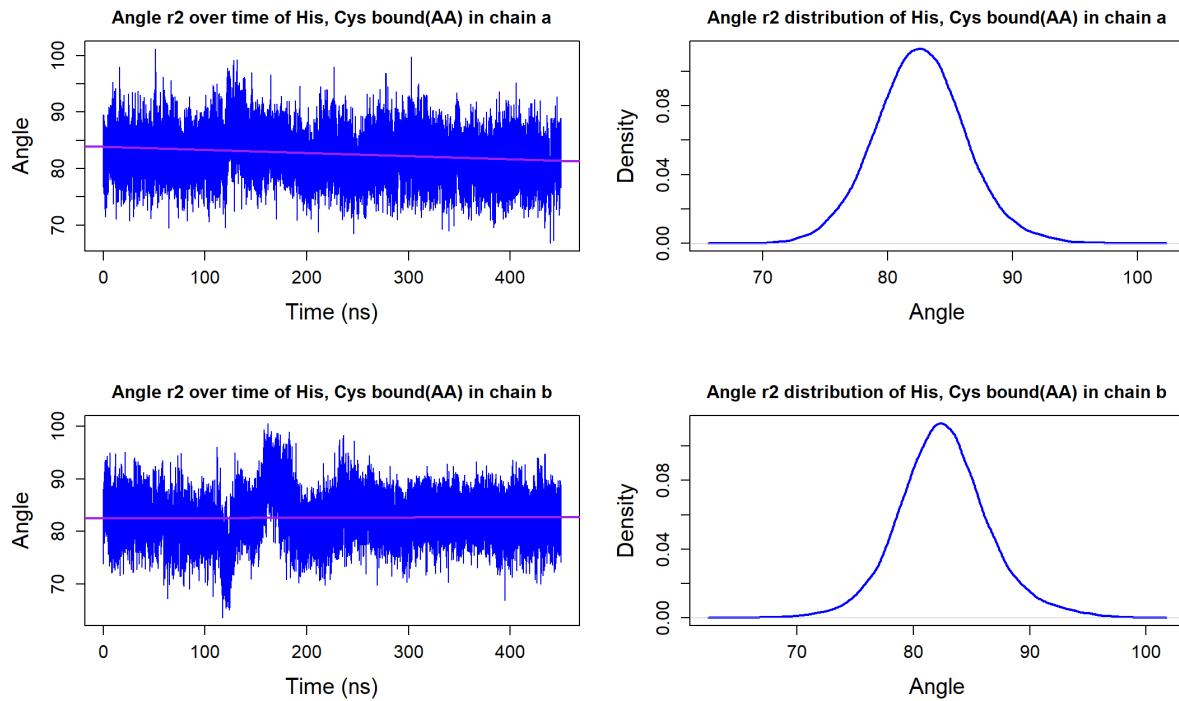


Figure 4.5: Rotational Dynamics of α -Helix 8 in His-Cys Bound State (AA) for Chains a and b: Evolution of angle r_2 (left) and its distribution (right) over time. Angle r_2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

4.2.2 His, CO bound state (EE and EE')

State EE represents the scenario where CO replaces the Cys52-SG binding to the Heme- Fe^{2+} instead (Figure 1.14). After performing 1000ns of MD simulations, a stable state EE was obtained, with an RMSD value of approximately 3.8 Å at the end of simulations (Figure 6.3a). To further examine the conformational changes resulting from breaking the salt bridge between Cys52-SG and Arg266-NH1, RMSF analysis was conducted and compared to the RMSF plot of state AA (Figure 6.3b). Generally, the four chains overlapped quite well, except for two loop regions (residues 52 to 65 and residues 290 to 310), which exhibited even more pronounced fluctuations in state EE. However, no significant enhancement of motions was observed in the α -helix 8 regions (residues 259 to 271) in both chains of state EE.

State EE' also reached a stable conformation, with an RMSD value of approximately 4.5 Å at the end of simulations (Figure 6.4a). Similar to state EE, the RMSF analysis of state EE' (Figure 6.4b) exhibited the same pattern, with the two loop regions becoming more flexible, while the rest of the regions did not undergo any significant conformational changes.

The distance and angle measurements provided further insights into the behavior of the region of interest (Heme- α helix 8-PLP). As expected, the distance h_1 , representing the salt bridge between Cys52-SG and Arg266-NH1, showed random fluctuations in both chain a and chain b of state EE (Figure 4.6a) and EE' (Figure 4.7a), confirming the absence of the salt bridge.

In state EE, the four hydrogen bonds in both chain a and chain b (Figure 4.6b) were highly stable, consistent with our expectations and observations in His, Cys bound state (AA) from Figure 4.3b. However, in state EE', the hydrogen bonding pattern was different from the other states (Figure 4.7b). Specifically, in chain a, Thr257-OG1 formed a hydrogen bond with PLP-O3P, while Thr260-OG1 formed a hydrogen bond with PLP-O1P. In chain b, Thr257-OG1 formed a hydrogen bond with PLP-O1P, while Thr260-OG1 formed a hydrogen bond with PLP-O2P. Despite the differences in hydrogen bonding partners, the "phosphate binding loop" remained intact, and the region exhibited some flexibility.

The distance changes between specific residues in the Heme- α helix 8-PLP region provided valuable insights into the dynamics and interactions within this region.

In state EE, the distance (h4) between Asn149-ND2 and PLP119-O showed an interesting behavior. After 750ns, the distance increased beyond the hydrogen bond range, indicating a breaking of the hydrogen bond interaction. However, towards the end of the simulation in chain a, the distance returned to within the hydrogen bond range, suggesting a reformation of the hydrogen bond (Figure 4.6c). The hydrogen bond between Ser349-OG and PLP119-N1 (h5) remained stable throughout the simulation (Figure 4.6d).

In chain b of state EE, the hydrogen bond between Asn149-ND2 and PLP119-O broke between 0 to 80ns and reformed later. There was another short breaking of this hydrogen bond around 780 to 810 ns (Figure 4.6c). Similarly, the hydrogen bond between Ser349-OG and PLP119-N1 (h5) showed a longer time hydrogen bond breaking between 560 and 820 ns (Figure 4.6d).

The distance (*t*) between the center of geometry (COG) of the heme and the COG of α -helix 8 was measured over time to observe the behavior of α -helix 8. Generally, the translation of α -helix 8 was limited, with the distance (*t*) increasing by about 1 Å and decreasing by approximately 0.5 Å over the course of the simulation (Figure 6.5a).

The rotation measurements (*r*1 and *r*2) provided further insights into the motions of α -helix 8 and the impact of the salt bridge breaking between Cys52-SG and Arg266-NH1.

Angle *r*1 measured whether Arg266 rotated due to the salt bridge breaking, and angle *r*2 measured whether the end of α -helix 8, specifically residue Thr260, rotated together with Arg266.

As shown in Figure 4.8, in chain a, both Arg266 and Thr260 started to rotate after 250ns, which might have played a role in the subsequent hydrogen bond breaking between Asn149-ND2 and PLP119-O after 750ns. On the other hand, in chain b, the rotation of both Arg266 and Thr260 occurred after approximately 750ns, coinciding with the two hydrogen bond-breaking events (h4 and h5).

Notably, the initial angle *r*2 in chain b was close to 90 degrees, which was 10 degrees larger than that in chain a (Figure 4.10). This initial difference in angle might explain the early hydrogen bond breaking observed in chain b.

In state EE', a noticeable hydrogen bond-breaking event was observed between 520ns and 650ns in chain a, specifically for those two hydrogen bonds (shown in Figure 4.7c and 4.7d). However, in chain b, no significant hydrogen bond breaking was observed during the simulation.

For chain a, the α -helix 8 moved approximately 1 Å away from the Heme group (Figure 6.5b). Notably, no rotation of residue Arg266 in chain a was observed (Figure 4.9). However, residue Thr260 began to rotate from 180ns and continued to rotate until the end of the simulations (Figure 4.11). The initial angle *r*2 for residue Thr260 in chain a was approximately 80 degrees, which increased to almost 110 degrees during the simulation. Upon further inspection of the structure, it was discovered that there was no significant movement of residue Thr260. Instead, it was observed that the Heme group underwent a flipping motion of approximately 30 degrees (shown in Figure 4.21b), which explains the peculiar behavior observed in the distance measurements. This highlights a limitation in the distance measurements analysis, as the assumption that the position of the Heme group remains unchanged over time may not always hold true. The bond between the Heme group and His65 still exists, but as residue His65 is located in the loop region, it may exhibit greater flexibility and movement, leading to the observed flipping motion of the Heme group.

In summary, the analysis of His, CO bound state (EE and EE') provided valuable insights into the dynamics of the CBS protein and its interactions with the Heme and PLP groups. Key findings include:

- RMSF analysis showed enhanced flexibility in two loop regions (residues 52 to 65 and residues 290 to 310) in both states EE and EE', while the rest of the protein remained relatively stable compared to the initial state AA.

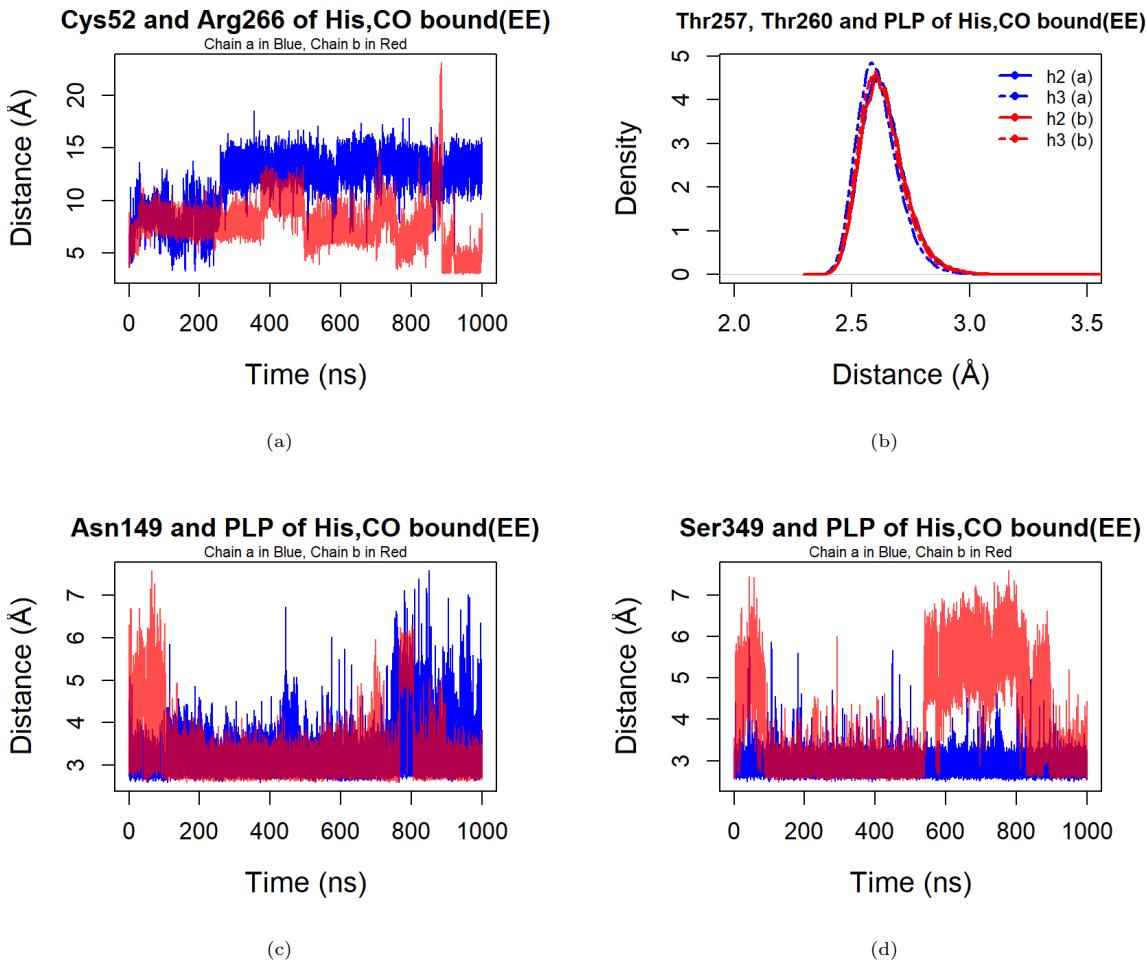


Figure 4.6: Analyzing Distances in His, CO Bound State (EE) for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1; (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

- The protonation of the sulfur atom of Cys52 led to the breaking of hydrogen bonds between 52Cys and 54Tyr, as well as 52Cys and 266Arg, allowing Cys52 to exhibit increased flexibility.
- The phosphate binding loop (residues 256 to 260) formed eight stable hydrogen bonds with the phosphate group of PLP. However, the position of the three oxygen atoms in the phosphate group may undergo occasional rotations.
- The α -helix 8 did not undergo significant translation away from the Heme group, with a limited movement of approximately 1 Å.
- Hydrogen bond-breaking events were observed between Asn149-ND2 and PLP119-O, as well as Ser349-OG and PLP119-N1. However, these hydrogen bonds tended to reform during the simulations, because MD simulations can not simulate actual internal hydrogen jumping event that to protonate the 119PLP-O.
- The rotation of α -helix 8, starting from Arg266 and subsequently Thr260, may be correlated with the breaking of hydrogen bonds between Asn149-ND2 and PLP119-O and Ser349-OG and PLP119-N1. The rigid structure of the α -helix 8 and PLP via strong hydrogen bonds in the phosphate binding loop likely contributes to this behavior.

It is worth noting that the analysis of state EE' was complicated by the flipping motion of the Heme group, which limited the interpretation of the results.

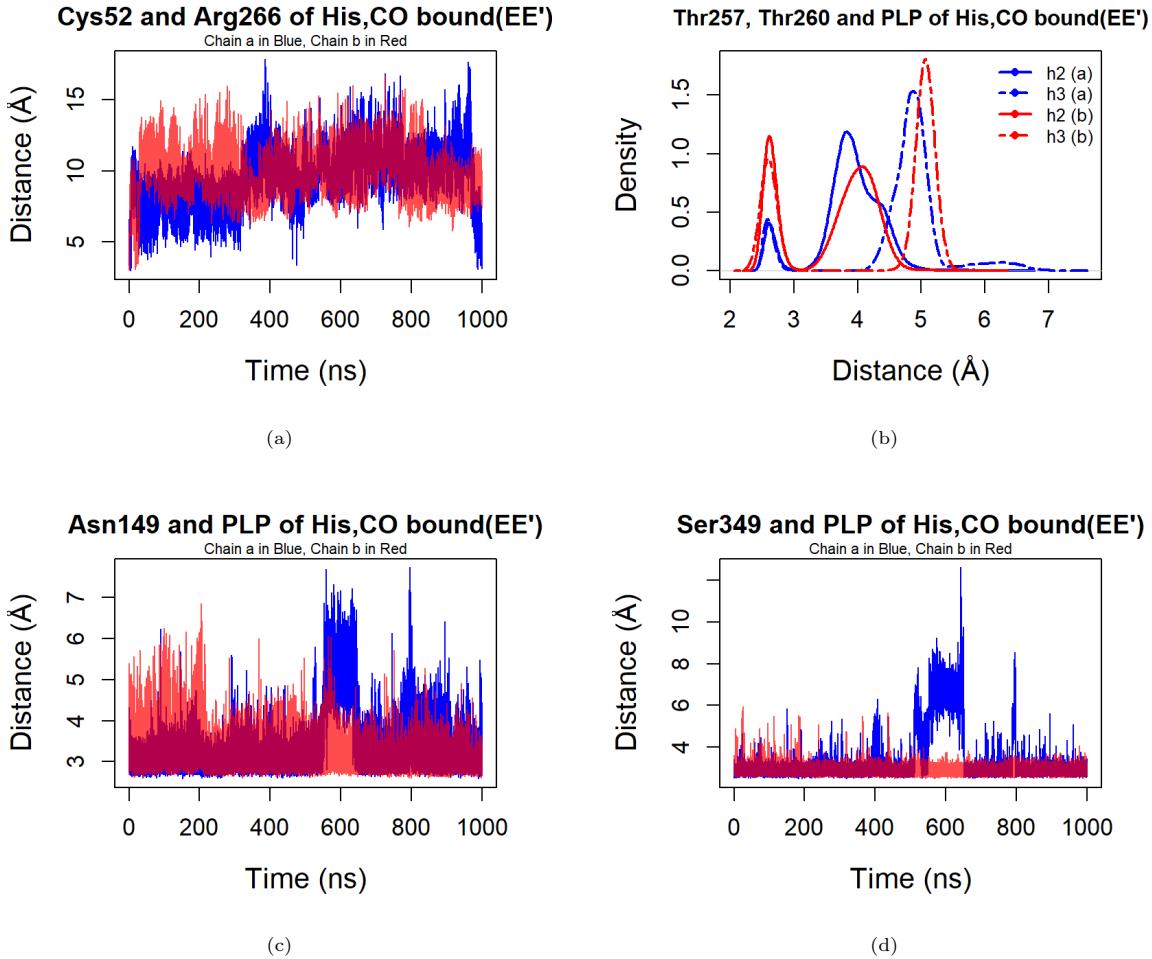


Figure 4.7: Analyzing Distances in His, CO Bound State (EE') for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1: (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

4.2.3 His, NO bound state (FF and FF')

State FF, which was not observed in experiments, was hypothesized to be a potential scenario where, after Cys52 decoordinates from the Heme group, NO could bind to the Heme-Fe²⁺ when the environmental concentration of NO is high. After 1000ns of simulations, both FF and FF' reached stable states with RMSD values of approximately 4.5 Å and 4 Å (Figure 6.6a and 6.7a), respectively. Therefore, the breaking of the salt bridge between 52Cys and 266Arg did not lead to significant conformational changes in the protein compared to the initial state AA (Figure 6.1a). However, RMSF analysis of state FF (Figure 6.6b) revealed some distinct peaks not observed in previous results. Notably, the region between β 5 and α 6 (residues 190 to 200) displayed a higher RMSF of 5.5 Å in chain a of FF, whereas in the normal state AA, this region exhibited fluctuations with only around 2.5 Å RMSF. Additionally, increased flexibility was observed in residues 90 to 100 and residues 320 to 380 in chain a of FF. These unique peaks suggest that significant events may have occurred throughout the simulations, and the underlying reasons for the increased flexibility in these regions will be discussed in the subsequent "Dynamical Cross-Correlation Matrix" section. Interestingly, no significant peaks were observed in the α -helix 8 region (residues 259

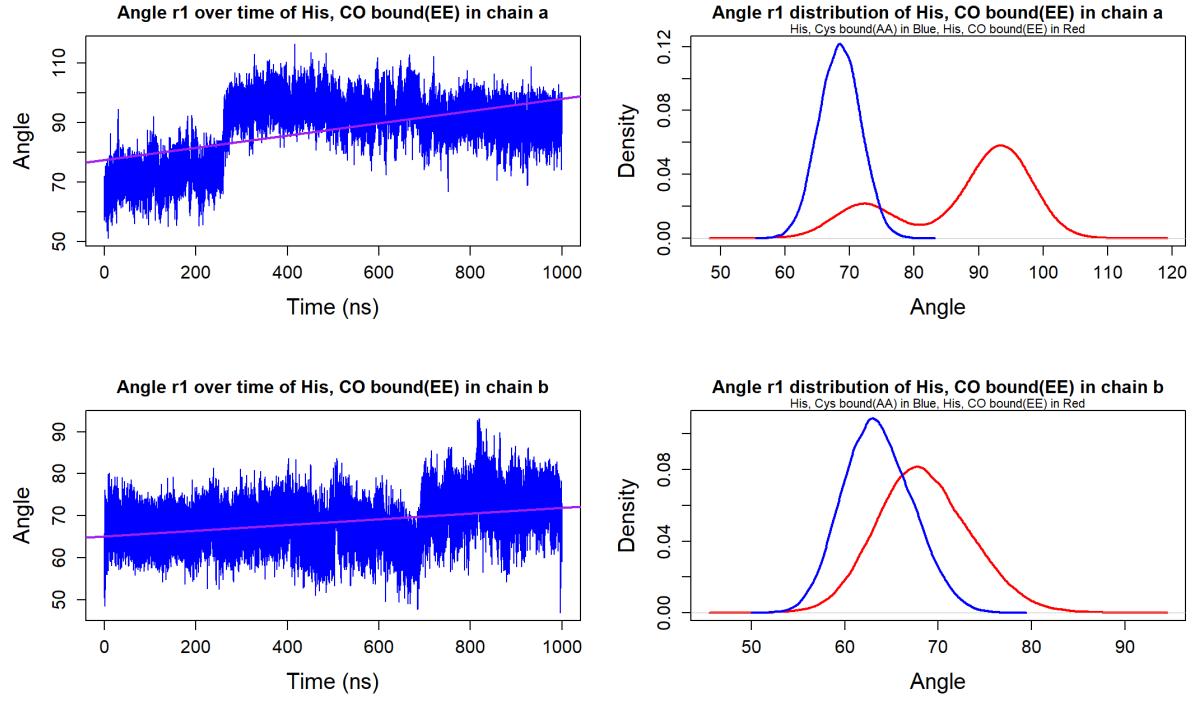


Figure 4.8: Rotational Dynamics of α -Helix 8 in His, CO Bound State (EE) for Chains a and b: Evolution of angle r_1 (left) and its distribution (right) over time. Angle r_1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

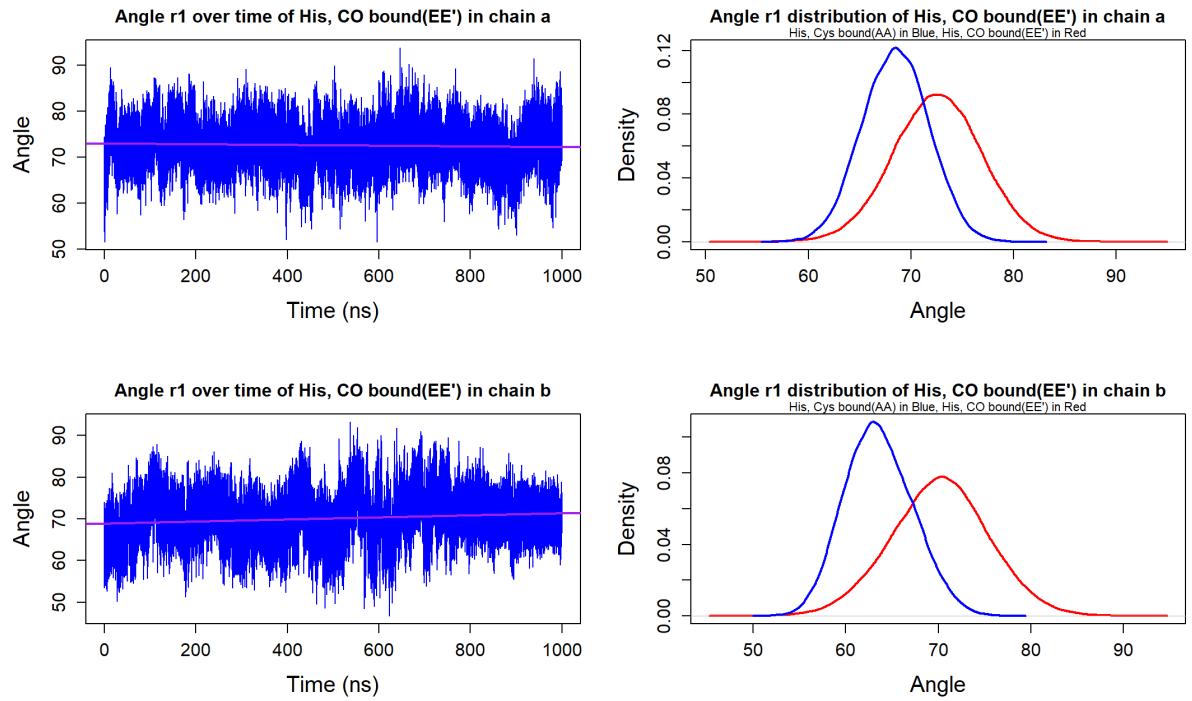


Figure 4.9: Rotational Dynamics of α -Helix 8 in His, CO Bound State (EE') for Chains a and b: Evolution of angle r_1 (left) and its distribution (right) over time. Angle r_1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

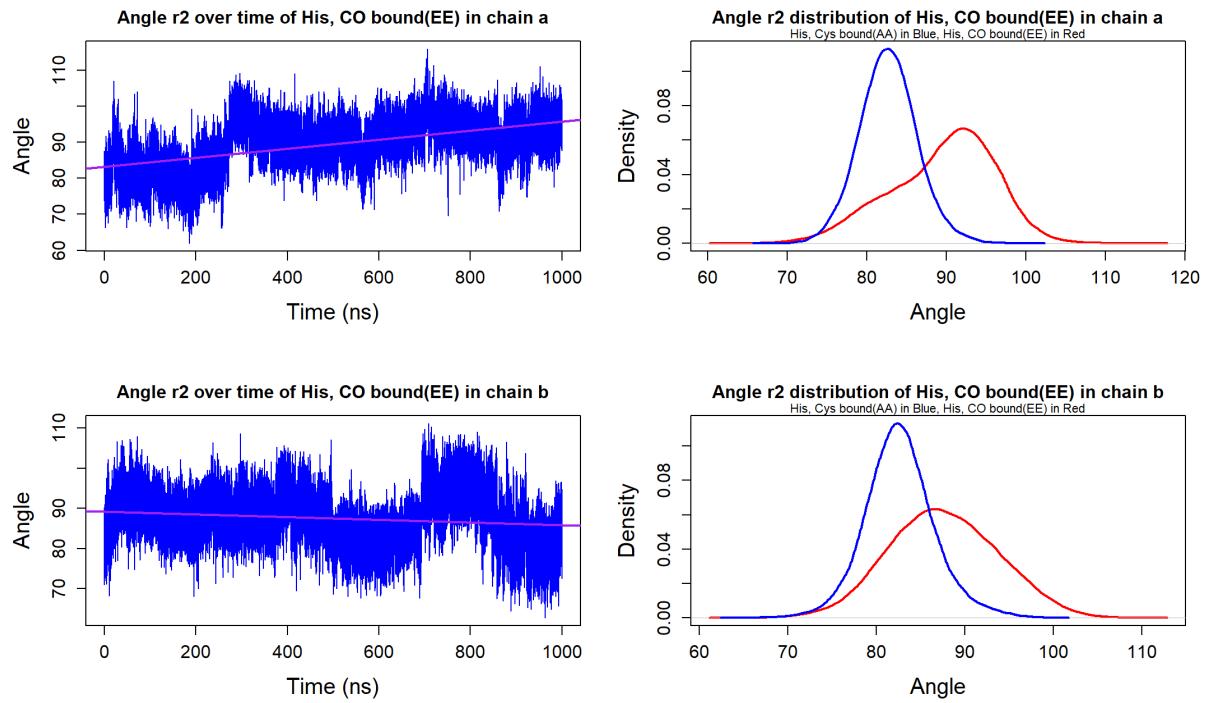


Figure 4.10: Rotational Dynamics of α -Helix 8 in His, CO Bound State (EE) for Chains a and b: Evolution of angle r_2 (left) and its distribution (right) over time. Angle r_2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

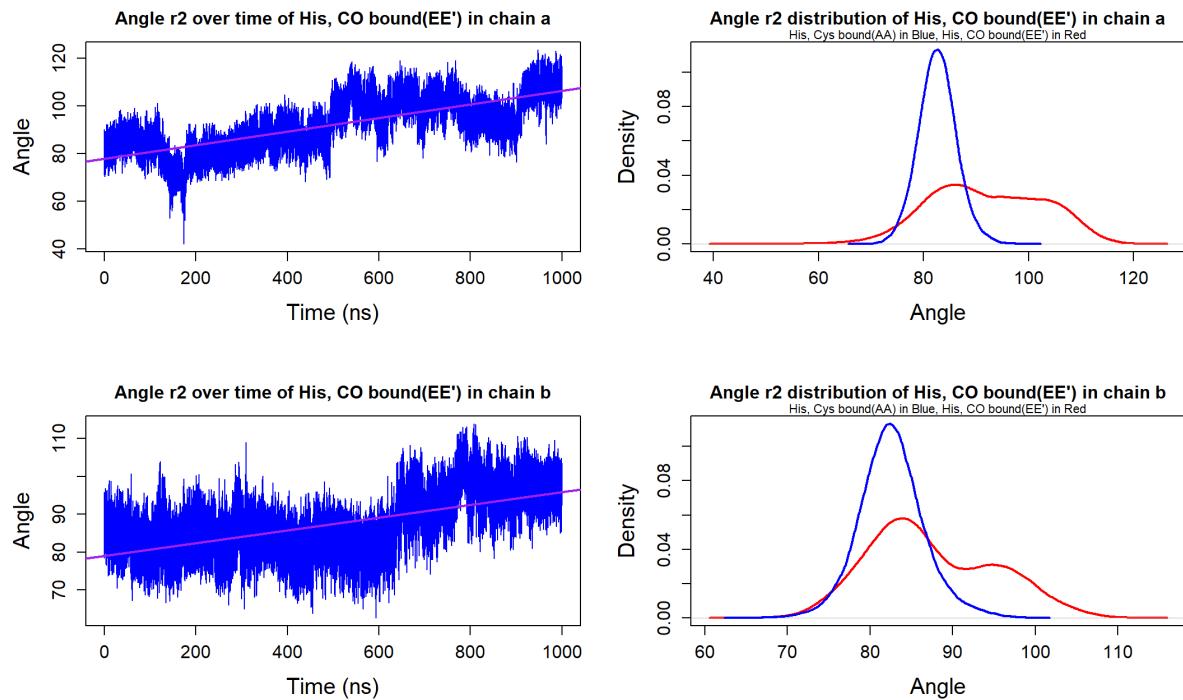


Figure 4.11: Rotational Dynamics of α -Helix 8 in His, CO Bound State (EE') for Chains a and b: Evolution of angle r_2 (left) and its distribution (right) over time. Angle r_2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

to 271) in both chains of FF. However, the RMSF result of FF' did not provide any new information (Figure 6.7b).

Distance analysis was performed for both FF and FF' to gain a deeper understanding of the Heme- α helix 8-PLP region. Starting with the distance (h1) between Cys52-SG and Arg266-NH1 (Figure 4.12a and 4.13a), a similar pattern was observed as in His, CO bound state (EE and EE'), where the breaking of the salt bridge allowed Cys52 to exhibit increased flexibility. The stability of the four hydrogen bonds in state FF was evident (Figure 4.12b). However, in FF', a wider distribution was observed in chain b (Figure 4.13b). Upon examining the structures over time, it was discovered that a similar event occurred as in chain b of state EE', where from 200 to 600ns, Thr257 formed a hydrogen bond with PLP119-O1P (instead of O3P) and Thr260 formed a hydrogen bond with PLP119-O2P (instead of O1P). However, after 600ns, these hydrogen bonds reverted to their primary configuration. It further supports the notion that the position of the three oxygen atoms in the phosphate group of PLP may undergo occasional rotations.

The distance measurements in state FF revealed significant hydrogen bond-breaking events between Asn149-ND2 and PLP119-O (Figure 4.12c), as well as Ser349-OG and PLP119-N1 (Figure 4.12d). In both chains (a and b), the hydrogen bonds (h4 and h5) were observed to break at around 750ns and did not reform throughout the remaining 1000ns simulations. This indicates that the breaking of these hydrogen bonds was stable and persisted over time. However, in chain b, the hydrogen bond-breaking events were not as pronounced as in chain a, with occasional reforming of the hydrogen bonds observed.

Interestingly, despite the hydrogen bond-breaking events, there was no significant translation observed in both chains (a and b) (Figure 6.8a). Instead, the angles (r1 and r2) changed over time, providing insight into the dynamics of the α -helix 8. In chain a, Arg266 started to rotate in a clockwise manner after the salt bridge with Cys52 was disrupted, leading to an increase in the angle r1 from 70° to 90° (Figure 4.14). The rotation motion then propagated to Thr260, resulting in a climbing trend of angle r2 from 750ns (Figure 4.16), potentially explaining the hydrogen bond-breaking events at Asn149 and Ser349.

In chain b, there was a significant angle (r1) jumping from 65° to 90° at 600ns (Figure 4.14), but this was attributed to the reference used, which was the Heme plane rather than Arg266 (shown in Figure 4.21c). This observation highlights the complexity of interpreting certain results in the context of the Heme- α helix 8-PLP region dynamics.

In state FF', further analysis of the distance measurements provided additional insights into the dynamics of the Heme- α helix 8-PLP region. The hydrogen bond between Asn149-ND2 and PLP119-O exhibited distinct behaviors in both chains (a and b) (Figure 4.13c). In chain a, the hydrogen bond broke briefly around 650ns to 700ns and remained stable for the rest of the 1000ns simulations. However, in chain b, the hydrogen bond was broken at the early stage (around 180ns) and reformed for the remaining simulation time. The hydrogen bond between Ser349-OG and PLP119-N1 showed contrasting behaviors in the two chains. In chain b, the hydrogen bond remained consistently broken throughout the 1000ns simulations, while in chain a, no significant breaking event was observed (Figure 4.13d).

Similar to state FF, no translation was observed in both chains of FF', indicating minimal movement of the Heme group away from the α -helix 8 (Figure 6.8b). The rotation of Arg266 in chain a of FF' showed similarities to that in chain a of state FF, where it underwent a clockwise rotation after the salt bridge with Cys52 was disrupted (Figure 4.15). However, no significant rotation was observed for Thr260 in chain a of FF' after 1000ns of MD simulations (Figure 4.17), suggesting that no major conformational changes occurred at Asn149 and Ser349 in this state. In chain b, neither Arg266 nor Thr260 exhibited significant rotational motion.

In summary, His, NO bound state (FF and FF') can be considered as fictitious configurations that were not observed in experiments. From a computational perspective, they can be treated as two "replicates" of His, CO bound state, given the similarity in force field parameters for CO and NO. The RMSF analysis of state FF revealed intriguing findings, demonstrating increased flexibility in regions other than α -helix 8, which warrants further investigation in the "Dynamical Cross-Correlation Matrix" section.

Moreover, the analysis of hydrogen bond-breaking events in Asn149 and Ser349 of state FF after 750ns provides valuable insights. The rotation of Arg266, followed by Thr260, likely contributes to the rotation of the entire PLP group, leading to the breaking of these hydrogen bonds. However, the rotation of the Heme group poses challenges in interpreting some results, and the flexibility of the three oxygen atoms in the phosphate group of PLP further adds complexity to the dynamics.

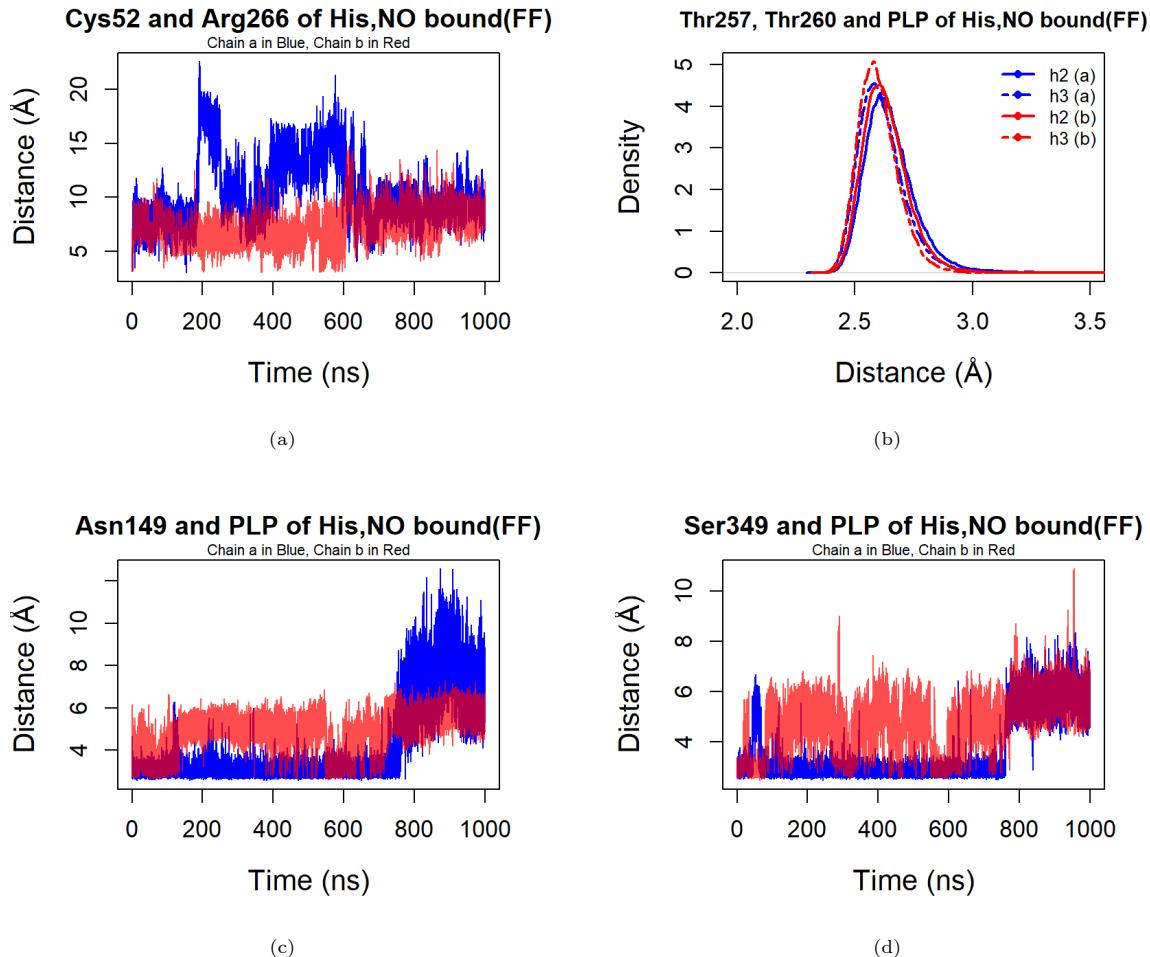


Figure 4.12: Analyzing Distances in His, NO Bound State (FF) for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1: (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

4.2.4 NO bound state (CC)

State CC (Figure 1.14) explores the scenario where both Cys52 and His65 are decoordinated from the Heme group, which may provide insights into the role of His65 in enzyme activity inhibition. Notably, NO bound state (CC) exhibits the highest flexibility, as evidenced by the RMSD (Figure 6.9a) and RMSF (Figure 6.9b) results. Particularly, the Heme binding domain undergoes significant movement upon losing its bond to the Heme group, while the rest of the protein remains relatively stable.

Both Cys52 (Figure 4.18a) and His65 (Figure 6.10) display increased flexibility after losing their bonds with the Heme group. Additionally, the rotation of the three oxygen atoms in the phosphate group of PLP occurs in chain b, leading to Thr257 forming a hydrogen bond with PLP119-O2P (instead of O3P), and Thr260 forming a hydrogen bond with PLP119-O3P (instead of O1P) after 50ns, remaining stable

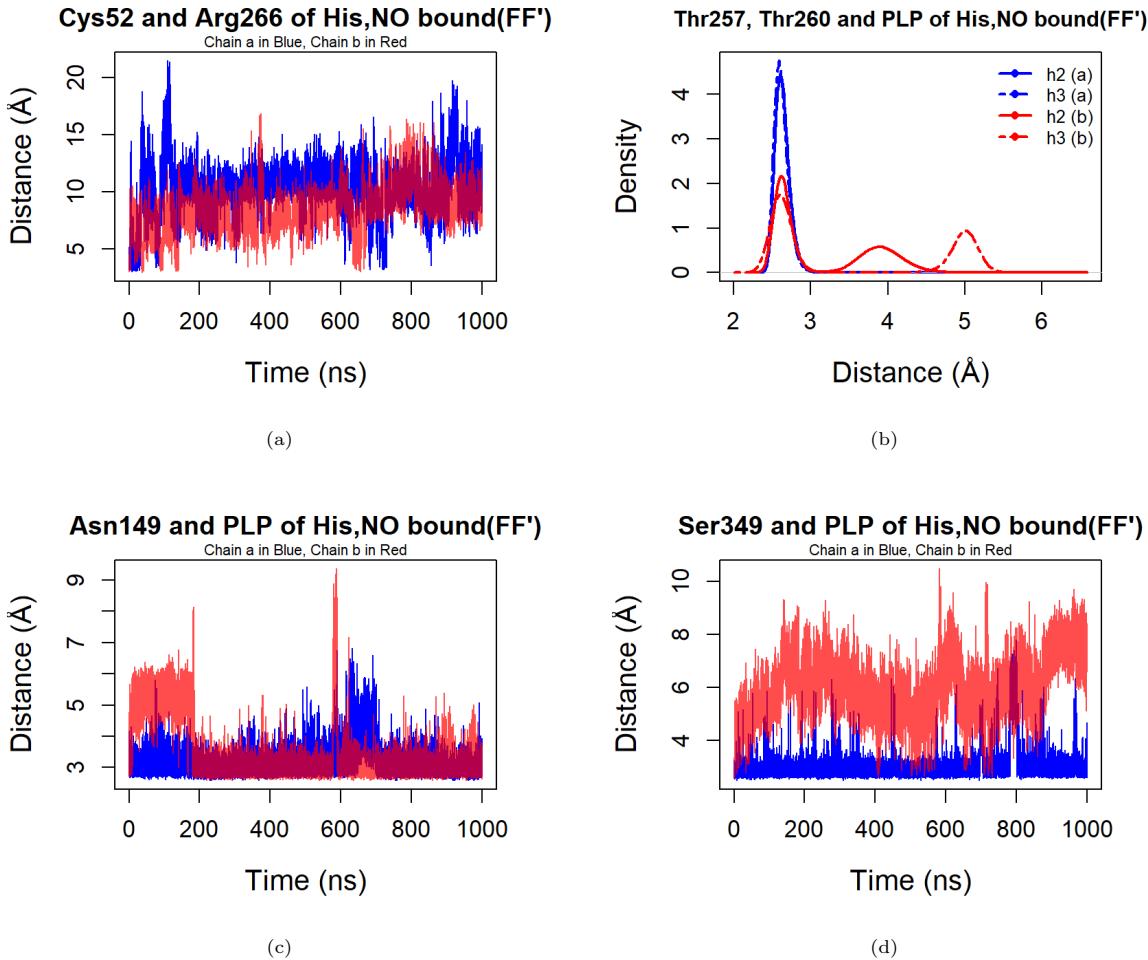


Figure 4.13: Analyzing Distances in His, NO Bound State (FF') for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1; (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

throughout the simulations. Consequently, the distribution of distances (h2 and h3) shifts significantly (Figure 4.18b).

The hydrogen bond between Asn149-ND2 and PLP119-O undergoes breaking between 200ns to 900ns in chain a, while in chain b, it appears to be broken most of the time based on distance measurements over time (Figure 4.18c). Moreover, the hydrogen bond between Ser349-OG and PLP119-N1 remains stable in chain a but may experience a breaking event between 400ns to 600ns, followed by reforming the hydrogen bond (Figure 4.18d).

In state CC, the translation measurements (Figure 6.11) show that both α helices 8 remain in their original locations. However, since the Heme groups are no longer bonded, they are prone to rotation. In chain a of CC, the Heme group undergoes a clockwise rotation of almost 70° (shown in Figure 4.21d), making the interpretation of angles r1 and r2 challenging. Fortunately, in chain b, the Heme group remains relatively stable in its original location after overlapping with the initial structure.

Interestingly, the rotation results (Figure 4.19 and 4.20) for chain b show a decreasing trend in both angles, which is a new observation compared to previous states. At the first 150ns of simulations, angle r2 jumps from 90° to 100° , possibly explaining the rotation of the three oxygen atoms in the phosphate

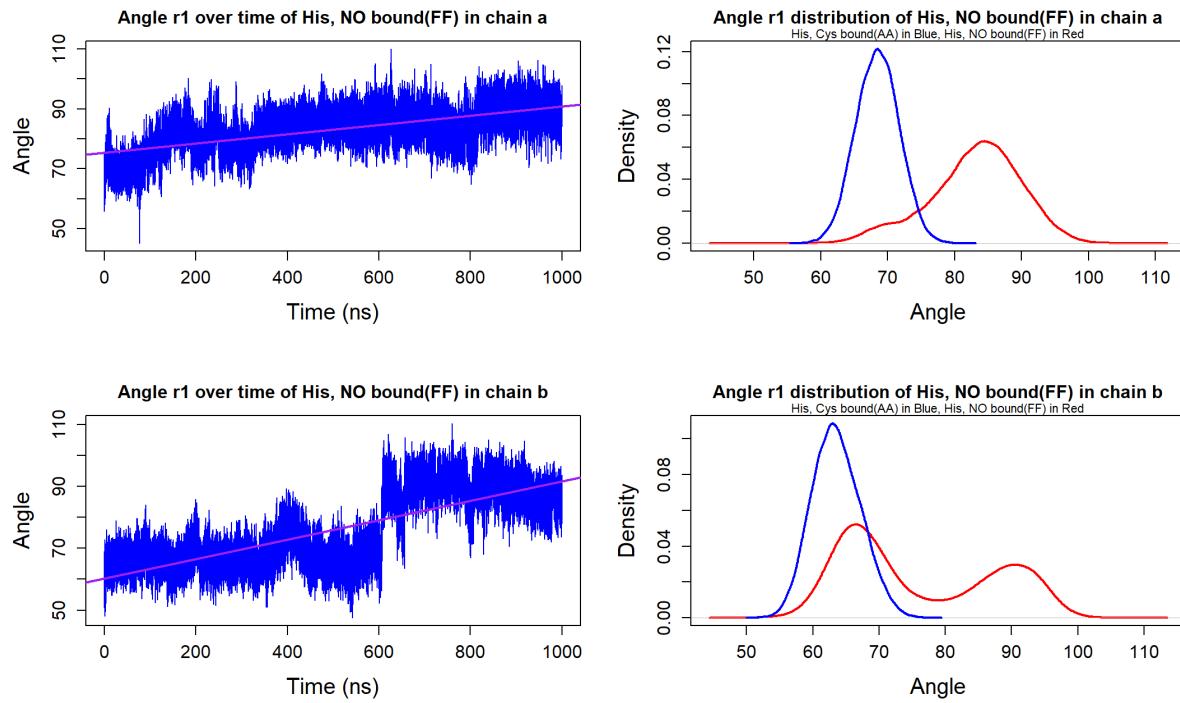


Figure 4.14: Rotational Dynamics of α -Helix 8 in His, NO Bound State (FF) for Chains a and b: Evolution of angle r_1 (left) and its distribution (right) over time. Angle r_1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

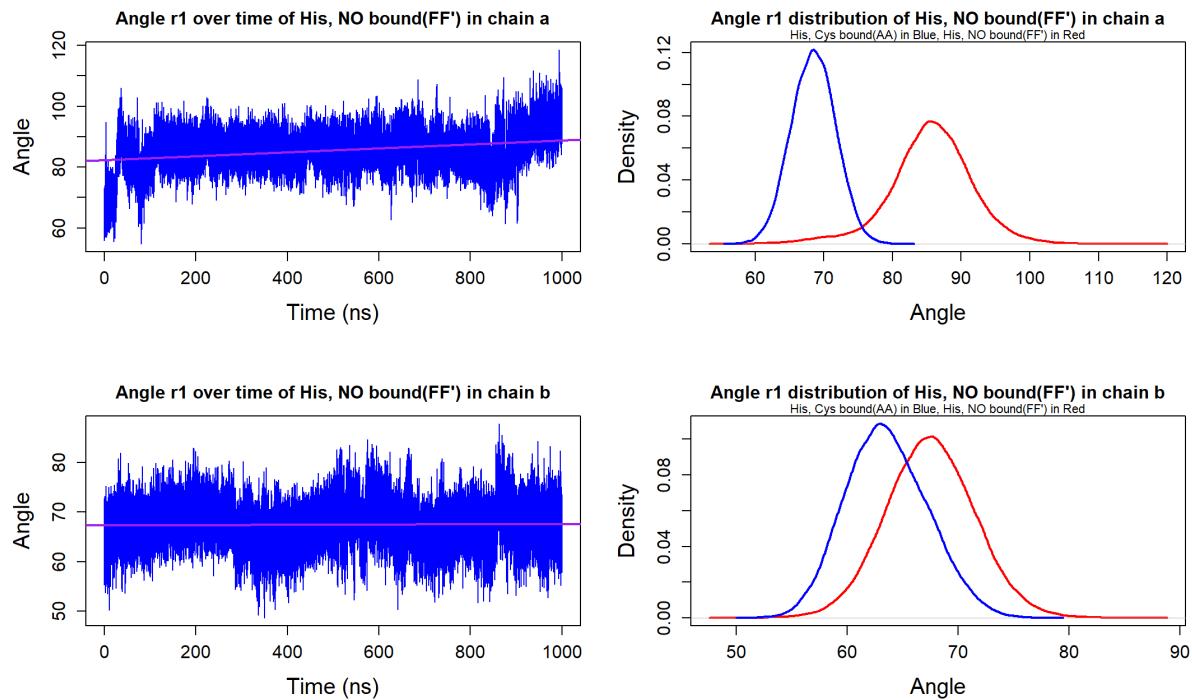


Figure 4.15: Rotational Dynamics of α -Helix 8 in His, NO Bound State (FF') for Chains a and b: Evolution of angle r_1 (left) and its distribution (right) over time. Angle r_1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

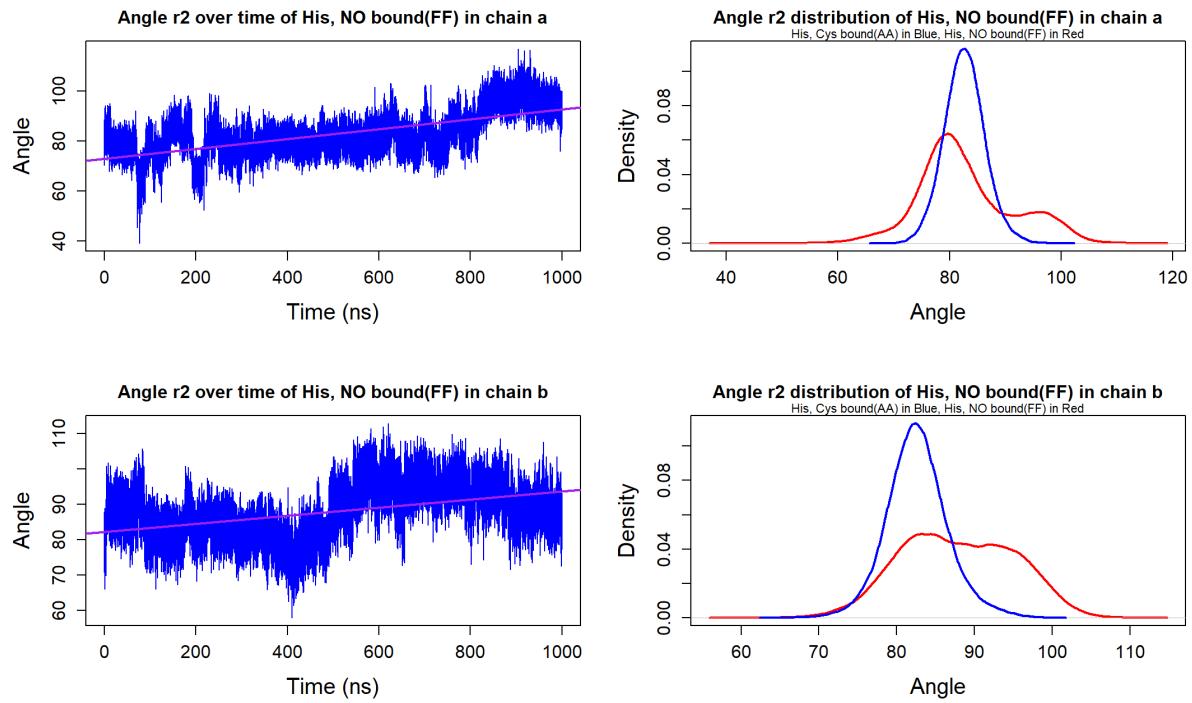


Figure 4.16: Rotational Dynamics of α -Helix 8 in His, NO Bound State (FF) for Chains a and b: Evolution of angle r_2 (left) and its distribution (right) over time. Angle r_2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

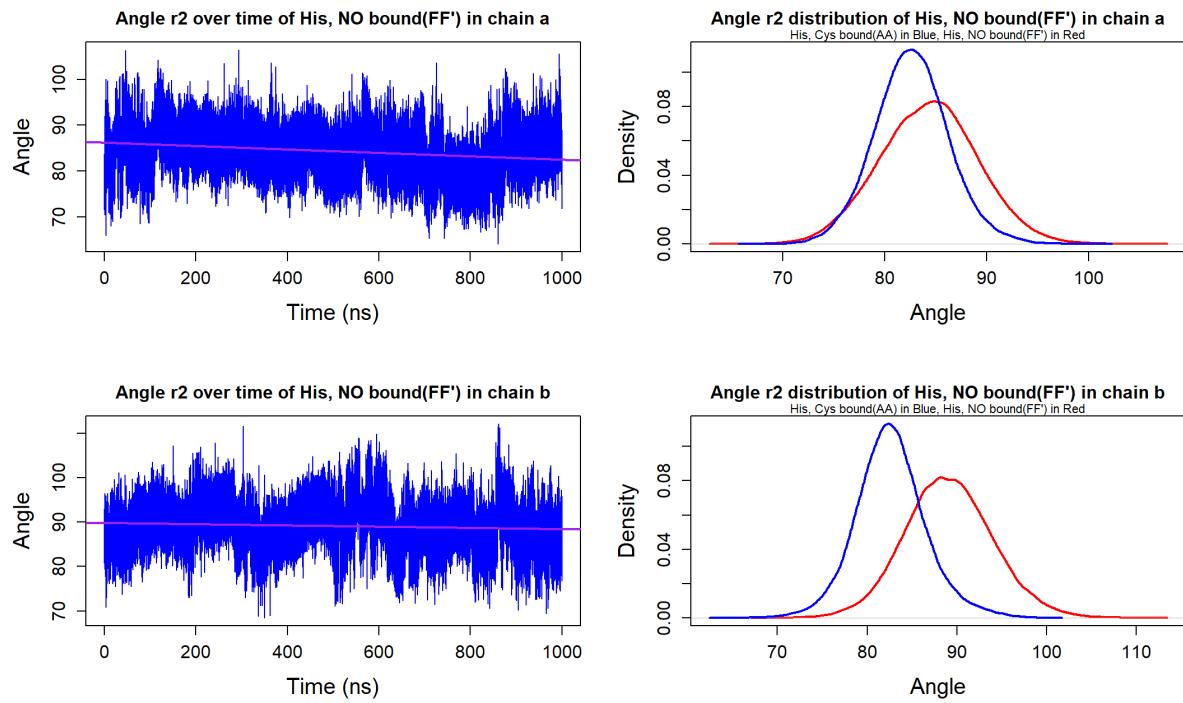


Figure 4.17: Rotational Dynamics of α -Helix 8 in His, NO Bound State (FF') for Chains a and b: Evolution of angle r_2 (left) and its distribution (right) over time. Angle r_2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

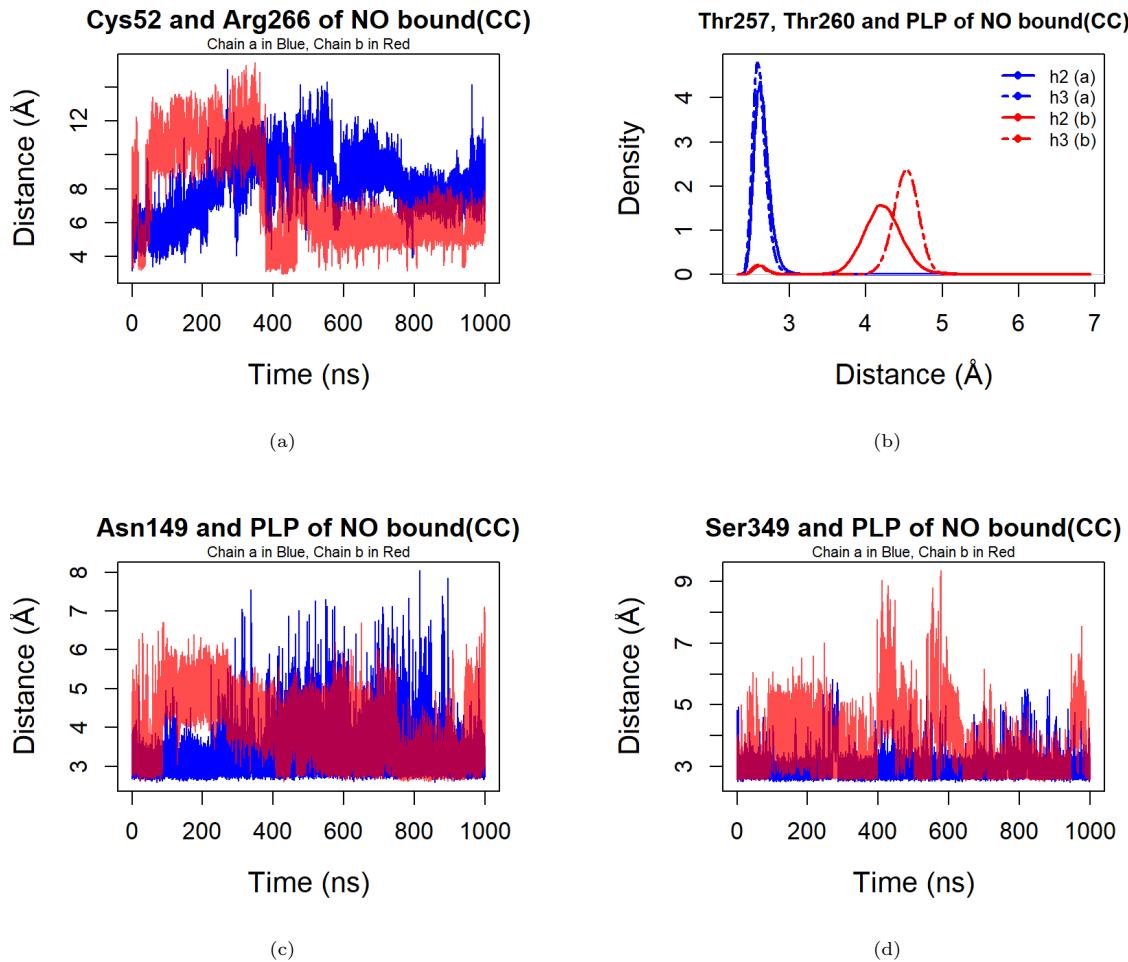


Figure 4.18: Analyzing Distances in NO Bound State (CC) for Both Chains: (a) Dynamic Interaction between 52CYS-SG and 266Arg-NH1; (b) Density Distribution of Key Distances in Chain a and Chain b: h2: Distance between 257Thr-OG1 and 119PLP-O2P (Chain a) h2: Distance between 257Thr-OG1 and 119PLP-O3P (Chain b) h3: Distance between 260Thr-OG1 and 119PLP-O3P (Chain a) h3: Distance between 260Thr-OG1 and 119PLP-O1P (Chain b) (c) Evolving Distance between 149Asn-ND2 and 119PLP-O. (d) Dynamic Distance Analysis: 349Ser-OG to 119PLP-N1 Over Time. Tracked over time for Chain a (blue) and Chain b (red)

group. Subsequently, the angle stays relatively stable for the rest of the simulation. However, the changes in angles r1 and r2 over time, in this case, are difficult to correlate with the hydrogen bond-breaking events at Asn149 and Ser349.

In summary, the analysis of NO bound state (CC) reveals that both Cys52 and His65 can become more flexible after losing their bonds with the Heme group. However, the specific role of the decoordination of His65 in enzyme activity inhibition has not been clearly elucidated based on the current results. The observation of the rotation of the three oxygen atoms in the phosphate group of PLP is consistent with previous findings.

The rotation of the Heme group in chain a presents challenges in interpreting the results for chain b. In chain b, the angle (r1) between the normal of the Heme group and the vector from the Fe^{2+} on the Heme to the CA of Arg266 decreases from 75° to 55° in the first 400ns, which is an abnormal behavior compared to previous observations. Similarly, the angle r2 between the normal of the Heme group and the vector from the Fe^{2+} on the Heme to the CA of Thr260 jumps from 90° to 100° in the first 150ns, which is also unusual. The behavior of angles r1 and r2 over time in chain b is challenging to correlate with the hydrogen bond-breaking events at Asn149 and Ser349, as observed in previous states.

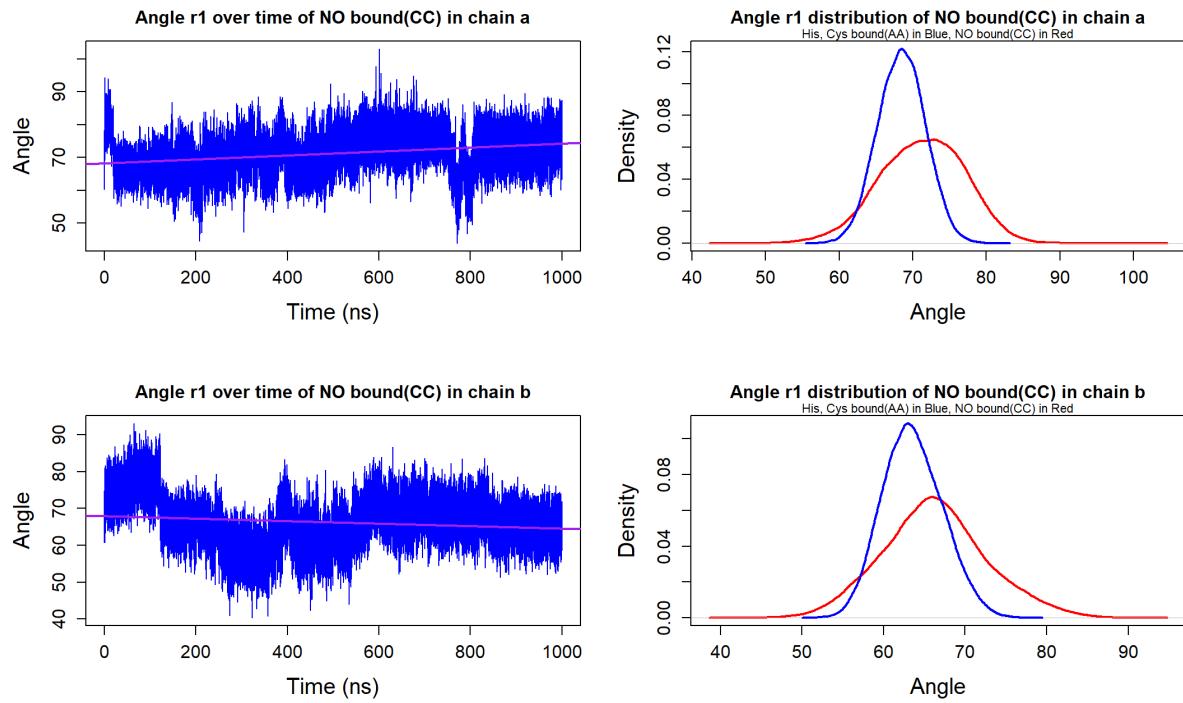


Figure 4.19: Rotational Dynamics of α -Helix 8 in NO Bound State (CC) for Chains a and b: Evolution of angle r1 (left) and its distribution (right) over time. Angle r1 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Arg266-CA.

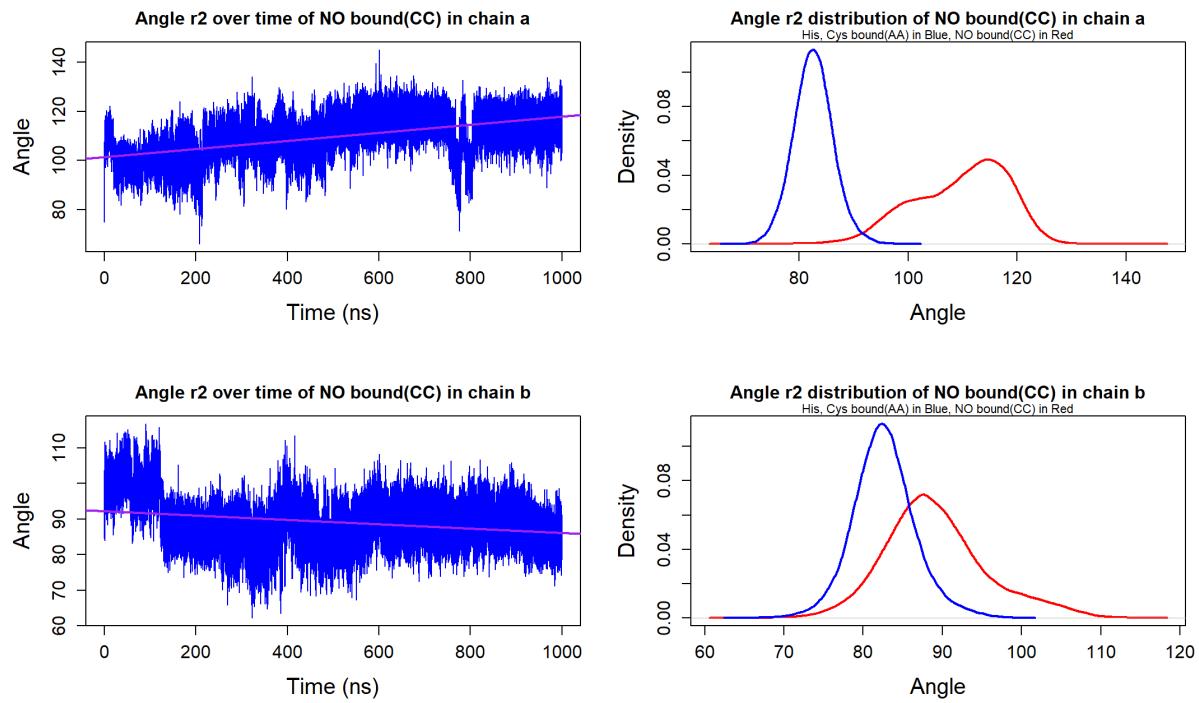


Figure 4.20: Rotational Dynamics of α -Helix 8 in NO Bound State (CC) for Chains a and b: Evolution of angle r2 (left) and its distribution (right) over time. Angle r2 represents the angle between the Heme plane normal and the vector extending from Heme Fe to Thr260-CA.

4.2.5 Discussion

In summary, the analysis of His, CO bound state (EE and EE'), His, NO bound state (FF and FF'), and NO bound state (CC) compared to the initial His, Cys bound state supports the hypothesis [53] that the protonation of Cys52-SG after its decoordination from the Heme group breaks the salt bridge between Cys52-SG and Arg266-NH1, allowing Cys52-SG and Arg266-NH1 to move more freely. According to previous studies, the inhibitory effects of NO and CO on the enzyme activity can be rapidly reversed by oxidation from O₂, leading to the restoration of the iron to its ferric state[81]. Our simulation results (Figure 4.6a, 4.7a, 4.12a, 4.13a and 4.18a) demonstrate that Cys52 exhibits significant movement, with a displacement of up to 10 Å away from the iron. This implies that it may take a considerable amount of time for Cys52 to return to its original position and undergo deprotonation by O₂.

The breaking salt bridge indicates that α -helix 8 may experience a slight displacement from the Heme group, possibly smaller than 1 Å, although this observation is not entirely conclusive.

Although His, CO bound state, His, NO bound state, and NO bound state share different configurations, the regulation mechanism should be the same. Based on all results demonstrated in Section 4.2, it is sure that the salt bridge breaking enables α -helix 8 free to rotate, with Arg266 rotating in a clockwise manner, transmitting the motion to Thr260 via internal hydrogen bonds within the helix. This rotation subsequently affects the PLP group, leading to the breaking of hydrogen bonds between Asn149-ND2 and PLP119-O, as well as Ser349-OG and PLP119-N1, through strong interactions within the phosphate binding loop and phosphate group of PLP. However, the hydrogen bonds somehow will break due to the fluctuation maybe and tend to reform quickly after breaking (since hydrogen jumping event can not be simulated during the simulations), making the interpretation of results more challenging.

The naive assumption of the Heme group's plane normal remaining stable throughout the simulations was a limitation in accurately measuring the rotation of Arg266 and Thr260. As observed in chain a of state EE', chain b of state FF, and chain a of state CC, the Heme group experienced significant flipping due to the high variability of the loop regions to which it is attached (Figure 4.21). This flipping of the Heme group can introduce uncertainties and complexities in interpreting the rotation angles of Arg266 and Thr260.

The proposed regulation mechanism is primarily based on the hypothesis [53] and is supported by multiple MD simulations, especially those from chain a of His, NO bound state (FF). In this case, the salt bridge between Cys52-SG and Arg266-NH1 is broken, leading to the immediate and continuous clockwise rotation of Arg266. The rotational motion then propagates to Thr260, causing it to rotate in a similar manner. Consequently, the PLP also undergoes rotation, resulting in the breaking of the two hydrogen bonds between PLP and Asn149, as well as PLP and Ser349 after 750ns of simulation. However, there are instances where no clear causal relationships are observed. For example, in chain b of His, CO bound state (EE), the rotation of both Arg266 and Thr260 occurs around 680ns, while the breaking of the hydrogen bond between PLP and Ser349 starts from 560ns, indicating that the rotation happened after the breaking event.

In some cases, such as chain b of His, CO bound state (EE'), and both chains of His, NO bound state (FF'), specific residues did not exhibit rotation during the 1ms simulations. This may be attributed to the fact that 1ms might not be sufficient for rotation to occur in these cases. Similarly, the diverse results obtained from His, CO bound state, His, NO bound state, and NO bound state despite starting from the same His, Cys bound state (AA) may be due to different initial structures after energy minimization, NVT, and NPT steps. Additionally, the random initialization of initial velocities for the simulations may contribute to different behavior among different runs. Conducting more replicates and applying enhanced sampling methods may provide a better understanding of their behaviors.

An interesting observation is the increased flexibility in certain regions (residues 192 to 198, 92 to 103, and 328 to 367) in chain a of His, NO bound state (FF), which is not observed in other states. These regions may be highly correlated with the complete breaking of hydrogen bonds between Asn149-ND2 and PLP119-O, and between Ser349-OG and PLP119-N1. To further understand their correlation, dynamical cross-correlation analysis was conducted.

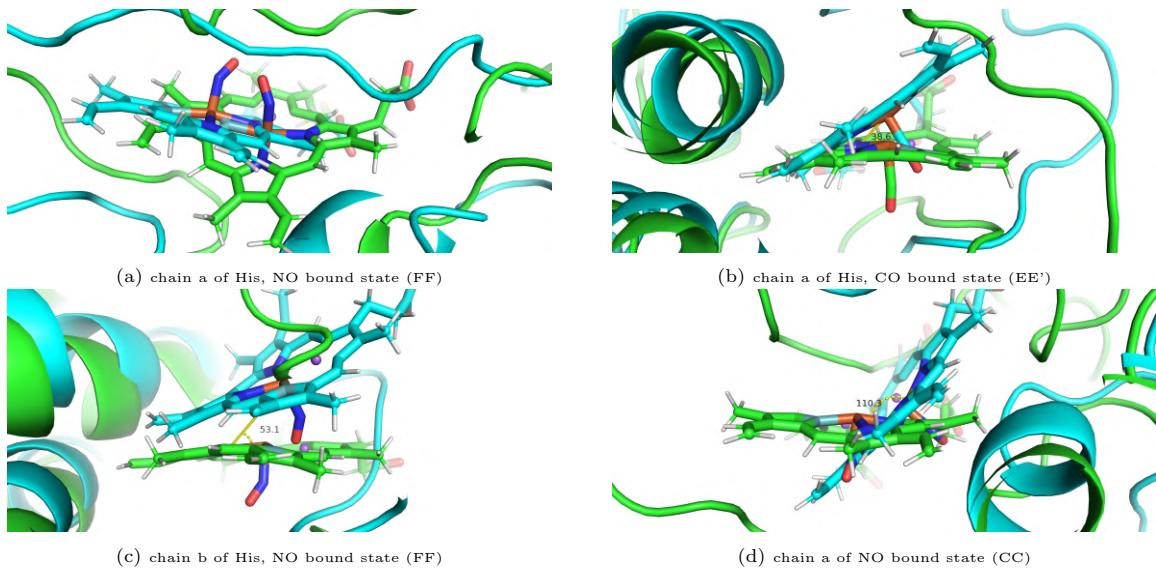


Figure 4.21: Comparing Heme group overlap at the start (Green) and end (Cyan) of 1000ns MD simulations. (a) Represents a representative reference configuration from state FF for chain a, showing minimal flipping. (b) The Heme group rotated 38.6 degrees compared to the initial structure. (c) The Heme group rotated 53.1 degrees compared to the initial structure. (d) The Heme group rotated 69.7 degrees compared to the initial structure

4.3 Dynamical cross correlation analysis

The dynamical cross-correlation analysis was performed on the full trajectory of His, Cys bound state (AA) and the last 200 ns trajectory of His, NO bound state (FF). In Figure 4.22a and 4.22b, both chains of the CBS protein were plotted in the same graph. However, the residues numbering was automatically renumbered from 1, rather than starting from residue 43 as in our case. To understand the results, the corresponding residue IDs in the DCCM plot and the PDB file were listed in the Table 4.1.

In state AA, two highly positively dynamic correlated blocks were observed, especially in chain b (Figure 4.22a). One block spans from residue No. 410 to 520, and the other from residue No. 545 to 680. These blocks are not continuous but show significant dynamic correlations. In chain a, the first block shrinks to residues No. 100 to 180, and the second block splits into two smaller blocks with residues No. 180 to 225 and No. 280 to 340. The correlations in these regions are generally not very high, as indicated by the colors in the figure.

After introducing the numbering to the Heme group and breaking the salt bridge between Cys52-SG and Arg266-NH1, as well as the two hydrogen bonds between Asn149-ND2 and PLP119-O and between Ser349-OG and PLP119-N1, these positively correlated regions in His, NO bound state (FF) show enhanced dynamical correlations. The color in the DCCM plot indicates higher correlation values, represented by darker blue colors, which are above 75% (Figure 4.22b).

Table 4.1: Corresponding Secondary Structure in PDB File

Residue No. in DCCM plot	Residues ID in PDB file	Corresponding Secondary structure
99–107	141–149	β 3, and the start of α 3 (Asn149)
150–156	192–198	Two β turns between β 5 and α 5
50–61	92–103	α 1
286–325	328–367	α 9 and α 10 with a β turn (Ser349)

The dynamical correlation analysis further explored the relationship between residues 192 to 198 (two beta turns between β 5 and α 5) and the region where Asn149 is located (β 3 and the start of α helix 3). It was observed that the motion of the two beta turns is highly correlated with β 4 (residues 164 to 170) and shows some correlation with β 3, where Asn149 is connected (Figure 4.23a and 4.23b). Although the

dynamical correlation is not extremely high, it supports the assumption that the distinct peaks observed in the RMSF of state FF of chain a for residues 192 to 198 (Figure 6.6b) are related to the hydrogen bond-breaking event at Asn149. The motion of β 3, where Asn149 connects to, is also correlated with the motion from residues 143 to 223, with some parts showing higher correlations that are enhanced in state FF (Figure 4.23a and 4.23b), especially residues 163 to 169 (β 4) and α helix 5 and β 6 (from residues 198 to 222).

Similarly, the RMSF result of state FF of chain a (Figure 6.6b) revealed another two peaks for residues 92 to 103 and residues 328 to 367, where Ser349 is located. The dynamical cross-correlation matrix (DCCM) analysis (Figure 4.24a and 4.24b) confirmed positive dynamical correlations, especially for residues 290 to 300 with residues 50 to 60. This suggests that the motion of α helix 1 and the region containing α helices 9 and 10 with a β turn are linked, which may be influenced by the breaking of the hydrogen bond between Ser349-OG and PLP119-N1.

To gain further insight into the structural changes, mode vectors were used to represent the moving direction of these residues between the starting and ending frames of the 1000ns simulations. It is evident that all these α helices and β sheets complexes (residues 143 to 223) shifted upwards, particularly those loop regions (Figure 4.25a). This can be explained by the loss of the hydrogen bond between Asn149-ND2 and PLP119-O, which leads to a decrease in constraints and allows the complex to move upward to some extent. The enhanced motion of residues 192 to 198 and the correlated movements of nearby regions provide additional evidence of the dynamic changes induced by the hydrogen bond-breaking event at Asn149. However, it's important to note that this significant motion might be an artifact of the simplified MD simulations, as the Bateman module on the top of the catalytic core was removed, and the basal state, where two CBS motifs are directly above the α helices and β sheets complexes, may be more relevant. In the basal state, these complexes might be limited in their motion by the presence of the CBS motifs. Nevertheless, the observed motion of the "Asparagine loop" (Figure 1.6 (c)), residues 146 to 152, moving away from the PLP could have important implications for enzyme activity. This motion may reduce the likelihood of interactions between the loop and the substrate serine and hinder the stabilization of the intermediate aminoacrylate [52] during the reaction mechanism (Figure 1.10), potentially influencing the protein's catalytic activity.

Additionally, another motion was observed in two residues regions: one around α helix 1 (residues 92 to 103) and the other around α helices 9 and 10 with a β turn (residues 328 to 367), which moved downwards simultaneously (Figure 4.25b). This movement can also be attributed to the breaking of the hydrogen bond between Ser349-OG and PLP119-N1. Block 4 (residues 347 to 350) in Figure 1.6 is responsible for stabilizing the PLP by holding its pyridine ring, especially during the aldimine interconversion process[52]. However, with the hydrogen bonds between Ser349-OG and PLP119-N1 broken, the proper positioning and stabilization of the PLP will be disrupted: the enzyme may have difficulty transitioning through the different reaction intermediates during the catalytic cycle, potentially hindering its ability to efficiently perform its enzymatic function.

Overall, the dynamical cross-correlation analysis provides valuable insights into the interplay of motions within the CBS protein and highlights the structural changes induced by hydrogen bond-breaking events.

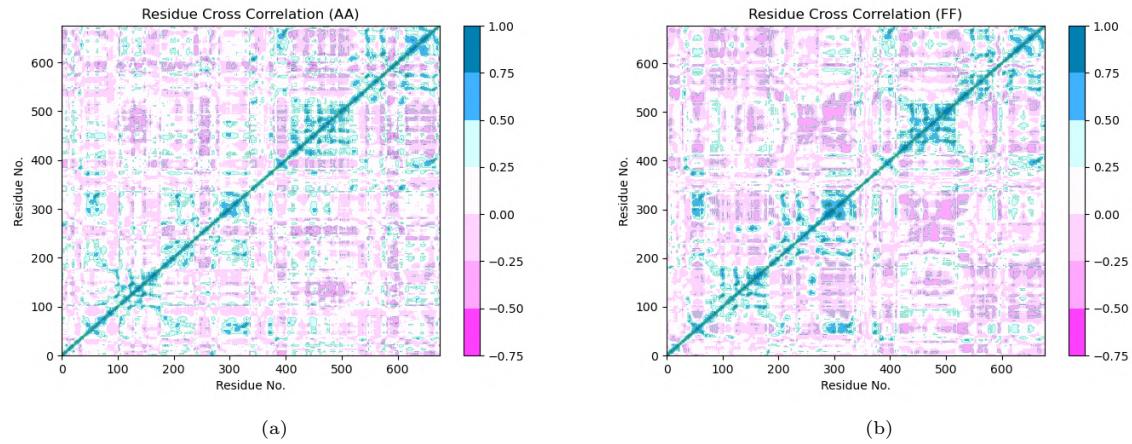


Figure 4.22: DCCM comparison between AA (a) and FF (b) for chain a (residues No. 1 to 339) and chain b (residues No. 340 to 678).

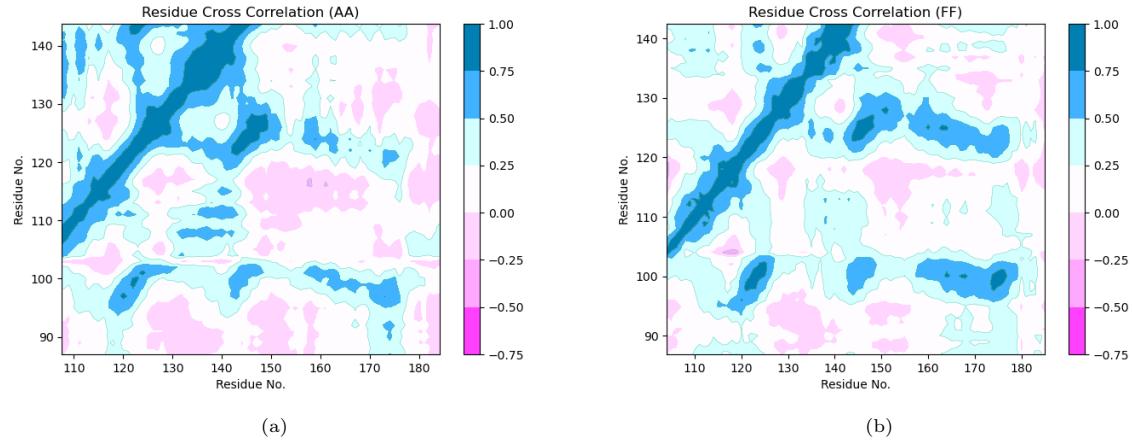


Figure 4.23: Zoomed-in DCCM comparison between AA (a) and FF (b) for chain a. x-axis: residues No. 110 to 180, y-axis: residues No. 90 to 140

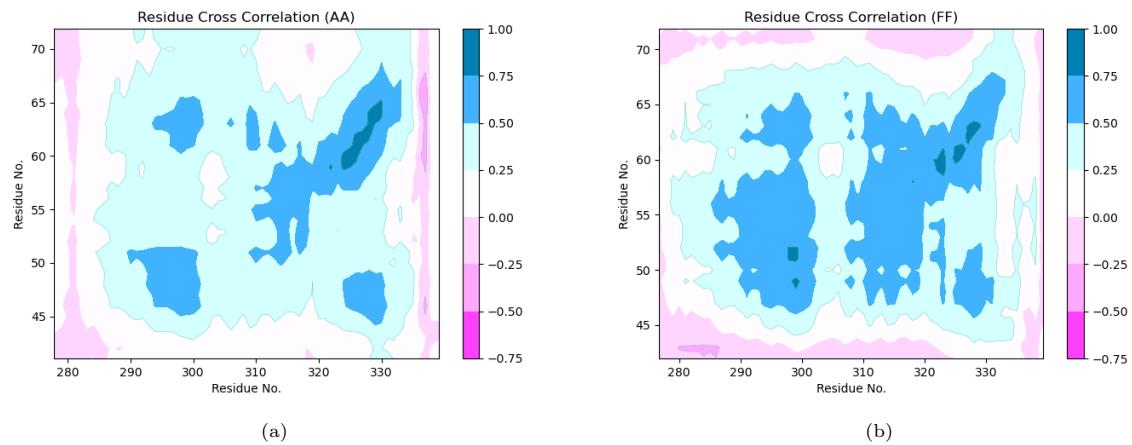


Figure 4.24: Zoomed-in DCCM comparison between AA (a) and FF (b) for chain a. x-axis: residues No. 280 to 339, y-axis: residues No. 45 to 70

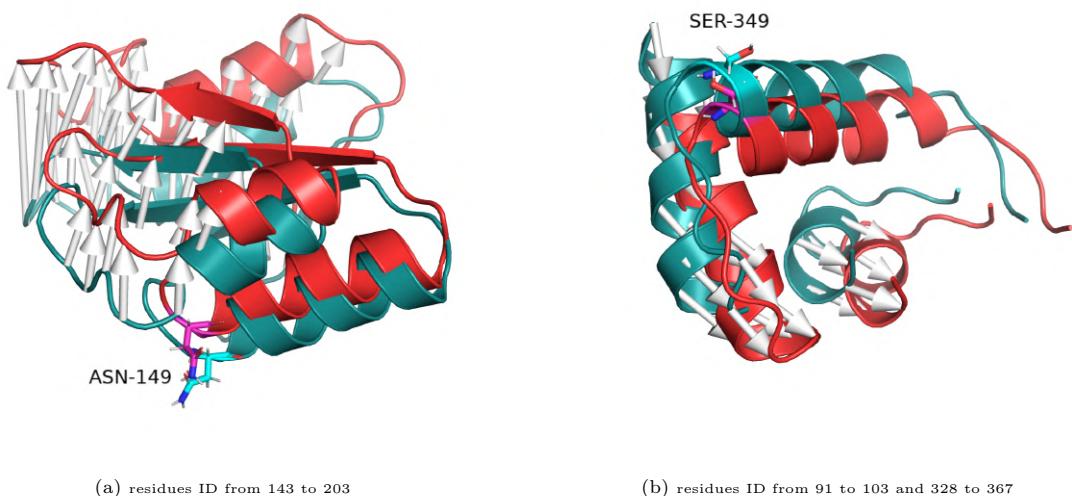


Figure 4.25: Motion of α helices and β sheets complex upon hydrogen bond breaking: (a) Asn149-ND2 and PLP119-O, (b) Ser349-OG and PLP119-N1

Chapter 5

Conclusion

In this comprehensive study, we employed both multiple sequence alignment (MSA) and classical molecular dynamics (MD) simulations to investigate the regulation mechanism of NO/CO on human Cystathionine beta synthase (CBS). Our primary hypothesis was that a bidirectional communication exists between the Heme group and the catalytic core, PLP, via α helix 8. The MSA results revealed a high conservation of proposed residues involved in the communication between the Heme and PLP among CBS proteins from various organisms, including insects and mammals, suggesting their potential roles in regulating enzyme activity.

To explore the regulatory effects of NO/CO on CBS, different states of the enzyme were subjected to extensive 1ms MD simulations. Through various analytical methods, we gained deeper insights into the regulatory mechanism (Figure 5.1). Our analysis indicated that upon decoordination from the Heme group, Cys52's sulfur atom becomes exposed to the solvent, leading to its prompt protonation. This protonation event causes the disruption of the salt bridge between Cys52-SG and Arg266-NH1, resulting in the release of constraints on α helix 8, where Arg266 is located. Subsequently, the rotation motion initiated at Arg266 propagates along the helix until it reaches the end, where Thr260 and surrounding residues form multiple hydrogen bonds with the phosphate group of PLP. This rotational motion induces a corresponding rotation of PLP, breaking the hydrogen bonds between Asn149-ND2 and PLP119-O and between Ser349-OG and PLP119-N1. Consequently, PLP119-O becomes available to accept the internal hydrogen from PLP119-NZ, leading to the deactivation of PLP into its enolimine tautomer.

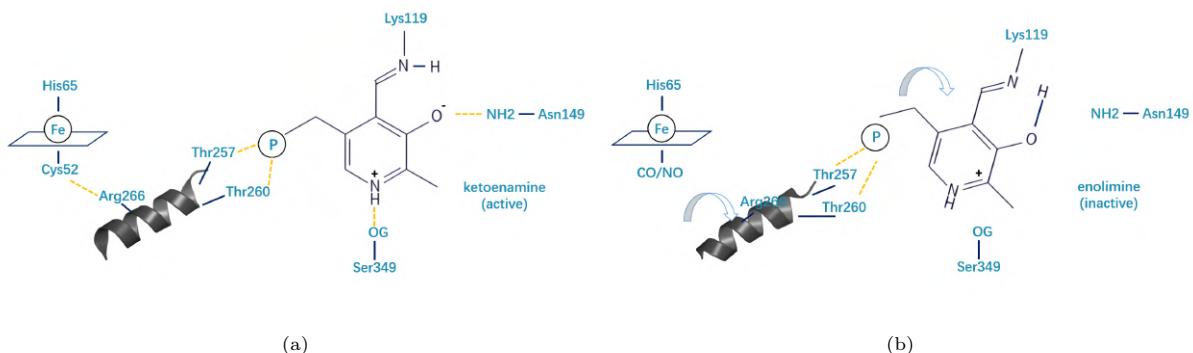


Figure 5.1: Revealing the underlying regulation mechanism begins with the initial state (a), where the Heme-Fe binds to both the His and Cys residues. (b): Upon introducing CO/NO to the Heme group, Cys52 undergoes decoordination and protonation due to its proximity to the solvent, breaking the salt bridge interaction between 52CYS-SG and 266Arg-NH1. This frees Cys52 and Arg266, initiating rotational movement. Arg266's rotation drives the rotation of α -helix 8 and PLP, aided by strong hydrogen bonds from Thr257 and Thr260. Consequently, key hydrogen bonds, Asn149-ND2 and PLP119-O, and Ser349-OG and PLP119-N1, break. This enables PLP119-O to accept hydrogen from PLP119-NZ, deactivating PLP into its enolimine tautomer form—a finely orchestrated process dictating the system's dynamic behavior.

The breaking of the hydrogen bond between Asn149-ND2 and PLP119-O also triggers an upward motion

of a complex comprising multiple helices and sheets (residues 143 to 203). Within this complex, the "Asparagine loop" (residues 146 to 152), which is responsible for interacting with the substrate serine, may lose its proper interaction capability. Similarly, the motion of Block 4 (residues 347 to 350), known to stabilize PLP by holding its pyridine ring during aldimine interconversion, shows a downward motion upon the hydrogen bond between Ser349-OG and PLP119-N1 breaking. As a result, these residues may fail to adequately stabilize PLP in the correct position for effective interaction with substrates, thus affecting the aldimine interconversion process.

In conclusion, our study has provided valuable insights into the complex regulatory mechanism of NO/CO on CBS. The bidirectional communication between the Heme group and PLP, mediated by α helix 8, is a key player in modulating enzyme activity. However, it is important to acknowledge that not all MD simulations fully support this proposed regulation mechanism, as the breaking of the two hydrogen bonds (between Asn149-ND2 and PLP119-O, and between Ser349-OG and PLP119-N1) is not always directly induced by the rotation of α -helix 8. The reformation of hydrogen bonds during MD simulations, due to the limitations in simulating internal hydrogen jumping, presents challenges in interpreting the results accurately.

To better characterize the elusive motion of α helix 8, more advanced quantitative methods are required to accurately describe the rotation of α helix 8 or even PLP. Advanced computational techniques and enhanced sampling methods could provide a clearer understanding of the regulation mechanism.

In addition, conducting more replicates or employing enhanced sampling methods will be crucial to further confirm or explore new insights into the CO/NO regulation mechanism on CBS.

Chapter 6

Appendix

6.1 Abbreviation

- AdoMet: S-adenosyl methionine
- CO: Carbon monoxide
- CSE: Cystathionase
- CP: cysteine-proline
- COM: center of mass
- COG: center of geometry
- DCCM: Dynamic cross-correlation matrix
- H₂S: Hydrogen sulfide
- hCBS: human Cystathionine-beta-synthase
- hcy: homocysteine
- MD simulations: Molecular dynamics simulations
- NO: Nitric oxide
- PBC: Periodic Boundary Conditions
- PLP: pyridoxal 5'-phosphate
- QM/MM: Quantum mechanics/Molecular Mechanics
- RMSD: root-mean square deviation
- RMSF: root-mean square fluctuation
- SASA: solvent accessibility surface area
- SNPs: single nucleotide polymorphisms

6.2 Supplementary Figures

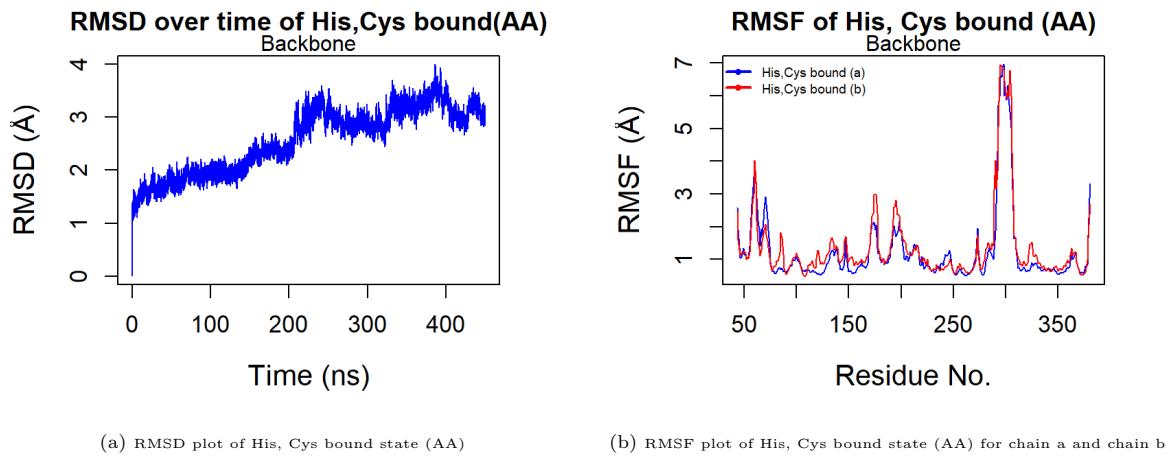


Figure 6.1: General analysis of His, Cys bound state (AA) for chain a and chain b

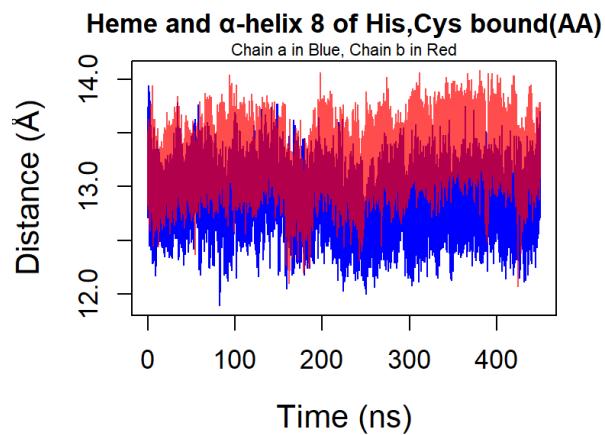
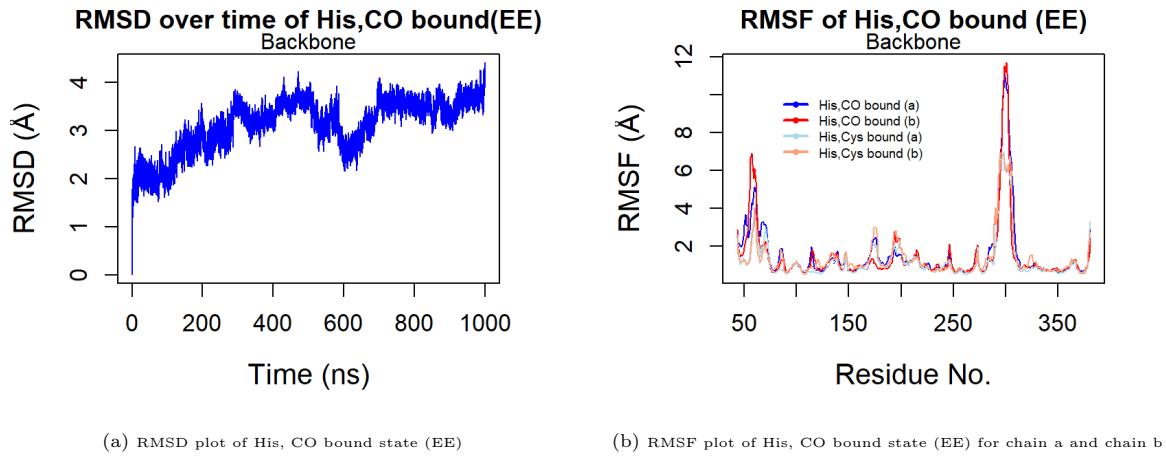
Figure 6.2: Translation measurements of α helix 8 in His, Cys bound state (AA) for chain a and chain b

Figure 6.3: General analysis of His, CO bound state (EE) for chain a and chain b

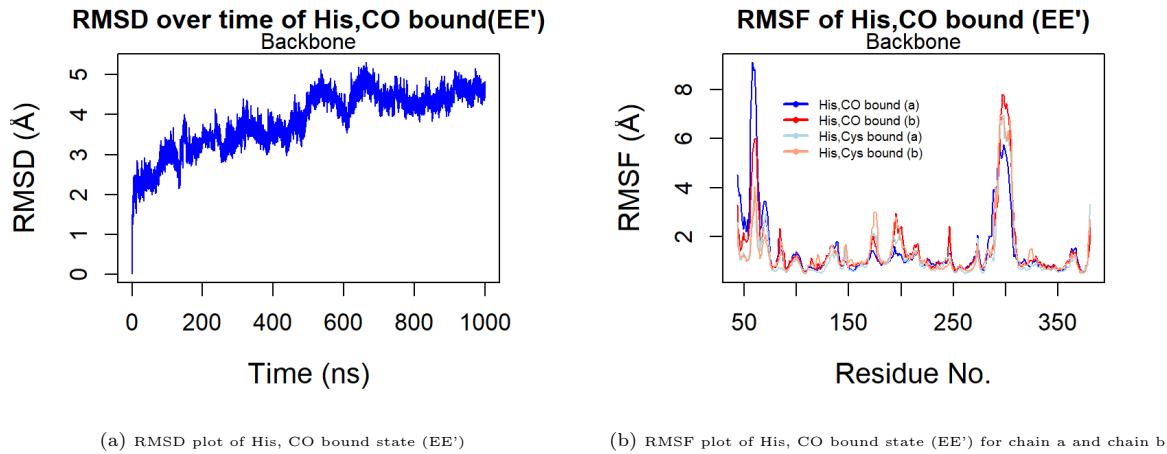


Figure 6.4: General analysis of His, CO bound state (EE') for chain a and chain b

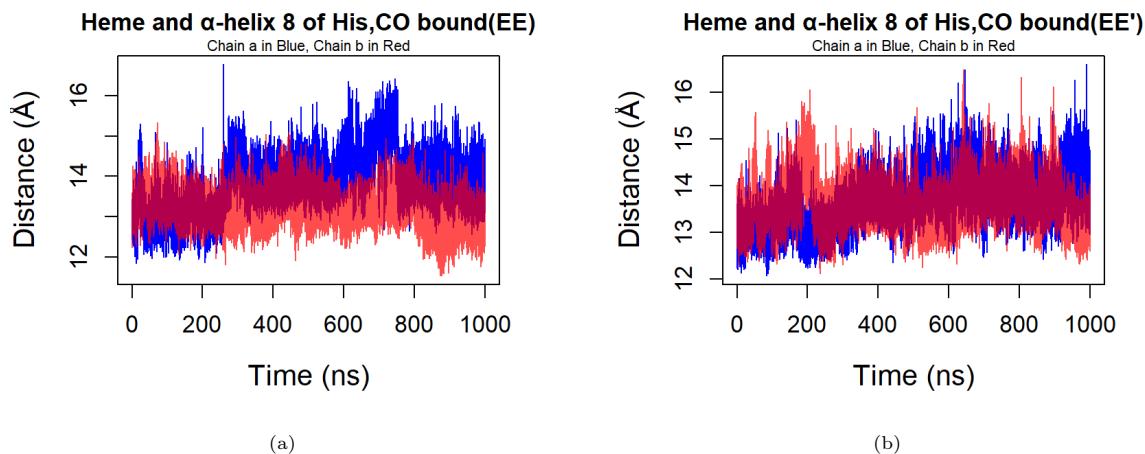
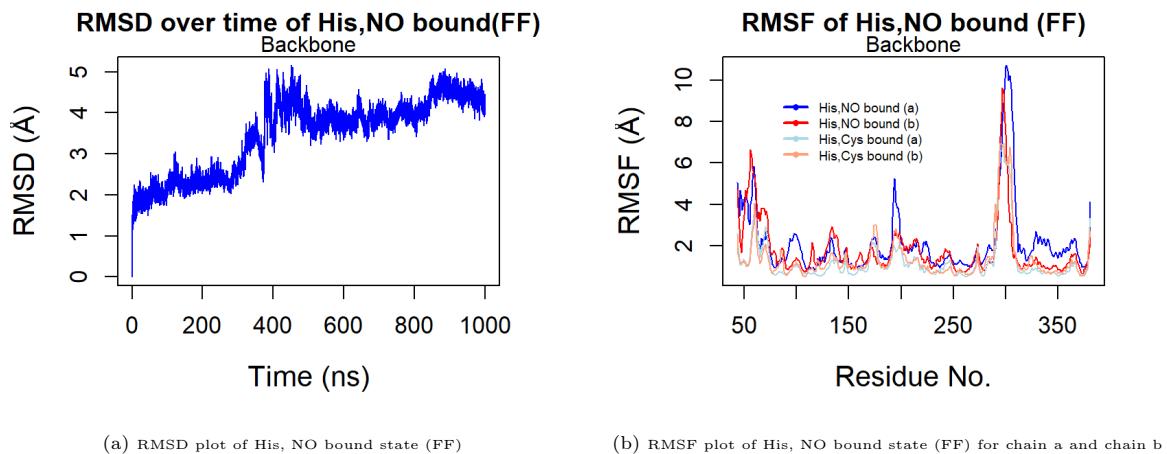
Figure 6.5: Translation measurements of α helix 8 in His, CO bound state EE (a) and EE' (b) for chain a and chain b

Figure 6.6: General analysis of His, NO bound state (FF) for chain a and chain b

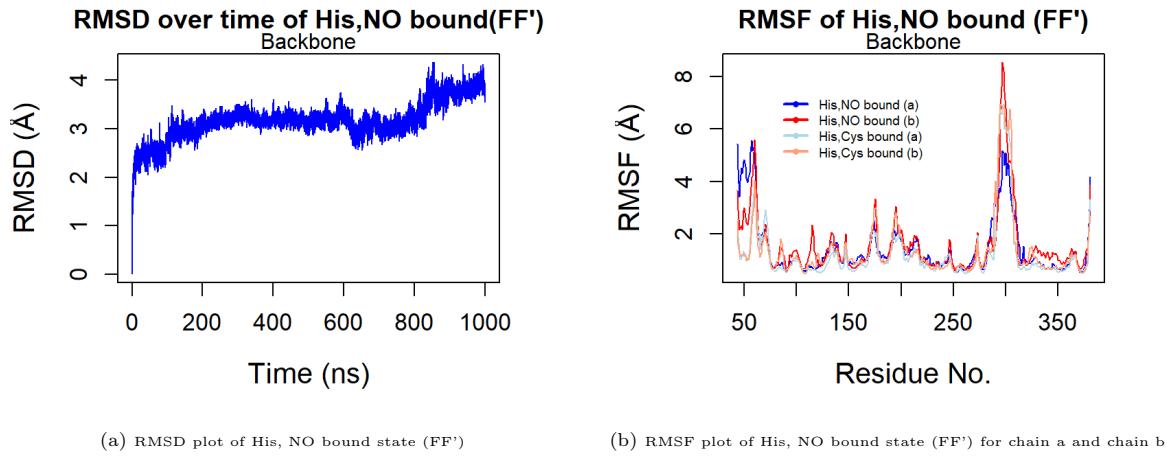


Figure 6.7: General analysis of His, NO bound state (FF') for chain a and chain b

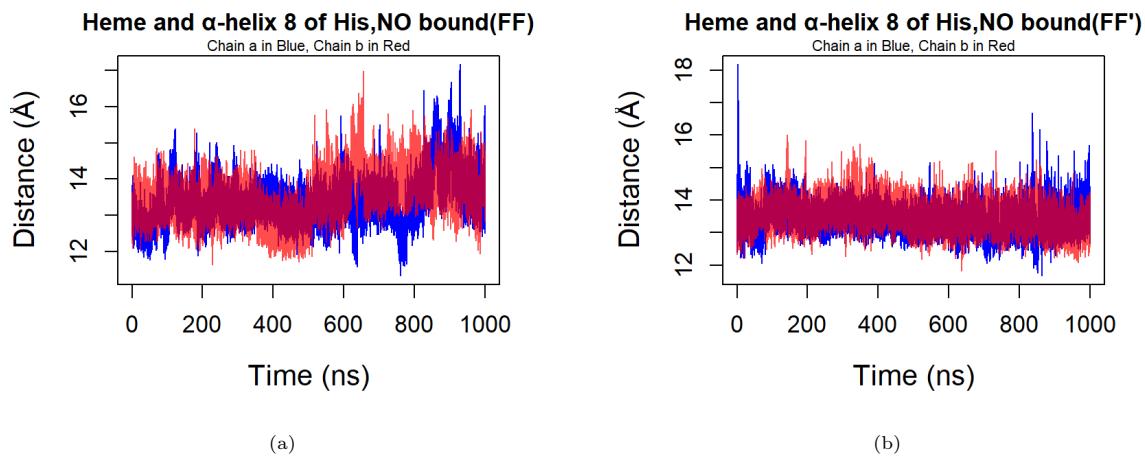
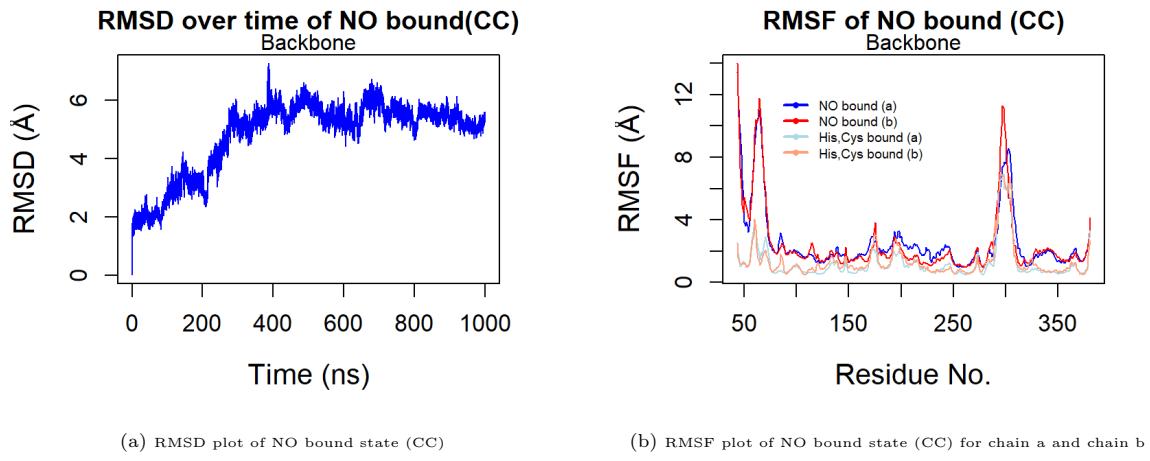
Figure 6.8: Translation measurements of α helix 8 in His, NO bound state FF (a) and FF' (b) for chain a and chain b

Figure 6.9: General analysis of NO bound state (CC) for chain a and chain b

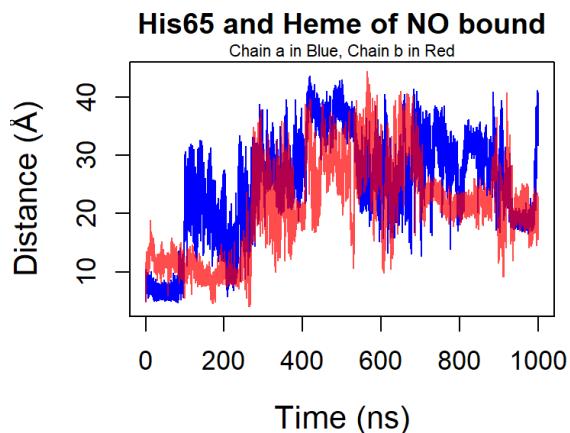


Figure 6.10: Evolving Distance between His65-NE2 and Heme-Fe. Tracked over time for Chain a (blue) and Chain b (red)

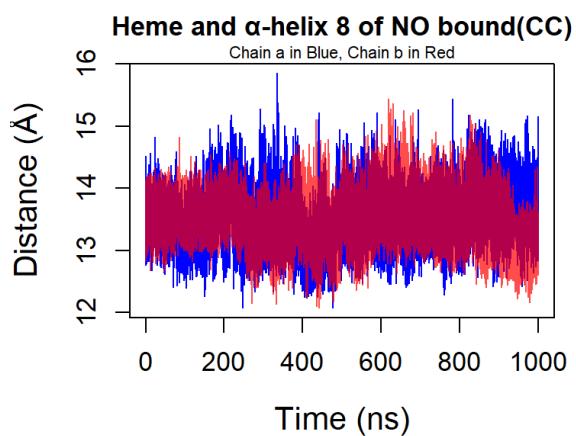


Figure 6.11: Translation measurements of α helix 8 in NO bound state (CC) for chain a and chain b

Bibliography

- [1] Karim Zuhra, Fiona Augsburger, Tomas Majtan, and Csaba Szabo. Cystathionine- β -synthase: Molecular Regulation and Pharmacological Inhibition. *Biomolecules*, 10(5):697, May 2020. Number: 5 Publisher: Multidisciplinary Digital Publishing Institute.
- [2] Hyung-Ok Lee, Liqun Wang, Yin-Ming Kuo, Sapna Gupta, Michael J. Slifker, Yue-sheng Li, Andrew J. Andrews, and Warren D. Kruger. Lack of global epigenetic methylation defects in CBS deficient mice. *Journal of Inherited Metabolic Disease*, 40(1):113–120, 2017. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1007/s10545-016-9958-5>.
- [3] Liqun Wang, Xulin Chen, Baiqing Tang, Xiang Hua, Andres Klein-Szanto, and Warren D. Kruger. Expression of mutant human cystathionine beta-synthase rescues neonatal lethality but not homocystinuria in a mouse model. *Human Molecular Genetics*, 14(15):2201–2208, August 2005.
- [4] Karine Robert, Johnny Nehmé, Emmanuel Bourdon, Gérard Pivert, Bertrand Friguet, Claude Delcayre, Jean-Maurice Delabar, and Nathalie Janel. Cystathionine β Synthase Deficiency Promotes Oxidative Stress, Fibrosis, and Steatosis in Mice Liver. *Gastroenterology*, 128(5):1405–1415, May 2005.
- [5] Sapna Gupta, Jirko Kühnisch, Aladdin Mustafa, Sarka Lhotak, Alexander Schlachterman, Michael J. Slifker, Andres Klein-Szanto, Katherine A. High, Richard C. Austin, and Warren D. Kruger. Mouse models of cystathionine β -synthase deficiency reveal significant threshold effects of hyperhomocysteinemia. *The FASEB Journal*, 23(3):883–893, 2009. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1096/fj.08-120584>.
- [6] Sapna Gupta and Warren D. Kruger. Cystathionine Beta-Synthase Deficiency Causes Fat Loss in Mice. *PLOS ONE*, 6(11):e27598, November 2011. Publisher: Public Library of Science.
- [7] Chin-Yi Tsai, Meng Teng Peh, Wei Feng, Brian William Dymock, and Philip Keith Moore. Hydrogen Sulfide Promotes Adipogenesis in 3T3L1 Cells. *PLOS ONE*, 10(3):e0119511, March 2015. Publisher: Public Library of Science.
- [8] Jéssica Lamberty Faverzani, Tatiane Grazieli Hammerschmidt, Angela Sitta, Marion Deon, Moacir Wajner, and Carmen Regla Vargas. Oxidative Stress in Homocystinuria Due to Cystathionine β -Synthase Deficiency: Findings in Patients and in Animal Models. *Cellular and Molecular Neurobiology*, 37(8):1477–1485, November 2017.
- [9] S. Dayal, T. Bottiglieri, E. Arning, N. Maeda, M. R. Malinow, C. D. Sigmund, D. D. Heistad, F. M. Faraci, and S. R. Lentz. Endothelial dysfunction and elevation of S-adenosylhomocysteine in cystathionine beta-synthase-deficient mice. *Circulation Research*, 88(11):1203–1209, June 2001.
- [10] Gary L. Baumbach, Curt D. Sigmund, Teodoro Bottiglieri, and Steven R. Lentz. Structure of Cerebral Arterioles in Cystathionine β -Synthase-Deficient Mice. *Circulation Research*, 91(10):931–937, November 2002. Publisher: American Heart Association.
- [11] Hong Wang, XiaoHua Jiang, Fan Yang, John W. Gaubatz, Lang Ma, Mark J. Magera, XiaoFeng Yang, Peter B. Berger, William Durante, Henry J. Pownall, and Andrew I. Schafer. Hyperhomocysteinemia accelerates atherosclerosis in cystathionine beta-synthase and apolipoprotein E double knock-out mice with and without dietary perturbation. *Blood*, 101(10):3901–3907, May 2003.

- [12] Utpal Sen, Pushpakumar B. Sathnur, Sourav Kundu, Srikanth Givvmani, Denise M. Coley, Paras K. Mishra, Natia Qipshidze, Neetu Tyagi, Naira Metreveli, and Suresh C. Tyagi. Increased endogenous H₂S generation by CBS, CSE, and 3MST gene therapy improves ex vivo renovascular relaxation in hyperhomocysteinemia. *American Journal of Physiology. Cell Physiology*, 303(1):C41–51, July 2012.
- [13] Sumit Kar, Hamid R. Shahshahan, Tyler N. Kambis, Santosh K. Yadav, Zhen Li, David J. Lefer, and Paras K. Mishra. Hydrogen Sulfide Ameliorates Homocysteine-Induced Cardiac Remodeling and Dysfunction. *Frontiers in Physiology*, 10, 2019.
- [14] Weimin Zhao, Joseph Fomusi Ndisang, and Rui Wang. Modulation of endogenous production of H₂S in rat tissues. *Canadian Journal of Physiology and Pharmacology*, 81(9):848–853, September 2003. Publisher: NRC Research Press.
- [15] Tomas Majtan, Jakub Krijt, Jitka Sokolová, Michaela Křížková, Maria A. Ralat, Jana Kent, Jesse F. Gregory, Viktor Kožich, and Jan P. Kraus. Biogenesis of Hydrogen Sulfide and Thioethers by Cystathionine Beta-Synthase. *Antioxidants & Redox Signaling*, 28(4):311–323, February 2018. Publisher: Mary Ann Liebert, Inc., publishers.
- [16] Kristian K. Jensen, Neil S. Geoghegan, Lan Jin, Tom G. Holt, Qi Luo, Lorraine Malkowitz, Weihua Ni, Shuo Quan, M. Gerard Waters, Aiwu Zhang, Heather H. Zhou, Kang Cheng, and Ming-juan Luo. Pharmacological activation and genetic manipulation of cystathionine beta-synthase alter circulating levels of homocysteine and hydrogen sulfide in mice. *European Journal of Pharmacology*, 650(1):86–93, January 2011.
- [17] Poulami Basu, Natia Qipshidze, Utpal Sen, Srikanth Givvmani, Charu Munjal, Paras K. Mishra, and Suresh C. Tyagi. Chronic hyperhomocysteinemia causes vascular remodelling by instigating vein phenotype in artery. *Archives of Physiology and Biochemistry*, 117(5):270–282, December 2011. Publisher: Taylor & Francis _eprint: <https://doi.org/10.3109/13813455.2011.599844>.
- [18] Sounik Saha, Prabir K. Chakraborty, Xunhao Xiong, Shailendra Kumar Dhar Dwivedi, Soumyajit Banerjee Mustafi, Noah R. Leigh, Ramani Ramchandran, Priyabrata Mukherjee, and Resham Bhattacharya. Cystathionine β -synthase regulates endothelial function via protein S-sulphydrylation. *The FASEB Journal*, 30(1):441–456, 2016. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1096/fj.15-278648>.
- [19] Alexandra Latini, Paula Juliana Seadi Pereira, Réjean Couture, Maria Martha Campos, and Sébastien Talbot. Oxidative Stress: Neuropathy, Excitability, and Neurodegeneration. *Oxidative Medicine and Cellular Longevity*, 2019:e2715326, January 2019. Publisher: Hindawi.
- [20] Jun-Feng Wang, Yu Li, Jin-Ning Song, and Hong-Gang Pang. Role of hydrogen sulfide in secondary neuronal injury. *Neurochemistry International*, 64:37–47, January 2014.
- [21] Weifang Rong, Hideo Kimura, and David Grundy. The Neurophysiology of Hydrogen Sulfide. *Inflammation & Allergy - Drug Targets (Formerly Current Drug Targets - Inflammation & Allergy)*, 10(2):109–117, April 2011.
- [22] Ramasamy Tamizhselvi, Philip K. Moore, and Madhav Bhatia. Hydrogen sulfide acts as a mediator of inflammation in acute pancreatitis: in vitro studies using isolated mouse pancreatic acinar cells. *Journal of Cellular and Molecular Medicine*, 11(2):315–326, 2007. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1582-4934.2007.00024.x>.
- [23] Lindsei K. Sarna, Yaw L. Siow, and Karmin O. The CBS/CSE system: a potential therapeutic target in NAFLD? *Canadian Journal of Physiology and Pharmacology*, 93(1):1–11, January 2015. Publisher: NRC Research Press.
- [24] Sarathi Mani, Wei Cao, Lingyun Wu, and Rui Wang. Hydrogen sulfide and the liver. *Nitric Oxide*, 41:62–71, September 2014.
- [25] Ruben Esse, Apolline Imbard, Cristina Florindo, Sapna Gupta, Eoin P. Quinlivan, Mariska Davids, Tom Teerlink, Isabel Tavares de Almeida, Warren D. Kruger, Henk J. Blom, and Rita Castro. Protein arginine hypomethylation in a mouse model of cystathionine β -synthase deficiency. *The FASEB Journal*, 28(6):2686–2695, 2014. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1096/fj.13-246579>.

- [26] Julien Hamelet, Christophe Noll, Clémentine Ripoll, Jean-Louis Paul, Nathalie Janel, and Jean-Maurice Delabar. Effect of hyperhomocysteinemia on the protein kinase DYRK1A in liver of mice. *Biochemical and Biophysical Research Communications*, 378(3):673–677, January 2009.
- [27] Sudhakar Veeranki, Lee J. Winchester, and Suresh C. Tyagi. Hyperhomocysteinemia associated skeletal muscle weakness involves mitochondrial dysfunction and epigenetic modifications. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1852(5):732–741, May 2015.
- [28] San-Qiao Yang, Li Jiang, Fang Lan, Hai-jun Wei, Ming Xie, Wei Zou, Ping Zhang, Chun-Yan Wang, Yu-Rong Xie, and Xiao-Qing Tang. Inhibited Endogenous H₂S Generation and Excessive Autophagy in Hippocampus Contribute to Sleep Deprivation-Induced Cognitive Impairment. *Frontiers in Psychology*, 10, 2019.
- [29] Ruili Yang, Cunye Qu, Yu Zhou, Joanne E. Konkel, Shihong Shi, Yi Liu, Chider Chen, Shiyu Liu, Dawei Liu, Yibu Chen, Ebrahim Zandi, Wanjun Chen, Yanheng Zhou, and Songtao Shi. Hydrogen Sulfide Promotes Tet1- and Tet2-Mediated Foxp3 Demethylation to Drive Regulatory T Cell Differentiation and Maintain Immune Homeostasis. *Immunity*, 43(2):251–263, August 2015.
- [30] Akiko Ichinohe, Takeshi Kanaumi, Sachio Takashima, Yasushi Enokido, Yasuo Nagai, and Hideo Kimura. Cystathionine beta-synthase is enriched in the brains of Down's patients. *Biochemical and Biophysical Research Communications*, 338(3):1547–1550, December 2005.
- [31] Jerzy Belłowski, Grażyna Wójcicka, and Anna Jamroz-Wiśniewska. Hydrogen sulfide in the regulation of insulin secretion and insulin sensitivity: Implications for the pathogenesis and treatment of diabetes mellitus. *Biochemical Pharmacology*, 149:60–76, March 2018.
- [32] Carmen Fernández-Rodríguez, Iker Oyenarte, Carolina Conter, Irene González-Recio, Reyes Núñez-Franco, Claudia Gil-Pitarch, Iban Quintana, Gonzalo Jiménez-Osés, Paola Dominici, María Luz Martínez-Chantar, Alessandra Astegno, and Luis Alfonso Martínez-Cruz. Structural insight into the unique conformation of cystathionine β -synthase from Toxoplasma gondii. *Computational and Structural Biotechnology Journal*, 19:3542–3555, January 2021.
- [33] Alexander A. Baykov, Heidi K. Tuominen, and Reijo Lahti. The CBS Domain: A Protein Module with an Emerging Prominent Role in Regulation. *ACS Chemical Biology*, 6(11):1156–1163, November 2011. Publisher: American Chemical Society.
- [34] Tomas Majtan, Angel L. Pey, Roberto Fernández, José A. Fernández, Luis A. Martínez-Cruz, and Jan P. Kraus. Domain Organization, Catalysis and Regulation of Eukaryotic Cystathionine Beta-Synthases. *PLOS ONE*, 9(8):e105290, August 2014. Publisher: Public Library of Science.
- [35] June Ereño-Orbea, Tomas Majtan, Iker Oyenarte, Jan P. Kraus, and Luis Alfonso Martínez-Cruz. Structural basis of regulation and oligomerization of human cystathionine β -synthase, the central enzyme of transsulfuration. *Proceedings of the National Academy of Sciences*, 110(40):E3790–E3799, October 2013. Publisher: Proceedings of the National Academy of Sciences.
- [36] Thomas J. McCorvie, Jolanta Kopec, Suk-Joon Hyung, Fiona Fitzpatrick, Xidong Feng, Daniel Termine, Claire Strain-Damerell, Melanie Vollmar, James Fleming, Jay M. Janz, Christine Bulawa, and Wyatt W. Yue. Inter-domain Communication of Human Cystathionine β -Synthase: STRUCTURAL BASIS OF S-ADENOSYL-L-METHIONINE ACTIVATION *. *Journal of Biological Chemistry*, 289(52):36018–36030, December 2014. Publisher: Elsevier.
- [37] Suneeta Devi, Khaja Faisal Tarique, Mohammad Farhan Ali, Syed Arif Abdul Rehman, and Samudrala Gourinath. Identification and characterization of Helicobacter pylori O-acetylserine-dependent cystathionine β -synthase, a distinct member of the PLP-II family. *Molecular Microbiology*, 112(2):718–739, 2019. _eprint: <https://onlinelibrary.wiley.com/doi/10.1111/mmi.14315>.
- [38] Suneeta Devi, Syed A. Abdul Rehman, Khaja F. Tarique, and Samudrala Gourinath. Structural characterization and functional analysis of cystathionine β -synthase: an enzyme involved in the reverse transsulfuration pathway of *Bacillus anthracis*. *The FEBS journal*, 284(22):3862–3880, November 2017.

- [39] Yasuyuki Matoba, Tomoki Yoshida, Hisae Izuhara-Kihara, Masafumi Noda, and Masanori Sugiyama. Crystallographic and mutational analyses of cystathionine β -synthase in the H2S-synthetic gene cluster in *Lactobacillus plantarum*. *Protein Science*, 26(4):763–783, 2017. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/pro.3123>.
- [40] Daniela Marciano, Marianela Santana, and Cristina Nowicki. Functional characterization of enzymes involved in cysteine biosynthesis and H2S production in *Trypanosoma cruzi*. *Molecular and Biochemical Parasitology*, 185(2):114–120, October 2012.
- [41] Roman Vozdek, Aleš Hnízda, Jakub Krijt, Marta Kostrouchová, and Viktor Kožich. Novel structural arrangement of nematode cystathionine β -synthases: characterization of *Caenorhabditis elegans* CBS-1. *Biochemical Journal*, 443(Pt 2):535–547, April 2012.
- [42] Yupeng Tu, Cheryl A. Kreinbring, Megan Hill, Cynthia Liu, Gregory A. Petsko, Christopher D. McCune, David B. Berkowitz, Dali Liu, and Dagmar Ringe. Crystal Structures of Cystathionine β -Synthase from *Saccharomyces cerevisiae*: One Enzymatic Step at a Time. *Biochemistry*, 57(22):3134–3145, June 2018. Publisher: American Chemical Society.
- [43] Markos Koutmos, Omer Kabil, Janet L. Smith, and Ruma Banerjee. Structural basis for substrate activation and regulation by cystathionine beta-synthase (CBS) domains in cystathionine β -synthase. *Proceedings of the National Academy of Sciences*, 107(49):20958–20963, December 2010. Publisher: Proceedings of the National Academy of Sciences.
- [44] Paula Giménez-Mascarell, Tomas Majtan, Iker Oyenarte, June Ereño-Orbea, Juraj Majtan, Jaroslav Klaudiny, Jan P. Kraus, and Luis Alfonso Martínez-Cruz. Crystal structure of cystathionine β -synthase from honeybee *Apis mellifera*. *Journal of Structural Biology*, 202(1):82–93, April 2018.
- [45] A. Dong, V. Kery, J. Matsuura, M. C. Manning, J. P. Kraus, and J. F. Carpenter. Secondary structure of recombinant human cystathionine beta-synthase in aqueous solution: effect of ligand binding and proteolytic truncation. *Archives of Biochemistry and Biophysics*, 344(1):125–132, August 1997.
- [46] Amit Kumar, Amelie Wißbrock, Nishit Goradia, Peter Bellstedt, Ramadurai Ramachandran, Diana Imhof, and Oliver Ohlenschläger. Heme interaction of the intrinsically disordered N-terminal peptide segment of human cystathionine- β -synthase. *Scientific Reports*, 8(1):2474, February 2018. Number: 1 Publisher: Nature Publishing Group.
- [47] Kei Segawa, Miki Watanabe-Matsui, Kengo Tsuda, Toshitaka Matsui, Mikako Shirouzu, Kazuhiko Igarashi, and Kazutaka Murayama. Biophysical characterization of heme binding to the intrinsically disordered region of Bach1. *European biophysics journal: EBJ*, 48(4):361–369, May 2019.
- [48] Nina Frank, Vladimir Kery, Kenneth N. Maclean, and Jan P. Kraus. Solvent-Accessible Cysteines in Human Cystathionine β -Synthase: Crucial Role of Cysteine 431 in S-Adenosyl-l-methionine Binding. *Biochemistry*, 45(36):11021–11029, September 2006. Publisher: American Chemical Society.
- [49] Erez M. Bublil, Tomas Majtan, Insun Park, Richard S. Carrillo, Helena Hůlková, Jakub Krijt, Viktor Kožich, and Jan P. Kraus. Enzyme replacement with PEGylated cystathionine β -synthase ameliorates homocystinuria in murine model. *The Journal of Clinical Investigation*, 126(6):2372–2384, June 2016.
- [50] Aaron T. Smith, Yang Su, Daniel J. Stevens, Tomas Majtan, Jan P. Kraus, and Judith N. Burstyn. Effect of the Disease-Causing R266K Mutation on the Heme and PLP Environments of Human Cystathionine β -Synthase. *Biochemistry*, 51(32):6360–6370, August 2012.
- [51] V. Kery, L. Poneleit, J. D. Meyer, M. C. Manning, and J. P. Kraus. Binding of pyridoxal 5'-phosphate to the heme protein human cystathionine beta-synthase. *Biochemistry*, 38(9):2716–2724, March 1999.
- [52] Irene González-Recio, Carmen Fernández-Rodríguez, Jorge Simón, Naroa Goikoetxea-Usandizaga, María Luz Martínez-Chantar, Alessandra Astegno, Tomas Majtan, and Luis Alfonso Martínez-Cruz. Current Structural Knowledge on Cystathionine β -Synthase, a Pivotal Enzyme in the Transsulfuration Pathway. In *Encyclopedia of Life Sciences*, pages 453–468. John Wiley & Sons, Ltd, 2020. _eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/9780470015902.a0028966>.

- [53] Colin L. Weeks, Sangita Singh, Peter Madzelan, Ruma Banerjee, and Thomas G. Spiro. Heme Regulation of Human Cystathionine β -Synthase Activity: Insights from Fluorescence and Raman Spectroscopy. *Journal of the American Chemical Society*, 131(35):12809–12816, September 2009.
- [54] Markus Meier, Miroslav Janosik, Vladimir Kery, Jan P Kraus, and Peter Burkhard. Structure of human cystathione b-synthase: a unique pyridoxal 5c-phosphate-dependent heme protein. *The European Molecular Biology Organization Journal*, page 7, 2001.
- [55] Aaron P. Landry, Joseph Roman, and Ruma Banerjee. Structural perspectives on H₂S homeostasis. *Current Opinion in Structural Biology*, 71:27–35, December 2021.
- [56] A. Bateman. The structure of a domain common to archaeabacteria and the homocystinuria disease protein. *Trends in Biochemical Sciences*, 22(1):12–13, January 1997.
- [57] G. Bukovska, V. Kery, and J. P. Kraus. Expression of Human Cystathionine β -Synthase in Escherichia coli: Purification and Characterization. *Protein Expression and Purification*, 5(5):442–448, October 1994.
- [58] June Ereño-Orbea, Tomas Majtan, Iker Oyenarte, Jan P. Kraus, and Luis Alfonso Martínez-Cruz. Structural insight into the molecular mechanism of allosteric activation of human cystathionine β -synthase by S -adenosylmethionine. *Proceedings of the National Academy of Sciences*, 111(37), September 2014.
- [59] Angel L. Pey, Tomas Majtan, Jose M. Sanchez-Ruiz, and Jan P. Kraus. Human cystathionine β -synthase (CBS) contains two classes of binding sites for S -adenosylmethionine (SAM): complex regulation of CBS activity and stability by SAM. *Biochemical Journal*, 449(1):109–121, January 2013.
- [60] Tomas Majtan, Angel L. Pey, June Ereño-Orbea, Luis Alfonso Martínez-Cruz, and Jan P. Kraus. Targeting Cystathionine Beta-Synthase Misfolding in Homocystinuria by Small Ligands: State of the Art and Future Directions. *Current Drug Targets*, 17(13):1455–1470, September 2016.
- [61] S. Taoka, M. West, and R. Banerjee. Characterization of the heme and pyridoxal phosphate cofactors of human cystathionine beta-synthase reveals nonequivalent active sites. *Biochemistry*, 38(9):2738–2744, March 1999.
- [62] Shinichi Taoka and Ruma Banerjee. Characterization of NO binding to human cystathionine b-synthase: Possible implications of the effects of CO and NO binding to the human enzyme. *Journal of Inorganic Biochemistry*, page 7, 2001.
- [63] Laishram R. Singh, Xulin Chen, Viktor Kožich, and Warren D. Kruger. Chemical chaperone rescue of mutant human cystathionine β -synthase. *Molecular Genetics and Metabolism*, 91(4):335–342, August 2007.
- [64] Mrinalini Puranik, Colin L. Weeks, Dorothee Lahaye, Ömer Kabil, Shinichi Taoka, Steen Brøndsted Nielsen, John T. Groves, Ruma Banerjee, and Thomas G. Spiro. Dynamics of Carbon Monoxide Binding to Cystathionine β -Synthase *. *Journal of Biological Chemistry*, 281(19):13433–13438, May 2006. Publisher: Elsevier.
- [65] Carmen Gherasim, Pramod K. Yadav, Omer Kabil, Wei-Ning Niu, and Ruma Banerjee. Nitrite Reductase Activity and Inhibition of H₂S Biogenesis by Human Cystathionine β -Synthase. *PLOS ONE*, 9(1):e85544, January 2014. Publisher: Public Library of Science.
- [66] João B. Vicente, Henrique G. Colaço, Marisa I.S. Mendes, Paolo Sarti, Paula Leandro, and Alessandro Giuffrè. NO• Binds Human Cystathionine β -Synthase Quickly and Tightly. *Journal of Biological Chemistry*, 289(12):8579–8587, March 2014.
- [67] C. George Priya Doss, B. Rajith, R. Magesh, and A. Ashish Kumar. Influence of the SNPs on the structural stability of CBS protein: Insight from molecular dynamics simulations. *Frontiers in Biology*, 9(6):504–518, December 2014.

- [68] Sapna Gupta, Simon Kelow, Liqun Wang, Mark D. Andrade, Roland L. Dunbrack, and Warren D. Kruger. Mouse modeling and structural analysis of the p.G307S mutation in human cystathione β -synthase (CBS) reveal effects on CBS activity but not stability. *Journal of Biological Chemistry*, 293(36):13921–13931, September 2018.
- [69] Aashish Bhatt and Ehesan Ali. Understanding the role of R266K Mutation in Cystathione β -synthase (CBS) Enzyme: An In Silico study. *Journal of Biomolecular Structure and Dynamics*, page 17, 2021.
- [70] Pramod Kumar Yadav, Peter Xie, and Ruma Banerjee. Allosteric Communication between the Pyridoxal 5-Phosphate (PLP) and Heme Sites in the H2S Generator Human Cystathione β -Synthase *. *Journal of Biological Chemistry*, 287(45):37611–37620, November 2012. Publisher: Elsevier.
- [71] Hans C. Andersen. Molecular dynamics simulations at constant pressure and/or temperature. *The Journal of Chemical Physics*, 72(4):2384–2393, July 2008.
- [72] Jing Huang and Alexander D. MacKerell. CHARMM36 all-atom additive protein force field: validation based on comparison to NMR data. *Journal of Computational Chemistry*, 34(25):2135–2145, September 2013.
- [73] Brad A. Wells and Alan L. Chaffee. Ewald Summation for Molecular Simulations. *Journal of Chemical Theory and Computation*, 11(8):3684–3695, August 2015. Publisher: American Chemical Society.
- [74] Manolo Gouy, Stéphane Guindon, and Olivier Gascuel. SeaView version 4: A multiplatform graphical user interface for sequence alignment and phylogenetic tree building. *Molecular Biology and Evolution*, 27(2):221–224, February 2010.
- [75] Xavier Robert and Patrice Gouet. Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Research*, 42(Web Server issue):W320–324, July 2014.
- [76] Qiang Wang, Adrian A. Canutescu, and Roland L. Dunbrack. SCWRL and MolIDE: Computer programs for side-chain conformation prediction and homology modeling. *Nature protocols*, 3(12):1832–1847, 2008.
- [77] Jumin Lee, Xi Cheng, Jason M. Swails, Min Sun Yeom, Peter K. Eastman, Justin A. Lemkul, Shuai Wei, Joshua Buckner, Jong Cheol Jeong, Yifei Qi, Sunhwan Jo, Vijay S. Pande, David A. Case, Charles L. III Brooks, Alexander D. Jr. MacKerell, Jeffery B. Klauda, and Wonpil Im. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *Journal of Chemical Theory and Computation*, 12(1):405–413, January 2016. Publisher: American Chemical Society.
- [78] K. Vanommeslaeghe and A. D. Jr. MacKerell. Automation of the CHARMM General Force Field (CGenFF) I: Bond Perception and Atom Typing. *Journal of Chemical Information and Modeling*, 52(12):3144–3154, December 2012. Publisher: American Chemical Society.
- [79] Mark James Abraham, Teemu Murtola, Roland Schulz, Szilárd Pál, Jeremy C. Smith, Berk Hess, and Erik Lindahl. GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX*, 1-2:19–25, September 2015.
- [80] Mustafa Tekpinar, Bertrand Neron, and Marc Delarue. Correction to “Extracting Dynamical Correlations and Identifying Key Residues for Allosteric Communication in Proteins by correlationplus”. *Journal of Chemical Information and Modeling*, 61(11):5720–5720, November 2021. Publisher: American Chemical Society.
- [81] Sebastián Carballal, Peter Madzelan, Carlos F. Zinola, Martín Graña, Rafael Radi, Ruma Banerjee, and Beatriz Alvarez. Dioxygen reactivity and heme redox potential of truncated human cystathionine beta-synthase. *Biochemistry*, 47(10):3194–3201, March 2008.