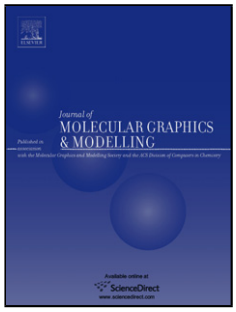


Title: Structural and dynamical aspects of *Streptococcus gordonii* FabH through molecular docking and MD simulations

Author: Amen Shamim Sumra Wajid Abbasi Syed Sikander Azam

PII: S1093-3263(15)30003-6
DOI: http://dx.doi.org/doi:10.1016/j.jmgm.2015.05.013
Reference: JMG 6553

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 11-12-2014
Revised date: 21-5-2015
Accepted date: 22-5-2015

Please cite this article as: Amen Shamim, Sumra Wajid Abbasi, Syed Sikander Azam, Structural and dynamical aspects of Streptococcus gordonii FabH through molecular docking and MD simulations, Journal of Molecular Graphics and Modelling http://dx.doi.org/10.1016/j.jmgm.2015.05.013

Structural and dynamical aspects of *Streptococcus gordonii* FabH through molecular docking and MD simulations

Amen Shamim, Sumra Wajid Abbasi, Syed Sikander Azam*
Computational Biology Lab, National Center for Bioinformatics, Quaid-i-Azam University, Islamabad 45320, Pakistan
*Corresponding author: Tel.: +92-51-90644130; E-mail: ssazam@qau.edu.pk; <ABS-Highlights
HEAD>syedazam2008@gmail.com

Highlights

- To overcome multidrug resistance, FAB pathway was blocked in *Streptococcus gordonii* by molecular docking approach.
- Active site tunnel has a conserved catalytic triad during entire simulation.
- Dimeric docked complex remained stable as compare to undocked protein during MD simulation.
- This strategy facilitates to develop drug which would target merely the pathogen's system, without harming the biological and metabolic pathways of host.

Abstract

β-Ketoacyl-ACP-synthase III (FabH or KAS III) has become an attractive target for the development of new antibacterial agents which can overcome the multidrug resistance. Unraveling the fatty acid biosynthesis (FAB) metabolic pathway and understanding structural coordinates of FabH will provide valuable insights to target *Streptococcus gordonii* for curing oral infection. In this study, we designed inhibitors against therapeutic target FabH, in order to block the FAB pathway. As compared to other targets, FabH has more interactions with other proteins, located on the leading strand with higher codon adaptation index value and associated with lipid metabolism category of COG. Current study aims to gain *in silico* insights into the structural and dynamical aspect of *S. gordonii* FabH via molecular docking and molecular dynamics (MD) simulations. The FabH protein is catalytically active in dimerization while it can lock in monomeric state. Current study highlights two residues Pro88 and Leu315 that are close to each other by dimerization. The active site of FabH is composed of the catalytic triad formed by residues Cys112, His249, and Asn279 in which Cys112 is involved in acetyl transfer, while

His249 and Asn279 play an active role in decarboxylation. Docking analysis revealed that among the studied compounds, methyl-CoA disulfide has highest GOLD score (82.75), binding affinity (-11kcal/mol) and exhibited consistently better interactions. During MD simulations, the FabH structure remained stable with the average RMSD value of 1.7 Å and 1.6 Å for undocked protein and docked complex respectively. Further, crucial hydrogen bonding of the conserved catalytic triad for exhibiting high affinity between the FabH protein and ligand is observed by RDF analysis. The MD simulation results clearly demonstrated that binding of the inhibitor with *S. gordonii* FabH enhanced the structure and stabilized the dimeric FabH protein. Therefore, the inhibitor has the potential to become a lead compound.

Key words: Drug design, Homology modelling, Molecular docking, Molecular dynamics simulation.

Introduction

Streptococcus gordonii resistance against majority of antibiotics has necessitated the search for new therapeutic antibacterial agents with novel targets [1]. The identification of compounds with novel mode of action towards a new *S. gordonii* target enzyme would be an effective solution. *S. gordonii* belongs to the viridians group which is involved in biofilm formation [2]. Biofilm is an important component of human dental plaque by virtue of its ability to adhere to tooth surfaces [3]. Bacterial biofilms are source of chronic infections because they show resistance to antibiotics and disinfectant chemicals [4]. *S. gordonii* causes bacterial endocarditis and septic arthritis by entering into the blood stream usually after oral trauma [5]. Major surveys of septic arthritis demonstrate that its predominant causative organisms are streptococci. Anaerobic organisms rarely cause septic arthritis, but are more common when there is a history of penetrating trauma [6]. Treatment for septic arthritis was undertaken by penicillin, ampicillin, cefotaxime, erythromycin, clindamycin and vancomycin. Gentamycin (5mg/kg/day) and rifampicin (900mg/day) were added to the treatment of endocarditis [5]. These antibiotics are effective in early stages but multi-drug resistance (MDR) complicates the treatment and leads to high failure rate. Identification of novel drug targets is the key step to combat these infections. Highlighting the drug targets in *S. gordonii*, fatty acid biosynthesis (Fab) has proven alluring target for causing oral infection [7]. Notably, β-Ketoacyl-ACP-synthase III (FabH or KAS III)

(EC 2.3.1.180) a condensing enzyme [8], plays an essential and regulatory role in bacterial fatty acid elongation cycles [7, 9-11]. Moreover, it characterizes a promising target for the design of novel therapeutic drugs through various kinds of compounds [12, 13] .

FabH is highly conserved among these pathogens; *Escherichia coli* [8], *Plasmodium falciparum* [14], *Mycobacterium tuberculosis* [15], *Staphylococcus aureus* [16] and *Streptococcus pneumonia* [17]. In above mentioned pathogens, FabH catalyses a direct condensation reaction between the acyl coenzyme A (CoA) primers used to initiate fatty acid biosynthesis [15, 18, 19]. In E.coli FabH, the active site is consisted of a catalytic triad in which Cys112 residue is adjacent to His244 and Asn274 [20]. The residue Cys112 is required to attack the acetyl-CoA substrate to form the acetyl-FabH intermediate while His244 and Asn274 play an active role in decarboxylation [21].

The medicinal importance of FabH has led to increased research interest and focus on the development of new antibiotics. Drugs that act against the fatty acid biosysthesis pathway block the enoyl-ACP reductase (FabI) component such as triclosan [22] and isoniazid [23]. Fungal antibiotic cerulenin inhibits both FabB and FabF condensing enzymes by covalent modification of the active-site sulfhydryl [24, 25]. Another natural product inhibits all condensing enzymes including FabH is thiolactomycin [26, 27]. Platensimycin is potent target of *Staphylococcus aureus and Enterococcus*, FabH and FabF [28]. However, inspected pathogen has resistance against currently used antibiotics that are an optimal regimen to cure the *S. gordonii* infections that has not been established yet.

Structural studies, shedding light to investigate underlying molecular mechanisms that may contribute to FabH inhibition have been explored in the current work. Structural, architectural and dynamical properties of protein were observed and recapitulated, that prompted a thorough investigation and hence is the central theme of this study.

Materials and Methods

To elucidate the entire comparative modelling, we selected a target protein that had been crystallized with an experimental structure which was available to serve as a template. The complete workflow (Figure 1) represents various analyses and selection conditions followed in

this study. The amino acid sequence of FabH from *S. gordonii* was retrieved from uni-Prot (ID: A8AYW5; residues 1-324 aa).

The tetrameric crystal structure of *Enterococcus faecalis* (strain ATCC 700802 / V583) FabH (uniprot ID: Q820T1; PDB ID 3il5; residues 1-321 aa) served as template in which 2-({[4-bromo-3-(diethylsulfamoyl)phenyl]carbonyl}amino) benzoic acid involved as a ligand [29].

The monomer structure used for this study corresponds to the FabH monomer. The initial sequence alignment of *S. gordonii* FabH was performed by align2d function of Modeller9.11 [30]. Target-template alignment and structural coordinates were used to build 5 initial models by contentment of spatial restraints, subjected to energy minimization. Online modelling servers Swiss-model [31], I-TASSER [32], Mod-base [33], 3DJigSaw [34], (PS)2 [35], Phyre [36], M4T [37], Raptorx [38], HHpred [39] and ESyPred3D [40] were utilized to predict the *S. gordonii* FabH structure. Afterwards, the best model was chosen from the comparison among MODELLER and online modelling servers. The most reliable model of *S. gordonii* protein FabH came out through MODELLER, which was then visualized through CCP4MG version 2.7.3 [41]. Additional assessment was done by energy objective function (GA341, DOPE and molpdf score). Once a final model is selected, it can be further refined by minimization. The minimization consisted of 750 steps of steepest descent and conjugate gradient using a dielectric constant of 1. In minimization, the Amber parameters were used to remove possible close contacts between amino acids in the model and to relax the structure with the tripos force field (TFF) using UCSF Chimera [42]. To probe the protein-protein interactions, dimeric form of FabH was generated using PatchDock [43] methodology. Dimer assessment was done by energy minimization and quality check tools. The stereochemical properties of dimer structure were examined by PROCHECK [44] and ProSA [45]. The environmental preference for each residue was assessed by the program VERIFY-3D [46] and the distribution of non-bonded atoms in the neighborhood of each atom was analysed by ERRAT [47]. Discrete optimization potential energy score was evaluated by Modeller9.11. The symmetry of target and template proteins was verified with the root mean square deviation (RMSD) values and Q-score. There were no violations detected by either method. Secondary structure was predicted using PSIPRED [48]. It is an empirical technique based on analysis of relative frequencies of each amino acid in helices,

turns and beta sheets in a protein. Reliability of structure was assessed by model evaluation servers ModEval [49], Modfold server[50] and Qmean [51].

To predict active site, multiple sequence alignment and consensus generation of critical residues in the target protein corresponding to the several pathogens was carried out, using Web Logo [52]. Sequence conservation was calculated at a particular position in the alignment in which this equation is used, R_{seq} as the difference between the maximum possible entropy and the entropy of the observed symbol distribution:

$$R_{seq} = S_{max} - S_{obs} = \log_2 N - (-\sum_{n=1}^{N} p_n \log_2 p_n)$$

For ligand preparation, total 282 synthetic compounds selected from literature [13, 16, 17, 53-64] that were assayed against *E.coli*, *M. tuberculosis*, *S. pneumonia*, *P. falciparum*, *S. aureus*, *Cuphea lanceolata*, *Enterococcus faecalis*, *Haemophilus influenzae* and *Francisella tularensis* subsp. holarctica. The selection criteria of inhibitors were: (i) all selected compounds must be reported inhibitors against FabH (ii) Selected inhibitors must have reported IC50 values (iii) ADME properties of a compound were checked on the basis of their molecular properties such as molecular weight (MW), hydrogen bond donor (HD), hydrogen bond acceptor (HA), lipophilicity (LP), and polar solvent accessibility (PSA). For docking analysis, the conformational energies of inhibitors were minimized, using UCSF Chimera [42].

Several Dock Scores and binding affinities implemented in main stream molecular docking programs that were estimated in this study. The scoring functions GoldScore, ChemScore and ChemPLP were assessed from GOLD [65] which utilizes genetic algorithm to explore the rotational flexibility of receptor hydrogens and ligand conformational flexibility [66]. In GOLD, docking was carried out using the wizard with default parameters population size (100), selection- pressure (1.1), number of operations (10,000), number of islands (1), niche size (2) and operator weights for migrate (0), mutate (100), and crossover (100) were applied. The active site of a 10 Å radius sphere was defined by selecting an active site residue of protein (Arg215). Default genetic algorithm settings were used for all calculations and a set of 10 solutions were saved for each ligand. GOLD was used with GoldScore fitness function, which is a molecular mechanics like function and has been optimized for the calculation of binding positions of ligand. In order to calculate the binding affinities of selected inhibitors, docking was further

carried out by AutoDock/Vina [67]. In file preparation step, such as pdbqt files for protein and ligands preparation and grid box creation, AutoDock Tools (ADT) [68] were employed. The polar hydrogens, united atom Kollman charges, solvation parameters and fragmental volumes were assigned to the protein. AutoGrid is used for the preparation of the grid map using a grid box with a number of points in xyz of 60*60*60 box, which encloses the ligand. The box spacing was 1 Å and grid center was designed at dimensions (x,y,z): -193.334, 172.566, 38.081, which covered the entire anticipated pocket. Finally, AutoDock vina was executed using protein and ligand information along with grid box properties in the configuration file. The Pearson correlation coefficient between the known binding affinities and the binding scores produced by the given scoring function was measured by R version 3.0.2 [69]. The visual inspection of protein and ligand interactions was carefully done using VMD software [70].

Molecular dynamics simulations were performed using the AMBER 10 [71]. The solvated protein was subjected to energy minimization, heating, equilibration and production run using AMBER algorithm. Sander module was used to obtaining these MD simulation trajectories while analyses were done by PTRAJ module [72].

In first step, ligand library was generated through antechamber program. For ligands, the general AMBER force field (GAFF) [73] was chosen and for enzyme the ff03.r1 force field. A pre-equilibrated elementary TIP3P tetrahedral box of 84.29 Å × 101.40 Å × 93.51 Å was utilized to completely immerse the protein which added about 18813 water molecules. Total volume of box was 799353.099 Å^3 and overall compactness of the system was 0.851 g/cm^3. 26 Na+ ions were added to neutralize the system. The pre-processing of docked protein was carried out by the graphical interface of LEaP [74]. Docked and undocked protein was subjected to energy minimization to remove any steric clashes before undergoing MD simulation. The 1500 iterations were employed in which 750 steps using steepest decent and followed by 750 steps of conjugate gradient. SHAKE algorithm was employed to constrained the hydrogen bond lengths of docked and undocked proteins [75]. Periodic boundary condition was applied to keep the temperature constant, a non-bonded cutoff value was set to 8 Å [76]. To collect the statistically accurate results total 50 nanoseconds (ns) simulation was performed in which analysis files were saved after every 0.5 picoseconds (ps) time.

MD trajectories of docked complex were analysed using the five classical measures such as Root Mean Square Deviation (RMSD), Root Mean Square Fluctuations (RMSF) and Radial Frequency Distribution (RFD). In post simulation analysis of trajectories, PTRAJ module was used in which following properties were analysed root mean square deviation (RMSD), root mean square fluctuation (RMSF), radius of gyration and beta factor (B-factor).

The RMSD was calculated for the Cα atoms using the starting structure as reference frame and check the deviation of the coordinate of given set of atom in a time interval. The fluctuations of Cα atoms were measured for each residue from its mean position. The disorderedness in system is measured by temperature dependent vibration of an atom. It is calculated similar to RMSF but by multiplying the squared fluctuation with (8/3) π^2. Overall shape compactness or expansion of the system was calculated by radius of gyration. The radial distribution function is plotted between HG of enzyme with S47 and S48 of inhibitor. It is a prominent analytical tool in order to describe the structural features of a system.

Results and Discussion

This study unveils a therapeutic target that paves a way to open a new horizon for drug discovery and could develop treatments for diseases that so far remain intractable. Current drug discovery process is characterized by the identification of potential therapeutic candidate in *S. gordonii* and can be accomplished by a number of ways in short span. FabH, potential drug target has more interactions with other proteins and located on leading strand with higher codon adaptation index value as compared to other targets and associated with lipid metabolism (I category) of COG. Hence, designing inhibitors against FabH to block FAB pathway represents an attractive strategy which will damage the growth and survival of *S. gordonii*.

Model building framework and benchmarking

Reliable three-dimensional model of *S. gordonii* FabH can provide valuable insights into basic principles of molecular recognition and aid to find inhibitors against it. The urgency to solve these impediments is as prodigious as ever. Therefore, we built the homology models of FabH monomer by Modeller9.11 (Table 1) and different online web servers (Table 2). To reflect the duplication, FabH monomer has internal pseudo-symmetry in its structure. However, there is

barely any sequence similarity between the two domains of FabH protein. Intriguingly, sequences and peripheral structures have diverged significantly throughout evolution. We highlighted widely divergent α/β motifs of FabH that contain a clear internal duplication between N and C terminal domains via secondary structure prediction. Predicting the secondary structure of model protein demonstrated that it has more alpha helices as compared to beta sheets (Figure 2). Statistical analysis showed 35.49% helices, 23.77% beta sheets, and 10.19% turns. Each generated model of *S. gordonii* FabH monomer was scrutinized on the basis of DOPE assessment score (Figure 3) and ramachandran plot. Minimized model of *S. gordonii* FabH has lower optimized energy than template (indicated the higher stability of predicted model FabH). The best model in terms of the stereochemical quality showed overall G-factor value of -0.15 which indicates that geometry of the model corresponds to high probability confirmation with 92.5% residues in the core region of the ramachandran plot. The number of residues in allowed, generously allowed and disallowed regions were 5.8%, 1.7% and 0.0%, respectively.

Quality check

The theoretically generated structure was subjected over the backbone fragment of chain-A of *S. gordonii* FabH to estimate the root mean square deviation (RMSD) and Q-score. The RMSD and Q-score of 0.719 Å and 0.87 Å respectively, suggested a reliable 3D structure. The stereochemical quality of the homology models was evaluated before and after energy refinements. The quality and structural features of non-bonded interactions between different atoms was identified to be 79.1 and the Z score values were -8.36, calculated using different online servers.

Evaluation of dimer interface

The *S. gordonii* FabH exerts its catalytic function as a dimer, which could potentially be exploited in developing new strategies for inhibitor design. Previous studies revealed that crystal structure of *E. coli* FabH is catalytic active in dimerization while it can lock FabH in monomeric state [77]. Current study highlighted two residues Pro88 and Leu315 that are far apart in the primary sequence and the tertiary structure of *S. gordonii* FabH. However, within the functional dimer, these two residues are not only brought into the immediate neighbourhood of each other

by virtue of dimerization (Figure 4); they adjacent to but not in the immediate vicinity of the active site cleft. The active site tunnel is principally created by the consensus generator in which extended loop regions are stabilized by interactions involving a number of conserved, charged and polar residues (Figure 5). Binding cavity of *S. gordonii* FabH is found to be consistent with a reported structure of this enzyme in different species that consists of a deep active site tunnel. Consensus of FabH enzyme from eight different bacterial species: *Escherichia coli*, *Pseudomonas aeruginosa*, *Haemophilus infuenzae*, *Thermus thermophiles*, *Staphylococcus aureus*, *Mycobacterium tuberculosis*, *Spinacia oleracea* and *Entrococcus faecalis* inferred the most pronounced region; first loop (L1 of the N-terminal domain and L9 of the C-terminal domain and the last one (L7 of the N- and L15 of the C-terminal domain) that has functional importance in *S. gordonii*. These regions of FabH protein were also conserved and have functional importance in different bacteria [20, 29]. Four alpha helices, L1α2, L7α2, Cα1 and Cα2 and one beta strand, L9β2 formed the entrance of tunnel whereas four basic residues, aromatic and conserved residues surround the entrance. Cys112 is an important residue of pocket, which lies at the bottom of the tunnel and is positioned precisely at the N terminus of helix Nα3.

In *S. gordonii* FabH, the active site is composed of catalytic triad formed by residues Cys112, His249, and Asn279 (Figure 6b) that was also reported in FabH protein of other bacteria in which Cys112 is involved in acetyl transfer, while His249 and Asn279 have an active role in decarboxylation [20, 77, 78]. Residues His249 and Asn279 are present close to Cys112 in the long active site tunnel; such that distance between the CαH249−CαN279 atoms,

CαN279−CαC112 atoms and CαH249−Cα C112 is ~8 Å, ~10 Å and ~6 Å respectively. The per

residue RMSD plot over the 40 ns simulation time exhibits that the catalytic residues His249, Asn279, Cys112 have minimal fluctuations in architecture of protein that make it conserved throughout the simulation. Previous studies on *E. coli* FabH mentioned the mechanism of conserved active site in which proton is transferred between the thiolate hydrogen of Cys112 and Nα nitrogen atom of His244 which generates the thiolate anion that attacks on carbonyl carbon

of acetyl CoA [21]. In *S. gordonii* FabH, distance between thiolate hydrogen in Cys112 (-SH) and the nitrogen atom in His249 (N) was calculated. Interestingly, our analysis illustrates active site tunnel remains conserved without undergoing large conformational changes. These Observations clearly show that His249, Asn279 and Cys112 are key active residues required for the condensation reaction and Cys112 are specifically involved in the transacylation whereas His249 and Asn279 are required for the decarboxylation (Figure 6a).

Binding site mapping

Comparative assessment of four different scoring functions was conducted to predict the preferred orientation of bounded molecules in dimmer FabH. The poses of the ligands were sorted using the best GOLD fitness scores and most favourable binding affinities of Autodock vina for each ligand. To identify the best protein ligand complex, Pearson correlation coefficients were estimated between the docking scores (GOLDScore, ChemScore, ChemPLP). Results delineated that Goldscore values are positively correlated with ChemPLP and negatively correlated with Chemscore. A comparison of the predicted binding energies and Goldscores of protein ligand complexes were demonstrated a strong correlation between the docking results as shown in Figure 7. Further validation was done by ADME properties that were well within the acceptable range thus consolidating their lead potential. For the estimation of connection between the experimental bioactivities and docking calculations, a squared correlation coefficient was calculated and found as acceptable with the values of 0.751 and 0.526 respectively (Supplemental Figure 1). The docking results of top 10 compounds are tabulated in Supplementary Table 1 in which GOLD fitness score and binding energy of methyl-CoA disulfide is 82.75 and -11kcal/mol respectively, so it is ranked as most active compound (Figure 8). The best hit, methyl-CoA disulfide has also been reported to be a potent competitive inhibitor of choline acetyltransferase by Currier and Mautner. Currier and Maunter proposed that the methyl-CoA disulfide forms a tight complex with the enzyme by reaction with thiol group of CoA [79, 80]. Critical evaluation of the nature of binding interaction of methyl-CoA disulfide revealed very strong contacts with the following protein residues Asp219, Arg254, Arg215, Arg37, Lys258, Asn279, Asn252, His249, Cys112, Met212, Gly152, Phe218, Phe157, Phe311, Thr38, Gly312, Ala251, Ile217, Thr38, Leu156, Leu189, Glu214 and Ser36 (Figure 8a,b). Methyl-CoA disulfide bound compactly at the binding tunnel of FabH. It formed a mixed

disulfide bond with thiol group of Cys112 at the bottom of a tunnel to inhibit the FabH (Figure 9). Previous studies of *Escherichia coli* FabH (ecFabH) and *Mycobacterium tuberculosis* FabH mention that methyl-CoA disulfide shows rapid inhibition of one monomer of FabH through formation of a methyl disulfide conjugate with Cys112 [81]. Additionally, close contacts of inhibitor with the side chain atoms of basic residues (Arg254, Arg215, Arg37 and Lys258) and main chain atoms of aromatic residue (Phe218) flanked the active site entrance. The bond with main chain atoms of Gly152 and Gly312 is critical to obstruct access to the active site. The results revealed prominent hydrogen bonds with the ligand. This can be observed in the 2D interaction images and details are provided in Supplementary Table 2. The presence of the hydrogen bond interaction leads to higher stability of the protein-ligand complex and provides a specificity of interaction between protein and ligand which is fundamental aspect of molecular recognition [82]. Hydrogen bonds also contribute towards the formation of a concrete and stable ligand-receptor complex structure.

Molecular Dynamics Simulation

Simulation results were attained in order to check stability of dimeric structure of *S. gordonii* FabH. Different examples are present in which time dependent molecular dynamics have been applied on docked complexes to study the protein-protein interactions, conformational fluctuations, structural, architectural changes and dynamical properties of proteins [83, 84]. MD simulation results were analysed by plotting different time dependent graphs and by processing trajectory files. These results allowed us to speculate the conformational variations and stability of protein structure with inhibitor during entire simulation. In order to examine conformational variations of *S. gordonii* FabH within hydrated environment, five physical properties of the system are analysed in which root mean square deviation (RMSD), root mean square fluctuation (RMSF), B-factor, radius of gyration, radial distribution function (RDF), potential energy and total energy plots are included.

RMSD of the atomic positions was calculated to determine the stability of protein system. The Cα atoms of undocked and docked complex of *S. gordonii* FabH were followed over the 10 ns and 40 ns time period, respectively. The undocked dimer structure of FabH protein remained stable during the entire simulation with the average RMSD value of 1.7 Å and a maximum value

of 2.4 Å, while the docked complex showed some fluctuations at 32-39 ns and had average RMSD of 1.6 Å, with maximum value of 2.1 Å (Figure 10). The RMSD of the system stays well

below 2 Å [83]. Superimposed RMSD graphs of undocked and docked complex FabH protein

illustrate stable behavior while docked complex has less RMSD value as compared to undocked protein (Figure 11).

Throughout the simulation, no drastic increase in RMSD value has been observed that indicates the overall stability of system. Stability of protein structure indicates that inhibitor is tightly bound with FabH protein and no conformational changes have been observed. Hence, selected inhibitor may serve as lead compound for further screening of FabH protein.

Root mean square fluctuation (RMSF) was calculated to check the average fluctuation of the docked and undocked protein residues over time. The undocked dimeric protein of FabH showed minimal fluctuations for most of residues throughout the simulation (Figure 12b). Fluctuated residues of undocked protein were mostly present in loop regions. Residues that had large fluctuations in undocked protein of FabH were 345-366, 469-480 and 560-570 with highest peak of 3.7 Å at 10^{th} ns. The fluctuations in FabH docked complex were analysed at 40^{th} ns that shows relatively minimal fluctuations at different positions (Figure 12a). Active site region is specified in Figure 12d by 1-19 numbers that indicates different peaks of RMSF plot. In the active site cleft residues Arg215, Asp219, Phe218, Ile217 and Thr38 had a highest fluctuation rate greater than 1 Å while Cys112, Gly312, Phe311, Asn279, Arg254, His249, Asn252, Ala251, Met212, Leu156 and Phe157 were within 1 Å (Figure 12c) suggesting the high stability and conservation of the binding pocket. The fluctuations of binding pocket architecture were calculated by variation in bond distances during the 40^{th} ns simulation to understand the dynamics of the catalytic residues. Inhibitor of *S. gordonii* FabH has disulfide bridge that S-S bridge makes a H-bond with His249, Ala251, Asn279, Cys112, Met212, His249 and Gly312 in first nanosecond.

Furthermore the per residue RMSF plot illustrates that catalytic triad of *S. gordonii* FabH distance between His249 and ligand S47 remains around 2.5 Å for the first to forty nanosecond that is shortest bond among all interactions. The distance of FabH Asn279 and ligand S47 in dimer structure remains 3.6 Å in first nanosecond but it is drastically increased in 40th ns simulation. The residue Cys112 and ligand S48 remains 3.7 Å in first nanosecond but it is lowered for the rest of the simulation. Interestingly our analysis inferred that fluctuations and differences in distances exhibit that active site tunnel has undergone slight changes in which Cys112 and His249 came closer to each other however Cys112 and Asn279 drifted away and create changes in architecture of the active site cleft (Figure 6c,d). Peak 5 contains Leu156 residue that is α-helix in 1 to 29 ns after that it converted into a loop in 30 to 40 ns (Figure 13).

Fluctuations of active site residues are in the range of 0.3 to 1.4 Å suggesting the high stability of the region while average RMSF value over the 40th ns turns out to be 0.7 Å that confirms the stability of system. All other peaks of active site do not show any fluctuations in architecture of protein structure. Other than active site residues 22-30, 198-206 and 519-527 had high fluctuations with highest peak of 4.6 Å. However, these residues are not present in active site cleft and dimer interface (Figure 14). After omitting these residues we again estimate RMSD, this graph shows that no drastic increase in RMSD value has been observed (Supplemental Figure 2). Hence residues of docked complex that showed high fluctuation rate are not catalytically active and do not effect the active site cleft. These results confirmed conservation of active site cleft and indicated that docked complex protein is in stabilizing form.

The polypeptide backbones and side chains of proteins are monitored as a function of time to assess the stability of protein and dynamic disorderedness of a system. B-factor depicts complete coherence with the RMSD and RMSF values. In RMSD and RMSF, active site cleft remains stable in entire time dependent behavior with minimal fluctuations and same in case of B-factor. The oscillations were seen at three points that show maximum number of fluctuations in comparison of whole protein. The B-factor calculated from RMSF values shows the significant fluctuations indicating the disorderedness of C-alpha atoms. The B-factor was investigated over the course of the simulation in the undocked and docked *S. gordonii* FabH protein in which residues from 350-360, 470-475 and 570-575 are found to fluctuate by a greater degree (Figure 15a, b). These residues are not part of active site cleft and dimer interface. However, the overall

average value of undocked and docked complex of *S. gordonii* FabH protein is 23.7 Å and 29.7 Å respectively (Figure 15c). In present study, average value of active site residues not exceed from 30 Å that determines the stability of ligand binding within the active site cleft (Figure 15d). The relative stability of RMSF with B-factor values also validated correct ligand selection.

Another indicator of stability of protein is radius of gyration which describes the overall shape and compactness of a protein. It is observed as RMSD between protein's center of gravity and its ends. The radius of gyration for the undocked and docked complex during 10 and 40 ns, respectively, measured which is indicated that overall system remains stable throughout the simulation time. The average value of undocked and docked complex in FabH protein is 20.25 Å (Figure 16a) and 19.9 Å (Figure 16b), respectively, which confirmed the stability of protein.

Radial distribution graphs provide an indirect estimation of the location, on the surface or in the interior, of different residue types [85]. It is a criterion for finding the distribution of atoms, molecules or other species around target specially interacting residues that are crucial for the stability of ligand during the ligand binding process [86]. Hydrogen bond interactions were calculated to check the stability of binding pocket residues (Arg215, Asn252, Asn279, His249 and Cys112) that efficiently trapped FabH inhibitor which remained tightly bound over the course of simulation. The RDF between the H-atoms of His249, Cys112, and Asn279 was calculated during the 40 ns simulation to understand the dynamics of the catalytic residues. In His249, largest peak of the RDF graph before simulation appears at 2.5 Å having a g(r) value of 0.37 whereas the largest peak after simulation appears at 2.7 Å having a g(r) value of 0.47. After MD simulations, the peak narrowed, suggesting that during the MD simulation inhibitor spend sufficient time with His249 residue of protein to exhibit activity (Figure 17). For Asn279, largest peak of the RDF graph before simulation appears at 3.6 Å having a g(r) value of 0.29 whereas the largest peak after simulation appears at 2.3 Å having a g(r) value of 0.36. A comparison of peaks in Figure 18 illustrated that the peak after MD simulation is more pronounced in magnitude and narrower in width than that of one before simulation, rendering that O atom of ligand moves toward protein's active site. The rotation as shown in Figure 19 resulted in bringing the ligand in vicinity of the protein residue Cys112 which resulted in the formation of a hydrogen bond. The interaction diagram taken at both the start and end of simulation gives evidence of residue involved in bond formation due to movement of the ligand. Relatively

higher fluctuation in the hydrogen bond of residue Cys112 suggests that the hydrogen bond is less stable as compared to the other two hydrogen bonds formed between catalytic triad residues and ligand. Active site residue Arg215 has highest peak at 3.3 Å and g(r) value of 0.21 before simulation while the after simulation has highest peak at 3.8 Å and g(r) value of 0.30. It was observed that peak after simulation was high but was distorted, probably because of steric hindrance effect whereas it got narrow and defined, indicating the shortening of hydrogen bond distance and increasing probability of finding ligand around Arg215 (Figure 20). Residue Asn252 has largest peak at 2.1 Å and g(r) value of 0.26 before simulation whereas after simulation largest peak is at 2.1 Å with g(r) value of 0.25. This indicates that peak before simulation and after simulation was high but distorted possibly due to steric hindrance effect. After MD simulations, the peak narrowed suggesting that during the MD simulation the Asn252 and H protons spend sufficient time close to one another to exhibit activity (Figure 21). RDFs for other active site residues Phe218, Phe213 and Arg37 suggested fidelity of the composed system which is a contributing factor in the formation of stable complexes. Overall, RDF results are coherent with aforementioned MD simulation and docking results.

Current study demonstrated that undocked dimer structure of FabH protein remained stable during the entire simulation. It showed that homolog model structure is rationally accurate and analogous to template. While in docked complex we observed that pattern of simulation graph does not support any extensive domain shifts within the structural framework of the protein-ligand complex. Active site tunnel has a conserved catalytic triad and dimeric docked complex that remains stable as compared to undocked protein. Thus, It signifies that the ligand placement well complemented within the active site and does not destabilize the protein.

Our finding consolidates aforementioned studies, proposing that predicted ligand (methyl-CoA disulfide) acts as a potent inhibitor that target exclusively the pathogen’s system, without harming the biological and metabolic pathways of host.

Conclusion

Current study highlighted a potential drug target (FabH) that is key to oral infection in which methyl-CoA disulfide acts as a potent inhibitor. The expression of FabH protein in bacteria

makes it an ideal target for the identification of novel antibacterial agents. Our findings, afford a comprehensive insight into inhibitor design against this protein whereby preferred orientation can be used to calculate binding affinities for a wide range of inhibitors. The entire course of simulation indicates that protein has achieved equilibration and reached the lower energy state. The stability of FabH protein with bound to inhibitor methyl-CoA disulfide can be inferred from the current results, as compared the undocked protein. The selected inhibitor has high affinity to inhibit dimeric structure of FabH. Estimated correlation coefficient values and important interactions affirm the inhibition of FabH activity and provide a structural framework for the design of novel drugs.

Acknowledgements

Authors are highly grateful to Higher Education Commission (HEC) Pakistan for financial support.

Author Information

Corresponding Author
*E-mail: ssazam@qau.edu.pk; syedazam2008@gmail.com

Funding

Grants are not available to support this research.

References

1. Raju, R.M., A.L. Goldberg, and E.J. Rubin, *Bacterial proteolytic complexes as therapeutic targets.* Nature Reviews Drug Discovery, 2012. **11**(10): p. 777-789.
2. Zheng, L., et al., *CcpA regulates biofilm formation and competence in Streptococcus gordonii.* Molecular Oral Microbiology, 2012. **27**(2): p. 83-94.
3. Kolenbrander, P.E., et al., *Bacterial interactions and successions during plaque development.* Periodontology 2000, 2006. **42**: p. 47-79.
4. Hoiby, N., et al., *Antibiotic resistance of bacterial biofilms.* International Journal of Antimicrobial Agents, 2010. **35**(4): p. 322-332.

5. Yombi, J.C., et al., *Streptococcus gordonii septic arthritis : two cases and review of literature.* Bmc Infectious Diseases, 2012. **12**.
6. Mathews, C.J., et al., *Management of septic arthritis: a systematic review.* Annals of the Rheumatic Diseases, 2007. **66**(4): p. 440-445.
7. Heath, R.J. and C.O. Rock, *The Claisen condensation in biology.* Natural Product Reports, 2002. **19**(5): p. 581-596.
8. Tsay, J.T., et al., *Isolation and Characterization of the Beta-Ketoacyl-Acyl Carrier Protein Synthase-Iii Gene (Fabh) from Escherichia-Coli K-12.* Journal of Biological Chemistry, 1992. **267**(10): p. 6807-6814.
9. Revill, W.P., et al., *Beta-ketoacyl acyl carrier protein synthase III (FabH) is essential for fatty acid biosynthesis in Streptomyces coelicolor A3(2).* Journal of Bacteriology, 2001. **183**(11): p. 3526-30.
10. Lai, C.Y. and J.E. Cronan, *beta-ketoacyl-acyl carrier protein synthase III (FabH) is essential for bacterial fatty acid synthesis.* Journal of Biological Chemistry, 2003. **278**(51): p. 51494-51503.
11. Li, Y.L., G. Florova, and K.A. Reynolds, *Alteration of the fatty acid profile of Streptomyces coelicolor by replacement of the initiation enzyme 3-ketoacyl acyl carrier protein synthase III (FabH).* Journal of Bacteriology, 2005. **187**(11): p. 3795-3799.
12. Azam, S.S. and A. Shamim, *An insight into the exploration of druggable genome of Streptococcus gordonii for the identification of novel therapeutic candidates.* Genomics, 2014. **104**(3): p. 203-2014.
13. Lee, P.J., et al., *Targeting the Fatty Acid Biosynthesis Enzyme, beta-Ketoacyl-Acyl Carrier Protein Synthase III (PfKASIII), in the Identification of Novel Antimalarial Agents.* Journal of Medicinal Chemistry, 2009. **52**(4): p. 952-963.
14. Prigge, S.T., et al., *The initiating steps of a type II fatty acid synthase in Plasmodium falciparum are catalyzed by pfACP, pfMCAT, and pfKASIII.* Biochemistry, 2003. **42**(4): p. 1160-1169.
15. Choi, K.H., et al., *Identification and substrate specificity of beta -ketoacyl (acyl carrier protein) synthase III (mtFabH) from Mycobacterium tuberculosis.* J Biol Chem, 2000. **275**(36): p. 28201-7.
16. He, X. and K.A. Reynolds, *Purification, characterization, and identification of novel inhibitors of the beta-ketoacyl-acyl carrier protein synthase III (FabH) from Staphylococcus aureus.* Antimicrobial Agents and Chemotherapy, 2002. **46**(5): p. 1310-1318.
17. Khandekar, S.S., et al., *Identification, substrate specificity, and inhibition of the Streptococcus pneumoniae beta-ketoacyl-acyl carrier protein synthase III (FabH).* J Biol Chem, 2001. **276**(32): p. 30024-30.
18. Choi, K.H., R.J. Heath, and C.O. Rock, *beta-ketoacyl-acyl carrier protein synthase III (FabH) is a determining factor in branched-chain fatty acid biosynthesis.* Journal of Bacteriology, 2000. **182**(2): p. 365-370.
19. Clough, R.C., et al., *Purification and Characterization of 3-Ketoacyl-Acyl Carrier Protein Synthase-Iii from Spinach - a Condensing Enzyme Utilizing Acetyl-Coenzyme-a to Initiate Fatty-Acid Synthesis.* Journal of Biological Chemistry, 1992. **267**(29): p. 20992-20998.

20. Davies, C., et al., *The 1.8 A crystal structure and active-site architecture of beta-ketoacyl-acyl carrier protein synthase III (FabH) from escherichia coli.* Structure, 2000. **8**(2): p. 185-95.
21. Qiu, X.Y., et al., *Refined structures of beta-ketoacyl-acyl carrier protein synthase III.* Journal of Molecular Biology, 2001. **307**(1): p. 341-356.
22. Heath, R.J., et al., *Broad spectrum antimicrobial biocides target the FabI component of fatty acid synthesis.* Journal of Biological Chemistry, 1998. **273**(46): p. 30316-30320.
23. Rozwarski, D.A., et al., *Modification of the NADH of the isoniazid target (InhA) from Mycobacterium tuberculosis.* Science, 1998. **279**(5347): p. 98-102.
24. Kauppinen, S., M. Siggaardandersen, and P. Vonwettsteinknowles, *Beta-Ketoacyl-Acp Synthase-I of Escherichia-Coli - Nucleotide-Sequence of the Fabb Gene and Identification of the Cerulenin Binding Residue.* Carlsberg Research Communications, 1988. **53**(6): p. 357-370.
25. Moche, M., et al., *Structure of the complex between the antibiotic cerulenin and its target, beta-ketoacyl-acyl carrier protein synthase.* Journal of Biological Chemistry, 1999. **274**(10): p. 6031-6034.
26. Jackowski, S., et al., *Acetoacetyl-Acyl Carrier Protein Synthase - a Target for the Antibiotic Thiolactomycin.* Journal of Biological Chemistry, 1989. **264**(13): p. 7624-7629.
27. Tsay, J.T., C.O. Rock, and S. Jackowski, *Overproduction of Beta-Ketoacyl-Acyl Carrier Protein Synthase-I Imparts Thiolactomycin Resistance to Escherichia-Coli K-12.* Journal of Bacteriology, 1992. **174**(2): p. 508-513.
28. Wang, J., et al., *Discovery of platencin, a dual FabF and FabH inhibitor with in vivo antibiotic properties.* Proc Natl Acad Sci U S A, 2007. **104**(18): p. 7612-6.
29. Gajiwala, K.S., et al., *Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme.* FEBS Lett, 2009. **583**(17): p. 2939-2946.
30. Eswar, N., et al., *Comparative protein structure modeling using Modeller.* Curr Protoc Bioinformatics, 2006. **Chapter 5**: p. Unit 5 6.
31. Arnold, K., et al., *The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.* Bioinformatics, 2006. **22**(2): p. 195-201.
32. Roy, A., A. Kucukural, and Y. Zhang, *I-TASSER: a unified platform for automated protein structure and function prediction.* Nature Protocols, 2010. **5**(4): p. 725-738.
33. Pieper, U., et al., *ModBase, a database of annotated comparative protein structure models, and associated resources.* Nucleic Acids Res, 2011. **39**: p. D465-D474.
34. Bates, P.A., et al., *Enhancement of protein modeling by human intervention in applying the automatic programs 3D-JIGSAW and 3D-PSSM.* Proteins-Structure Function and Genetics, 2001: p. 39-46.
35. Chen, C.C., J.K. Hwang, and J.M. Yang, *(PS)(2): protein structure prediction server.* Nucleic Acids Res, 2006. **34**: p. W152-W157.
36. Kelley, L.A. and M.J.E. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server.* Nature Protocols, 2009. **4**(3): p. 363-371.
37. Fernandez-Fuentes, N., et al., *M4T: a comparative protein structure modeling server.* Nucleic Acids Res, 2007. **35**: p. W363-W368.
38. Kallberg, M., et al., *Template-based protein structure modeling using the RaptorX web server.* Nature Protocols, 2012. **7**(8): p. 1511-1522.

39. Soding, J., *Protein homology detection by HMM-HMM comparison.* Bioinformatics, 2005. **21**(7): p. 951-960.
40. Lambert, C., et al., *ESyPred3D: Prediction of proteins 3D structures.* Bioinformatics, 2002. **18**(9): p. 1250-1256.
41. McNicholas, S., et al., *Presenting your structures: the CCP4mg molecular-graphics software.* Acta Crystallographica Section D-Biological Crystallography, 2011. **67**: p. 386-394.
42. Pettersen, E.F., et al., *UCSF chimera - A visualization system for exploratory research and analysis.* Journal of Computational Chemistry, 2004. **25**(13): p. 1605-1612.
43. Duhovny, D., R. Nussinov, and H.J. Wolfson, *Efficient unbound docking of rigid molecules.* Algorithms in Bioinformatics, Proceedings, 2002. **2452**: p. 185-200.
44. Laskowski, R.A., et al., *Procheck - a Program to Check the Stereochemical Quality of Protein Structures.* Journal of Applied Crystallography, 1993. **26**: p. 283-291.
45. Wiederstein, M. and M.J. Sippl, *ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins.* Nucleic Acids Res, 2007. **35**(Web Server issue): p. W407-10.
46. Eisenberg, D., R. Luthy, and J.U. Bowie, *VERIFY3D: Assessment of protein models with three-dimensional profiles.* Macromolecular Crystallography, Pt B, 1997. **277**: p. 396-404.
47. Colovos, C. and T.O. Yeates, *Verification of Protein Structures - Patterns of Nonbonded Atomic Interactions.* Protein Science, 1993. **2**(9): p. 1511-1519.
48. McGuffin, L.J., K. Bryson, and D.T. Jones, *The PSIPRED protein structure prediction server.* Bioinformatics, 2000. **16**(4): p. 404-405.
49. Eramian, D., et al., *How well can the accuracy of comparative protein structure models be predicted?* Protein Science, 2008. **17**(11): p. 1881-1893.
50. McGuffin, L.J., *The ModFOLD server for the quality assessment of protein structural models.* Bioinformatics, 2008. **24**(4): p. 586-587.
51. Benkert, P., M. Biasini, and T. Schwede, *Toward the estimation of the absolute quality of individual protein structure models.* Bioinformatics, 2011. **27**(3): p. 343-350.
52. Crooks, G.E., et al., *WebLogo: A sequence logo generator.* Genome Research, 2004. **14**(6): p. 1188-1190.
53. Castillo, Y.P. and M.A.C. Perez, *Bacterial beta-ketoacyl-acyl carrier protein synthase III (FabH): An attractive target for the design of new broad-spectrum antimicrobial agents.* Mini-Reviews in Medicinal Chemistry, 2008. **8**(1): p. 36-45.
54. Zhang, H.J., et al., *Synthesis, molecular modeling and biological evaluation of beta-ketoacyl-acyl carrier protein synthase III (FabH) as novel antibacterial agents.* Bioorganic & Medicinal Chemistry, 2011. **19**(15): p. 4513-4519.
55. Li, Z.L., et al., *Design, synthesis and biological evaluation of urea derivatives from o-hydroxybenzylamines and phenylisocyanate as potential FabH inhibitors.* Bioorganic & Medicinal Chemistry, 2011. **19**(15): p. 4413-4420.
56. Wen, L., et al., *Functional expression of Francisella tularensis FabH and FabI, potential antibacterial targets.* Protein Expression and Purification, 2009. **65**(1): p. 83-91.
57. Lv, P.C., et al., *Design, synthesis, and structure-activity relationships of pyrazole derivatives as potential FabH inhibitors.* Bioorganic & Medicinal Chemistry Letters, 2010. **20**(15): p. 4657-4660.

58. Lv, P.C., et al., *Design, synthesis and biological evaluation of novel thiazole derivatives as potent FabH inhibitors.* Bioorganic & Medicinal Chemistry Letters, 2009. **19**(23): p. 6750-6754.
59. Nie, Z., et al., *Structure-based design, synthesis, and study of potent inhibitors of beta-ketoacyl-acyl carrier protein synthase III as potential antimicrobial agents.* Journal of Medicinal Chemistry, 2005. **48**(5): p. 1596-1609.
60. Yang, Y.S., et al., *Discovery and modification of sulfur-containing heterocyclic pyrazoline derivatives as potential novel class of beta-ketoacyl-acyl carrier protein synthase III (FabH) inhibitors.* Bioorganic & Medicinal Chemistry Letters, 2012. **22**(14): p. 4619-4624.
61. Alhamadsheh, M.M., et al., *Synthesis and biological evaluation of thiazolidine-2-one 1,1-dioxide as inhibitors of Escherichia coli beta-ketoacyl-ACP-synthase III (FabH).* Bioorganic & Medicinal Chemistry Letters, 2007. **17**(4): p. 879-883.
62. Al-Balas, Q., et al., *Identification of 2-Aminothiazole-4-Carboxylate Derivatives Active against Mycobacterium tuberculosis H37Rv and the beta-Ketoacyl-ACP Synthase mtFabH.* Plos One, 2009. **4**(5).
63. He, X., et al., *1,2-dithiole-3-ones as potent inhibitors of the bacterial 3-ketoacyl acyl carrier protein synthase III (FabH).* Antimicrobial Agents and Chemotherapy, 2004. **48**(8): p. 3093-3102.
64. Jones, A.L., et al., *Beta-ketoacyl-acyl carrier protein synthase III from pea (Pisum sativum L.): properties, inhibition by a novel thiolactomycin analogue and isolation of a cDNA clone encoding the enzyme.* Planta, 2003. **216**(5): p. 752-61.
65. Jones, G., et al., *Development and validation of a genetic algorithm for flexible docking.* J Mol Biol, 1997. **267**(3): p. 727-48.
66. Jones, G., et al., *Development and validation of a genetic algorithm for flexible ligand docking.* Abstracts of Papers of the American Chemical Society, 1997. **214**: p. 154-COMP.
67. Trott, O. and A.J. Olson, *AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading.* J Comput Chem, 2010. **31**(2): p. 455-61.
68. Morris, G.M., et al., *AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility.* Journal of Computational Chemistry, 2009. **30**(16): p. 2785-91.
69. Berry, D., *R Programming for Bioinformatics.* Journal of the Royal Statistical Society Series a-Statistics in Society, 2009. **172**: p. 696-696.
70. Humphrey, W., A. Dalke, and K. Schulten, *VMD: visual molecular dynamics.* J Mol Graph, 1996. **14**(1): p. 33-8, 27-8.
71. Weiner, P.K. and P.A. Kollman, *Amber - Assisted Model-Building with Energy Refinement - a General Program for Modeling Molecules and Their Interactions.* Journal of Computational Chemistry, 1981. **2**(3): p. 287-303.
72. Roe, D.R. and T.E. Cheatham, *PTRAJ and CPPTRAJ: Software for Processing and Analysis of Molecular Dynamics Trajectory Data.* Journal of Chemical Theory and Computation, 2013. **9**(7): p. 3084-3095.
73. Wang, J.M., et al., *Automatic atom type and bond type perception in molecular mechanical calculations.* Journal of Molecular Graphics & Modelling, 2006. **25**(2): p. 247-260.

74. Case, D.A., et al., *The Amber biomolecular simulation programs.* Journal of Computational Chemistry, 2005. **26**(16): p. 1668-1688.
75. Andersen, H.C., *Rattle - a Velocity Version of the Shake Algorithm for Molecular-Dynamics Calculations.* Journal of Computational Physics, 1983. **52**(1): p. 24-34.
76. Berendsen, H.J.C., et al., *Molecular-Dynamics with Coupling to an External Bath.* Journal of Chemical Physics, 1984. **81**(8): p. 3684-3690.
77. Ramamoorthy, D., E. Turos, and W.C. Guida, *Identification of a New Binding Site in E. coli FabH using Molecular Dynamics Simulations: Validation by Computational Alanine Mutagenesis and Docking Studies.* J Chem Inf Model, 2013. **53**(5): p. 1138-1156.
78. Qiu, X.Y., et al., *Crystal structure of beta-ketoacyl-acyl carrier protein synthase III - A key condensing enzyme in bacterial fatty acid biosynthesis.* Journal of Biological Chemistry, 1999. **274**(51): p. 36465-36471.
79. Currier, S.F. and H.G. Mautner, *Evidence for a Thiol Reagent Inhibiting Choline-Acetyltransferase by Reacting with Thiol-Group of Coenzyme-a Forming a Potent Inhibitor.* Biochemical and Biophysical Research Communications, 1976. **69**(2): p. 431-436.
80. Currier, S.F. and H.G. Mautner, *Interaction of Analogs of Coenzyme-a with Choline-Acetyltransferase.* Biochemistry, 1977. **16**(9): p. 1944-1948.
81. Alhamadsheh, M.M., et al., *Alkyl-CoA disulfides as inhibitors and mechanistic probes for FabH enzymes.* Chemistry & Biology, 2007. **14**(5): p. 513-524.
82. Panigrahi, S.K., *Strong and weak hydrogen bonds in protein-ligand complexes of kinases: a comparative study.* Amino Acids, 2008. **34**(4): p. 617-633.
83. Azam, S.S., R. Uddin, and A. Wadood, *Structure and dynamics of alpha-glucosidase through molecular dynamics simulation studies.* Journal of Molecular Liquids, 2012. **174**: p. 58-62.
84. Azam, S.S., et al., *Role of N-acetylserotonin O-methyltransferase in bipolar disorders and its dynamics.* Journal of Molecular Liquids, 2013. **182**: p. 25-31.
85. Donohue, J., *Radial Distribution Functions of Some Structures of the Polypeptide Chain.* Proc Natl Acad Sci U S A, 1954. **40**(6): p. 377-81.
86. Azam, S.S. and S.W. Abbasi, *Molecular docking studies for the identification of novel melatoninergic inhibitors for acetylserotonin-O-methyltransferase using different docking routines.* Theoretical Biology and Medical Modelling, 2013. **10**.

Table 1: Analysis of modeller structures for prediction of *Streptococcus gordonii* FabH model.

Modeller	Most favoured regions	Additional allowed regions	Generously allowed regions	Disallowed regions	Non-glycine and non-proline residues	Total number of residues	G-Factors	Molpdf score	DO sc
Model 1	271	19	2	0	292	324	-0.12	1908.45	-3
	92.80%	6.50%	0.70%	0.00%	100.00%				
Model 2	268	21	3	0	292	324	-0.11	2043.00	-3
	91.80%	7.20%	1.00%	0.00%	100.00%				
Model 3	266	23	2	1	292	324	0.36	1972.91	-3
	91.10%	7.90%	0.70%	0.3%*	100.00%				
Model 4	270	17	5	0	292	324	-0.15	2121.79	-3
	92.50%	5.80%	1.70%	0.00%	100.00%				
Model 5	269	19	3	1	292	324	-0.1	2028.01	-3
	92.10%	6.50%	1.00%	0.3%*	100.00%				
Minimized	238	49	3	2	292	324	-0.39	--	-3
Model 4	81.5%*	16.80%	1.00%	0.7%*	100.00%				
Dimer Model 4	476	98	6	4	584	648	-0.39	--	-7
	81.5%*	16.80%	1.00%	0.7%*	100.00%				

Table 2: Characteristics validating the homology model of *Streptococcus gordonii* FabH predicted from o

Online Web Server	Most favoured regions	Additional allowed regions	Generously allowed regions	Disallowed regions	Non-glycine and non-proline residues	Total number of residues	G-
Swiss Model	249	36	4	2	291		
	85.6%*	12.40%	1.40%	0.7%*	100.00%	323	-0.
(PS)2 Modeller	101	7	2	0	110		
	91.80%	6.40%	1.80%	0.00%	100.00%	126	-0.
(PS)2 Ramp	263	24	3	2	292		
	90.10%	8.20%	1.00%	0.7%*	100.00%	324	-0.
I-tasser Model 1	252	30	4	6	292		
	86.3%*	10.30%	1.40%	2.1%*	100.00%	324	-0.
Phyre	251	38	3	0	292		
	86.0%*	13.00%	1.00%	0.00%	100.00%	324	-0.
Modbase (Model 1)	264	21	3	0	288		
	91.70%	7.30%	1.00%	0.00%	100.00%	320	-0.
Modbase (Model 2)	259	27	3	1	290		
	89.3%*	9.30%	1.00%	0.3%*	100.00%	322	-0.
M4T Server	268	18	6	0	292		
	91.80%	6.20%	2.10%	0.00%	100.00%	324	-0.
	266	23	3	0	292	324	-0.

Raptorx	266	23	3	0	292	324	-0.
3D Jigsaw	230	51	7	3	291		
	79.0%**	17.50%	2.40%	1.0%*	100.00%	323	-0.
	258	23	6	3	290		
	89.0%*	7.90%	2.10%	1.0%*	100.00%		

Figure 1. The schematic workflow is illustrating the complete hierarchy of docked and undocked protein analysis.

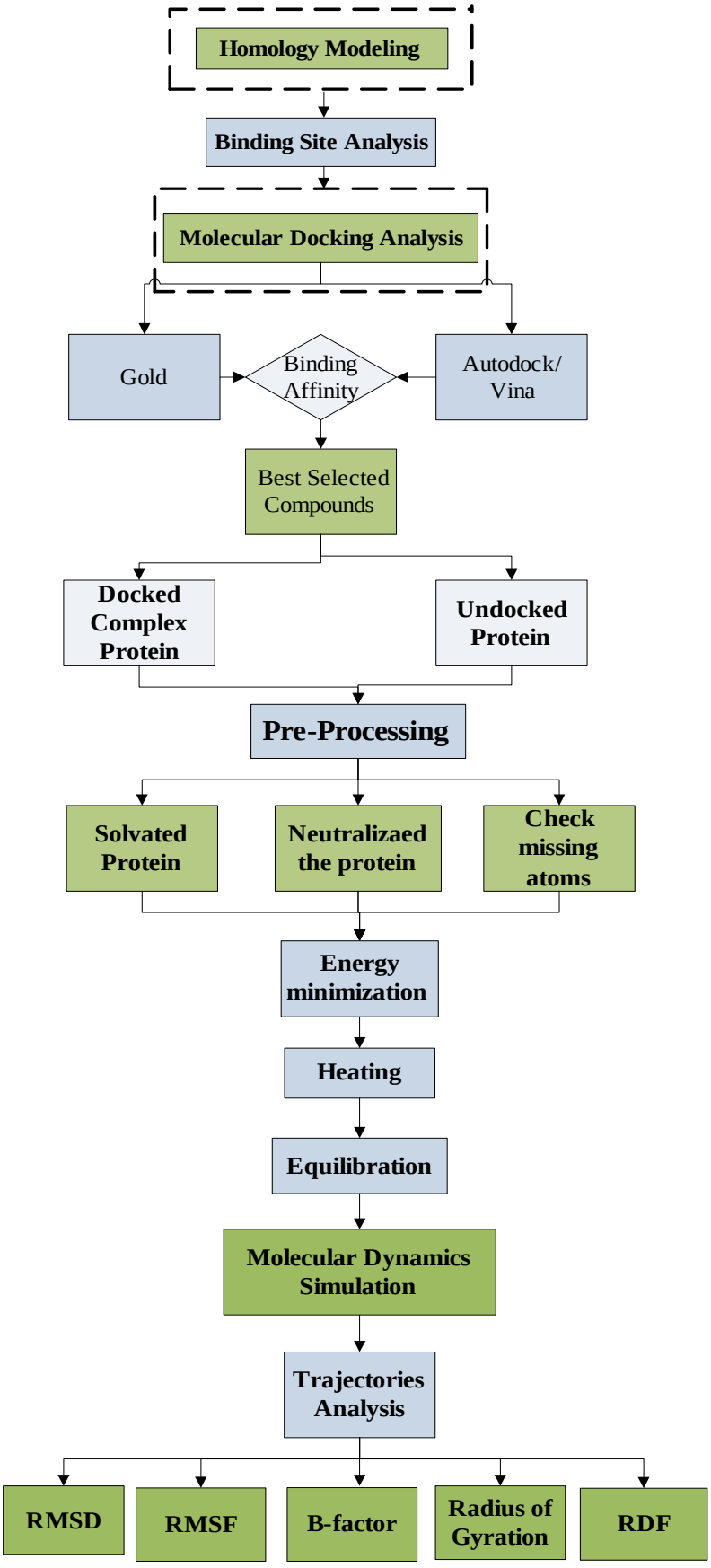


Figure 2. Depiction of secondary structure element of FabH protein.

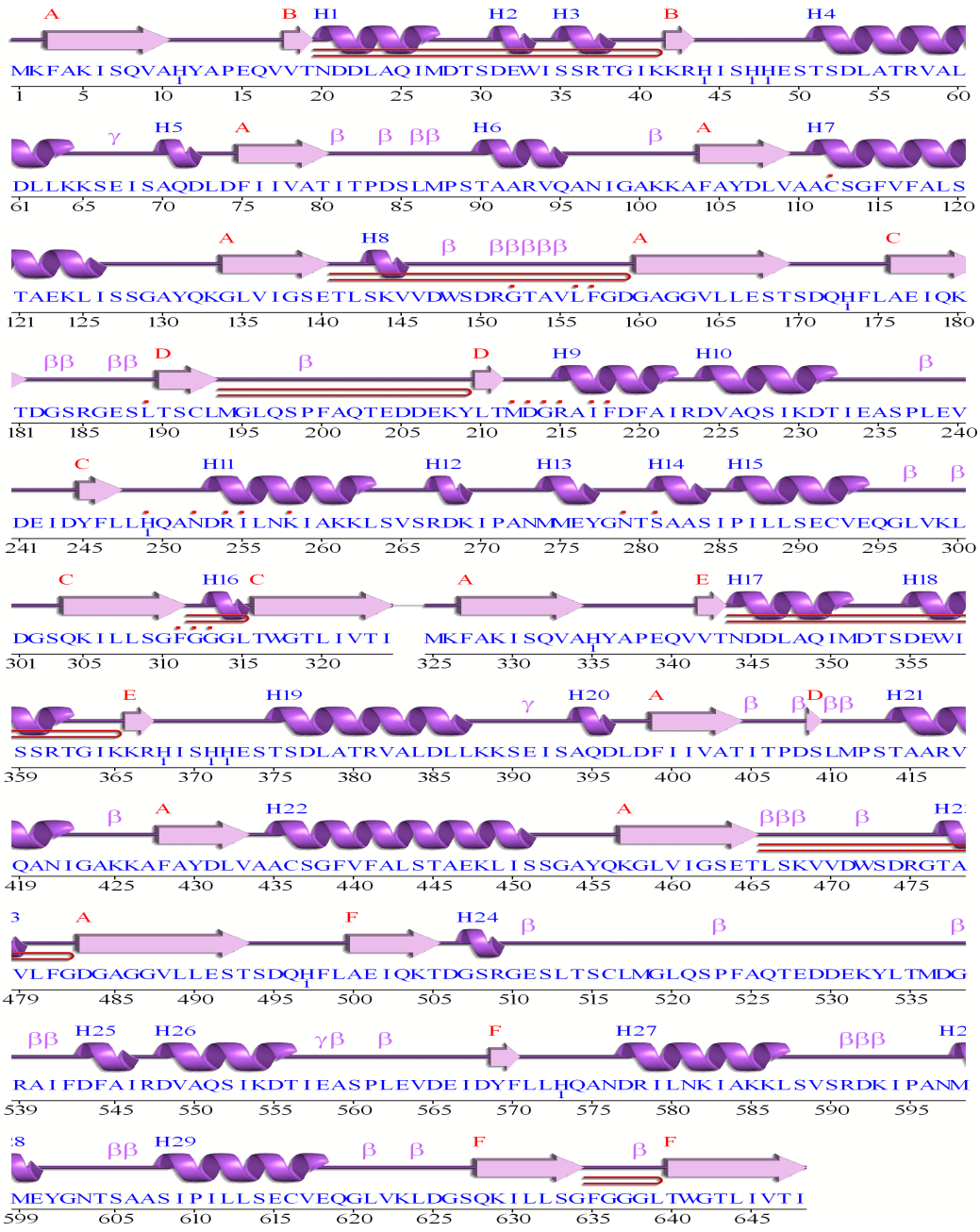


Figure 3. DOPE plot of minimized model of *S. gordonii* FabH and template *E. faecalis* FabH (PDB id: 3il5) in which protein's residue numbers are on the x-axis and the DOPE score is plotted on the y-axis.

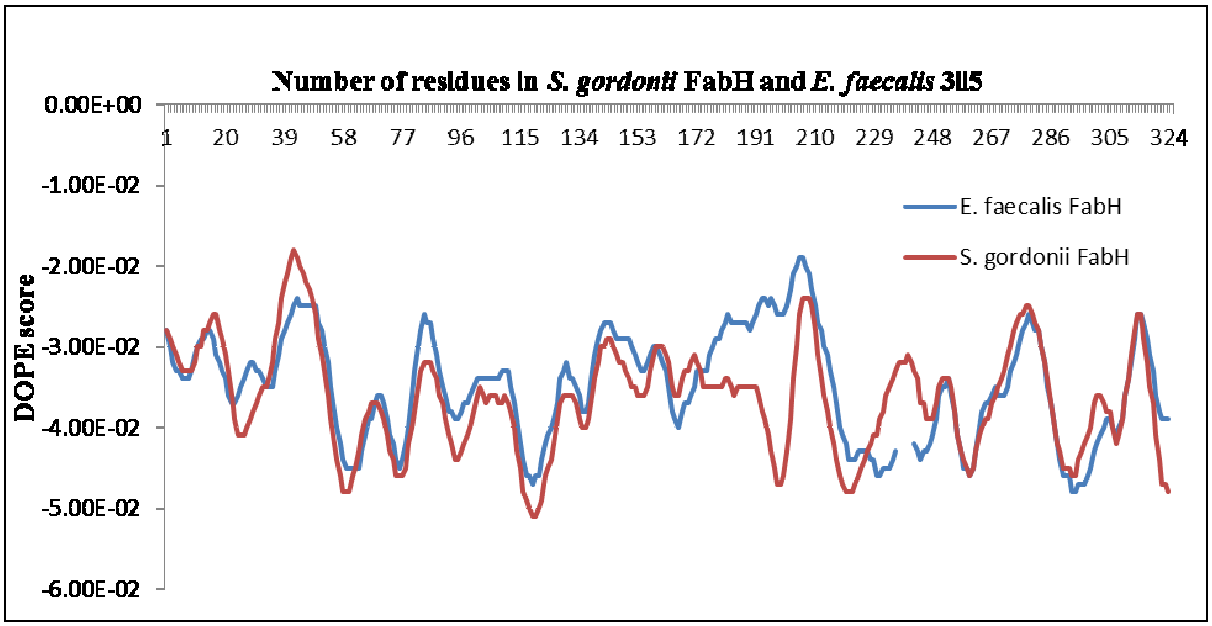


Figure 4. Pro 88 and Leu 315 of *S. gordonii* FabH protein residues are found in immediate vicinity of each other by virtue of dimerization.

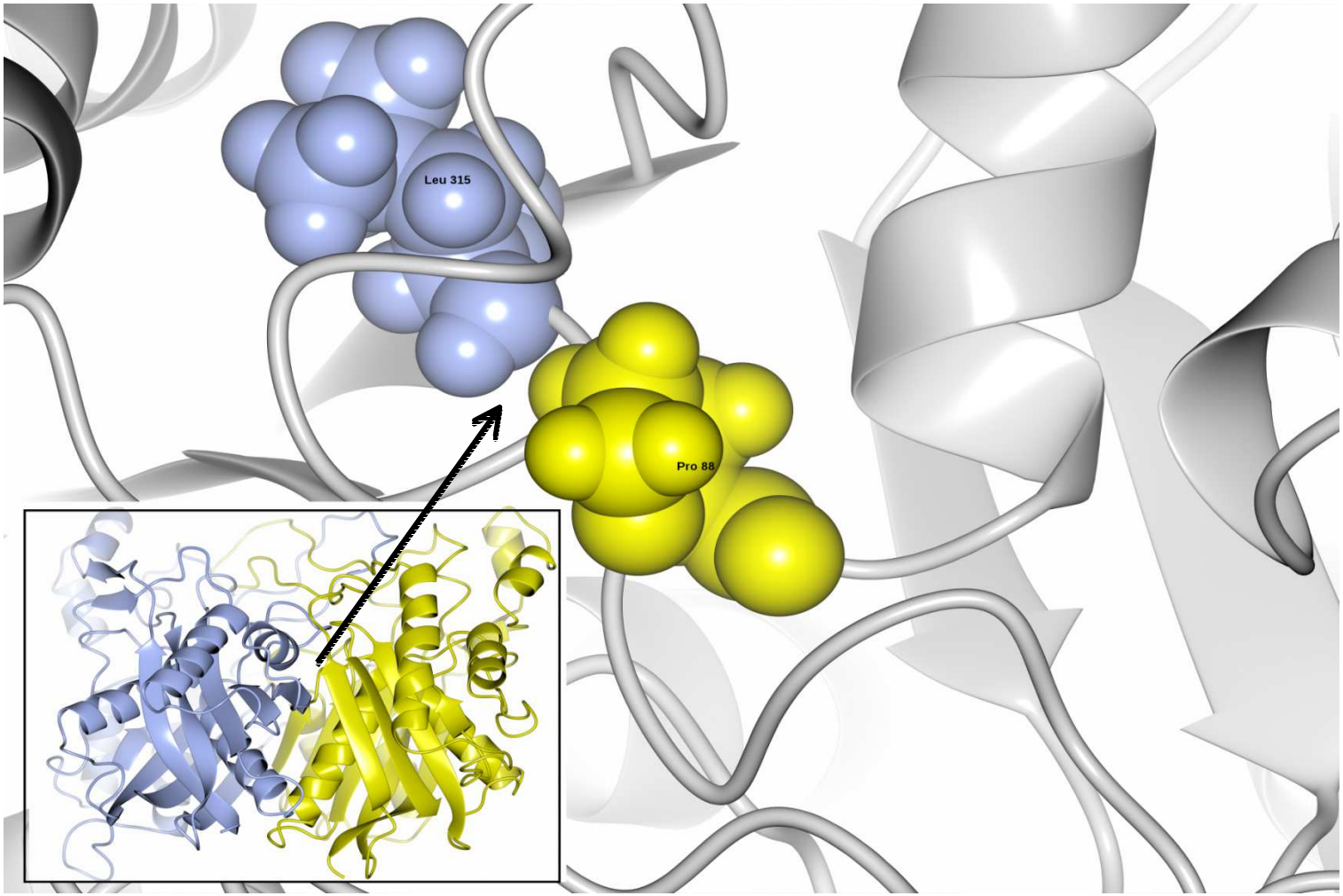


Figure 5. Docked pose of FabH, highlighting the structure conservation and most active residues of protein's binding pocket.

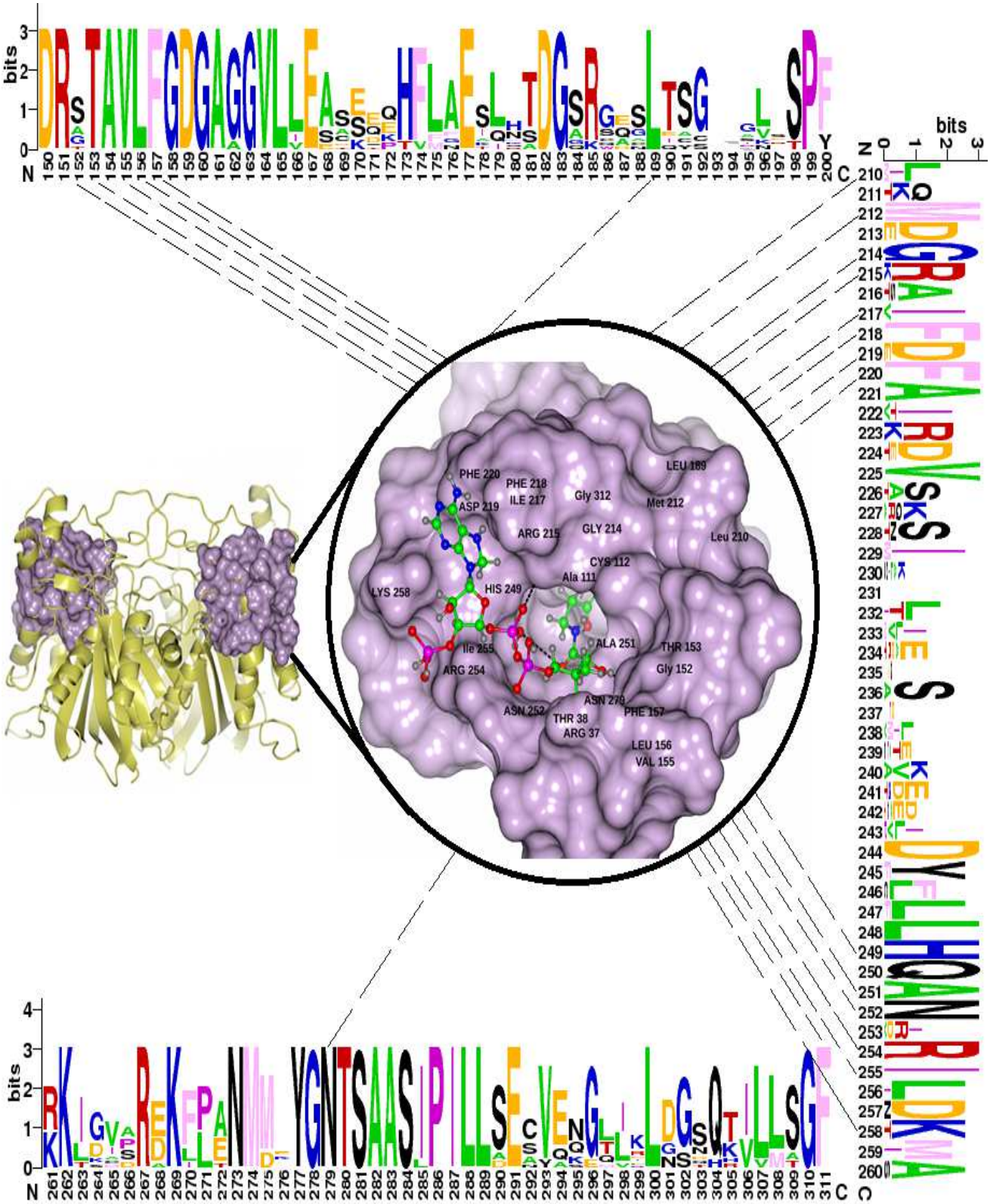
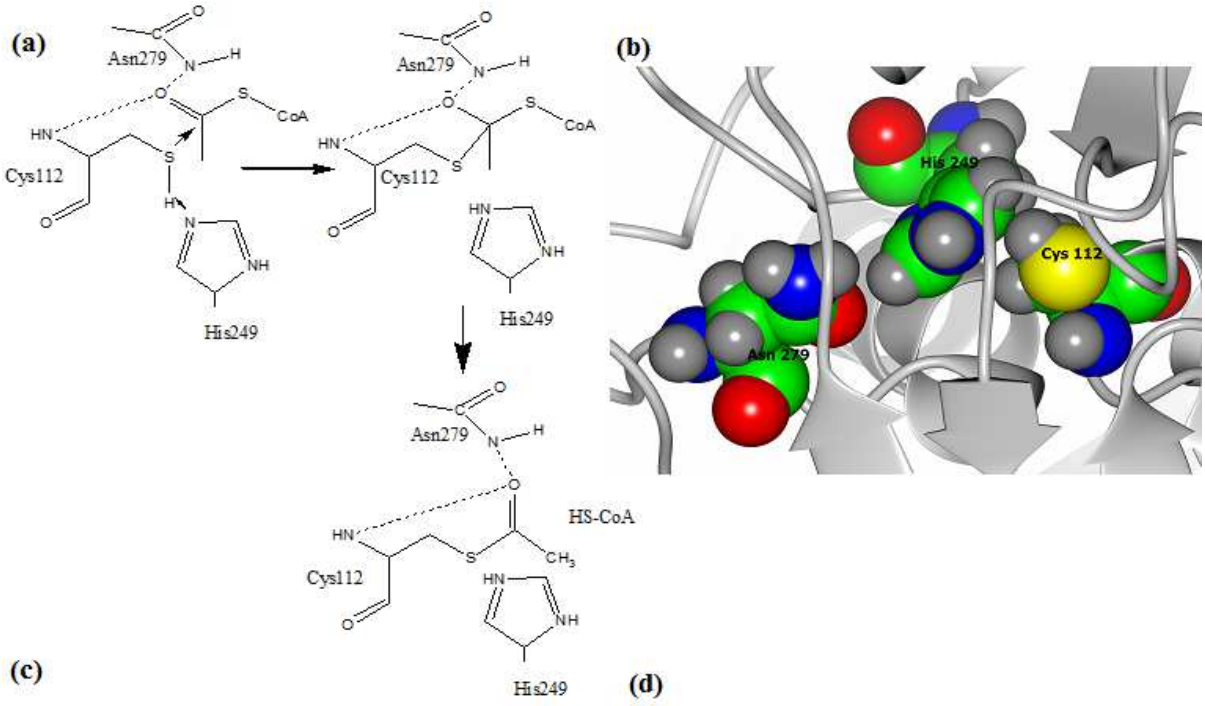


Figure 6. (a) Proposed mechanism of condensation reaction in *S.gordonii* FabH between His249, Asn279 and Cys112**.** (b) Representation of the catalytic triad (Cys112, Asn279 and His249) here active site tunnel of FabH represented as ribbons and showing the catalytic triad in CPK. (c) Showing the distance of catalytic triad with disulfide bond of inhibitor at first nanosecond. (d) Distance of catalytic triad at 40th ns.



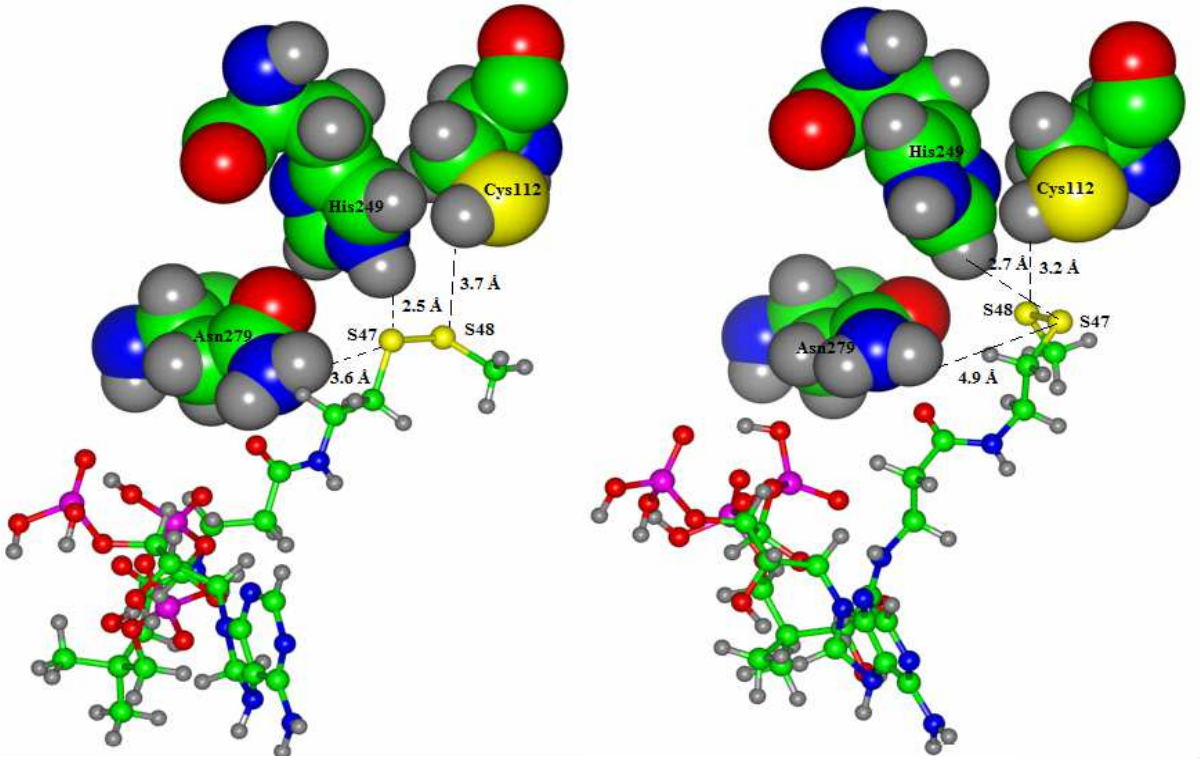


Figure 7. Correlation coefficient is calculated between GOLD fitness score and Autodock/vina binding energy. Red spot indicated the lead compound that is methyl-CoA disulfide.

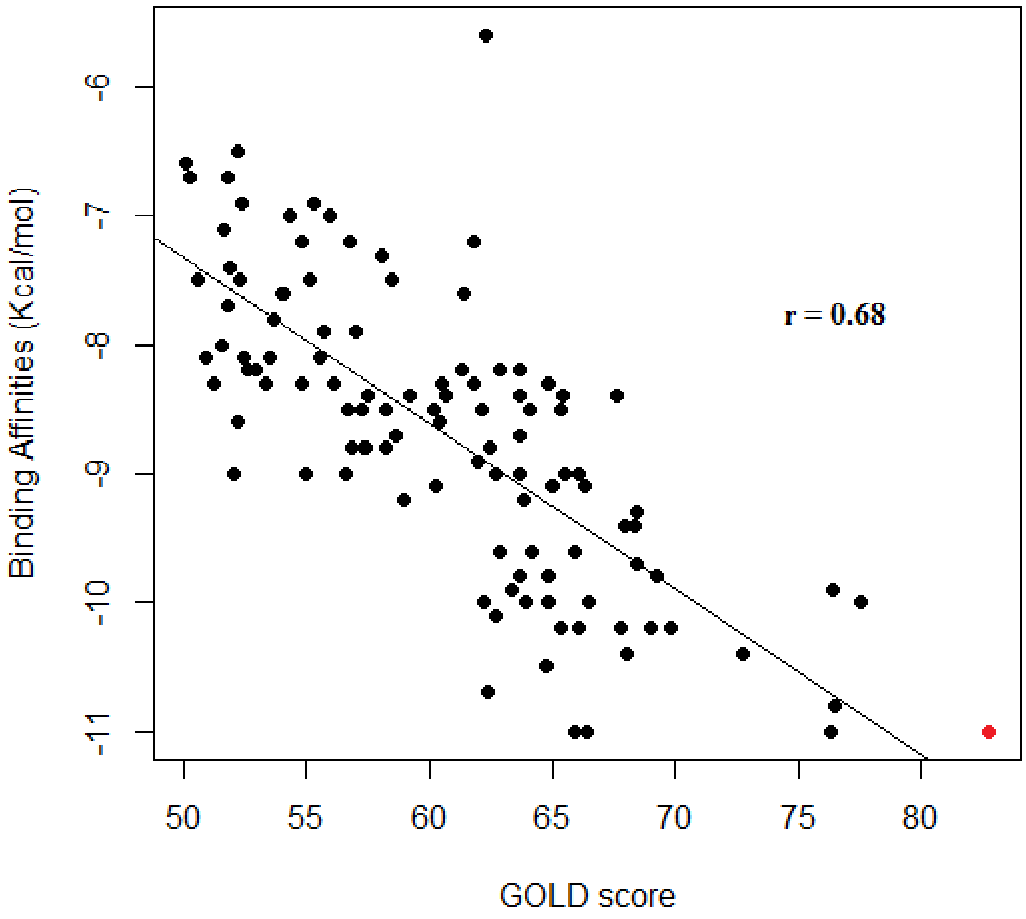
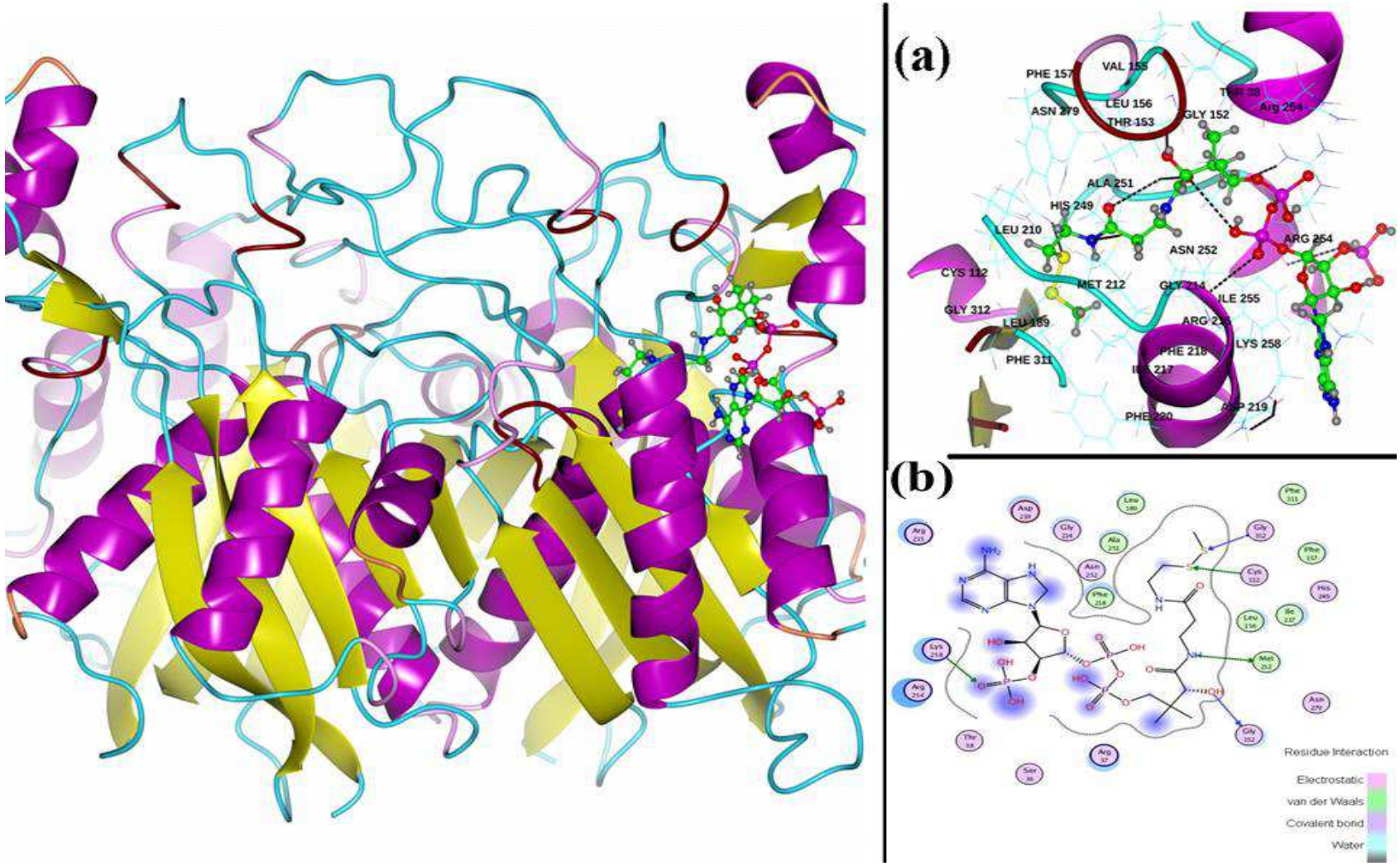


Figure 8. Docked poses: Binding mode of top ranked docked poses by (a) GOLD, For clarity, only interacting important residues are displayed. Inhibitors are represented in ball and stick style and protein residues in new ribbon style (b) 2D depiction of interaction diagram of compound methyl-CoA disulfide in the *S. gordonii* FabH interface binding site.



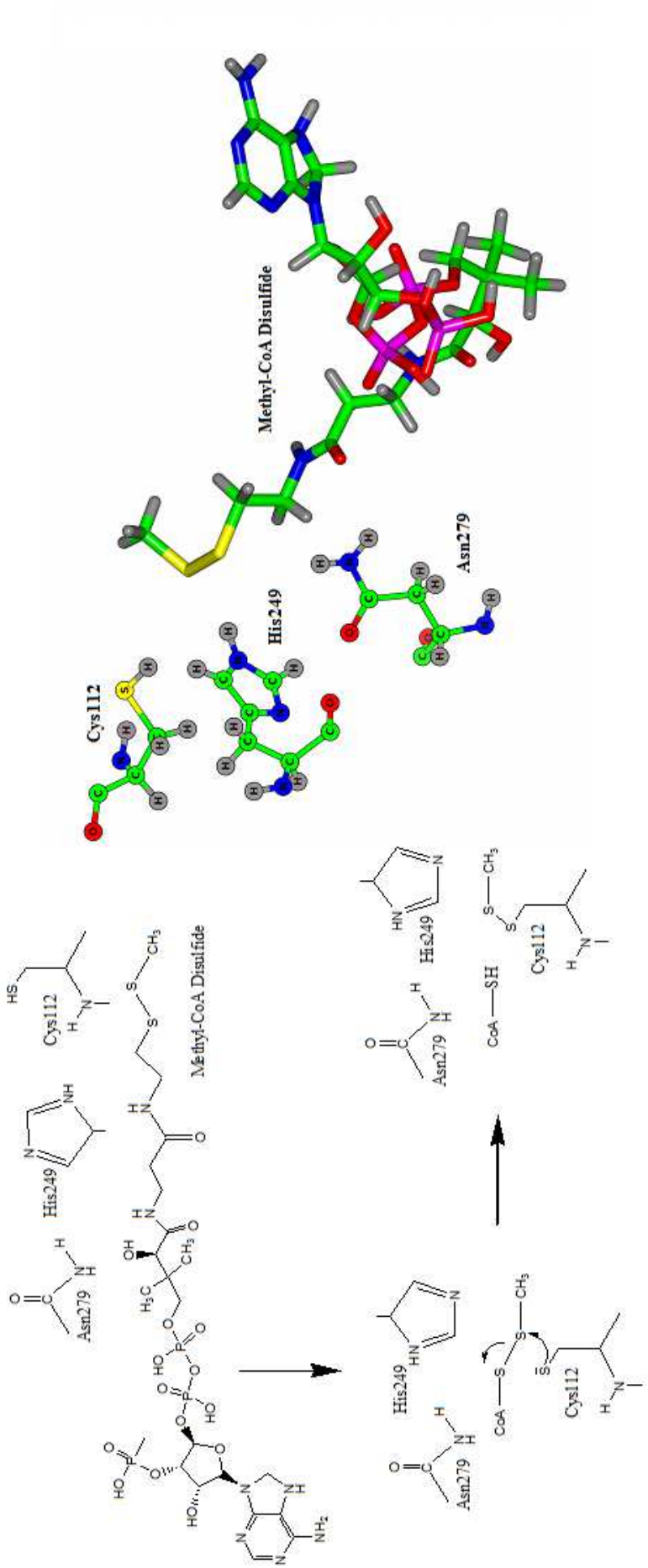


Figure 9. Inhibition of FabH by methyl-CoA disulfide

Figure 10. Root Mean Square Deviation graphs of docked complex of *S. gordonii* FabH, those parts which have high drastic changes are zoomed. (a) Graph represents docked complex over the 40 ns MD simulation. (b) Zoomed part of RMSD from 32 to 39 ns.

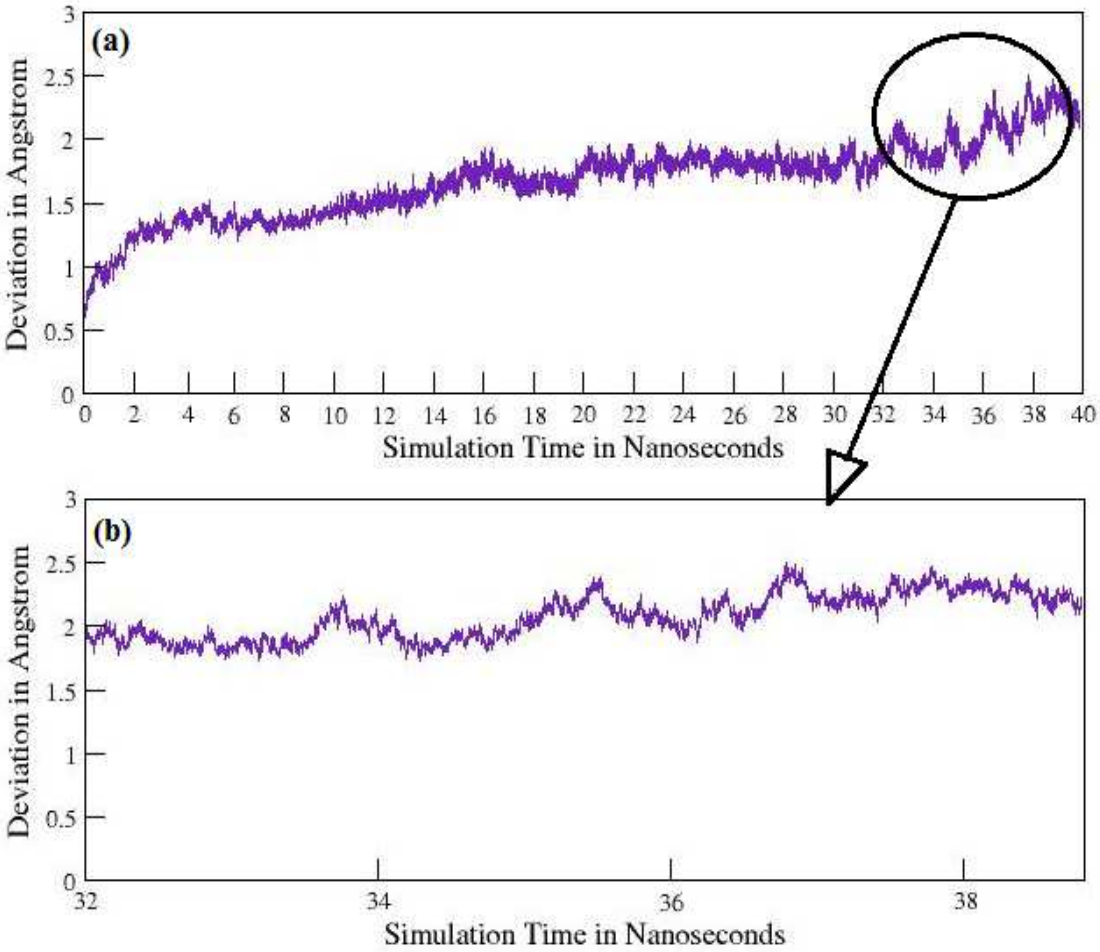
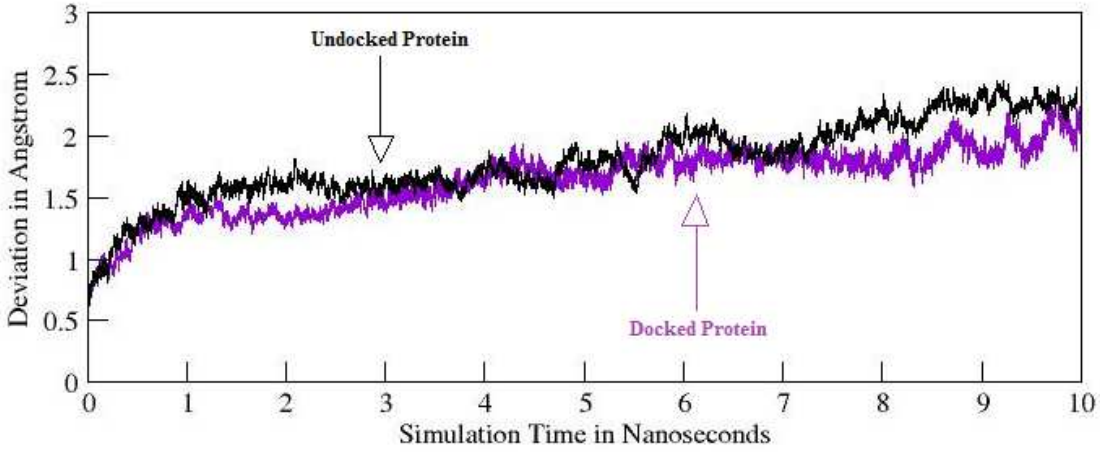


Figure 11. Superimposed RMSD for the undocked and docked complex of *S. gordonii* FabH over 10 ns MD simulation.



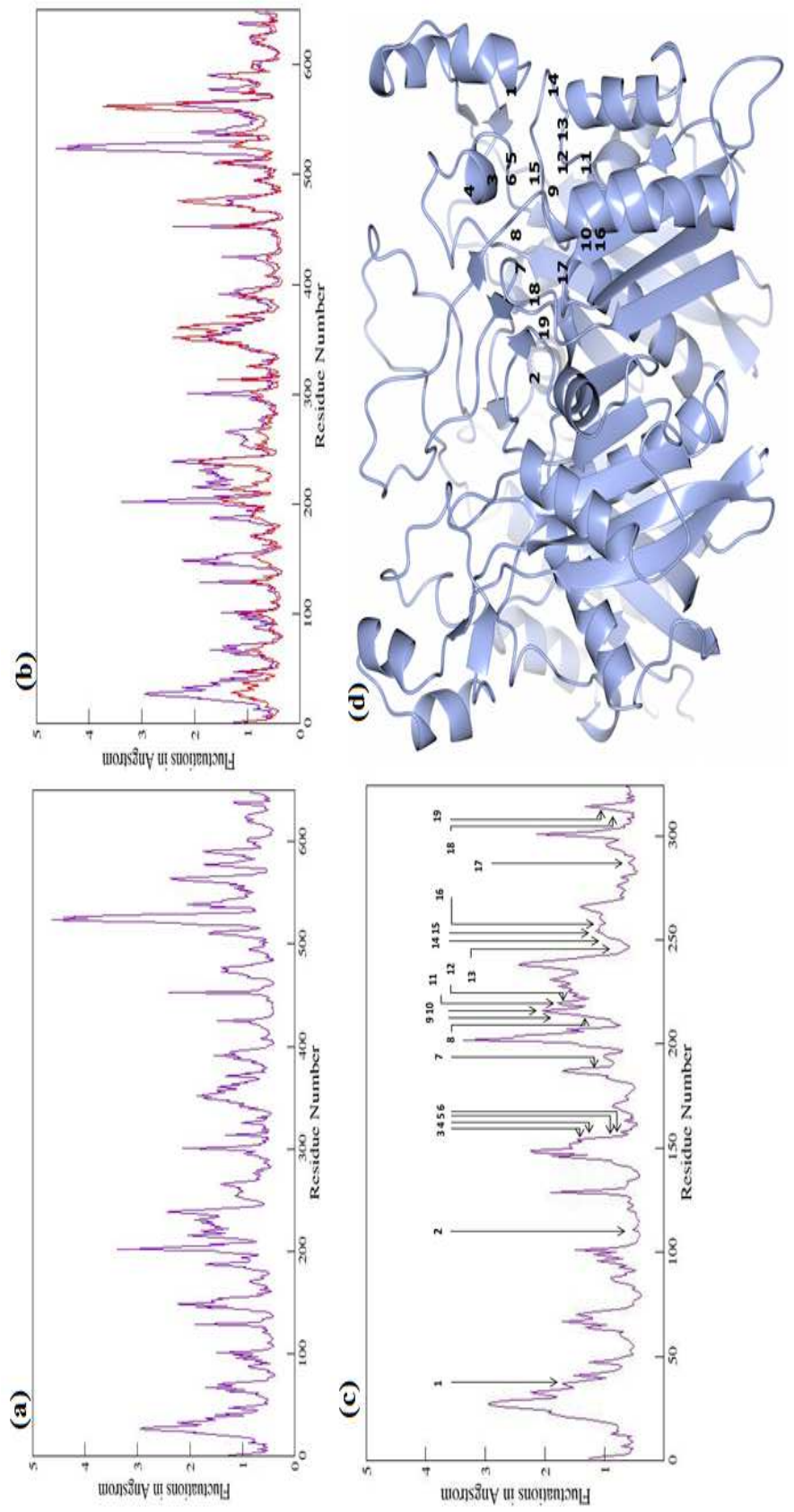


Figure 12. (a) RMSF plot of dimeric structure FabH over 40 ns. (b) Superimposed RMSF at 10 ns between docked and undocked dimeric structure of FabH. (c) Representing active site residues with arrows in RMSF plot. (d) FabH dimer structure in which active site residues are labeled, corresponding to the RMSF plot.

Figure 13. Active site residue 156 exhibits large fluctuations at 40 ns in which red, blue and yellow color representing 1st, 30th and 40th ns respectively.

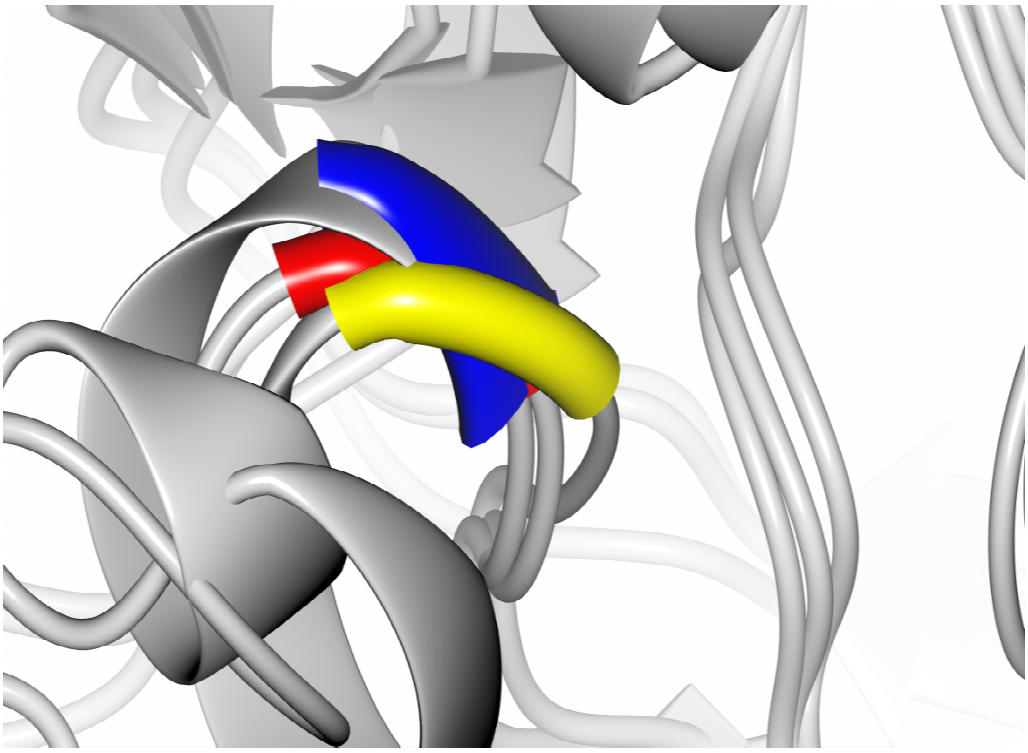
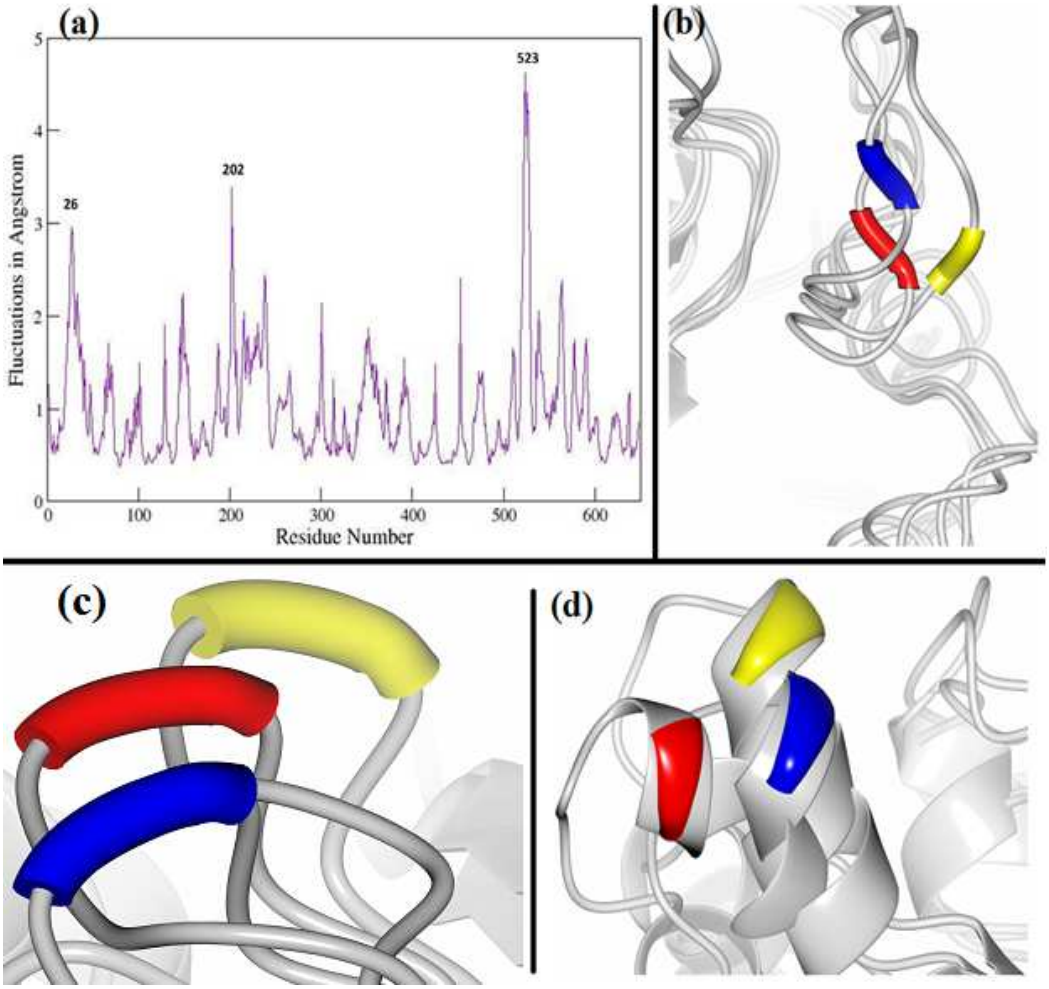


Figure 14. (a) RMSF plot of dimeric docked structure of *S. gordonii* FabH, representing three high peaks (b) 202 (c) 523 (d) 26 these superimposed structures illustrate large fluctuations in C-alpha atoms of residues at 1st, 10th and 40th ns with red, blue and yellow color respectively.



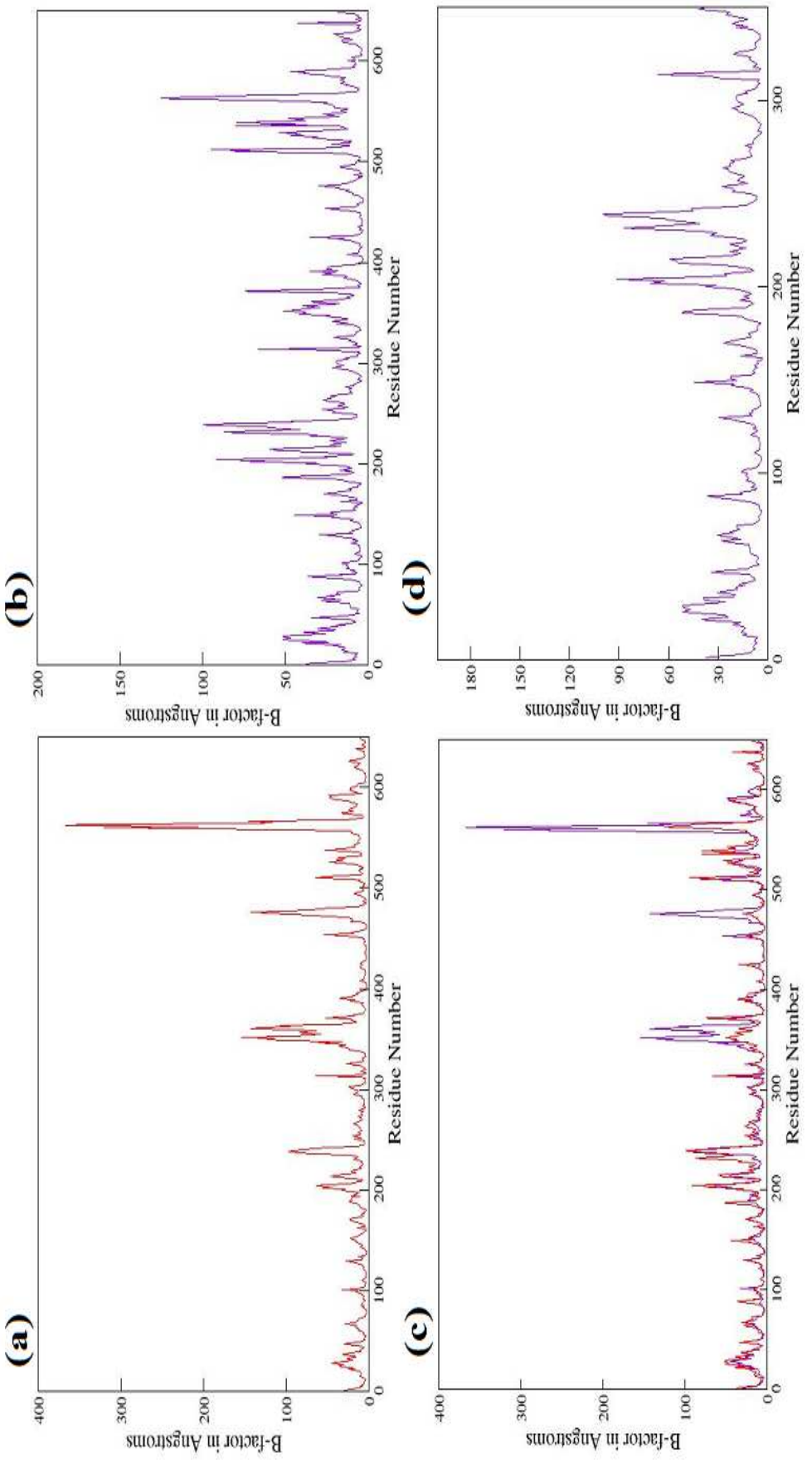
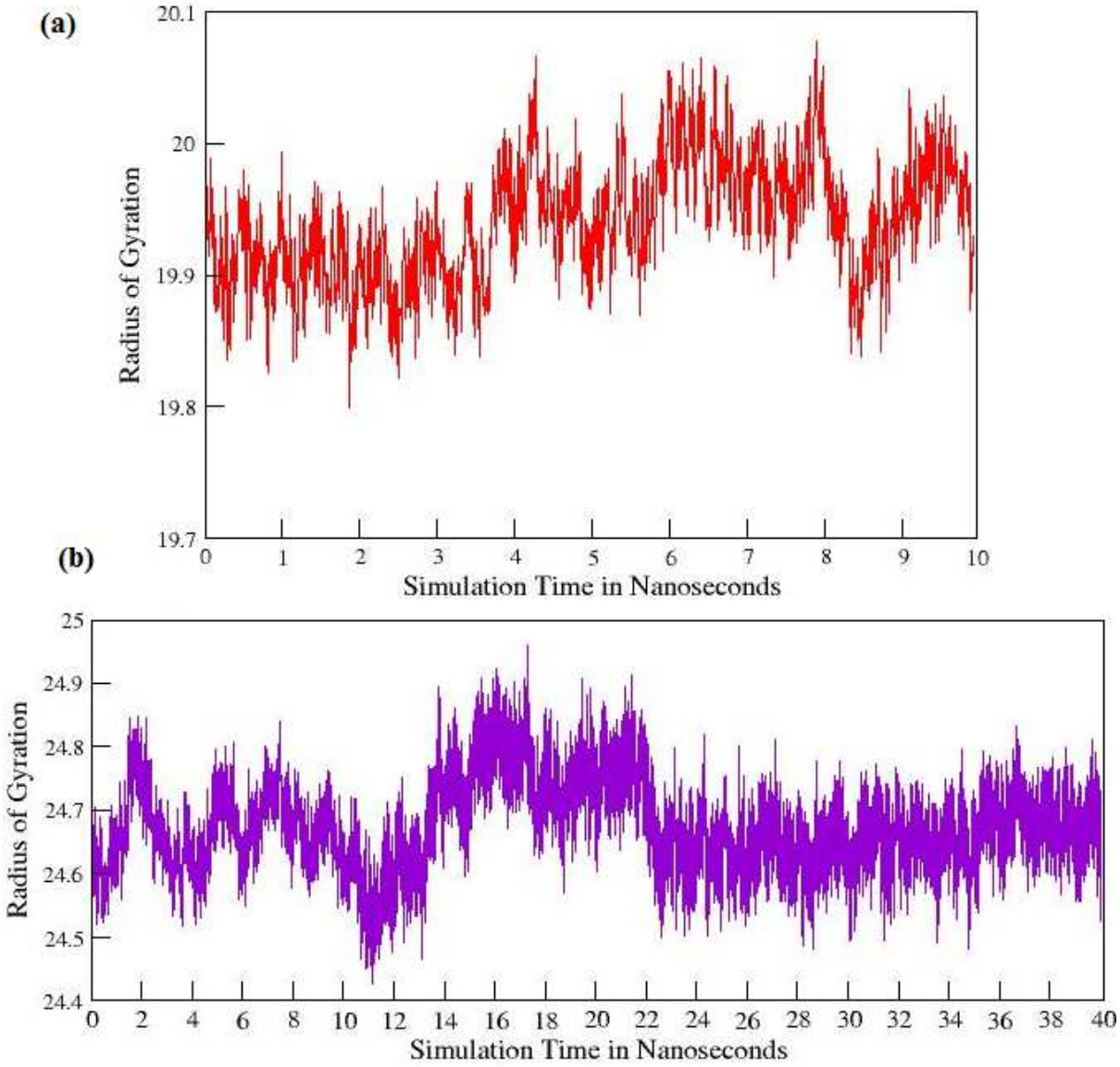


Figure 15. (a) B-factor values of FabH protein in undocked form (red color). (b) Graph representing B-factor of docked protein of FabH at 40 ns (magenta color). (c) Superimposed B-factor of undocked and docked FabH protein. (d) B-factor plot representing the fluctuations in active site residues of docked complex.

Figure 16. Radius of gyration measured for the FabH (a) Undocked protein of FabH at 10 ns. (b) Docked complex of FabH at 40 ns.



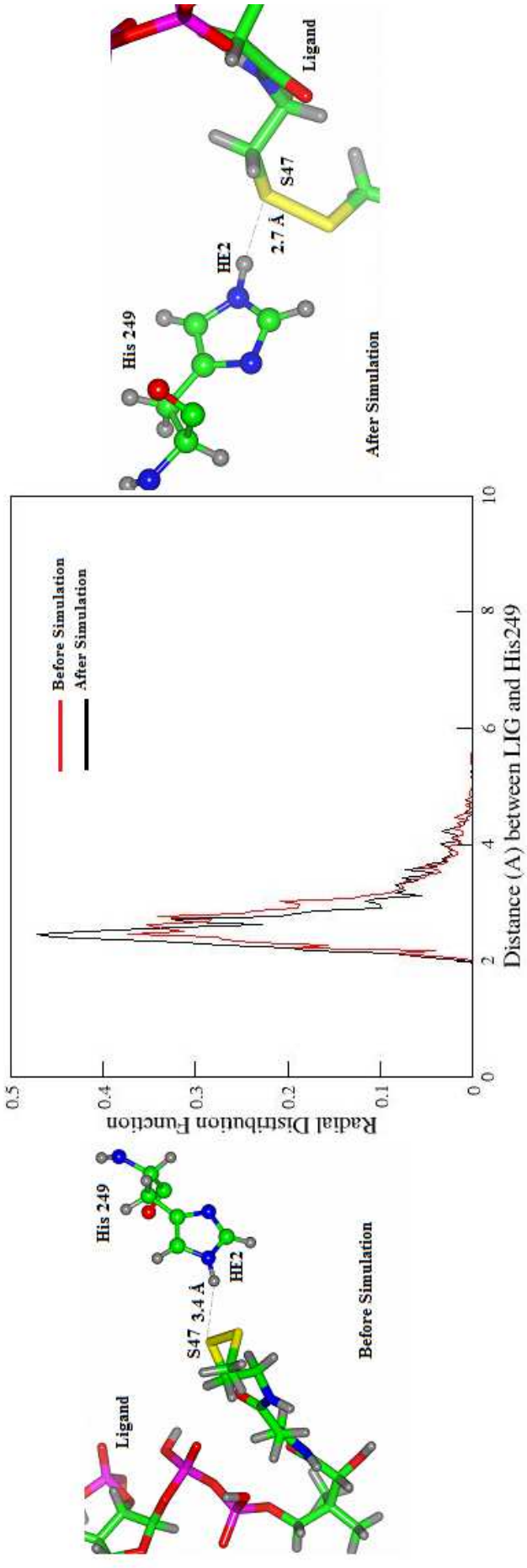


Figure 17. RDF graph of inhibitor and His249.

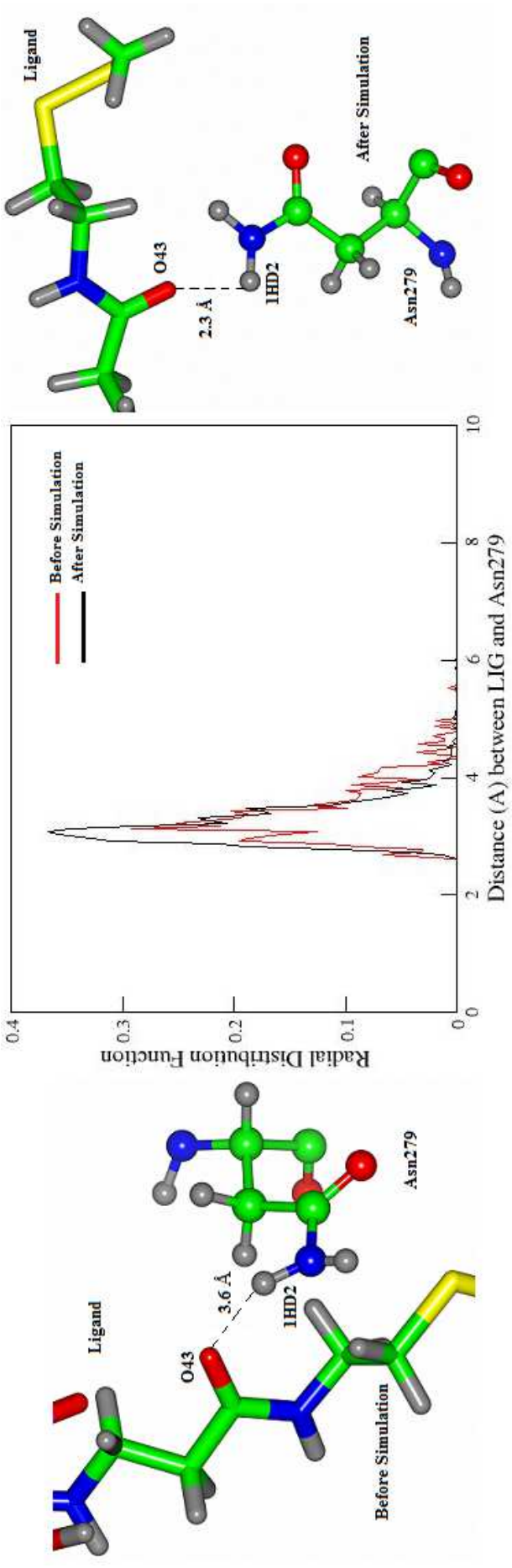


Figure 18. RDF graph of inhibitor and Asn279.

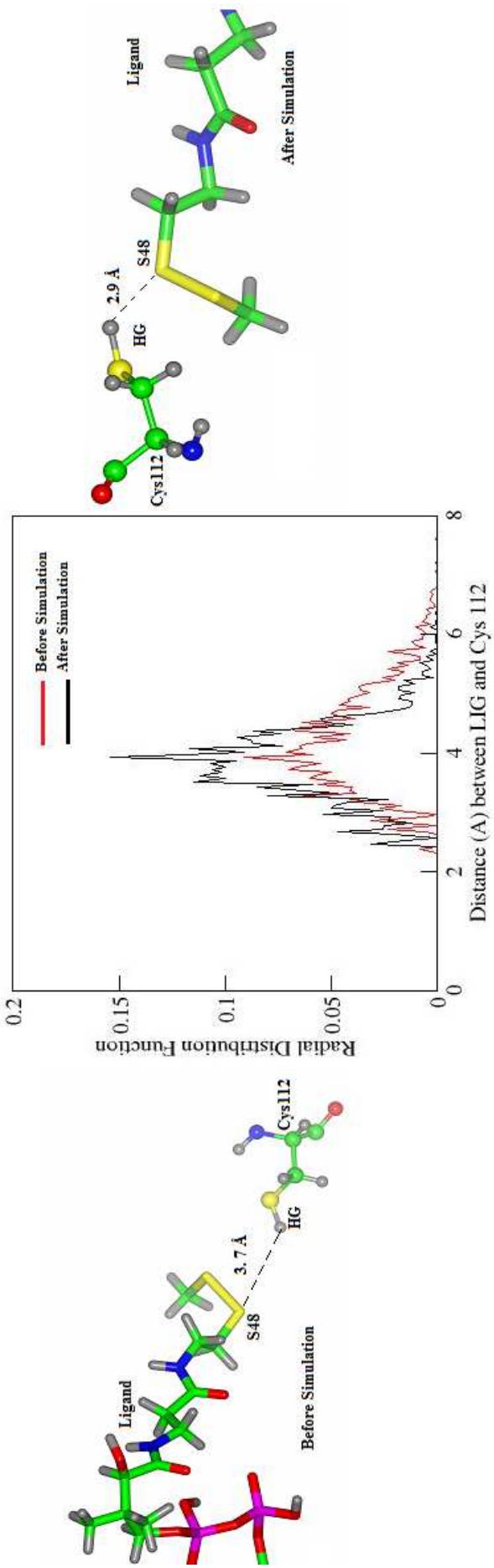


Figure 19. RDF graph of inhibitor and Cys112.

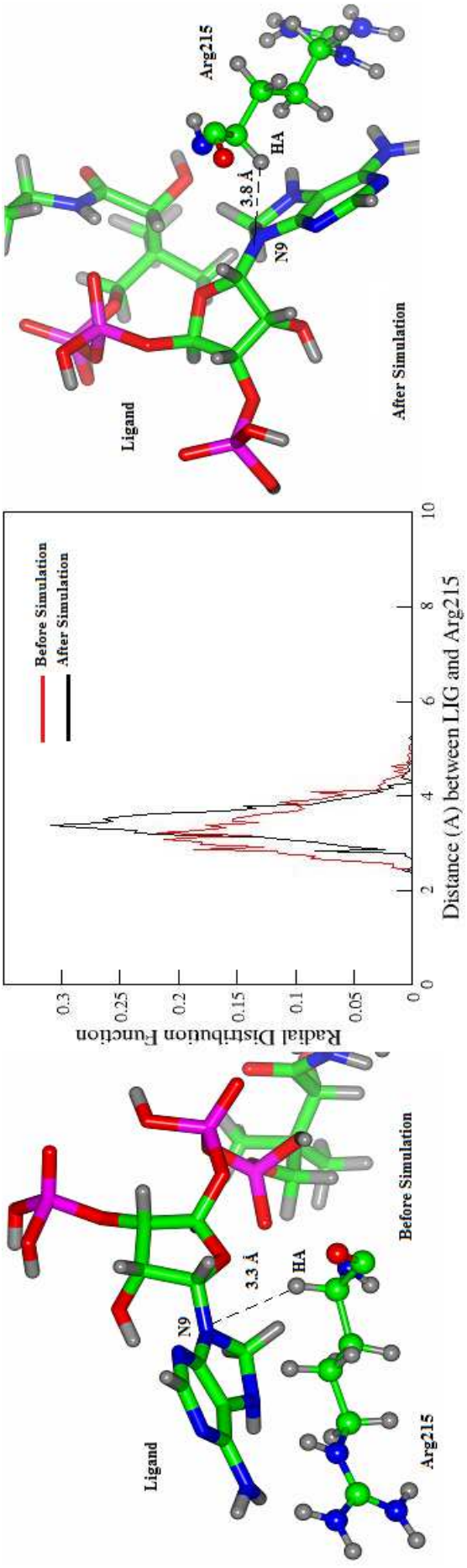


Figure 20. RDF graph of inhibitor and Arg215.

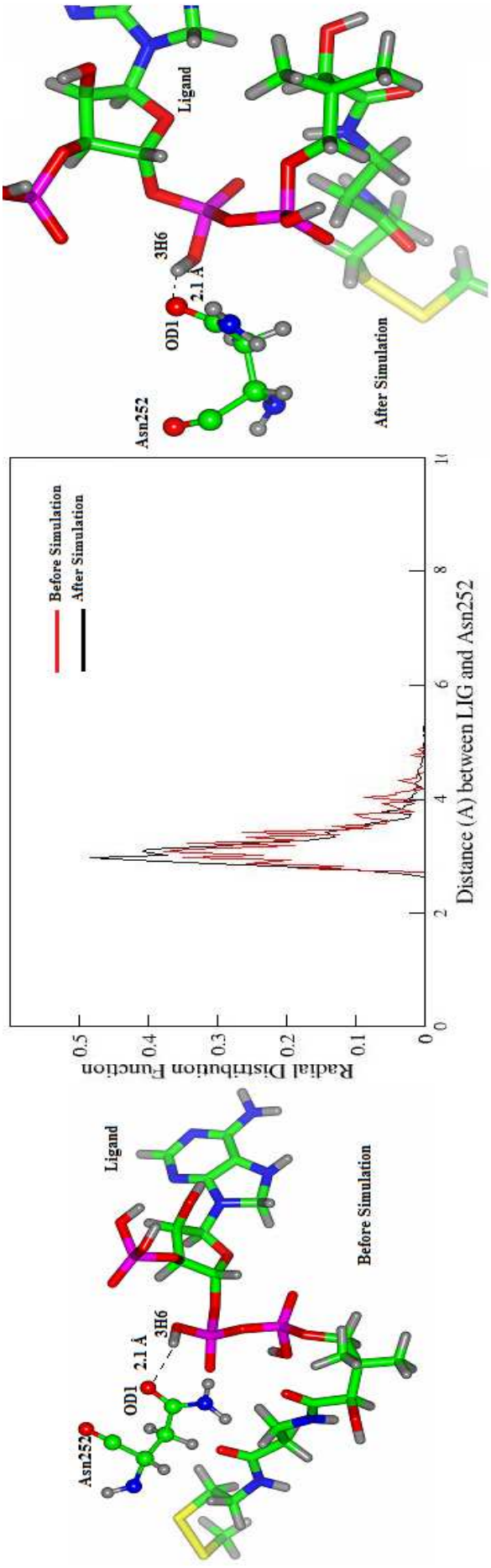


Figure 21. RDF graph of inhibitor and Asn252.