In silico screening of *W. salutaris* bioactive constituents reveal betulinic acid, ursolic acid acetate and eucosterol as potential *M. tuberculosis* antagonists targeting FabF

Nokukhanya Gumede
University of Kwazulu-Natal
Kgothatso E. Machaba
University of Kwazulu-Natal
Umar Ndagi
Federal University of Technology Minna
Hezekiel M. Kumalo
University of Kwazulu-Natal
Ndumiso N. Mhlongo (MhlongoN4@ukzn.ac.za)
University of KwaZulu-Natal - Howard College Campus

Research Article

Keywords: Mycobacterium tuberculosis, FabF, binding free energy, inhibitory potential, Warburgia salutaris

Posted Date: June 9th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-274959/v1

Abstract

Tuberculosis (TB) remains a long-standing burdening disease to control worldwide. The lengthy current TB treatment, which boasts with unbearable adverse effects, and frequent emergence of drug resistant strains of *M. tuberculosis* lays an increasing burden. This behests urgent discovery and development of alternative novel medicine to alleviate TB. In this report, *in silico* methods were applied to examine the propensity of *W. salutaris* active compounds as potential inhibitors of *M. tuberculosis* fatty acid biosynthesis protein (FabF). Thirteen compounds were virtually screened against FabF and subjected to molecular dynamics simulations and post-dynamics analyses to examine their inhibitory potential. Betulinic acid, ursolic acid and ursolic acid acetate had the best binding energies and hence the best inhibitory potential against FabF and desirable cytotoxicity profile. These compounds bind and interact with FabF active site residues to exert their inhibitory potential. Findings in this preliminary report warrant further experimental validation towards the development of these compounds as potential drugs targeting FabF in the treatment of tuberculosis.

Introduction

Mycobacterium tuberculosis (*Mtb*) is a Tuberculosis (TB) causing agent which, if not properly diagnosed, can be fatal. TB has been reported as one of ten greatest killer diseases globally which affected about 10 million people in 2017 (Who, 2019). TB is also considered a leading opportunistic disease and cause of death in HIV/AIDS patients as compromised immune system in HIV infected individuals enhances the growth of *Mtb* (Trinh *et al.*, 2015).

Mtb invades the human cells through encapsulation with a cell wall made up of long-chain fatty acids known as mycolic acids (Takayama et al., 2005). Encapsulation by mycolic acids effectively allows the bacterium to grow inside the microphage, hidden from the external environment. Mycolic acids are crucial for growth and survival of *Mtb*, therefore providing a potential underlying mechanism for *Mtb* pathogenesis (Gajiwala et al., 2009). As a result, *Mtb* fatty acid biosynthesis pathway provides a great focal point for the discovery of anti-tubercular drugs.

Unlike mammals with only type I fatty acid synthesis, mycobacteria have both type I and II systems (Choi et al., 2000). Type I synthesis involves a single multifunctional enzyme that carries out all the reactions of chain elongation. The type II pathway has each enzymatic activity found on a separate protein. The amino acid sequences and the active sites of type I and II fatty acid biosynthesis enzymes are completely different from each other (Gajiwala et al., 2009). This allows a possibility to design specific and potent inhibitors of enzymes from the type II pathway, with little or no effect on type I fatty acid biosynthesis enzyme. β-ketoacyl-acyl carrier protein synthase III (FabH) and β-ketoacyl-acyl carrier protein synthase II (FabF) are the proteins of fatty acid synthesis type II pathway (Daffé et al., 2017). FabH catalyses the initial cycle that involves condensation of malonyl-ACP and Acetyl-CoA, while FabF catalyses subsequent cycles of elongation that result to mycolic acids (Sridharan et al., 2008).

Currently, TB is treated with a 6 months multi-drug regimen starting with rifampicin (RIF), isoniazid (INH), pyrazinamide (PZA) and ethambutol (EMB) taken for 2 months; followed by RIF and INH for 4 months (Jnawali et al., 2013). These drugs mainly inhibit the biosynthesis of essential bacterial biomolecules including mycolic acids. However, inappropriate use of anti-TB drugs and failure to complete the full course of treatment have been reported to encourage the development of drug-resistant *Mtb* strains (Jnawali et al., 2013). Frequent emergence of drug-resistant strains compromises the already difficult anti-tubercular treatment (Chollet et al., 2016). Thiolactomycin (TLM) is an FDA approved antibiotic known to inhibit type II fatty acid biosynthesis and has shown great activity against a broad spectrum of pathogens (Kremer et al., 2000) both in vitro and in vivo (Douglas et al., 2002). The major problem with these synthetic drugs is that their manufacture is cost-effective and time consuming, and they are associated with drastic adverse effects on human health, thus they do not reach and pass preclinical stages. Hence, there is a great need for alternative sources of medicine to provide safe and affordable drugs.

Recently, the search for medicinal plants has increased drastically and numerous medicinal plants with bioactive compounds, *Daucus carota*, *Centella coriacea*, (Lawal et al., 2014), *Warburgia salutaris* (Tabuti et al., 2010). *W. salutaris* is traditionally used to treat various diseases such as diabetes (Msomi et al., 2019), fever, flu (Maroyi, 2013), flu and chest infections including TB (Botha et al., 2004). The crude extracts of *W. salutaris* has been reported to possess anti-mycobacterial activity (Madikane et al., 2007). Phytochemical screening of *W. salutaris* has led to the identification of bioactive compounds which include iso-mukaadial acetate (Nyaba et al., 2018), salutarisolide (Frum et al., 2005), muzigadial (Rabe & Staden, 2000), warburganal, mukaadial, isopolygodial, polygodial, betulinic acid, eucosterol, cinnamodial, ugandensdial, ursolic acid and Ursolic acid acetate (Maroyi, 2013) presented in Fig. 1.

These compounds have been reported to exhibit a wide range of biological activities including antioxidant (Soyingbe et al., 2018), antidiabetic, antimalarial, antimicrobial, antifungal (Soyingbe et al., 2018) and antibacterial properties (Maroyi, 2013). Among these compounds, presence of hydroxyl group in the A or B ring combined with the presence of either an acid, propenyl, hydroxymethyl or 2-hydroxypropyl in the E ring of the triterpene was reported to be the potential structures that can result in anti-tubercular activity (Wachter et al., 1999). Although anti-bacterial activity of these compounds have been reported in some species such as *Staphylococcus aureus*, *Escherichia coli*, *Klebsiella pneumoniae* and *Bacillus subtilis*, but to the extent of our knowledge there has not been any anti-tubercular properties reported on any of these compounds. Hence, channelling our interest in finding the active ingredient responsible for *W. salutaris*`s previously reported anti-mycobacterial activity. This study aims to explore the capacity of twelve *W. salutaris* compounds as potential inhibitors of *Mtb* FabF through the application of robust computational methods.

Materials And Methods

2.1 Collection of Materials

The X-ray crystal structure of FabF was obtained from the Protein Data Bank (Berman et al., 2002), with PDB code 2GP6 and a resolution of 2.4 Å (Sridharan et al., 2007). Thiolactomycin (TLM) was obtained from a FabH complex (PDB code 4C6U, 2.4 Å resolution) (Schiebel et al., 2013). Thirteen structures of *Warbugia salutaris* compounds; betulinic acid (BA), cinnamodial (CM), eucosterol (EUC), mukaadial (MUK), muzigadial (MUZ), polygodial (PG), isopolygodial (IPG), salutarisolide (ST), ugandensdial (UG) ursolic acid (UA), ursolic acid acetate (UAA) and warbuganal (WBG) were obtained from PubChem (Kim et al., 2019).

2.2 Processing of structures for molecular docking

The crystal structure of FabF exists as a homodimer with chain A and B. To reduce the computation cost, only chain A was considered for this study. Hydrogen atoms were removed from all ligands prior to energy minimization using MMFF94 force field in Avogadro (Hanwell et al., 2012) software. UCSF Chimera (Pettersen *et al*., 2004) was used to separate TLM from FabF and to add hydrogen atoms to the protein in preparation for subsequent molecular docking.

2.3 Molecular docking

AutoDock Vina (Trott *et al*., 2010) was applied to dock all ligands into FabF. Gasteiger (Steffen *et al*., 2010) partial charges were allocated during docking. AutoDock graphical user interface issued by MGL tools was utilised to outline the AutoDock atom types. The active site docking was carried out in which the grid box was drawn encapsulating all amino acids that participate in the binding of a ligand to FabF crystal structure. The grid box was analysed with the following grid parameters: x = 74 Å, y = 68 Å, and z = 60 Å for the dimension, while for the centre grid was x = -7.321 Å, y = 44.488 Å, and z = -0.282 Å with the space being 0.431 and exhaustiveness = 8. The Lamarckian genetic algorithm (Morris *et al*., 1998) was used to generate docked conformations in accordance with their docking energy in a descending order. The resultant complexes were saved using UCSF Chimera software.

2.4 Settings for molecular dynamics simulation

The top scoring complexes of each compound from molecular docking were prepared for molecular dynamics (MD) simulations. Avogadro and UCSF Chimera software were used to prepare both the ligand and the receptor. Preparations included separating the ligand from the receptor, adding hydrogen atoms to the ligand while deleting them from the receptor and correcting the bonds of the ligand where necessary.

2.5 MD simulations

Molecular dynamic (MD) simulation of all complexes were carried out using the Amber 14 software (Salomon-Ferrer *et al*., 2013). The Antechamber module was used to generate atomic partial charges for the all compounds utilising Amber force field (GAFF) and restrained electrostatic potential parameters (Wang *et al*., 2004). To describe the protein systems, Amber force field ff14SB was applied (Perez *et al*., 2015; Tan *et al*., 2015). Herein, hydrogen atoms and counter ions were added to the protein, using LEAP module implemented in Amber14. Before simulations were started, all systems were enclosed in a TIP3P

water box with a located 10 Å. Hence, the cubic periodic boundary conditions were implemented in all the systems. Particle-mesh Ewald approach (Salomon-Ferrer *et al*., 2013) was applied to treat a long-range electrostatic interactions with a cut-off distance of 12 Å. The initial energy minimization step of all systems were applied with a restraint potential. Unrestrained conjugated gradient minimization was carried out for all system using Sander module of Amber14. Canonical ensemble (NVT) MD simulations were performed for 50 ps; where the system was gradually heated from 0 to 300 K with harmonic restraints of 5 kcal/mol $Å^{-2}$ for solute atoms, with the Langevin thermostat (Johnson *et al*., 2012) at 1 ps random collision frequency. The energy minimization, heating and equilibration steps were performed using the standard parameters and a standard MD production run for 100 ns was carried out in all systems. After MD simulations, trajectories of all systems were saved and analysed using the CPPTRAJ and PTRAJ modules of the Amber14 suite.

2.6 Binding energy calculations

The binding free energy calculations of all systems were performed using Molecular Mechanics/Generalised-Born Surface Area (MM/PBSA), an important method that gives detailed information on the binding affinity between the ligand and receptor (Genheden and Ryde, 2015). Hence, MM/PBSA approach was carried out in the current study. The approach is based on receptor-ligand complex, after the MD simulation. For a 100 ns trajectory, 1000 snapshots were evaluated during this calculation. The following set of equations describes the binding free energy calculation:

$$\Delta G = \Delta E_{ele} + \Delta E_{vdW} + \Delta G_{pol} + \Delta G_{nonpolar} - T\Delta S \quad (1)$$

From the equation above, ΔE_{ele} and ΔE_{vdW} are electrostatic and van der Waals interactions of the ligands with the proteins in gas phase, respectively. ΔG_{pol} is the polar solvation free energy calculated using MM-PBSA program. The $\Delta G_{nonpolar}$ signifies nonpolar solvation free energy. The entropy contributions ($-T\Delta S$) of the binding free energies were calculated using normal mode calculation (Genheden and Ryde 2012) in the translational, rotational, and vibrational entropy components by conformational snapshot derived from MD simulations.

2.7 Per-residue energy decomposition analysis

The contribution of each residue to the total binding free energy profile was carried out for all amino acids that interact with the compound, using the MM/PSA method in Amber 14 software.

2.8.0 Post-MD analyses

Generated MD trajectories were analysed using robust *in silico* analytical methods including Root Mean Square Deviations (RMSD), Root Mean Square Fluctuations (RMSF), hydrogen-bond network profile, Dynamic Cross Correlation (DCC) and Principal Component Analysis (PCA) using CPPTRAJ module in Amber14. Visualisation of trajectories was performed using UCSF Chimera. The results were analysed and plots were generated using Origin software (Seifert, 2014).

2.8.1 Principal Component Analysis

An effective method of explaining the dynamic nature of proteins is through the essential dynamics analysis method, also known as PCA (Desdouits *et al.*, 2015; Amadei, Linssen, & Berendsen, 1993). In the current study, PCA was performed on C-α atoms on 1000 snapshots at 100 ps time interval utilising an in-house script. The first two principal components (PC1 and PC2) were calculated and matrices were carried out using Cartesian coordinates of Cα atoms. PC1 and PC2 correspond to first two eigenvectors of covariance matrices. Origin software was utilised to construct PCA plots.

2.9 Toxicity analysis

Selected ligands from MD simulation were subjected to toxicity analysis using the toxicity prediction tool ProTox-II server (Banerjee *et al.*, 2018). This method predicts lethal doses of compounds in LD_{50} in mg/kg of test subject and assign LD_{50} values according to the toxicity class. This server has been previously used successfully to assess the toxicity profiles of some anti-tubercular compounds (Cinaroglu and Timucin, 2019).

Results And Discussion

3.1 Stability of simulated complexes

Based on the calculated binding free energies of all the simulated compounds, only BA, EUC and UAA had the best binding affinities relative to TLM and these were selected for further post-MD analyses. To determine the structural stability of each receptor-ligand complex, RMSD of the above-mentioned systems was calculated and presented in Fig. 2.

The average RMSD was 1.4 Å and 1.15 Å for FabF-TLM and FabF-BA respectively; and 1.0 Å for both FabF-EUC and FabF-UAA. These figures account for system stability since a standard parameter that signify a stable system is a maximum RMSD of 2 Å and below (Carugo, 2001). Therefore, we can conclude that in the duration of 100 ns simulation, all the systems formed an acceptable stable conformations.

3.2 Protein flexibility during MD simulation

The RMSF of amino acids C-alpha atoms of the ligand-receptor complexes provide structural insights into the residual flexibility of different regions of the studied complexes. Amino acids determine the conformational features of the protein. However, conformational changes in proteins are brought by chemical or mechanical alterations in amino acid sequence and ligand binding. Studies have shown that conformational changes occur as a result of ligand-induced motion during ligand binding to the receptor. The RMSF of the simulated complexes is presented in Fig. 3.

In all systems, the region of residues 116–150 shows high fluctuation denoting high residue mobility in this region, ranging from 4.3–8.8 Å. Although binding affinity of FabF-TLM (-15.92 Kcal/mol) is lower than that of FabF-BA (-31.17 Kcal/mol), FabF-EUC (-26.27 Kcal/mol) and FabF-UAA (-29.52 Kcal/mol), FabF-TLM complex showed lower mobility in residues 211–229. However, binding of TLM to Fab-F

seemed to have favoured an increased fluctuation all other residues when compared FabF-BA, FabF-EUC and FabF-UAA. From these analyses, it can therefore be inferred that the binding affinity of FabF-TLM is associated with decreased residue stability. A decrease fluctuation was also observed in residues 100–115 (which include residues Gly 112 and Leu 113 which participate in ligand binding) in FabF-BA, FabF-EUC and FabF-UAA compared to FabF-TLM complex. This decrease may be associated with the presence of numerous cyclohexane rings in these compounds which limit the capacity of residues in the vicinity to fluctuate, hence resulting to the rigidity of FabF complexed with natural compounds. However, the binding of BA, EUC and UAA rigidified the active site residues of FabF, suggesting that natural compounds form stable interactions with FabF relative to TLM.

3.3 Binding free energy analysis

MM/PBSA method is a popular approach used to calculate the binding free energy of ligands to biological macromolecules (Hou *et al*., 2011). Table 1 presents the total binding free energy of twelve natural compounds and TLM calculated from 100 ns MD trajectories.

Table 1
MM/GBSA binding free energies of TLM and natural compounds complexed with *Mt*FabF.

Compound	ΔG_{bind} (kcal/mol)	ΔE_{vdW} (kcal/mol)	ΔE_{ele} (kcal/mol)	ΔG_{gas} (kcal/mol)	ΔG_{sol} (kcal/mol)
TLM	-15.9202 ± 4.4404	-22.0500 ± 5.6107	-2.1481 ± 2.6860	-24.1981 ± 6.7809	8.2779 ± 3.0942
BA	-31.1728 ± 6.4487	-38.0026 ± 5.8906	-4.4420 ± 3.2699	-42.4446 ± 7.9292	11.2718 ± 2.7018
CM	-18.0718 ± 2.8962	-25.0894 ± 3.0505	-3.4432 ± 4.0811	-28.5317 ± 4.4773	10.4599 ± 3.6055
EUC	-26.2682 ± 3.6290	-32.7515 ± 3.7638	-10.3945 ± 4.7202	-43.1460 ± 5.6583	16.8774 ± 4.4173
IPG	-15.6064 ± 2.8636	-19.9091 ± 3.1328	-2.3129 ± 3.2626	-22.2220 ± 4.4410	6.6156 ± 2.9014
MUK	-22.5655 ± 3.1652	-31.7458 ± 3.2762	-5.6550 ± 5.1102	-37.4008 ± 6.3812	14.8354 ± 4.3488
MUZ	-22.8485 ± 3.1598	-28.4570 ± 2.7704	-5.7993 ± 3.6567	-34.2563 ± 4.0622	11.4078 ± 3.2152
PG	-15.5411 ± 2.7571	-19.6059 ± 2.9233	-2.4456 ± 4.2311	-22.0515 ± 4.9567	6.5104 ± 3.4457
ST	-15.8650 ± 2.4398	-19.1223 ± 2.3916	-5.5324 ± 4.1361	-24.6548 ± 4.3845	8.7898 ± 2.9509
UG	-18.3874 ± 3.2511	-25.4859 ± 3.6624	-7.0943 ± 4.4396	-32.5793 ± 5.7879	14.1920 ± 4.5417
UA	-21.1635 ± 4.3291	-27.2815 4.1629	-3.4895 ± 11.6004	-30.7710 ± 11.3643	9.6074 ± 10.6324
UAA	-29.5221 ± 4.1225	-36.5090 ± 5.0241	-2.3505 ± 3.8459	-38.8595 ± 7.2532	9.3373 ± 4.4640
WBG	-16.6187 ± 3.2374	-22.4188 ± 3.8510	-3.1064 ± 3.6792	-25.5252 ± 5.1962	8.9065 ± 3.2374

ΔEele = electrostatic energy; ΔEvdw = van der Waals energy; ΔGbind = calculated total binding free energy; ΔGsol = solvation free energy; TLM = Thiolactomycin; BA = betulinic acid; CM = cinnamodial; EUC = eucosterol; IPG = isopolygodial; MUK = mukaadial; PG = polygodial; ST = salutarisolide; UG = ugandensdial; UA = ursolic acid; UAA = Ursolic acid; WBG = warbuganal.

As evident in Table 1, nine compounds had higher binding energies compared to TLM. Only three compounds with best binding energies, FabF-BA (-31.17 Kcal/mol), FabF-EUC (-26.27 Kcal/mol) and FabF-UAA (-29.52 Kcal/mol) compared to FabF-TLM (-15.92 Kcal/mol) were considered for further analysis. The difference in binding free energy between systems indicate that the force of interactions contributes higher energy in the binding of natural compounds to FabF as compared to the TLM. Electrostatic contribution to the total binding free energy was higher in FabF-BA (-4.4420 kcal/mol) and

FabF-EUC (-10. 3945) compared to FabF-UAA (-2.3505 Kcal/mol) and FabF-TLM (-2.1481 Kcal/mol). However, FabF-BA, FabF-EUC and FabF-UAA systems exhibited a relatively higher van der Waal energy (-38.0026, -32.7515, -36.5090 Kcal/mol respectively) contribution to the total binding free energy compared to the FabF-TLM system with van der Walls energy contribution of -22.0500 kcal/mol. Based on these results, vdW forces are potentially important binding forces between studied ligands and FabF. Hence, vdW contribution is comparatively higher than the electrostatic contribution as displayed in Table 1.

3.4 Per-residue energy decomposition analysis

Assessing energy contribution of individual active site residues to the total binding free energy was achieved by decomposing binding profile into the contributions of each active site amino acids in all systems. In addition, this also provided a detailed understanding of the protein dynamics on a degree of different binding forces. Figure 4 shows that major energy contributors in FabF-BA system were Phe207, Leu113 and Phe403, with − 1.821, -1.491 and − 1.583 kcal/mol, respectively; in FabF-EUC were Phe207, Leu113 and Phe403 with − 2.567, -2.238 and − 1.007 Kcal/mol respectively; in FabF-UAA were Phe207, Pro203, Met210, Val276, Leu113 and Gly206 with − 1.515, -1.637, -1.368, -1.281, -1.182 and − 1.016 Kcal/mol respectively; In FabF-TLM system, the major energy contributors were Phe207, Val346 and Phe403 contributing − 1.050, -0.687, -0.650 kcal/mol respectively. Phe207 has highest energy contribution to the total binding free energy, suggesting that it favoured binding of all the studied ligands to FabF. However, in FabF-BA and FabF-UAA systems (Fig. 4) more active site residues show a significant high binding free energy contribution relative to their contribution in FabF-EUC and FabF-TLM systems, signifying a reason for a high potency of BA and UAA over TLM in FabF inhibition.

3.5 Principal Component Analysis (PCA)

To further understand conformational motions for all sysrems, PCA calculation was carried out. The clustering system of PCA was preferred due to its ability to describe various conformational states during the simulation period by grouping molecular structure into a subset based on their conformational similarities (Wolf and Kirschner, 2013). The conformational changes of all systems were projected along the first two principal components (PC1 vs PC2) to give a better understanding of conformations. The scatter plot (Fig. 5) shows the dominant changes in motion across PC1 vs PC2 of natural compounds and TLM bound to *Mtb* FabF.

A recognisable separation of motion was observed in the FabF-UAA system showing a greater correlated motion along PC1 and PC2 components, while other systems showed a significantly smaller or interrelated movement. It is obvious from the PCA plot that FabF-UAA system appeared to occupy less space, meaning that the binding of UAA to the residues in the active site of FabF induced the residue dynamics that resulted into conformational rigidity, thereby enhancing ligand residue interaction in this system.

3.6 Dynamic Cross-Correlation

In order to further analyse the internal dynamics of in the binding of natural compounds and TLM to FabF and to assess the existence of associated motions, DCCM analysis was carried out on the position of the Cα atoms using the 100 ns MD trajectories. The highly positive correlated movement of specific residues are represented by yellow to red colored regions, whereas highly negative anti-correlated movement of specific residues is represented by blue to black colored regions. The diagonal part shows obvious correlated movements, while other regions rarely exhibit such highly correlated movements. Figure 6 shows the DCC matrixes of Cα atoms fluctuations of the four systems.

Overall DCCM analyses showed that the binding of BA, EUC, UAA and TLM alters the structural conformation of FabF as evidenced by changes in associated movements and dynamics. Positive correlated regions in plot A are more prominent in marked regions between residues 375–400 related to plot B, C and D. In FabF these residues are seen located away from the active site. Upon natural compound binding to FabF correlation increases relative to FabF-TLM system suggesting that natural compounds binding to FabF may have resulted in conformational changes in the protein. In relation to plot A, plot B, C and D exhibit highly positive correlated motion in residues between 200–250 whereas slightly correlated motion are seen between 250–275 region. These prominent correlated regions reside a majority of hydrophobic active cite residues hence binding of BA, EUC and UAA to FabF may have induced residue dynamics that result to more to more correlated motion among the active site residues. These results are in agreement with observed RMSF suggesting that BA, EUC and UAA are experimentally more potent than TLM since their binding to FabF exhibit higher binding affinity to the active site residues.

3.7 Hydrogen bond network profile

Throughout biological systems, hydrogen bonds play an important role in maintain the structural integrity of proteins, protein-ligand interactions and catalysis. This is because of their special and essential existence (Chen *et al.*, 2016). In order to investigate the impact of BA, EUC, UAA and TLM binding on FabF, evaluation of hydrogen bond distances between amino acid residues interacting with these natural compounds and TLM in the active site was monitored for 100-ns simulations (Table 2a-2d).

Table 2

a: Calculated properties of hydrogen bonds formed between TLM and FabF.

H-bond acceptor	H-bond donor	Frame number	Average distance (Å)	Average angle (degrees)
Thiolactomycin				
TLM_416@O11	GLU_200@N	7076	2.89	155
TLM_416@ONA	GLY_114@N	1575	2.88	149
LEU_113@O	TLM_416@O6	325	2.77	141
TLM_416@O11	SER_115@OG	291	2.77	157
TLM_416@O11	MET_275@N	165	2.87	158
TLM_416@ONA	SER_115@N	149	2.92	152
TLM_416@ONA	LEU_113@N	145	2.86	150
TLM_416@ONA	SER_115@OG	68	2.84	254
TLM_416@ONA	ARG_211@NE	55	2.84	253
TLM_416@O11	SER_115@N	30	2.88	148
GLY_112@O	TLM_416@O6	20	2.81	140
TLM_416@ONA	ARG_211@NH2	20	2.89	154
VAL_346@O	TLM_416@O6	15	2.88	138
TLM_416@O11	GLN_209@NE2	14	2.87	145
TLM_416@O11	GLN_209@NE2	9	2.88	151
TLM_416@ONA	GLN_209@NE2	8	2.85	155
TLM_416@O6	GLN_209@NE2	8	2.93	152
TLM_416@O11	HIS_274@NE2	7	2.93	152
PRO_203@O	TLM_416@O6	4	2.69	142
TLM_416@O11	GLY_114@N	4	2.91	144
TLM_416@ONA	ARG_211@NH1	4	2.96	143
MET_275@O	TLM_416@O6	2	2.68	148
TLM_416@O6	ARG_211@NH2	2	2.80	157
TLM_416@O11	PHE_207@N	1	2.95	150

Table 2
b: Calculated properties of hydrogen bonds formed between BA and FabF.

H-bond acceptor	H-bond donor	Frame number	Average distance (Å)	Average angle (degrees)
Betulinic acid				
GLU_200@O	BA _416@O2	7476	2.75	159
BA _416@O1	GLU_200@N	619	2.91	160
BA _416@O2	GLY_114@N	510	2.99	153
GLN_209@OE1	BA _416@O	429	2.77	158
BA _416@O2	GLU_200@N	190	2.92	152
GLY_112@O	BA _416@O2	139	2.81	144
VAL_165@O	BA _416@O2	138	2.85	143
BA _416@O	GLN_209@NE2	50	2.90	156
MET_275@O	BA _416@O	21	2.82	153
BA _416@O1	GLY_114@N	20	2.92	144
BA _416@O	GLN_209@NE2	17	2.88	154
BA _416@O1	ILE_204@N	16	2.92	141
GLN_209@NE2	BA _416@O	12	2.90	154
LEU_113@O	BA _416@O2	1	2.83	135
GLY_206@O	BA _416@O	1	2.91	155
BA_416@O1	PHE_207@N	1	2.98	142

Table 2c: Calculated properties of hydrogen bonds formed between EUC and FabF.

H-bond acceptor	H-bond donor	Frame number	Average distance (Å)	Average angle (degrees)
Eucosterol				
EUC _416@O4	EUC_416@O2	64411	2.71	156
EUC _416@O2	HIS_344@NE2	33189	2.88	158
EUC _416@O2	EUC_416@O4	30198	2.72	151
EUC _416@O4	HIS_344@NE2	9630	2.88	152
EUC _416@O1	PHE_207@N	2414	2.89	163
EUC _416@O3	PHE_207@N	994	2.89	164
GLY_112@O	EUC _416@O2	744	2.78	162
EUC _416@O1	ARG_211@NH2	648	2.83	158
MET_275@O	EUC _416@O4	517	2.73	161
EUC _416@O1	ARG_211@NH1	153	2.84	151
EUC _416@O1	ARG_211@NH2	152	2.81	148
EUC _416@O4	ARG_211@NE	130	2.89	155
EUC _416@O4	ARG_211@NH1	110	2.88	149
EUC _416@O1	ARG_211@NE	79	2.86	155
MET_275@O	EUC _416@O2	57	2.74	152
EUC _416@O4	ARG_211@NH2	48	2.87	155
EUC _416@O1	ARG_211@NH1	43	2.88	146
VAL_165@O	EUC _416@O4	36	2.83	148
EUC _416@O4	ARG_211@NH1	25	2.91	149
EUC _416@O4	ARG_211@NH2	4	2.80	148
ARG_211@NH1	EUC _416@O4	3	2.93	141
EUC _416@O4	ALA_167@N	1	2.93	136
LEU_113@N	EUC_416@O2	1	2.94	140

Table 2d: Calculated properties of hydrogen bonds formed between UAA and FabF.

H-bond acceptor	H-bond donor	Frame number	Average distance (Å)	Average angle (degrees)
Ursolic acid acetate				
MET_275@O	UAA_416@O3	85	2.84	147
UAA_416@O2	GLN_209@NE2	31	2.88	153
UAA_416@O2	GLN_209@NE2	9	2.83	150

TLM-FabF (Table 2a) system showed an increase in hydrogen bonds distance between residue atoms interacting with TLM. The increase is about 0.007 Å and 0.003 Å greater than that of FabF-EUC and FabF-UAA system respectively. These differences may be responsible for relative high potency exhibited by these natural compounds in inhibiting FabF compared to TLM. Frequent occurrence of shorter hydrogen bond distance in FabF-UA (2.8 Å) and FabF-EUC (2.86 Å) relative to FabF-TLM (2.9 Å) may also have resulted to the high potency of natural compounds over TLM (Fig. 7). Also the presence of multiple cyclohexane rings in natural compounds may have influenced ligand orientation in the active site region and decreased the average angle formation, thereby increasing ligand-residue interaction.

3.8 Predicted toxicity analysis

Some chemical compounds are promiscuous in nature, meaning that they may interact with biological receptors different from experimentally recognised receptors. Compound promiscuity reflects the molecular basis of the pharmacological effects therefore, in-depth assessment of toxicity of compounds is essential to establish the degree of promiscuity among compounds at different level of drug research (Hu *et al*., 2014). Toxicity of a compound defines a quality of it being poisonous and estimating the toxicity and biological activity of a compound provides a detailed knowledge of other properties of a drug that likely have not been anticipated (Ndagi *et al*., 2014). The ProTox-II server toxicity prediction tool has been previously used to efficiently assess toxicity profiles of some anti-tubercular compounds (Cinaroglu and Timucin, 2019). It assigns toxicity classes of compounds ranging from 1–6: One being the most lethal and six being the non-toxic class. In this study, this sever was used to investigate the toxicity profile of TLM, BA, EUC and UAA (Table 3).

Table 3
Toxicity analysis of virtual screening hits and thiolactomycin.

Compound	LD50 (mg/kg)	Toxicity class
Thiolactomycin	1689	4
Betulinic acid	2500	5
Eucosterol	238	3
Ursolic acid acetate	4800	5
LD50: median lethal dose		

The toxicity analyses showed that the investigated natural compounds are generally less toxic, with BA and UAA being less toxic compared to TLM. This was derived from their toxicity class obtained from the ProTox-II server toxicity prediction tool. LD50 is the dose at which 50% of the test subject die upon exposure to the compound. The LD50 are associated with toxicity class of the compound. These results indicated the potential of these ligands as plausible TB drug leads. These findings are in agreement with who reported that extracts of *W. salutaris* are not toxic to mammalian cell line and possess great anti-mycobacterial activity (Madikane et al., 2007, Oluwagbemigia *et al.*, 2018).

Conclusion

The discovery of compounds from natural sources has given hope to the development and designing of potent *M. tuberculosis* FabF inhibitors for the treatment of tuberculosis. Understanding conformational features of FabF induced by natural compounds, which could open alternative avenues in the treatment of TB is unexploited. In this report, computational approaches have been employed to provide an in-depth understanding of the influence of *W. salutaris* bioactive compounds. BA, EUC and UAA had the highest binding energies amongst other compounds compared to TLM, meaning they have a higher binding affinity to FabF and consequently a higher inhibitory potential. In addition, BA and UAA further boasts with another attractive character of a relatively low toxicity profile compared to TLM. The qualities of these three compounds warrant further experimental validation as they bear promising inhibitory properties against *M. tuberculosis* FabF. This report serves as a starting point to the development of these compounds as drugs to potentially replace TLM as better drugs with reduced toxicity profiles in the treatment of TB.

Declarations

Funding: The authors acknowledge the University of KwaZulu-Natal and the National Research Foundation (NRF) for financial assistance

Conflicts of interest/Competing interests: The authors declare no conflict of interest

Availability of data and material: The dataset generated during the current study are available from the corresponding author on reasonable request.

Code availability: N/A

Authors' contributions: Ndumiso N. Mhlongo designed the project, Nokukhanya Gumede ran calculations, Kgothatso E. Machaba analysed the data, Hezekiel M. Kumalo and Umar Ndagi proof-read the manuscript. All authors read and approved the final manuscript.

References

Berman, H. M., Battistuz, T., Bhat, T. N., Bluhm, W. F., Bourne, P. E., Burkhardt, K., Feng, Z., Gilliland, G. L., Iype, L., Jain, S., Fagan, P., Marvin, J., Padilla, D., Ravichandran, V., Schneider, B., Thanki, N., Weissig, H., Westbrook, J. D., & Zardecki, C. (2002). The protein data bank. *Acta Crystallographica Section D: Biological Crystallography*, *58*(6 I), 899–907. https://doi.org/10.1107/S0907444902003451

Botha, J., Witkowski, E. T. F., & Shackleton, C. M. (2004). The impact of commercial harvesting on Warburgia salutaris (' pepper-bark tree ') in Mpumalanga , South Africa. *Biodiversity and Conservation*, *3*, 1675–1698.

Choi, K., Kremer, L., Besra, G. S., & Rock, C. O. (2000). Identification and Substrate Specificity of -Ketoacyl (Acyl Carrier Protein) Synthase III (mtFabH) from Mycobacterium tuberculosis *. *Journal of Biological Chemistry*, *275*(36), 28201–28207. https://doi.org/10.1074/jbc.M003241200

Chollet, A., Stigliani, J., Rosalia, M., Giorgia, P., Lherbet, C., Constant, P., Quémard, A., Bernadou, J., Pratviel, G., & Bernardes-génisson, V. (2016). *Evaluation of the inhibitory activity of (aza) isoindolinone- type compounds : toward in vitro InhA action , Mycobacterium tuberculosis growth and mycolic acid biosynthesis*. *June*, 740–755. https://doi.org/10.1111/cbdd.12804

Daffé, M., Quémard, A., & Marrakchi, H. (2017). *Mycolic Acids : From Chemistry to Biology*. https://doi.org/10.1007/978-3-319-43676-0

Douglas, J. D., Senior, S. J., Morehouse, C., Phetsukiri, B., Campbell, I. B., Besra, G. S., & Minnikin, D. E. (2002). Analogues of thiolactomycin : potential drugs with enhanced anti-mycobacterial activity. *Microbiology*, *148*, 3101–3109.

Frum, Y., Viljoen, A. M., & Drewes, S. E. (2005). and drimane sesquiterpenoids. *South African Journal of Botany*, *71*(3&4), 447–449. https://doi.org/10.1016/S0254-6299(15)30119-8

Gajiwala, K. S., Margosiak, S., Lu, J., Cortez, J., Su, Y., Nie, Z., & Appelt, K. (2009). Crystal structures of bacterial FabH suggest a molecular basis for the substrate specificity of the enzyme. *FEBS Letters*, *583*(17), 2939–2946. https://doi.org/10.1016/j.febslet.2009.08.001

Hanwell, M. D., Curtis, D. E., Lonie, D. C., Vandermeersch, T., Zurek, E., & Hutchison, G. R. (2012). Avogadro : An advanced semantic chemical editor , visualization , and analysis platform. *Journal of Cheminformatics*, *4*, 17. https://doi.org/10.1186/1758-2946-4-17

Jnawali, Nath, H., Hwang, Chul, S., Park, Kil, Y., Kim, H., Lee, Seon, Y., Chung, Tae, G., Choe, Hyeon, K., & Ryoo, S. (2013). Characterization of mutations in multi- and extensive drug resistance among strains of Mycobacterium tuberculosis clinical isolates in Republic of Korea. *Diagnostic Microbiology and Infectious Disease*. https://doi.org/10.1016/j.diagmicrobio.2013.02.035

Kim, S., Chen, J., Cheng, T., Gindulyte, A., He, J., He, S., Li, Q., Shoemaker, B. A., Thiessen, P. A., Yu, B., Zaslavsky, L., Zhang, J., & Bolton, E. E. (2019). PubChem 2019 update: Improved access to chemical data. *Nucleic Acids Research*, *47*(D1), D1102–D1109. https://doi.org/10.1093/nar/gky1033

Kremer, L., Douglas, J. D., Baulard, A. R., Morehouse, C., Guy, M. R., Alland, D., Dover, L. G., Lakey, J. H., Jacobs, W. R., Brennan, P. J., Minnikin, D. E., & Besra, G. S. (2000). Thiolactomycin and Related Analogues as Novel Anti-mycobacterial Agents Targeting KasA and KasB Condensing Enzymes in Mycobacterium tuberculosis *. *Jornal of Biomolecular Chemistry*, *275*(22), 16857–16864. https://doi.org/10.1074/jbc.M000569200

Lawal, I. O., Grierson, D. S., & Afolayan, A. J. (2014). Phytotherapeutic Information on Plants Used for the Treatment of Tuberculosis in Eastern Cape Province , South Africa. *Evidence-Based Complementary and Alternative Medicine*, *2014*(Figure 1).

Madikane, V. E., Bhakta, S., Russell, A. J., Campbell, W. E., Claridge, T. D. W., Elisha, B. G., Davies, S. G., Smith, P., & Sim, E. (2007). Inhibition of mycobacterial arylamine N -acetyltransferase contributes to anti-mycobacterial activity of Warburgia salutaris. *Bioorganic & Medicinal Chemistry*, *15*, 3579–3586. https://doi.org/10.1016/j.bmc.2007.02.011

Maroyi, A. (2013). Warburgia salutaris (Bertol. f.) Chiov.: A multi-use ethnomedicinal plant species. *J. Med. Plants Res*, *7*, 53–60.

Msomi, N. Z., Shode, F. O., Pooe, O. J., Mazibuko-mbeje, S., & Simelane, M. B. C. (2019). Iso-Mukaadial Acetate from Warburgia salutaris Enhances Glucose Uptake in the L6 Rat Myoblast Cell Line. *Biomolecules*, *9*, 1–12.

Nyaba, Z. N., Murambiwa, P., Opoku, A., Mukaratirwa, S., Shode, F., & Simelane, M. (2018). Isolation, characterization, and biological evaluation of a potent anti-malarial drimane sesquiterpene from Warburgia salutaris stem bark. *Malar. J*, *17*, 296.

Rabe, T., & Staden, J. Van. (2000). *Isolation of an antibacterial sesquiterpenoid from Warburgia salutaris*. *73*, 171–174.

Sargsyan, K., Grauffel, C., & Lim, C. (2017). How Molecular Size Impacts RMSD Applications in Molecular Dynamics Simulations. *Journal of Chemical Theory and Computation*, *13*(4), 1518–1524. https://doi.org/10.1021/acs.jctc.7b00028

Schiebel, J., Kapilashrami, K., Fekete, A., Bommineni, G. R., Schaefer, M., Mueller, M. J., Tonge, P. J., Kisker, C., Schiebel, J., Kapilashrami, K., Fekete, A., Bommineni, G. R., & Schaefer, C. M. (2013). *Structural Basis for the Recognition of Mycolic Acid Precursors by KasA , a Condensing Enzyme and Drug Target from Mycobacterium Tuberculosis **. https://doi.org/10.1074/jbc.M113.511436

Soyingbe, O. S., Mongalo, N. I., & Makhafola, T. J. (2018). In vitro antibacterial and cytotoxic activity of leaf extracts of Centella asiatica (L .) Urb , Warburgia salutaris (Bertol . F .) Chiov and Curtisia dentata (Burm . F .) C . A . Sm - medicinal plants used in South Africa. *BMC Complementary and Alternative Medicine*, *18*, 315.

Sridharan, S., Wang, L., Brown, A. K., Dover, L. G., Kremer, L., Besra, G. S., & Sacchettini, J. C. (2008). NIH Public Access. *Journal of Molecular Biology*, *366*(2), 469–480. https://doi.org/10.1016/j.jmb.2006.11.006.X-Ray

Sridharan, S., Wang, L., Brown, A. K., Dover, L. G., Kremer, L., Besra, G. S., Sacchettini, J. C., & Eugène, P. (2007). *X-Ray Crystal Structure of Mycobacterium tuberculosis β -Ketoacyl Acyl Carrier Protein Synthase II (mt KasB)*. 469–480. https://doi.org/10.1016/j.jmb.2006.11.006

Tabuti, J. R. S., Kukunda, C. B., & Waako, P. J. (2010). Medicinal plants used by traditional medicine practitioners in the treatment of tuberculosis and related ailments in Uganda. *Journal of Ethnopharmacology*, *127*, 130–136. https://doi.org/10.1016/j.jep.2009.09.035

Takayama, K., Wang, C., Besra, G. S., Takayama, K., Wang, C., & Besra, G. S. (2005). *Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis Pathway to Synthesis and Processing of Mycolic Acids in Mycobacterium tuberculosis*. *18*(1). https://doi.org/10.1128/CMR.18.1.81

Wachter, G. A., Valcic, S., Flagg, M. L., Franzblau, S. G., Montenegro, G., Suarez, E., & Timmermann, B. N. (1999). Antitubercular activity of pentacyclic triterpenoids from plants of Argentina and Chile. *Phytomedicine*, *6*(5), 341–345. https://doi.org/10.1016/S0944-7113(99)80056-7

Who. (2019). *GLOBAL TUBERCULOSIS REPORT 2019*. 1–2.

Figures

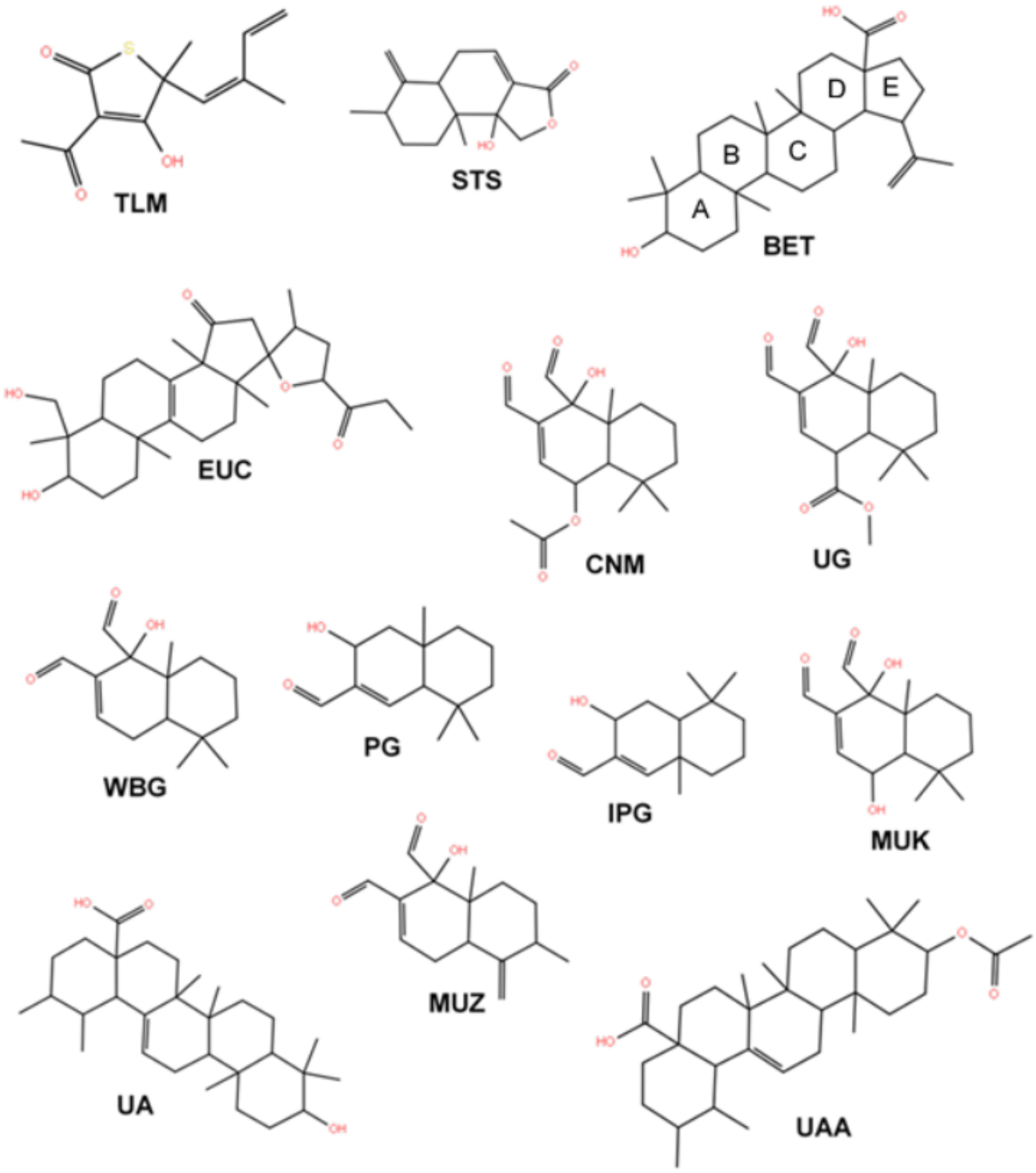

Figure 1

Chemical structures of thiolactomycin and natural compounds extracted from W. salutaris (Schrodinger, 2020). TLM= Thiolactomycin; BA= betulinic acid; CM= cinnamodial; EUC= eucosterol; IPG= isopolygodial; MUK= mukaadial; PG= polygodial; ST= salutarisolide; UG= ugandensdial; UA= ursolic acid; UAA= Ursolic acid; WBG= warburganal.

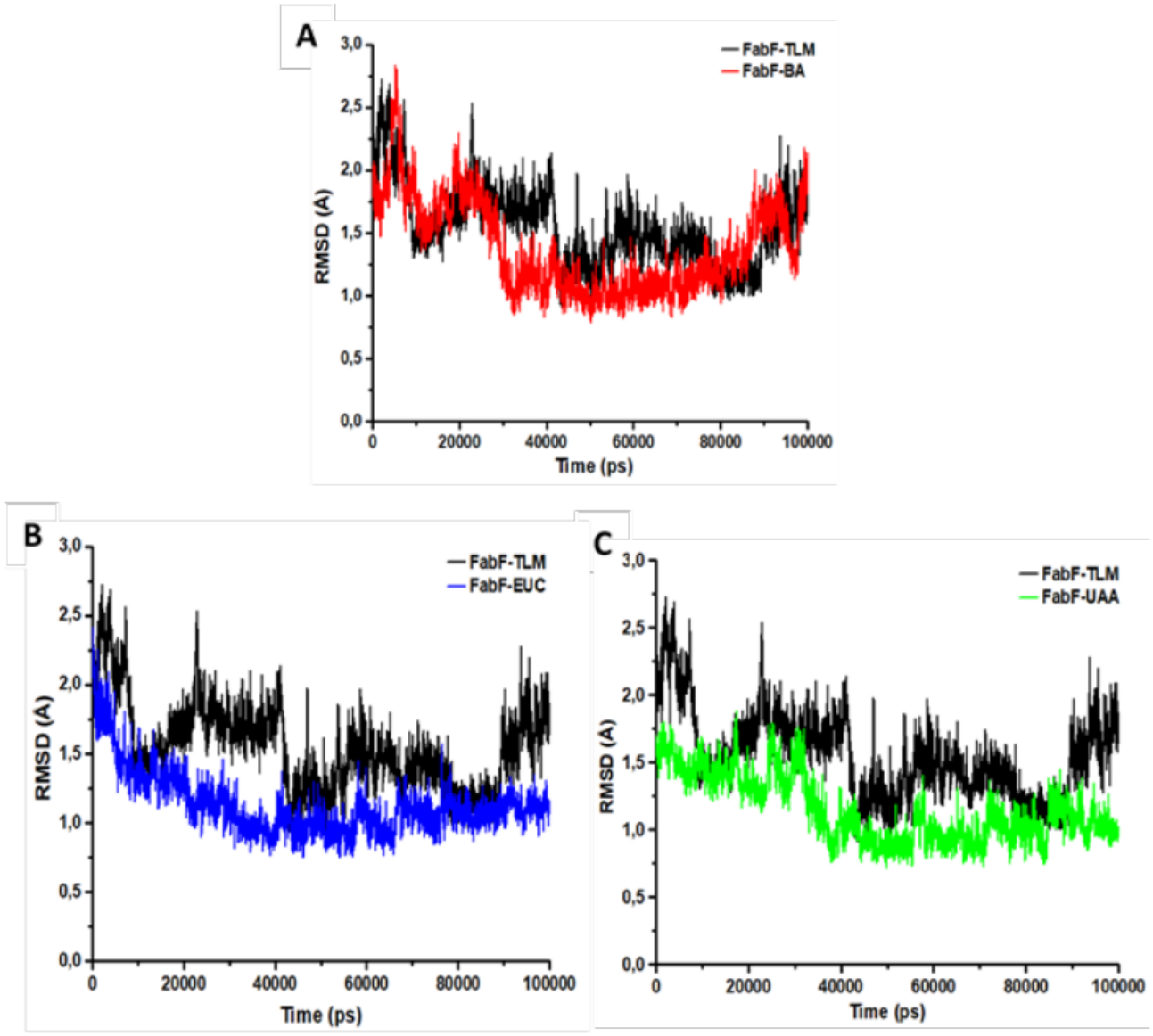

Figure 2

RMSD plots of FabF complexed with BA (A), EUC (B) and UAA (C) complexes compared with TLM (black).

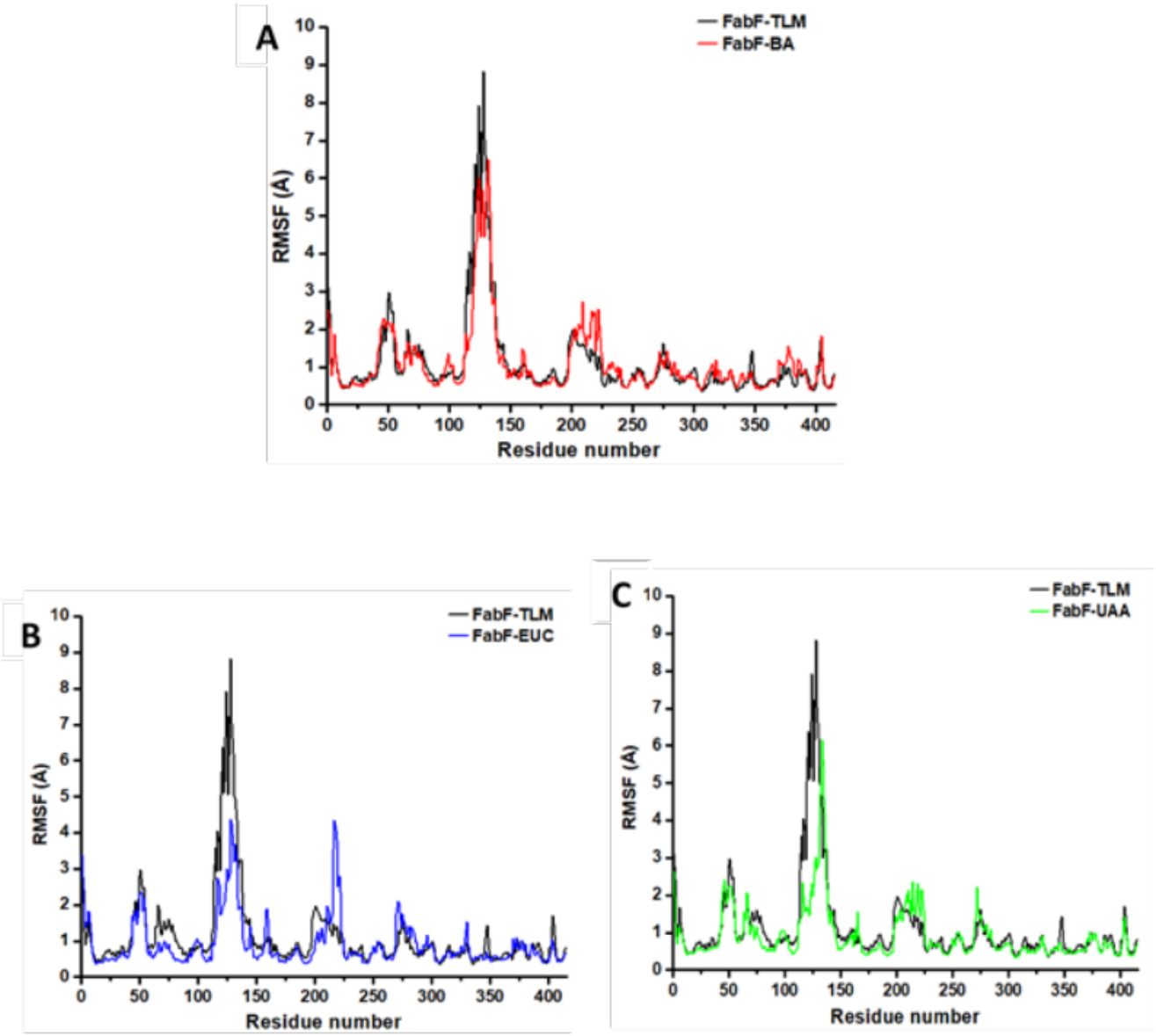

Figure 3

RMSF of FabF-BET (A), FabF-EUC (B) and FabF-UAA (C) complexes compared with FabF-TLM.

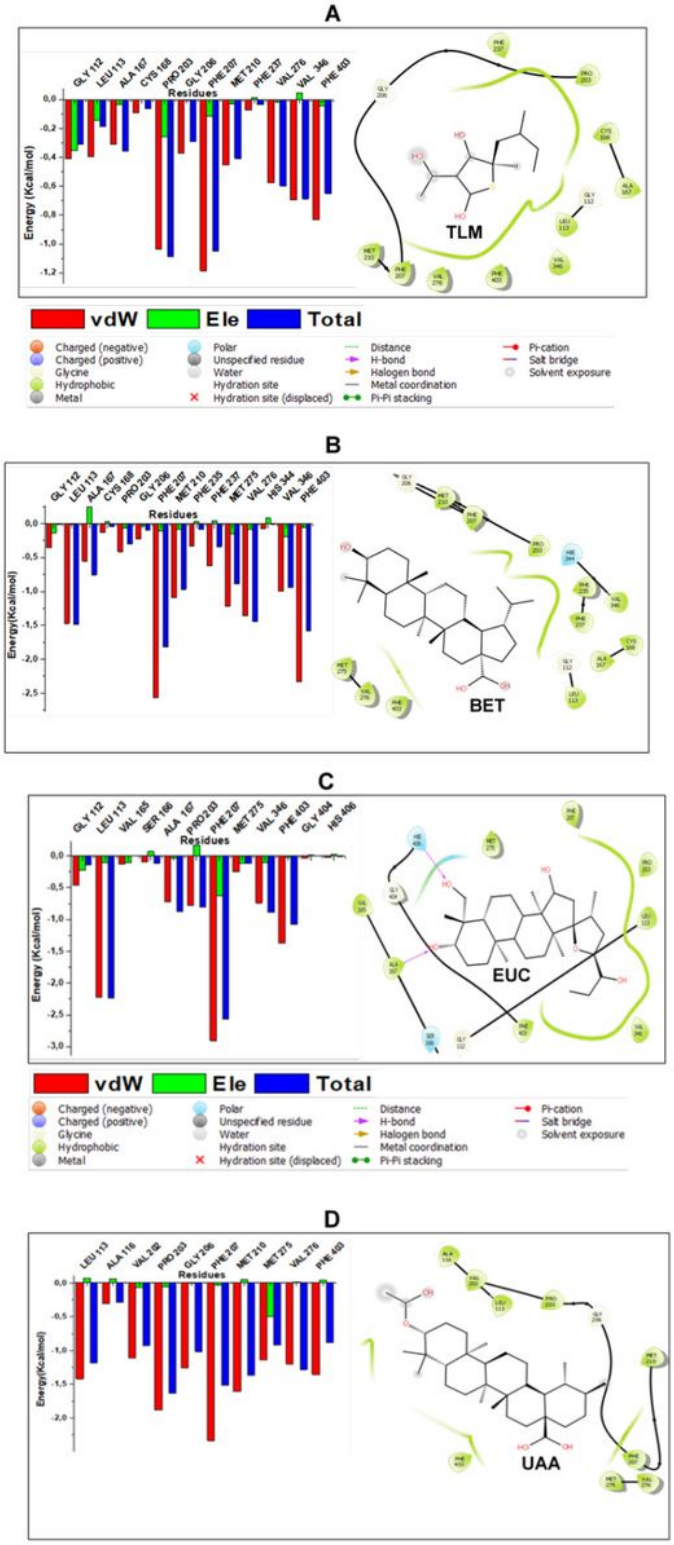

Figure 4

The per-residue energy decomposition analysis graph of FabF-TLM (A), FabF-BA (B), FabF-EUC (C) and FabF-UAA (D) systems with their 2D ligand-residue interaction network from fully minimised complex.

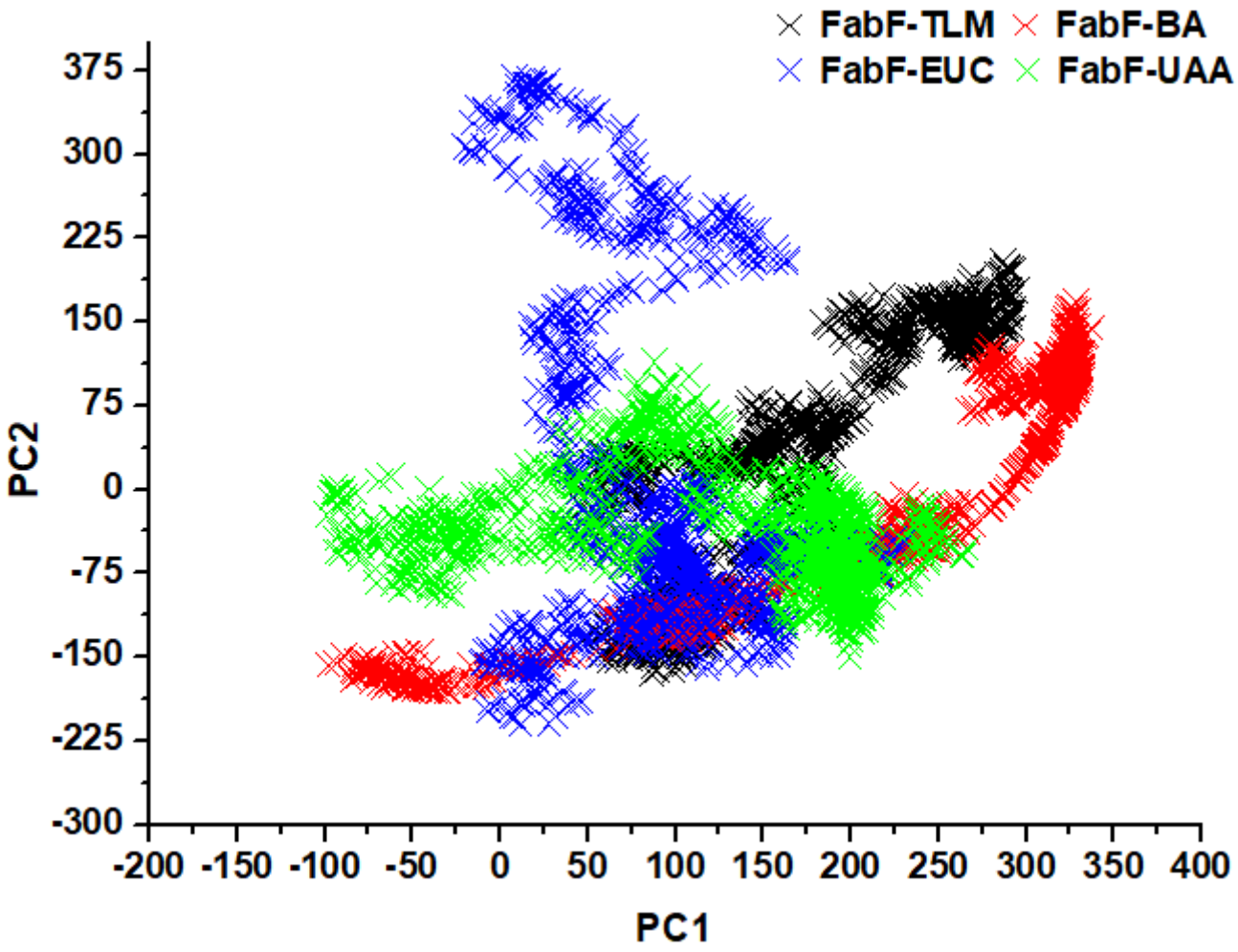

Figure 5

PCA projection of Cα atoms motion constructed by plotting the first two principal components (PC1 and PC2) in conformational space, FabF-TLM (black), FabF-BA (red), FabF-EUC (blue) and FabF-UAA (green).

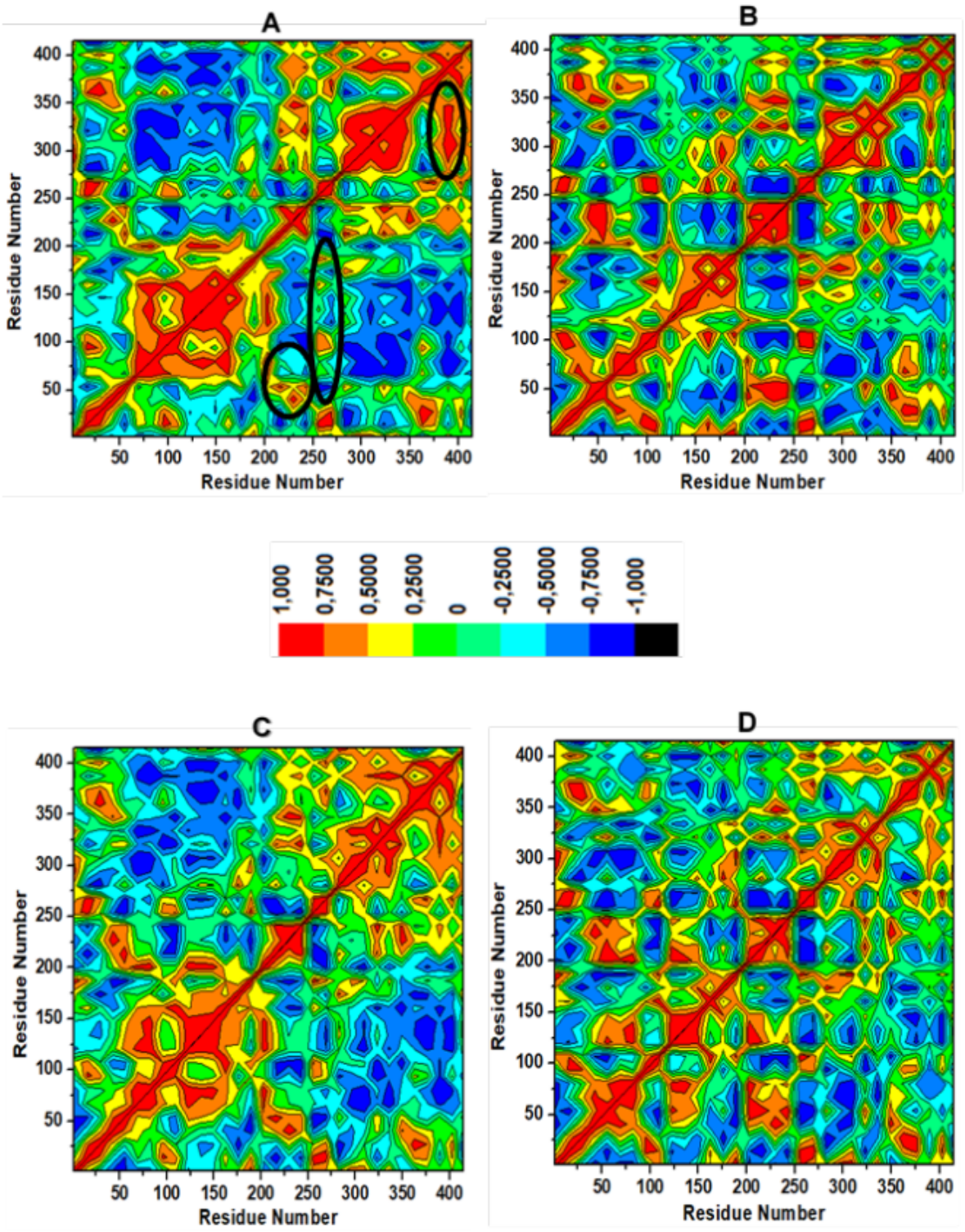

Figure 6

Cross-correlation matrices of Cα atom fluctuations in FabF-TLM (A), FabF-BA (B), FabF-EUC (C) and FabF-UAA (D) systems.

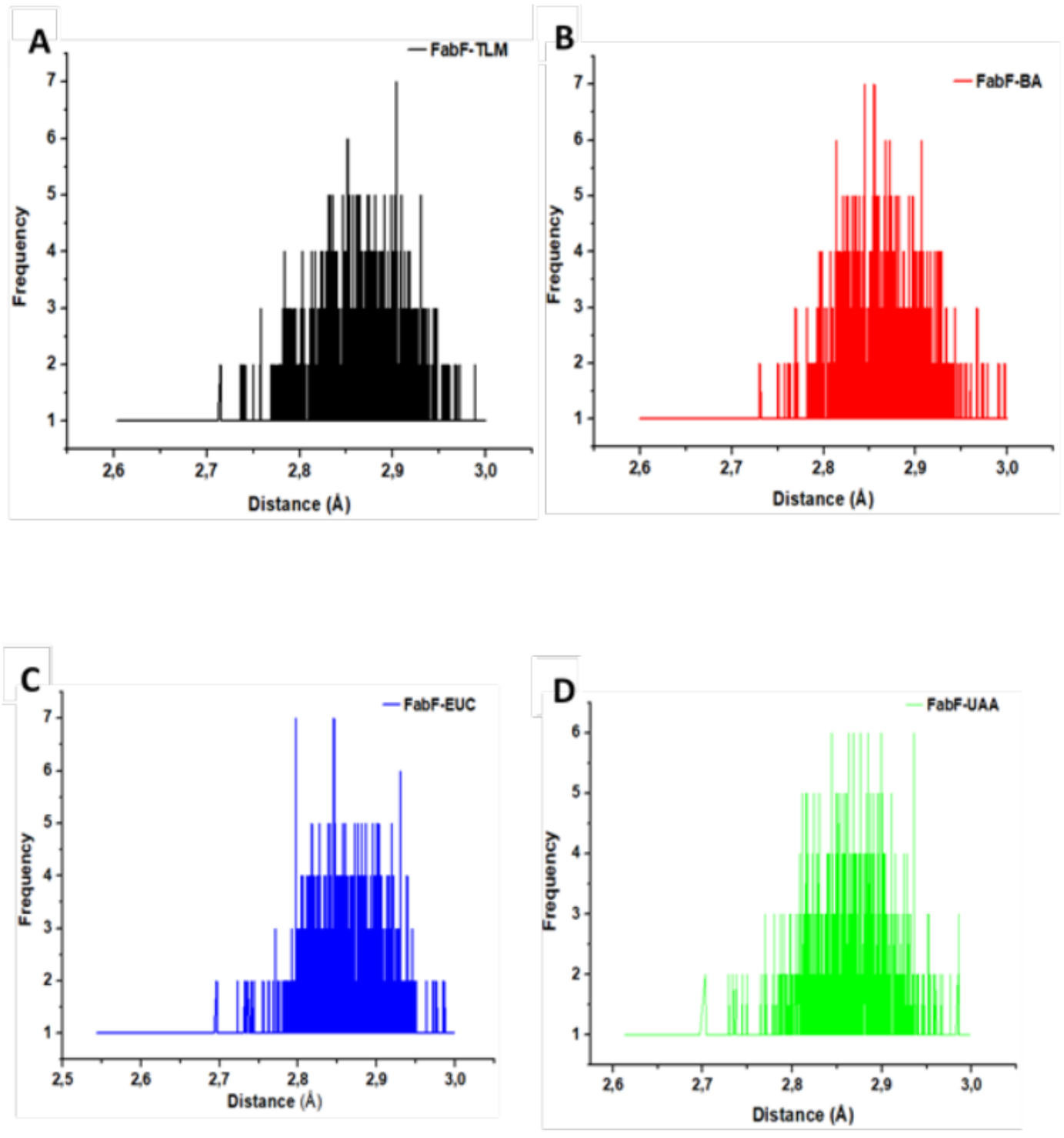

Figure 7

Occurrence of hydrogen bond distance between FabF-TLM (A), FabF-BA (B), FabF-EUC (C), FabF-UAA (D).