Computational Biology and Chemistry 98 (2022) 107658

Contents lists available at ScienceDirect

Computational Biology and Chemistry

journal homepage: www.elsevier.com/locate/cbac

Heterologous expression, biochemical characterisation and computational analysis of *Bacteroides fragilis* enolase

Erennur Ugurel [a], Sinem Kocer [b], Emrah Sariyer [c], Ozal Mutlu [d], Tugba Gul Inci [a], Osman Mutluhan Ugurel [a,e], Dilek Turgut-Balik [a,*]

[a] Yildiz Technical University, Faculty of Chemical and Metallurgical Engineering, Department of Bioengineering, Davutpasa Campus, 34210 / Esenler, Istanbul, Turkey
[b] Istanbul Yeni Yuzyil University, Faculty of Pharmacy, Department of Pharmaceutical Biotechnology, 34010/Cevizlibag, Istanbul, Turkey
[c] Artvin Coruh University, Vocational School of Health Services, Medical Laboratory Techniques, 08000 Artvin, Turkey
[d] Marmara University, Faculty of Arts and Sciences, Depatment of Biology, Goztepe Campus, 34722 / Kadikoy, Istanbul, Turkey
[e] Altınbas University, School of Engineering and Architecture, Department of Basic Sciences, 34217/ Bağcılar, Istanbul, Turkey

ARTICLE INFO

Keywords:
Bacteriodes fragilis
Enolase
Antibiotic resistance
Structure-based drug design
Enzyme kinetic
In silico studies

ABSTRACT

Bacteriodes fragilis is an anaerobic bacterium found in the human intestinal flora. In this study, *Bf*Eno was targeted with a structure-based drug design approach because inhibition of this enzyme may prevent both the aerobic and anaerobic pathways due to its role in the glycolytic pathway. First, the gene encoding *Bf*Eno was cloned, expressed and the protein produced over 95% purity. The K_m and V_{max} values of *Bf*Eno were determined as 314.9 μM and 256.2 μmol/min.mg, respectively. Drug-like chemicals were retrieved from the ZINC database for high-throughput virtual screening analyses. As a result of screening study, the ZINC91441604 has been proposed to bind to the active site of the enzyme and remain stable. The same compound exhibited weak binding to the human enolases than the bacterial enolase. Hence, ZINC91441604 may be proposed as a novel candidate for further *in vitro* and *in vivo* drug analysis towards the treatment of *B. fragilis* infections.

1. Introduction

Bacteroides fragilis is an anaerobic bacterium found in the human intestinal flora. It also has some pathogenic properties (Ulger Toprak et al., 2006). Although *B. fragilis* is not the most common bacterium among Bacteroides species, it is the most frequently isolated anaerobe in intestinal infections (Valguarnera and Wardenburg, 2020). Remarkably, toxin-secreting enterotoxigenic *B. fragilis* (ETBF) does not only cause human intestinal infection, furthermore has an effect on the development of colon cancer (Ulger Toprak et al., 2006; Majid and Andleeb, 2019). Despite, *B. fragilis* infections could be treated with antibiotics such as carbapenem, clindamycin and metronidazole (Ghotaslou et al., 2018), drug resistance to these antibiotics has been reported (Urbán et al., 2015; Nakano et al., 2011; Aldridge et al., 2001). Antibiotic resistance is one of the most urgent public health problems affecting human life in all aspects such as health, agriculture, social and economic futures (Center for Disease, 2019; Podolsky, 2018). Therefore, alternative novel drugs necessity has became emergence due to the increasing drug resistance (Roope et al., 2019). *In silico* drug screening is a very supportive method for *in vitro* studies with the advantages of saving time, being effective and economic in drug design studies. Metz et al. (2019). Structure-based drug design enables evaluation of the interaction between target molecule and ligands. This method also allows *in silico* drug screening on the target structure and analysis of new drug candidates (Singh et al., 2006).

Enolase is an essential enzyme involved in *B. fragilis*' anaerobic energy metabolism (Sijbrandi et al., 2005). Therefore it has been chosen as a target macromolecule for an alternative drug design. Enolase is a metalloenzyme, responsible for the conversion of 2-phosphoglyceric acid to phosphoenolpyruvate by binding the Mg^{+2} metal ion as a cofactor. It is also a multifunctional protein found in the cytoplasm of the cells (Díaz-Ramos et al., 2012).

In this study, *B. fragilis* enolase (*Bf*Eno) was cloned, expressed, and its kinetic parameters were determined following the purification of the protein. The 3-D structure of *Bf*Eno was generated by homology modeling and a compound library was screened virtually using molecular docking methods. The stability and interactions of the ligands were analysed by molecular dynamics simulations.

* Corresponding author.
E-mail address: Turkeydilekbalik@gmail.com (D. Turgut-Balik).

https://doi.org/10.1016/j.compbiolchem.2022.107658
Received 22 July 2021; Received in revised form 5 February 2022; Accepted 25 February 2022
Available online 26 February 2022

2. Materials and method

2.1. Cloning of BfEno into the pLATE31 expression vector

*Bf*Eno gene was cloned into the pLATE 31 vector using the aLICator LIC Cloning and Expression Kit 3 with C-terminal His-tag (Thermo Scientific, USA, #Cat. No. K1261) (Scientific, Thermo, 2013). Two specific primers for this system were designed; the forward primer *Bf*EnoLICF: 5′AGAAGGAGATATAACTATGAAAATAGAAAAAATTACAGGACG3' and the reverse primer *Bf*EnoLICR: 5′GTGGTGGTGATGGTGATGGCCGC CTTTTACTAC3' were used for amplification of the *Bf*Eno gene. PCR conditions were as follows; pre-denaturation at 95 °C for 5 min, denaturation at 94 °C for 1.5 min, binding at 50 °C for 2 min, elongation at 72 °C for 2 min. This reaction was carried out for 30 cycles and a final extension at 72 °C for 10 min was applied. *BfEno* in pLATE31 vector was then transformed into *E. coli* BL21(DE3) competent cells for expression (Sambrook et al., 1989).

2.2. Expression and purification of recombinant BfEno

*Bf*Eno was expressed and then purified using HisTALON® gravity columns (Clontech, Europe). Briefly, the recombinant *E. coli* BL21(DE3) cells were cultured in Luria-Bertani (LB) broth medium with the addition of 100 μg/mL ampicillin at 37 °C, 180 rpm. Afterwards, 500 μl of the culture including *BfEno* in *E. coli BL21(DE3)* inoculated into 50 mL LB broth medium and grown to OD_{600} 0.6. The cells were induced with 0.5 mM isopropyl β-D-1 thiogalactopyranoside (IPTG) addition and then incubated for 5 h at 30 °C. Further, cells were pelleted by centrifugation and pellets were resuspended in lysis buffer (50 mM Tris-HCl, pH 7,4) prior to sonication with 40% amplitude, 10 s pulses/10 s respite during 2 min on ice. The his-tagged *Bf*Eno was purified by affinity chromatography using HisTALON® gravity columns. Flowthrough, wash and elution fractions are obtained according to the manufacturer's instructions (Clontech, Europe). All the fractions were analysed by SDS-PAGE (Laemmli, 1970).

2.3. BfEno kinetic assay

Protein concentration was determined by UV spectrometer method. The absorbance of *Bf*Eno was obtained at 280 nm, and then Lambert-Beer Law equation was used to calculate the protein concentration. Molar absorption coefficient (ε) and molecular weight (MW) values of *BfEno* were identified by the ProtParam in ExPASy portal (Gasteiger et al., 2005).

Protein concentration (c) was calculated as follows:

$$c\ (mol / l) = (A_{280}) / (\varepsilon \times L)$$

$$c\ (mg / l) = (A_{280}) \times MW / (\varepsilon \times L)$$

A_{280} represents the absorbance and L defines the cuvette path length (cm) that is 1 cm in this study. According to these information, the concentration was calculated as follows (Grimsley and Pace, 2004);.

$$c\ (mol / l) = (A_{280}) / (\varepsilon)$$

$$c\ (mg / l) = (A_{280}) \times MW / (\varepsilon)$$

The kinetic parameters of *Bf*Eno were determined by measuring the increase of phosphoenolpyruvate (PEP) at 240 nm wavelength for 2 min. The measurements were carried out in 50 mM Tris-HCl (pH 7.4) and 1.5 mM $MgCl_2$ buffer at different 2-phosphoglyceric acid (2-PG) substrate concentrations (50–2000 μM) (Cayir et al., 2014; Yakarsonmez et al., 2019). *Bf*Eno activity assay was performed in duplicates at each substrate concentration. Specific activity of the enzyme was identified by using the reaction rate values, protein concentration and PEP's molar absorption coefficient ($\varepsilon PEP = 1400\ M^{-1}.\ cm^{-1}$) (Pal-Bhowmick et al., 2004). Then, steady-state kinetic graph was formed and the kinetic parameters (K_m *and* V_{max}) of *Bf*Eno were determined by using GraphPad Prism 6.0 program (GraphPad Software Ltd., La Jalla, USA).

2.4. Protein sequence analysis and structural modeling of BfEno

Bacteroides fragilis (Accession number: AKA51156), *Bacteroides vulgatus* (Accession number: RHH61676), *Fusobacterium nucleatum* (Accession number: EDK88978), *Enterococcus faecalis* (Accession number: CAK50216), *Streptococcus pneumoniae* (Accession number: UniProtKB/Swiss-Prot: Q97QS2.1), *Homo sapiens* (Accession number: UniProtKB/Swiss-Prot: P13929.5, muscle enolase), *Saccharomyces cerevisiae* (Accession number:UniProtKB/Swiss-Prot: P00924.3) enolase amino acid sequences were retrieved from the databases in FASTA format. These protein sequences were aligned using the MUSCLE program (http://www.ebi.ac.uk/Tools/msa/ muscle) (Edgar, 2004). Physicochemical properties of *BfEno* were analysed using the ProtParam tool in the ExPASy portal (Gasteiger et al., 2005).

Open and closed conformation models of *BfEno* were generated using MODELLER ver9.15 (Eswar et al., 2006). The template structures for both enzyme conformations were selected by using NCBI/BLAST server (http:// blast.ncbi.nlm.nih.gov/Blast.cgi). PSI-BLAST scanning was performed against the PDB (Protein Data Bank) proteins using the protein sequence of *BfEno* (Accession number: AKA51156). PDB file format of the template structures was obtained from the RSCB Protein Data Bank (www.rcsb.org). The FASTA format of *Bf*Eno amino acid sequence and the PDB file of template structures were entered into MODELLER ver9.15 program for building the models. Then, the best *Bf*Eno models were choosen according to the Discrete Optimized Protein Energy (DOPE) scores among 100 models. Root-mean-square deviation (RMSD) values were calculated using UCSF Chimera ver1.10.2 (Pettersen et al., 2004).

The open and closed conformation models of *Bf*Eno were simulated for 50 ns by MD methods for the optimization of the homology model and the determination of the thermodynamical equilibrium structure. TIP3PBOX water was added in an octahedral periodic box of 10 angstroms using AMBER14 (Case et al., 2012) tLeap. Following, Na^+ and Cl^- were added to the box to neutralize the system. Finally, the coordinate and topology files required for simulation were generated. Before starting the simulation, it is necessary to eliminate the bad interactions in the homology models and to minimize the high energies of the bonds between atoms. The first 1000 steps quickly reach the local minimum with the steepest descent algorithm (Fletcher and Powell, 1963) and then, the system was minimized with the conjugate gradient algorithm during 90,000 steps to reach the global minimum, effectively (Fletcher and Reeves, 1964). The systems were heated to 300 K for 0.4 ns with a 2 femtosecond steps at constant pressure with temperature control using Langevin thermostat to give the first velocity to the molecules in the system (Pastor et al., 1988). As the last step, the simulation production of the system was run using a weak couple temperature algorithm (Berendsen et al., 1984) at constant temperature and pressure, in 2 femtosecond steps for 50 ns. Throughout the simulation, the system was heated by controlling the bond strech and interactions with SHAKE mode (Ryckaert et al., 1977). Ewald summation method was used for the sum of the electrostatic energies in the system (Essmann et al., 1995). Grid size was specified for nfft1, nfft2, nfft3 commands and the larger the specified periodic box means the more accurate the calculation but the longer the calculation time.

These *Bf*Eno models were validated after 50 ns using molecular dynamic simulation. The quality of the open and closed conformation models was assessed by using ERRAT, Rampage, ProSA, ProQ, Verify3D, and QMEAN web-based programs (Colovos and Yeates, 1993; Lovell et al., 2003; Wiederstein and Sippl, 2007; Wallner and Elofsson, 2003; Benkert et al., 2008; Eisenberg et al., 1997).

2.5. Virtual screening and molecular dynamics simulation

Druggable sites of apo-*Bf*Eno structure were predicted by using DoGSiteScorer method from the ProteinPlus web portal (Volkamer et al., 2012; Fährrolfes et al., 2017) and SiteMap program (SiteMap, Schrödinger, LLC, New York, NY, 2017). Drug-like chemicals were retrieved from the ZINC database (Sterling and Irwin, 2015). Receptor structure and the chemical library were prepared using the Protein-PrepWizard and LigPrep, respectively, by applying the OPLS3 forcefield and protonation state at neutral pH (LigPrep, Schrödinger, LLC, New York, NY, 2017). The grid files of the target sites were generated using the Receptor Grid Generation panel. Virtual high-throughput screening was carried out by following the virtual screening workflow with the Glide program (Glide, Schrödinger, LLC, New York, NY, 2017). The first ten compounds with the best docking scores were identified by virtual screening among the database containing 6703 compounds.

Stability of *Bf*Eno-ligand complexes was simulated by molecular dynamic methods during 100 ns. First, the atomic force field parameters of ZINC91441604 (+2) and ZINC72431523 (+3) ligands were assigned using ANTECHAMBER tool in the AMBER package program. The force fields of the atoms were analysed using the parameter AMBER force field (gaff) and the AM1-BCC charge model (Jakalian et al., 2002). TIP3PBOX waters were added in an octahedral periodic box with a size of 10 angstroms *via* AMBER14 tLeap and Na^+ and Cl^- were added to the box in order to neutralize the system. Subsequently, the topology and coordinate files of the complexes were created. It is necessary to eliminate bad interactions in ligand-enzyme complexes and to minimize clashes between atoms. The first 1000 steps quickly reach the local minimum using the steepest descent algorithm (Fletcher and Powell, 1963) and then, in order to reach the global minimum effectively, the complexes are minimized with the conjugate gradient algorithm in 90,000 steps (Fletcher and Reeves, 1964). The system was heated to 300 K for 200 thousand steps at constant pressure with temperature control with Langevin thermostat to give the first velocity to the molecules (Pastor et al., 1988) in 2 femtosecond steps. As the last step, the simulation of the system was run using weak couple temperature algorithm (Berendsen et al., 1984) at constant temperature and pressure, in 2 femto second steps for 100 ns. Throughout the simulation, the system was heated by controlling the bond stretch and interactions with SHAKE mode (Ryckaert et al., 1977). Ewald summation method is used for the sum of the electrostatic energies in the system (Essmann et al., 1995).

In addition; the best-scored chemicals identified as the result of the docking on *Bf*ENO were used for comparative docking analysis with human enolases. The same compounds were docked on human enolases using the same methods described above. X-ray structures of human beta enolase (PDB ID: 2XSX_A) and human gamma enolase (PDB ID: 1TE6_B) (Chai et al., 2004) were collected from the RSCB PDB (www.rcsb.org). Open conformation of human alpha enolase was modelled by using the SWISS-MODEL (Waterhouse et al., 2018). Human gamma enolase (PDB ID: 1TE6_B) was used as template structure for modeling of human alpha enolase.

3. Results and discussion

3.1. Cloning of the gene encoding BfEno

The 1287 bp gene encoding *Bf*Eno was amplified (Fig. 1) and cloned into the pLATE31 vector using aLICator LIC cloning and expression Kit 3 (C-terminal His-tag) system. After transformation of *Bf*Eno into *E. coli BL21(DE3)* cells positive, colony selection was made by colony PCR method prior to DNA sequencing.

3.2. Expression and purification of recombinant BfEno

*Bf*Eno consisting 429 aa expressed in *E. coli BL21(DE3)* cells with 0.5 mM IPTG induction at 30 °C. The molecular weight of the protein was calculated at approximately 46.4 kDa by web-based ProtParam program. Recombinant his-tagged *Bf*Eno is purified by HisTALON® gravity columns using elution buffer containing 150 mM imidazole. *Bf*Eno elutions were analysed by SDS-PAGE and protein was obtained at approximately 46.4 kDa (Fig. 2). As seen in the Fig. 2, protein was purified over 95% purity and protein in elution 2 was used for kinetic analyses.

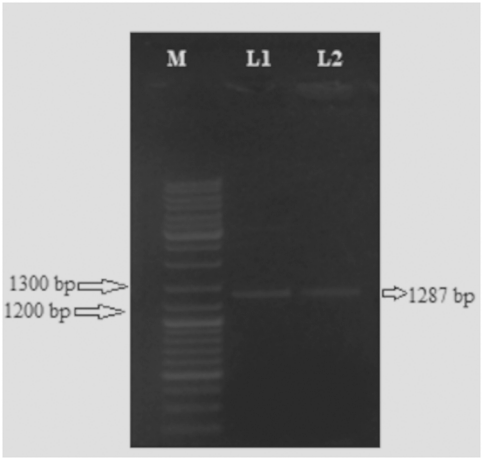

Fig. 1. Agarose gel analysis of *Bf*Eno M: Marker, L1: *Bf*Eno, L2: *Bf*Eno.

3.3. Determination of kinetic parameters of BfEno

The molar absorption coefficient (ε) and molecular weight (MW) values of *Bf*Eno were found to be 33,475 $M^{-1}.cm^{-1}$, 46,403 g / mol respectively, according to the ProtParam results. Pure enzyme was added to each activity reaction as 0.36 μg based on the calculations.

Reaction rate absorbances were measured at different substrate concentrations (50–2000 μM) at 240 nm to calculate K_m and V_{max} values for *Bf*Eno using Michaelis-Menten equation. Steady-state kinetic graph was generated by GraphPad Prism 6 (GraphPad Software Ltd., La Jalla, USA) program (Fig. 3). Despite K_m and V_{max} values of *Bf*Eno were identified lower in the study performed by Sijbrandi et al. (2005), these values were found higher in our hands as 314.9 μM and 256.2 μmol/min.mg, respectively. Protein made available for further drug discovery analysis after the production and characterisation of the kinetic parameters.

3.4. Protein sequence analyses of BfEno

Catalytic residues, loop regions, plasminogen binding sites and insertion sites on *Bf*Eno amino acid sequence were defined by protein

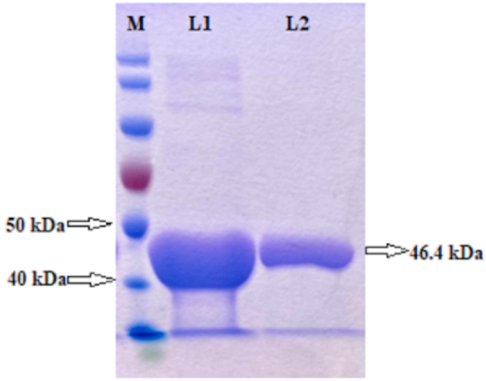

Fig. 2. SDS-PAGE gel analysis of purified recombinant *Bf*Eno. M: Protein Marker, L1: Elution 1, L2:Elution 2.

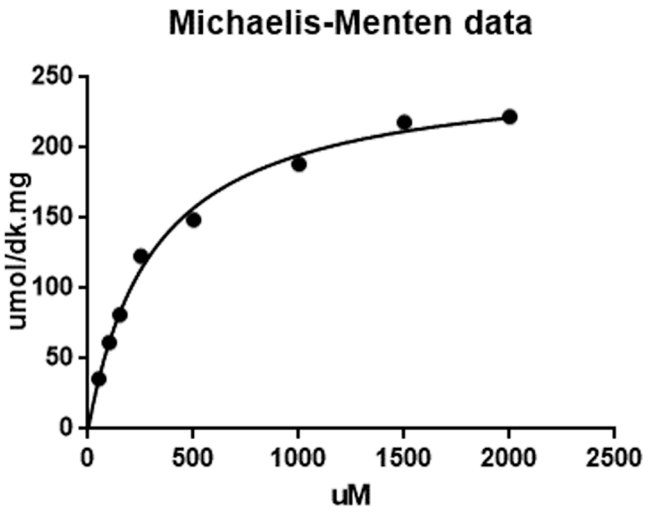

Fig. 3. Steady-state kinetics graph of *BfEno*.

sequence analysis. The sequence identity percentage of *Bf*Eno against *Bacteroides vulgatus, Fusobacterium nucleatum, Enterococcus faecalis, Streptococcus pneumoniae, Homo sapiens, Saccharomyces cerevisiae* enolases were found as 85.08%, 66.59%, 65.57%, 71.79%, 53.79% and 52.25% respectively (Fig. 4).

The crystal structure of *Saccharomyces cerevisiae* enolase was analysed previously (Larsen et al., 1996; Zhang et al., 1997) and therefore, residue numbering was configured based on *Saccharomyces cerevisiae* enolase. According to protein sequence analysis, amino acid residues bound to 2-PG (*Bf*Eno/Yeast enolase) were conserved as His154/159, Gln162/167, Lys340/345, His368/373, Arg369/374, Ser370/375, Lys391/396 (Larsen et al., 1996; Zhang et al., 1997). Glu163/168 and Glu204/211 amino acid residues that share a proton in the dehydration, were also identified. It was observed that Ser41/39 binds to the second Mg^{+2} ion and Asp241/246, Glu288/295 and Asp315/320 were also conversed to interact with the Mg^{+2} ion(Larsen et al., 1996; Zhang et al., 1997). In addition, Ala40/38 was found to be a residue that interacts with 2-PG (Zhang et al., 1997) (Fig. 4).

The tripeptide deletion on *Bacillus subtilis* enolase sequence was evaluated in the sequence analyses and it was observed that the loop between helix L and strand 1 is not found in the study of Brown et al. (1998). As shown in Fig. 4, this tripeptide deletion is also present between Gly134 and Gly135 on *B. fragilis* enolase amino acid sequence. It has also been suggested that the presence of these 2 Gly residues on *B. subtilis* enolase, similar to that in *Bf*Eno, may limit loop formation (Brown et al., 1998).

Enolase catalyzes a reaction in the glycolysis pathway and also the enzyme is present on the cell surface and plays a role as a plasminogen binding receptor (Díaz-Ramos et al., 2012). In the study of Bergmann et al., a specific sequence ($_{249}$FYDKERKVYD$_{257}$) on surface enolase from *Streptococcus pneumoniae* was observed as effective in host plasminogen binding (Bergmann et al., 2003). The $_{247}$FYHDGIYD$_{254}$ amino acid sequence on *Bf*Eno was found to be similar to the plasminogen binding sequence in this study. In addition, two amino acid deletions on the host (*Homo sapiens)* enolase sequence were also observed between His249

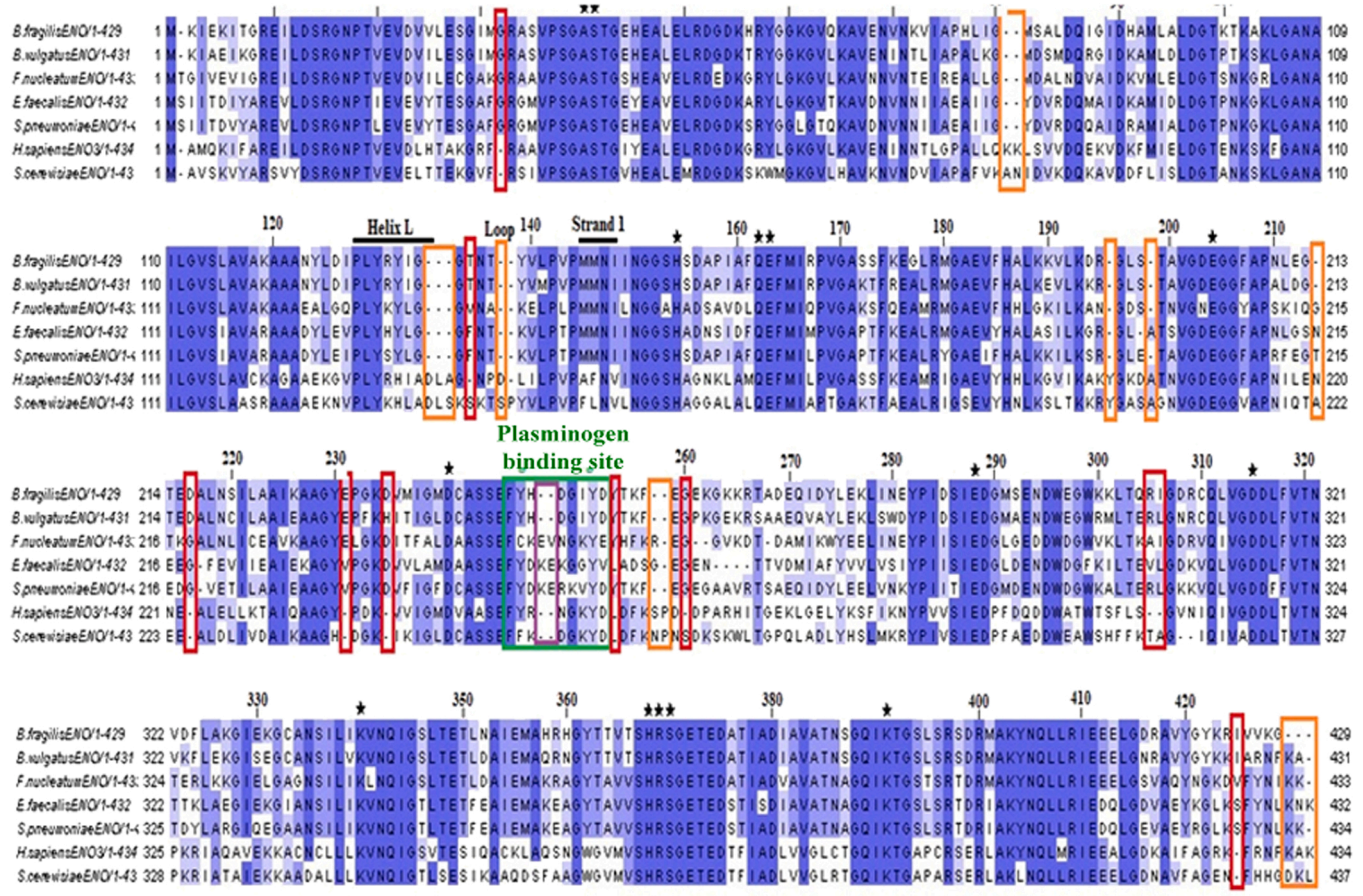

Fig. 4. Comparison of amino acid sequences of *B. fragilis, B. vulgatus, F. nucleatum, E. faecalis, S. pneumoniae, H. sapiens, S. cerevisiae* enolases using the MUSCLE program (http://www.ebi.ac.uk/Tools/ MS / muscle). Asterisk indicates catalytic residues. Regions in the red box indicate the insertion sites, regions in the orange box indicate the deletion regions, the site in the green box shows the plasminogen binding region and the site in the purple box shows the deletion in the plasminogen binding region.

and Asp250 residues on *Bf*Eno (Fig. 4) (Yakarsonmez et al., 2016).

A dipeptide insertion on *Bf*Eno was identified as Arg305 and Ile306 amino acid residues and also Gly32, Thr136, Asp216, Glu231, Asp235, Tyr255, Gly260, Ile425 amino acid residues were found as insertions on *Bf*Eno.

The physicochemical properties of *Bf*Eno were also determined in this study. ProtParam results showed that the enzyme was approximately 46.4 kDa, its theoretical pI was 5.43, the aliphatic index was 89.39, and the average hydropathicity was 0.258. In addition, *Bf*Eno has 61 negatively and 49 positively charged amino acid residues.

3.5. Structural modeling and validation

Group A *Streptococcal* enolase (PDB ID: 3ZLH_A) (Cork et al., 2015) was selected as the template structure for modelling of *BfEno* open conformation based on a percentage identity of 72%. *Staphylococcus aureus* enolase-PEP (PDB ID: 5BOE_A) (Wu et al., 2015) was also chosen as the template structure for modelling of *Bf*Eno closed conformation according to a percentage identity of 73%. Protein models were built by using the FASTA format of *Bf*Eno amino acid sequence and the PDB files of these template structures in MODELLER 9.15. The *BfEno* open conformation model was choosen based on DOPE score of -48700.19922 and the closed conformation model of *BfEno* was also selected according to DOPE score of -49377.01172.

The open conformation model was superimposed with 3ZLH_A template protein and the RMSD value was determined as 0.220 Å. According to the results of the closed conformation model superimposition with 5BOE_A template protein, the RMSD value was also identified as 0.223 Å. The RMSD value shows the distances between the atoms of the compared protein structures (Pettersen et al., 2004; Maiti et al., 2004) According to these values, *BfEno* models possessed high similarities with the template proteins.

Open and closed conformations of *Bf*Eno were modelled by MODELLER 9.15 were simulated for 50 ns by molecular dynamic methods. These models have been minimized to remove bad contacts and energies and were simulated to thermodynamically reach the equilibrium of the system. The system reached equilibrium after approximately 20 ns and the average RMSD value of open and closed conformations of *Bf*Eno were calculated as 1.768 and 1.721, respectively (Fig. 5a). Atomic fluctuation of each residue during 50 ns based on α-carbon was investigated by Bfactor plot. The fluctuation of the open and closed forms were given comparatively in the graph. The regions with the highest fluctuation in both molecules are between 30 and 40 and 240–270, and since the N-terminal regions did not have any secondary structure, the fluctuation was observed high (Fig. 5b). In general, similar regions fluctuated at a similar rate however, the region between 30 and 40 showed more fluctuation in the *BfEno* open conformation than in the closed conformation. This region is on the substrate binding site and interacts with Mg^{+2} which is the cofactor of the enzyme. The region between 240 and 270 has been defined as the plasminogen binding site and therefore showed high fluctuation. The closed form was more volatile and the open form has a more stable structure when these two graphs were analysed together.

The quality check of *Bf*Eno models was carried out by using ERRAT, RAMPAGE, ProSA, ProQ, QMEAN and Verif3D web-based programs (Colovos and Yeates, 1993; Lovell et al., 2003; Wiederstein and Sippl, 2007; Wallner and Elofsson, 2003; Benkert et al., 2008; Eisenberg et al., 1997) (Table 1).

Overall quality factors of the open and closed conformation models were found as 90.299 and 88.945, respectively. These high values indicated that the models were built with acceptable accuracy (Colovos and Yeates, 1993). Ramachandran plots of the open and closed conformation models were generated by using RAMPAGE program (Lovell et al., 2003). The favored region, allowed region and outlier region of the open conformation model was identified as 92.7%, 6.8% and 0.5%, respectively. Favored region, allowed region and outlier region of the closed conformation model were also found as 92%, 7.3%, 0.7%, respectively. According to these results, the presence of at least 90% of the amino acid residues in the favored region indicated that the protein models were constructed with high accuracy. Z-scores of the open and closed conformation models were determined as − 10.01 and − 9.61, respectively. These models were among the native proteins that were identified experimentally by X-ray crystallography. LGscores of the open and closed conformation models were calculated as 6.062 and 6.095, respectively. These scores were found as higher than 4 and the results indicated that the accuracy of both models were high (Wallner and Elofsson, 2003). QMEAN scores were determined to analyze protein model quality. The scores of the open and closed conformation models were found as − 2.17 and − 1.47, respectively. These negative values showed that the protein models were built with acceptable accuracy (Benkert et al., 2008). The residues of both *Bf*Eno models have averaged $^{3}D–^{1}D$ scores according to the results that were found as 93.56% and

Table 1
Validation results of the open and closed conformation models of *BfEno*.

Web-based program	The open conformation model of *Bf*Eno	The closed conformation model of *Bf*Eno
ERRAT (Overall quality factor)	90.299	88.945
RAMPAGE	Favored region (%): 92.7 Allowed region (%): 6.8 Outlier region (%): 0.5	Favored region (%): 92.0 Allowed region (%): 7.3 Outlier region (%): 0.7
PROSA Z-score	-10.01	-9.61
PROQ LGscore	6.062	6.095
QMEAN	-2.17	-1.47
Verify3D	93.56%	95.47%

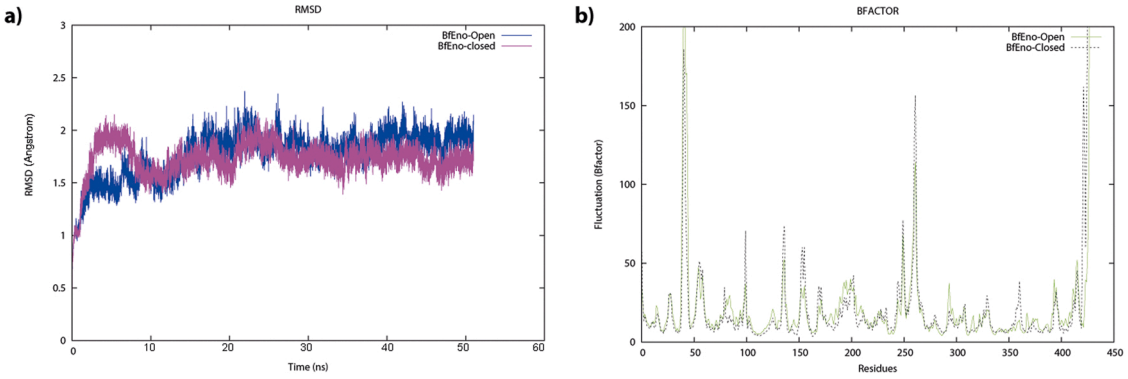

Fig. 5. RMSD and Bfactor plot of the open and closed conformation models for 50 ns MD simulation.

95.47% for the open and closed conformations. These scores were identified above 80% and both models were evaluated as good models by VERIFY3D (Eisenberg et al., 1997). Overall, the open and closed conformation models of *Bf*Eno were built as extremely good models according to the validation results.

3.6. Virtual screening and molecular dynamics simulation

DoGSiteScorer (Fig. 6A) and SiteMap (Fig. 6B) programs predicted the same sites as druggable pockets with the highest scores. One of these sites comprises the active site region (Site1) which has 0.84 drug score and 0.47 simple score, respectively based on DoGSiteScorer method. SiteMap program gave 0.985 Dscore (druggability score) to the same region. Sites with a higher Dscore, more probable to be druggable candidate (Halgren, 2009).

Another druggable site is Site2 located nearby the dimer interface region. DoGSiteScorer values were 0.86 (drug score)/0.44 (simple score), and Dscore were 0.978 for the Site2. Drug-like chemicals were screened for the interaction with the active site pocket (site1) and the dimer interface region (site2). Molecular docking results showed that the XP GScores of docked ligands were ranged from − 9.547 kcal/mol to − 10.867 kcal/mol in the active site (Table 2) (Fig. 7). Contributions of the H-bond and the LipophilicEvdW interactions were larger in the first four ligands, which have lower XP GScores than − 10 kcal/mol. However in site2, chemicals had higher scores (XP GScores from −5.624 to −6.855 kcal/mol) than the active site ligands (Table 3) (Fig. 7). LipophilicEvdW effects made a larger contribution to the overall XP GScore, while the H-bond and the electrostatic interactions were little in most ligands. Glide ligand efficiency (LE = (DockingScore)/N) is a metric that calculates docking score per non-hydrogen atom (N) and generally LE is used to select efficient hits within a pool of chemicals. When compared the ligand efficiency of the small chemicals that occupied the active site and Site2 region, the most promising results were taken from the active site ligands which all calculated LE values beter than the site2 ligands (Table2).

The top 10 ligands were also docked on human alpha, beta and gamma enolases for comparative docking analyses (Supplementary). The XP GScore of the best candidate compound (ZINC91441604) was calculated as − 8.296 kcal/mol, − 8.435 kcal/mol and − 7.007 kcal/mol for human alpha, beta and gamma enolases, respectively. The XP GScore of the best compound ZINC72431523 on the Site2 of *Bf*ENO was found as − 4.703 kcal/mol, − 4.845 kcal/mol, − 4.115 kcal/mol, for human alpha, beta and gamma enolases, respectively. Docking analysis showed that both compounds have better docking scores on *Bf*Eno than the host enolases.

A selected enolase inhibitor candidate ZINC91441604 in the active site and ZINC72431523 ligands in the second druggable site were analysed for their stability. Interactions of *Bf*Eno-ligand complexes were simulated by molecular dynamics methods during 100 ns. The RMSD of the open-enolase complex (AS-91441604) with the ligand ZINC91441604 in the active site and complex (P1S-72431523) with the ligand ZINC72431523 in the site2 was calculated as 1,37 and 1,26 during the 100 ns MD simulation, respectively. It was revealed from Fig. 8 that the RMSD value of AS-91441604 complex increases continuously and changes rapidly after an average of 80 ns. However, the average RMSD value of P1S-72431523 complex varies about 40 ns but it remained constant afterwards (Fig. 8a). It was concluded that P1S-72431523 complex was more stable than AS-91441604 complex when the mean values were compared.

The loop region between 248 and 269 was more fluctuated in both complexes from bfactor plots (Fig. 8b) because of the plasminogen binding region. The loop region 150–160 is known to be the catalytic region and therefore there was just a little more fluctuation in this region of both complexes. However, the ligand was docked to the active site in AS-91441604 complex, there was partially more fluctuation and therefore 235–240 region in the active site cavity showed more fluctuation. The region between 190 and 210 was stable in AS-91441604 complex but differed in P1S-72431523 complex. The ligand was located in the hydrophobic cavity formed by three β-strand structures of 1–40 and the α-helix structure of 390–400 in the complex P1S-72431523 and therefore more fluctuation was revealed due to the jam.

ZINC91441604 ligand docked to the active site of the enolase showed that it remained constant in the active site cavity after the 100 ns. However, the aromatic carbon ring of the ligand moved from the tight position to relax position where it was thermodynamically stable (Fig. 9). The meaning of the sudden change observed in the RMSD plot after 80 ns was assumed to be the time interval from which the ligand got away from its binding site.

The ligand in the P1S-72431523 complex appeared to be completely removed from the first position after simulation (Fig. 10). It was observed that the complex become stable after an average of 40 ns on the RMSD graph. Since the ligand was removed from the system, the enolase remained stable and the mean RMSD value was calculated low. According to these data, it is concluded that ZINC72431523 can not bind to the estimated second druggable binding site or remain stable. So, the second druggable region cannot be the ligand-binding site or the ligand have to be changed. As a result of *in silico* study, the ZINC91441604 has been proposed to bind to the active site of the enzyme and remain stable. The performed *in silico* study is expected to be the basis for further *in vitro* inhibition studies for enolase from *Bacteroides fragilis*.

4. Conclusion

Enolase is a vital enzyme because of catalyzing a reaction in the glycolysis pathway involved in the energy metabolism of *B. fragilis* and also possessing multiple biologic functions. Therefore, *Bf*Eno was

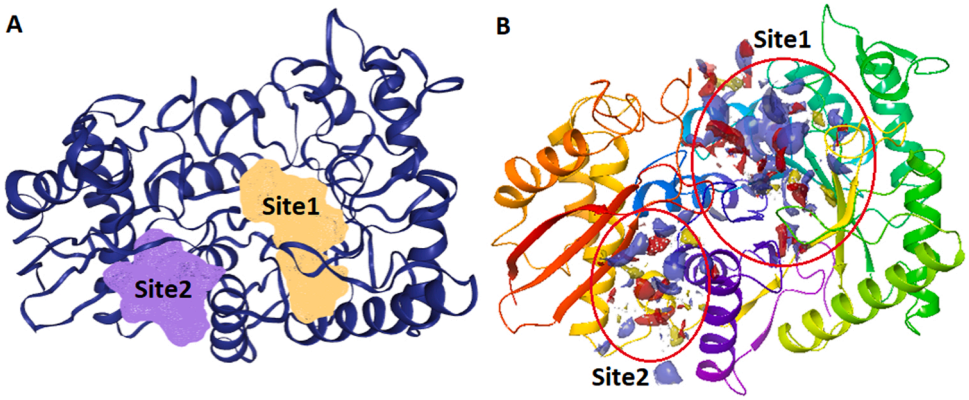

Fig. 6. Targeting sites of the *Bf*Eno. Predicted druggable sites from the DoGSiteScorer program (A) and SiteMap program (B).

Table 2
Virtual screening results of the top 10 ligands interacting with the active site region(Site1) of *BfEno*

Zinc ID	Glide emodela	XP GScorea	XP HBonda	XP Electroa	XP LipophilicEvdWa	Glide ligand efficiency
ZINC91441604	-63.610	-10.867	-3.705	-1.788	-2.321	-0.453
ZINC91782372	-63.808	-10.628	-3.862	-1.696	-2.108	-0.393
ZINC91519206	-68.078	-10.617	-3.836	-1.355	-2.364	-0.424
ZINC91476067	-63.590	-10.399	-3.520	-1.614	-2.199	-0.432
ZINC91915301	-55.155	-9.934	-3.400	-1.543	-1.900	-0.413
ZINC95354192	-61.410	-9.925	-2.616	-2.000	-1.162	-0.421
ZINC91616391	-58.025	-9.906	-3.268	-1.402	-2.192	-0.430
ZINC91876680	-55.770	-9.832	-3.360	-1.597	-1.739	-0.468
ZINC96166576	-57.674	-9.626	-2.150	-1.636	-1.575	-0.341
ZINC91596001	-58.656	-9.547	-2.998	-1.535	-2.060	-0.352

[a] -kcal/mol

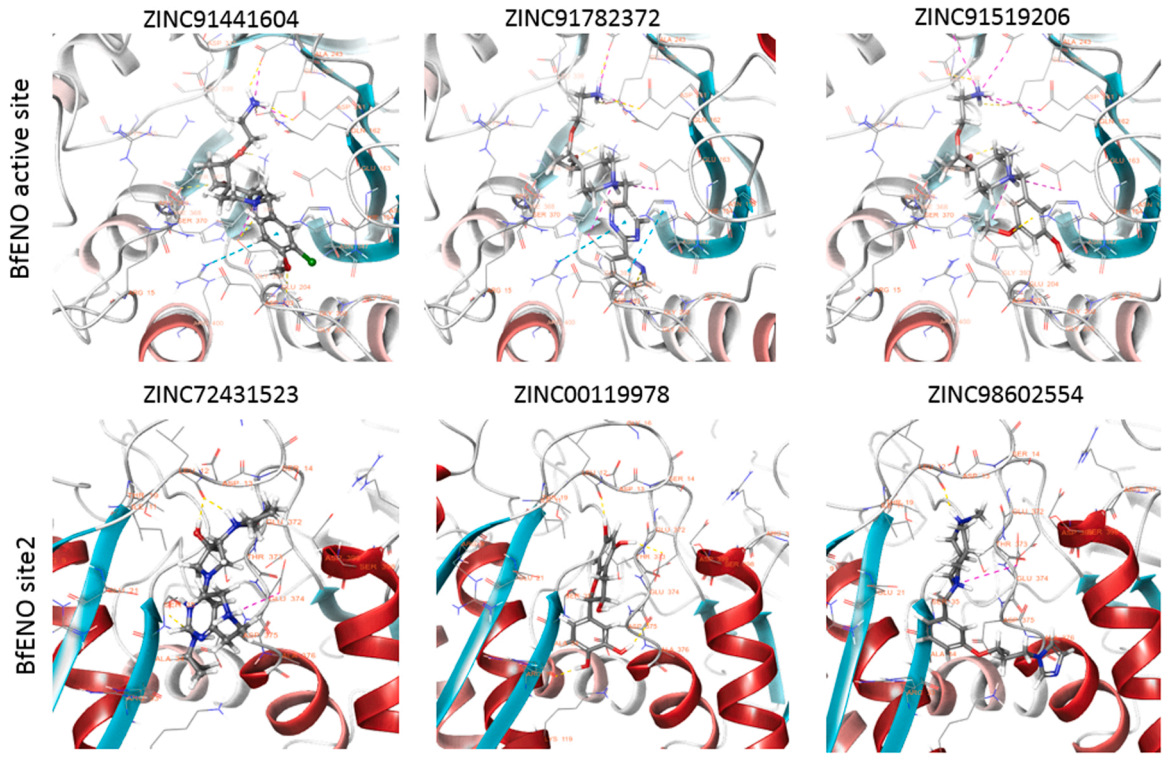

Fig. 7. Interaction of the top three chemicals at the druggable sites.

Table 3
Virtual screening results of the top 10 ligands interacting with the Site2 region of *Bf*Eno.

Zinc ID	Glide emodela	XP GScorea	XP HBonda	XP Electroa	XP LipophilicEvdWa	Glide ligand efficiency
ZINC72431523	-56.629	-6.855	-0.700	-1.243	-3.192	-0.206
ZINC00119978	-46.756	-6.505	-2.528	-1.236	-2.971	-0.310
ZINC98602554	-49.310	-6.372	-0.139	-0.744	-2.942	-0.233
ZINC96137519	-48.560	-6.201	-2.006	-0.805	-2.219	-0.246
ZINC91461604	-42.801	-6.059	-1.814	-0.883	-2.078	-0.249
ZINC98606074	-56.543	-5.972	-3.001	-1.664	-2.931	-0.210
ZINC91679285	-39.409	-5.911	-0.858	-0.777	-1.793	-0.256
ZINC72426939	-56.929	-5.896	-2.869	-1.487	-2.041	-0.243
ZINC91872963	-60.687	-5.801	-1.323	-0.753	-2.922	-0.215
ZINC71715090	-53.165	-5.624	-2.772	-1.350	-1.823	-0.193

[a] -kcal/mol

cloned, expressed and protein was obtained with over 95% purity in this study. Then, kinetic parameters of *Bf*Eno were determined and the enzyme was made available for the further drug analyzes. Within the scope of *in silico* studies; catalytic residues, loop regions, plasminogen binding sites and insertion sites on *Bf*Eno amino acid sequence were identified by using protein sequence analysis. The open and closed conformation models of *Bf*Eno were generated as extremely good models according to the validation results. As a result of virtual screening and MD Simulation studies, the ZINC91441604 has been proposed to bind to the active site of the enzyme and remain stable. The performed *in silico* study is expected to be the basis for *in vitro* inhibition studies for enolase from *Bacteroides fragilis*.

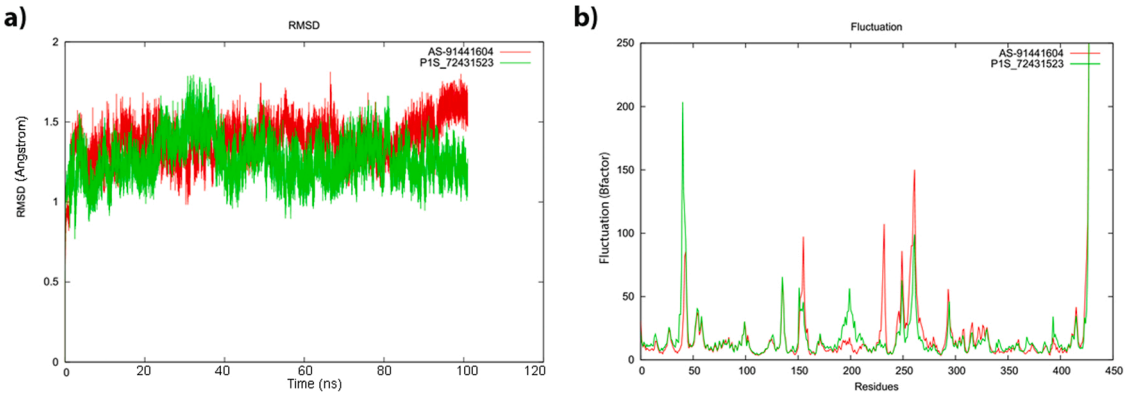

Fig. 8. **a)** RMSD plot of complex for 100 ns (Red: ZINC91441604 in the active site, Green: ZINC72431523 in second druggable site) **b)** Bfactor plot of complex for 100 ns (Red: ZINC91441604 in the active site, Green: ZINC72431523 in second druggable site).

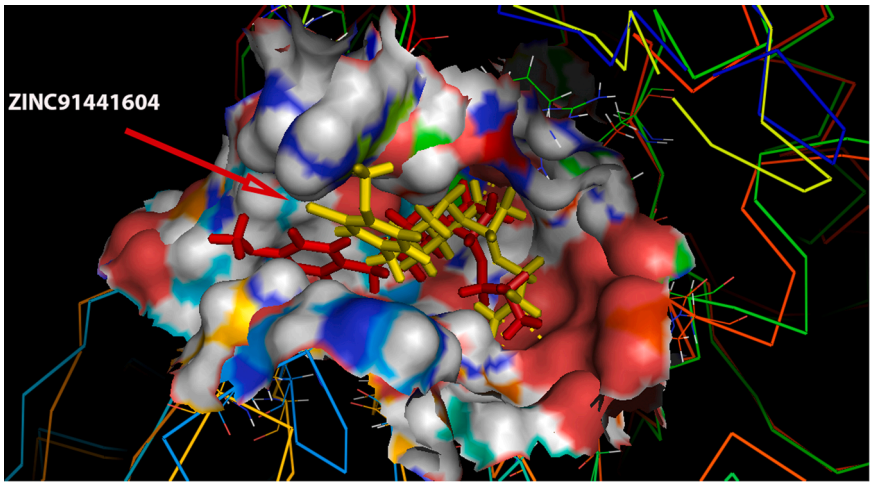

Fig. 9. Superimposition of *Bf*Eno-ZINC91441604 complex in the active site during 100 ns (Yellow: post-simulation, Red: pre-simulation).

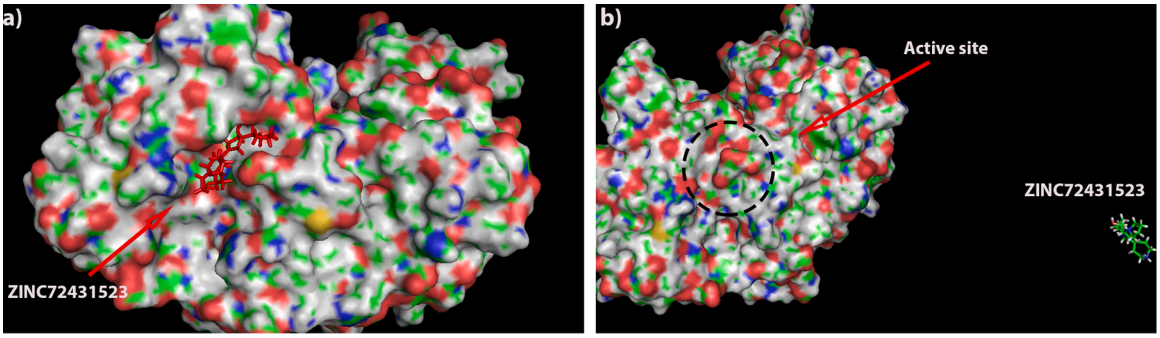

Fig. 10. Superimposition of *Bf*Eno - ZINC72431523 complex in the second druggable site during 100 ns (a: pre-simulation, b: post-simulation).

Funding

This research has been supported by Yıldız Technical University Scientific Research Projects Coordination Department, Turkey, Project Number: FBA-2018–3339.

CRediT authorship contribution statement

Erennur Ugurel: Conceptualization, Investigation, Writing – original draft. **Sinem Kocer:** Methodology, Software, Homology modelling, Writing – review & editing. **Emrah Sariyer:** Methodology, Software, Molecular dynamics simulation, Writing – review & editing. **Ozal Mutlu:** Methodology, Software, Molecular docking, Writing – review & editing. **Tugba Gul Inci:** Investigation, Writing – review & editing. **Osman Mutluhan Ugurel:** Investigation, Writing – review & editing. **Dilek Turgut-Balik:** Project administration, Conceptualization, Writing – review & editing.

Declarations

none.

Conflicts of interest/Competing interests.

The authors declare that they have no conflict of interest.

Acknowledgments

This research has been supported by Yıldız Technical University Scientific Research Projects Coordination Department. Project Number: FBA-2018–3339. The numerical calculations reported in this paper were fully/partially performed at TUBITAK ULAKBIM, High Performance and Grid Computing Center (TRUBA resources).

Ethics approval

Not applicable.

Consent to participate

Not applicable.

Consent for publication

Not applicable.

Availability of data and material

Not applicable.

Code availability

Not applicable.

Appendix A. Supporting information

Supplementary data associated with this article can be found in the online version at doi:10.1016/j.compbiolchem.2022.107658.

References

Ulger Toprak, N., Yagci, A., Gulluoglu, B., Akin, M., Demirkalem, P., Celenk, T., Soyletir, G., 2006. Clin. Microbiol. Infect. 12, 782–786.
Valguarnera, E., Wardenburg, J.B., 2020. J. Mol. Biol. 432, 765–785.
Majid, M., Andleeb, S., 2019. Sci. Rep. 9, 1–15.
Ghotaslou, R., Baghi, H.B., Alizadeh, N., Yekani, M., Arbabi, S., Memar, M.Y., 2018. Infect., Genet. Evol. 64, 156–163.
Urbán, E., Horváth, Z., Sóki, J., Lázár, G., 2015. Anaerobe 31, 55–58.
Nakano, V., Merino, V.R.C., Wexler, H.M., Avila-Campos, M.J., 2011. Clinics 66, 543–547.
Aldridge, K.E., Ashcraft, D., Cambre, K., Pierson, C.L., Jenkins, S.G., Rosenblatt, J.E., 2001. Antimicrob. Agents Chemother. 45, 1238–1243.
Center for Disease Control and Prevention, US Department of Health and Human Services CDC, (2019).
Podolsky, S.H., 2018. Palgrave Commun. 4 (1), 1–8.
Roope, L.S.J., Smith, R.D., Pouwels, K.B., Buchanan, J., Abel, L., Eibich, P., Butler, C.C., Tan, P.S., Walker, A.S., Robotham, J.V., Wordsworth, S., 2019. The challenge of antimicrobial resistance: What economics can contribute. Science 364 (6435), eaau4679. https://doi.org/10.1126/science.aau4679.
Metz, P., Boleij, A., Dutilh, B.E., Tjan, M.J., Wu, S., Pervaiz, M., Hermans, S., Shettigar, A., Sears, C.L., Ritschel, T., 2019. Front. Cell. Infect. Microbiol. 9, 364.
Singh, S., Malik, B.K., Sharma, D.K., 2006. Bioinformation 1, 314.
Sijbrandi, R., Den Blaauwen, T., Tame, J.R., Oudega, B., Luirink, J., Otto, B.R., 2005. Microbes Infect. 7, 9–18.
Díaz-Ramos, À., Roig-Borrellas, A., García-Melero, A., López-Alemany, R., 2012. J. Biomed. Biotechnol. (2012)).
Scientific, Thermo. "aLICator Ligation Independent Cloning and Expression System." (2013).
Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory Press.
Laemmli, U.K., 1970. nature 227, 680–685.
Gasteiger, E., Hoogland, C., Gattiker, A., Wilkins, M.R., Appel, R.D., Bairoch, A., 2005. Proteom. Protoc. Handb. 571–607.
Grimsley, G.R., Pace, C.N., 2004. Chapter 3 Curr. Protoc. Protein Sci.
Cayir, E., Erdemir, A., Ozkan, E., Topuzogullari, M., Bolat, Z.B., Akat, A., Turgut-Balik, D., 2014. Mol. Biotechnol. 56, 689–696.
Yakarsonmez, S., Danis, O., Mutlu, O., Topuzogullari, M., Sariyer, E., Yuce-Dursun, B., Turgut-Balik, D., 2019. Mol. Divers. 1–16.
Pal-Bhowmick, I., Sadagopan, K., Vora, H.K., Sehgal, A., Sharma, S., Jarori, G.K., 2004. Eur. J. Biochem. 271, 4845–4854.
Edgar, R.C., 2004. Nucleic Acids Res. 32, 1792–1797.
Eswar, N., Webb, B., Marti-Renom, M.A., Madhusudhan, M., Eramian, D., Shen, My, Pieper, U., Sali, A., 2006. Curr. Protoc. Bioinforma. 15, 1–5.6, 5.6.
Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. J. Comput. Chem. 25, 1605–1612.
Case, D., Darden, T., Cheatham III, T., Simmerling, C., Wang, J., Duke, R., Luo, R., Walker, R., Zhang, W., Merz, K., 2012. University of California,, San Francisco.
Fletcher, R., Powell, M.J., 1963. Comput. J. 6, 163–168.
Fletcher, R., Reeves, C.M., 1964. Comput. J. 7, 149–154.
Pastor, R.W., Brooks, B.R., Szabo, A., 1988. Mol. Phys. 65, 1409–1419.
Berendsen, H.J., Postma, Jv, van Gunsteren, W.F., DiNola, A., Haak, J., 1984. J. Chem. Phys. 81, 3684–3690.
Ryckaert, J.-P., Ciccotti, G., Berendsen, H.J., 1977. J. Comput. Phys. 23, 327–341.
Essmann, U., Perera, L., Berkowitz, M.L., Darden, T., Lee, H., Pedersen, L.G., 1995. J. Chem. Phys. 103, 8577–8593.
Colovos, C., Yeates, T.O., 1993. Protein Sci. 2, 1511–1519.
Lovell, S.C., Davis, I.W., Arendall III, W.B., De Bakker, P.I., Word, J.M., Prisant, M.G., Richardson, J.S., Richardson, D.C., 2003. Protein.: Struct., Funct., Bioinforma. 50, 437–450.
Wiederstein, M., Sippl, M.J., 2007. Nucleic Acids Res. 35, W407–W410.
Wallner, B., Elofsson, A., 2003. Protein Sci. 12, 1073–1086.
Benkert, P., Tosatto, S.C., Schomburg, D., 2008. Protein.: Struct., Funct., Bioinforma. 71, 261–277.
Eisenberg, D., Lüthy, R., Bowie, J.U., 1997. Methods Enzymol. 277, 396–404.
Volkamer, A., Kuhn, D., Grombacher, T., Rippmann, F., Rarey, M., 2012. J. Chem. Inf. Model. 52, 360–372.
Fährrolfes, R., Bietz, S., Flachsenberg, F., Meyder, A., Nittinger, E., Otto, T., Volkamer, A., Rarey, M., 2017. Nucleic Acids Res. 45, W337–W343.
Sterling, T., Irwin, J.J., 2015. J. Chem. Inf. Model. 55, 2324–2337.
Jakalian, A., Jack, D.B., Bayly, C.I., 2002. J. Comput. Chem. 23, 1623–1641.
Chai, G., Brewer, J.M., Lovelace, L.L., Aoki, T., Minor, W., Lebioda, L., 2004. J. Mol. Biol. 341, 1015–1021.
Waterhouse, A., Bertoni, M., Bienert, S., Studer, G., Tauriello, G., Gumienny, R., Heer, F. T., de Beer, T.A.P., Rempfer, C., Bordoli, L., 2018. Nucleic Acids Res. 46, W296–W303.
Larsen, T.M., Wedekind, J.E., Rayment, I., Reed, G.H., 1996. Biochemistry 35, 4349–4358.
Zhang, E., Brewer, J.M., Minor, W., Carreira, L.A., Lebioda, L., 1997. Biochemistry 36, 12526–12534.
Brown, C.K., Kuhlman, P.L., Mattingly, S., Slates, K., Calie, P.J., Farrar, W.W., 1998. J. Protein Chem. 17, 855–866.
Bergmann, S., Wild, D., Diekmann, O., Frank, R., Bracht, D., Chhatwal, G.S., Hammerschmidt, S., 2003. Mol. Microbiol. 49, 411–423.
Yakarsonmez, S., Cayir, E., Mutlu, O., Nural, B., Sariyer, E., Topuzogullari, M., Milward, M., Cooper, P., Erdemir, A., Turgut-Balik, D., 2016. Appl. Biochem. Microbiol. 52, 23–30.
Cork, A.J., Ericsson, D.J., Law, R.H., Casey, L.W., Valkov, E., Bertozzi, C., Stamp, A., Jovcevski, B., Aquilina, J.A., Whisstock, J.C., 2015. PLoS One 10, e0121764.
Wu, Y., Wang, C., Lin, S., Wu, M., Han, L., Tian, C., Zhang, X., Zang, J., 2015. Acta Crystallogr. Sect. D: Biol. Crystallogr. 71, 2457–2470.
Maiti, R., Van Domselaar, G.H., Zhang, H., Wishart, D.S., 2004. Nucleic Acids Res. 32, W590–W594.
Halgren, T.A., 2009. J. Chem. Inf. Model. 49, 377–389.