

# Molecular dynamics simulation (MDS) analysis of *Vibrio cholerae* ToxT virulence factor complexed with docked potential inhibitors

Ibrahim Torktaz, Ali Najafi\*, Reza Golmohamadi, Sorour Hassani

<sup>1</sup>Molecular Biology Research Center, Systems Biology and Poisonings institute, Baqiyatallah University of Medical Sciences, Tehran, IRAN; Ali Najafi; E-mail: najafi74@bmsu.ac.ir; najafi74@yahoo.com; Tel: 98-21-82482569; \*Corresponding author

Received February 11, 2018; Revised February 28, 2018; Accepted February 28, 2018; Published March 31, 2018

doi:10.6026/97320630014101

## Abstract:

The ToxT transcription factor mediates the transcription of the two major virulence factors in *Vibrio cholerae*. It has a DNA binding domain made of  $\alpha 4$ ,  $\alpha 5$ ,  $\alpha 6$ ,  $\alpha 7$ ,  $\alpha 8$ ,  $\alpha 9$  and  $\alpha 10$  helices that is responsible for the transcription of virulence genes. Therefore, it is of interest to screen ToxT against the ZINC ligand database containing data for a million compounds. The QSAR model identified 40 top hits for ToxT. Two target protein complexes with ligands Lig N1 and Lig N2 with high score were selected for molecular dynamics simulation. Simulation data shows that ligands are stable in the DNA binding domain of ToxT. Moreover, Lig N1 and Lig N2 passed pharmacological as well as ADME filters along with g-mmpbsa analysis with binding affinity of -199.831 kJ/mol for Lig N1 and -286.951 kJ/mol for Lig N2. Moreover, no Lipinski and PhysChem violations were identified. It is further observed that these compounds were not inhibitors of P-glycoprotein, CYP450 and renal organic cation transporters. The LD50 of 2.5804 mol/kg for Lig N1 and 2.7788 mol/kg for Lig N2 was noted with acceptable toxicity index.

**Keywords:** ToxT, *Vibrio cholerae*, QSAR, g-mmpbsa, molecular dynamics simulation

## Background:

*V. cholerae* is a gram-negative and facultative anaerobic pathogen that principally exists in aquatic locales. This is the causal agent of severe diarrhea by colonizing on the small intestine and secreting cholera toxin (CT). CT is a ribosylating toxin, which is responsible for the abundant diarrhea associated with cholerae. Secretion of CT by *V. cholerae* leads to a rise in cAMP levels in the host cells [1]. The elevation in cAMP concentrations results in the reduction adsorption of sodium, secretion of chloride ions into the lumen for naturalism and promotion of osmotic pressure, leading to voluminous secretion of water and electrolytes [2]. Toxin production occurs upon attachment of the bacteria onto the intestinal epithelium. Therefore, it is of interest to understand the regulatory mechanism of CT expression. Epidemic serotypes of *V. cholerae* are classical and El Tor where CT transcription is regulated by ToxT [3]. ToxT activates the transcription of *ctxAB* operon that encodes two CT subunits and the transcription of TCP operon encoding toxin-coregulated pilus genes [4]. Moreover, ToxT is an AraC family member, which includes the

binding domain into the DNA with helix-turn-helix motifs and activates the expression of a number of virulence genes including TCP and CT [5]. Transcription of ToxT is activated with four inner-membrane proteins namely *toxR*, *toxS*, *tcpP*, and *tcpH*. Two chromosomally encoded regulators; AphB and AphA activate the transcription of *tcpPH*. Yamasaki *et al.* have shown that red chili and one of its active compounds capsaicin inhibit CT production without affecting bacterial growth [6]. Hung *et al.* revealed that virstatin inhibited ToxT activity when ToxT was expressed under the control of a heterologous pBAD promoter [7]. Plecha *et al.* showed that unsaturated fatty acids inhibited the DNA binding of ToxT [8]. Two available commercial oral cholera vaccines, *ShanChol* and *Dukoral* are available. They provide an impound protection of >50% for at least two years in indigenous population. However, they are not currently licensed in the united states [9]. Many *V. cholerae* strains have become restraint to a range of antimicrobial agents including tetracycline and ampicillin [10]. Thus, there is an essential need to develop new drugs against cholera. Shakhnovich *et al.* reported that virstain

affects the ToxT activity of the *ctx* promoter by inhibition of ToxT dimerization [11]. Minto *et al.* revealed that malonate has the potential to inhibit the expression of disease-causing genes in *V. cholerae* through *ToxT* inhibition [12]. Therefore, it is of interest to screen ToxT against the ZINC ligand database containing data for a million compounds.

#### Methodology:

##### Ligand screening:

The 3D crystallography structure of ToxT from *V. Cholera* (PDB ID: 3GBG) was selected as the protein target in virtual screening model [13]. Molegro Virtual Docker (MVD) v 6.0 was used to calculate dock score and evaluate conformers. The DNA binding domain of ToxT identified and then the spherical virtual screening coordinate was located to this region. Nearly, 40000 drugs-like ligands were derived from subset 3\_p0.1 from standard in stock drug like category by ZINC database and were used for virtual screening.

ZINC is a free database containing various compounds for docking based screening [14]. Docking parameters were set as follow: grid resolution of 0.3 Å for all docking simulation, a maximum number of 1500 iterations and each of 10 independent runs were enforced on single populations of 50 individuals. An 18 Å radius was exactly set on the coordinates of X: 40, Y: 55 and Z: 36 to cover the entire DNA binding site including  $\alpha 4$ ,  $\alpha 5$ ,  $\alpha 6$ ,  $\alpha 7$ ,  $\alpha 8$ ,  $\alpha 9$  and  $\alpha 10$ . Non-polar hydrogen atoms were removed from the receptor structure and their partial charges were added to the corresponding carbon atoms. Moreover, flexible torsions of ligands were identified by MVD. MolDock based on guided differential evolution and PLANTS scores performed docking simulations. The best conformations were based on the lowest binding energy.

##### Pharmacophore model prediction and QSAR modeling:

We selected the top 40 successive hits for structural alignment in order to find structurally similar ligands. The structure of hits that retrieved from virtual screening was largely diverse and five ligands with most similar structures were selected for pharmacophore design. Mastero 10.2 from Schrodinger simulation package was used for developing pharmacophore models following QSAR. The extracted common pharmacophore model was used for screening a local database containing 1,000,000 lead like small molecules.

##### MD simulations:

GROMACS MD package 4.6.5 [15] was used to simulate the ToxT-ligand(s) interactions in a dynamic environment. The PRODRG server that is based on the GROMOS force field was employed to generate topology files for ligands [16]. GROMOS96 force field with a modified version of the 53a6.ff force field was utilized throughout the simulation study. The complexes of ToxT-ligand(s) were merged in a cubic-shaped box with the minimum distance of one nm between the protein surface and

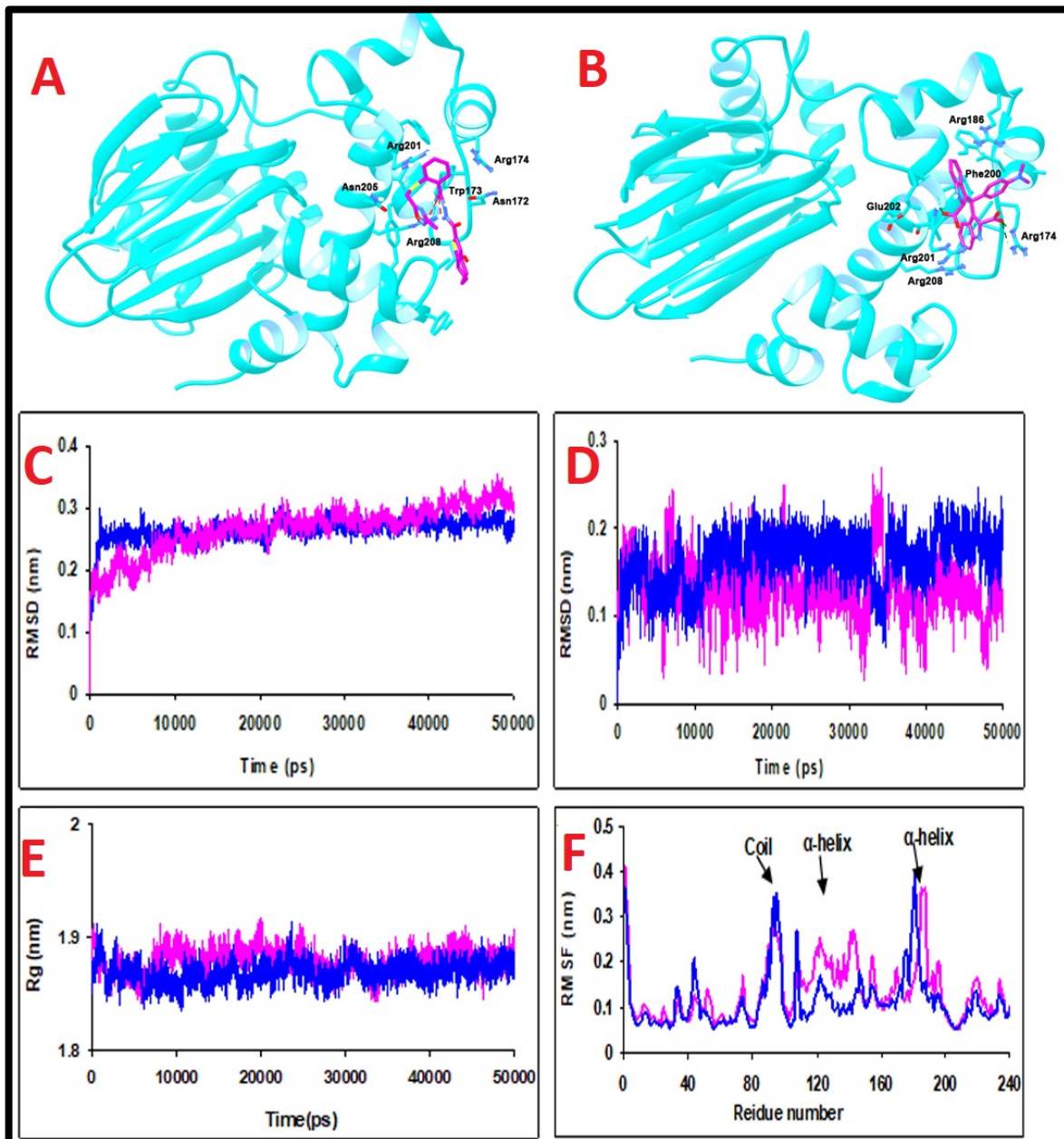
the box walls, followed by solvation in a simple point charge water molecules. Periodic boundary conditions were assigned the system was neutralized for energy minimization employing steepest descent algorithm with a tolerance of 1000 kJ/mol/nm. Equilibrations with harmonic restraints on the coordinates of the complexes atoms were performed after convergence. Parinello-Rahman barostat and modified Berendsen thermostat were applied to keep the pressure and temperature constant at 1 bar and 300 K, respectively. Particle-Mesh Ewald method was used to calculate Long-range electrostatic interactions. The MD runs were carried out in the NPT ensemble (50 ns) for each system. The binding affinities were calculated by *g\_mmpbsa* package [17].

#### Result and discussion

##### Molecular docking study:

The top 46 successive hits with the most efficient binding affinity are shown in Figure 2. Ligands that indicated Lipinski violation were excluded from the further study. The binding affinity of top 46 ligands was mostly in the same range but the structures indicated large variations. We used lig No# 1, 2, 3, 5 and 41 for a common pharmacophore model and QSAR prediction taking into consideration molecular similarity. The model was used for screening the ZINC database containing 1,000,000 lead like small molecules and the top hits were retrieved for a simulated docking study in an environment containing water molecules and neutralizing ions. Ligands, Lig N1 and Lig N2 showed the lowest binding energy with the receptor. The second docking screen with QSAR matched hits for new inhibitors of the regulatory protein of pathogenic operons in *V. cholerae* for these two ligands. Figure 1A shows Lig N1 interacts with ToxT by two strong H-bonds (receptor as a donor) with Arg208 (2.59 and 3.12Å with the energy of -2.5 and -2.45 kcal/mol, respectively). Moreover, it includes steric interactions with surrounded residues of Arg174, Trp173, Asn172, Asn205, Gly231, and Thr170 and  $\pi$ - $\pi$  interaction with Phe200.

The residues of Arg201 and Arg174 are involved in hydrogen bonds with Lig N2. One H-bond with Arg201 (2.91 with the energy of -2.07 kcal/mol) and two H-bonds with Arg174 (2.9454Å and 2.54Å with energy -2.5 and -0.89 kcal/mol, respectively) are observed. Lig N2 includes electrostatic interactions with Glu202 and Arg186, van der Waals interactions with Trp173, Phe200, Arg208, Ile204 and  $\pi$ - $\pi$  interaction with Trp175 (Figure 1B). Arg residue was the important residue involved in hydrogen bonding of receptor with ligands. The resulted data with scoring results showed that Lig N2 has more affinity to the receptor. Moreover, the RMSD of LigN1-receptor and LigN2-receptor is depicted in Figure 1C and Figure 1D, respectively. It showed that Lig N1 is more stable than Lig N2 in the DNA binding domain of ToxT. In addition, The  $R_g$  of both ligands (Figure 1E) indicated no significant drift. RMSF data (Figure 1F) indicated large fluctuations of segments in the coil regions (95-98 and 183-185) and  $\alpha$ -helixes (90-94 and 185-189).



**Figure 1:** (A) Lig N1 interacts with ToxT by two strong H-bonds (receptor as donor) with Arg208 (2.59 and 3.12Å with energy of -2.5 and -2.45 kcal/mol, respectively). Moreover, it includes steric interactions with surrounded residues of Arg174, Trp173, Asn172, Asn205, Gly231, and Thr170 and  $\pi$ - $\pi$  interaction with Phe200. (B) Lig N2 involved in hydrogen bindings with Arg201 and Arg174. one H-bonds with Arg201 (2.91 with energy of -2.07 kcal/mol) and two H-bond with Arg174 (2.9454Å and 2.54Å with energy -2.5 and -0.89 kcal/mol, respectively). Moreover, Lig N2 include electrostatic interactions with Glu202 and Arg186, van der Waals interactions with Trp173, Phe200, Arg208, Ile204 and  $\pi$ - $\pi$  interaction with Trp175. (C) RMSD value increases for ToxT -Lig N1 complex until; 0.28 nm, stayed around this value for; four ns, increases for a short while, and then stabilizes at; 0.32 nm. (D) RMSD value for ToxT-Lig N2 indicated that although the ligand is stable in the DNA Binding domain of ToxT, it has variation in RMSD, which means that its conformational alteration of the complex during MD simulation is much. (E) Rg of two protein-ligand complexes fluctuates around a stable value of 1.88 to 1.9 nm and does not show any significant drift. This low and nearly constant value of Rg approves good conformational stability and folding of two systems. (F) RMSF data indicated large fluctuations of segments belonging to the coils (95-98 and 183-185) and  $\alpha$ -helices (90-94 and 185-189). While these features are common to both complexes, the ToxT structure exhibited a more fluctuations in segments of coil (125-127 and 146-146) and  $\alpha$ -helix (128-145) when it binds to Lig N1. The RMSF values for residues (43-46), (94-97), and (173-183) are significant in ToxT-Lig N2 complex.



Ligand No#	Smiles
1	<chem>O[C@@H](CSc1nnc(-c2cccc2)n1Cc1cccc1)c1ccc2OCC(=O)Nc2c1</chem>
2	<chem>C[C@H](CN1CCc2c(C1)c(nn2Cc1cccc1)C(=O)N1CCCC1)Cc1ccc2OCOc2c1</chem>
3	<chem>O[C@@H](C1=C(O)C(=O)N([C@@H]1c1cccc1)c1nnc(SCc2ccc(F)cc2)s1)c1cccc1</chem>
4	<chem>CCc1nc(no1)-c1cc(c(C)s1)S(=O)(=O)N1CCC[C@H](C1)C(=O)Nc1ccc(C)o1</chem>
5	<chem>Fc1cccc(c1)-c1noc(n1)-c1cccn2c1nn(CC(=O)Nc1cc(Cl)cc(Cl)c1)c2=O</chem>
6	<chem>Cc1nn(Cc2ccc(F)cc2)c2sc(cc12)C(=O)N[N]1=C(O)N[C@](C)(C1=O)c1cccc1</chem>
7	<chem>O=S(=O)(N1CCCCC1)c1cccc(c1)-c1nnc(SCc2nnc(o2)-c2cccs2)o1</chem>
8	<chem>Cc1nc(SCc2nnc(o2)-c2c(C)onc2-c2cccc2)n(C2CCCC2)c1C</chem>
9	<chem>COc1ccc(cc1)[C@H]1[C@H](O)[C@H](NN1C(=O)CN[C@H](C)c1cccs1)c1c[nH]c2cccc12</chem>
10	<chem>Cc1c(nc2c(C)cccn12)[C@@H](O)C1=C(O)C(=O)N(Cc2ccc2)[C@@H]1c1cccc(c1)[N+][[O-]]=O</chem>
11	<chem>Cc1cc(C)n(CCNc2cn(nc2-c2ccc3OCOc3c2)-c2cccc2)n1</chem>
12	<chem>Cc1ccc(CNC(=O)CCc2nc(N3CCCC3)c3c(noc3n2)-c2ccc(F)cc2)o1</chem>
13	<chem>NCc1c2CCCCc2sc1NC(=O)CSc1nnc(-c2ccncc2)n1Cc1cccc1</chem>
14	<chem>C[C@H](Sc1nnc(CN2CCCC2)n1Cc1cccc1)c1nc2sc(C)C(c2c(=O)[nH]1</chem>
15	<chem>OC[C@H](NC(=O)c1cn(Cc2cccc2)c2cccc12)c1cccc1</chem>
16	<chem>CCOc1cccc(c1)[C@H]1N(Cc2cccn2)C(=O)C(O)=C1[C@H](O)c1nc2c(C)cccn2c1C</chem>
17	<chem>CO[C@@H](O)[C@@H]1C[C@H](CN1C1Cc2cccc2C1)NC(=O)c1cccc(Cn2nc(C)cc2C)c1</chem>
18	<chem>Cc1ccc(SCc2c(nnn2-c2nonc2N)C(=O)N\N=C\c2cccc(Cl)c2)cc1</chem>
19	<chem>Cc1c(nc2c(C)cccn12)[C@@H](O)C1=C(O)C(=O)N(Cc2ccc2)[C@@H]1c1ccc(C)cc1</chem>
20	<chem>Cc1c(nc2c(C)cccn12)[C@H](O)C1=C(O)C(=O)N(Cc2ccc2)[C@@H]1c1cccc1</chem>
21	<chem>Cc1ccc(CNC(=O)c2cc(NC(=O)CC3CCCC3)cc3nc([C@H]4CCCCO4)n(C)c23)cc1</chem>
22	<chem>CCO[C@H](O)c1[nH]c2ccc(OC)cc2c1NC(=O)CN1C[C@]2(C)C[C@H]1CC(C)(C)C2</chem>
23	<chem>Cc1onc(c1C(=O)N[C@@H](S)NNC(=O)c1cc([nH]n1)-c1cccc1)-c1cccc1</chem>
24	<chem>CSc1nn(CC(=O)NCc2ccc(C)cc2)c(N)c1-c1nc(no1)-c1cccc1</chem>
25	<chem>Cc1c(nc2c(C)cccn12)[C@H](O)C1=C(O)C(=O)N(Cc2cccn2)[C@H]1c1cccc(O)c1</chem>
26	<chem>CN1[CH][CH]N[C@@H]1[C@H]1[C@H]2N[CH]N[C@@H]2CCN1Cc1cn(CC(O)=N)c2cccc12</chem>
27	<chem>C[C@@]12C[C@@H](CC(C)(C)C1)N(CN1C(=O)C(=N)NC(=O)c3ccc3)c3cccc13)C2</chem>
28	<chem>Cc1cccc(c1)-c1nnc(SCC(=O)N2CCN(CC2)c2nccn2)n1Cc1cccc1</chem>
29	<chem>Cc1ccc(CN2C[C@H](CC2=O)C(=O)NNC(=O)CSc2nc3cccc3s2)cc1</chem>
30	<chem>OC(O)[C@@H](N1CCN(CC1)c1nccn1)c1cn(Cc2cccc2)c2ccc(cc12)C(O)O</chem>
31	<chem>Cc1sc(cc1S(=O)(=O)NCc1ccc2OCOc2c1)-c1nc(no1)-c1cccc(F)c1</chem>
32	<chem>OC1CCN(CC1)C(=O)c1nn(Cc2cccc2)c2CCN(Cc12)C(=O)c1cc2cccc2[nH]1</chem>
33	<chem>Cc1cc2nc(SCC(=O)NCc3ccc4OCOc4c3)n(Cc3ccc3)c(=O)c2[nH]1</chem>
34	<chem>Cc1c(nc2c(C)cccn12)[C@H](O)C1=C(O)C(=O)N(Cc2cccn2)[C@@H]1c1cccc1F</chem>
35	<chem>Fc1cccc(c1)-c1noc(n1)-c1cccn2c1nn(CC(=O)N[C@H]1CCc3cccc13)c2=O</chem>
36	<chem>O=C(NCc1nc(no1)-c1cccc1)[C@@H]1CCCN1S(=O)(=O)c1ccc2cccc2c1</chem>
37	<chem>O[C@H](CSc1nnc(CN2CCCC2)n1-c1cccc1)c1c[nH]c2cccc12</chem>
38	<chem>CCOc1ccc2nc(sc2c1)N1[C@H](C([C@@H](O)c2cccs2)=C(O)C1=O)c1ccc(C)o1</chem>
39	<chem>CCOc1ccc2nc(sc2c1)N1[C@H](C([C@@H](O)c2cccs2)=C(O)C1=O)c1ccc(O)c1</chem>
40	<chem>Cc1nsc1Cn1nc(C(=O)NC2CC2)c2CN(Cc12)C(=O)Cn1ncc2cccc12</chem>
41	<chem>OC(O)[C@H](CC1=c2cccc2=NC1)NS(=O)(=O)c1ccc2[nH]c(=O)c3cccc1c23</chem>
42	<chem>CC1=C[C@H](CN2CCC(CC2)n2cc(CN3CCC4(CCCC4)C3)nn2)NN1</chem>
43	<chem>OC(O)[C@@H](CC1=c2cccc2=NC1)NS(=O)(=O)c1ccc2[nH]c(=O)c3cccc1c23</chem>
44	<chem>CCc1sc(cc1C)C(=O)NNC(=O)[C@H](NC(=O)c1cccc1)c1cccc1</chem>
45	<chem>Fc1cccc(NC(=O)CSc2nnc(Cc3cc(=O)[nH]c(=O)[nH]3)n2-c2ccc(Cl)cc2)c1</chem>
46	<chem>O[C@@H](CSc1nnc(-c2cccc2)n1Cc1cccc1)c1ccc2OCC(=O)Nc2c1</chem>

**Figure 2:** The results of top successive hits retrieved from virtual screening among 40,000 drug like small molecules from 3\_p0.1 subset by a drug like category of Zinc database. MolDock and PLANTS are computational scores indicating binding affinity and are not relevant to chemical units.

### Result of Drug-likeness and toxicity evaluation of the ligands:

We checked the drug-likeness value of the top two virtual screening hits by FAF drugs [18] and admetSAR [19]. The results from FAF-Drug and MolSoft indicated that Lig N1 was fitted into Lipinski acceptable area. Moreover, the oral absorption estimation indicated that Lig N1 has acceptable hydrogen bond

donor/acceptor, rotatable bonds, molecular weight and hydrophobicity. The ligand encountered RO5 of drug-likeness without infringement. Lig N2 showed no Lipinski violation. In addition, the oral absorption estimation indicated that hydrogen bond donor/acceptor, rotatable bonds, molecular weight, and hydrophobicity are in an acceptable range. The results of

AdmetSAR revealed that Lig N1 is non-inhibitor and non-substrate of P-glycoprotein, CYP450 and renal organic cation transporter. It was observed that these hits are not carcinogen and not toxic in ADMES assay. The LD50 of rat acute toxicity value for Lig N1 was predicted equally to 2.5804 mol/kg, which means that it is not toxic. Lig N2 was predicted as non-inhibitor of P-glycoprotein and renal organic cation transporter but it has predicted to be the substrate of P-glycoprotein with the probability of 0.6381. Moreover, it was predicted to be non-inhibitor of CYP450 and not any carcinogenicity was predicted for it. Interestingly, it was predicted to be toxic in AMES assay. The LD50 of rat acute toxicity value for Lig N2 was predicted as 2.7788 mol/kg that puts Lig N2 in non-toxic chemicals category. From these two stages of docking screens and QSAR screening, Lig N1 and Lig N2 have selected to be taken into MD simulations as the next and final step in this evaluation process.

#### *g-mmpbsa analysis:*

*g-mmpbsa* package was used to calculate the binding affinity of ligands. By calculating potential energy in the vacuum, van der Waals, electrostatic interactions and net non-bonded potential energy between the protein and ligands were calculated. An average binding energy equal to -199.831 +/- 37.261 kJ/mol was achieved for Lig N1 and -286.951 +/- 48.611 kJ/mol for Lig N2 respectively. The binding affinities indicate that by attaching lig N1 and lig N2 to the DNA binding domain of ToxT, with the most probability, the protein cannot attach to the target DNA sequence.

#### Conclusion:

The molecular docking analysis of Lig N1 & N2 with the DNA binding domain of ToxT is shown in this report. The ligands docked with ToxT complex were stable over 50 ns molecular dynamics simulation. It is also noted that no violation of Lipinski rules and PhyChem were observed in the simulated target-ligand complexes.

#### Acknowledgment:

We appreciate Baqiyatallah University of Medical Sciences and Iranian national institute of elites for supporting this project.

#### Conflict of interest:

There is no conflict of interest.

#### References:

- [1] Ramamurthy T *et al.* *Microbes Infect.* 2003, **5**:4 [PMID: 12706446].
- [2] Zahid MS *et al.* *Appl Environ Microbiol.* 2010, **76**:13 [PMID: 20472740].
- [3] DiRita VJ *et al.* *Proc Natl Acad Sci U S A.* 1996, **93**:15 [PMID: 8755590].
- [4] Hsiao A *et al.* *Infect Immun.* 2009, **77**:4 [PMID: 19168737].
- [5] Jang J *et al.* *Microbiology.* 2011, **157**:5 [PMID: 21330435].
- [6] Yamasaki S *et al.* *Indian J Med Res.* 2011, **133** [PMID: 21415500].
- [7] Hung DT *et al.* *Science.* 2005, **310**:5748 [PMID: 16223984].
- [8] Plecha SC & Withey JH. *J Bacteriol.* 2015, **197**:10 [PMID: 25733618].
- [9] Ali M *et al.* *Lancet Infect Dis.* 2017. [PMID: 28161570].
- [10] Dubief B *et al.* *Fish Shellfish Immunol.* 2017, **60** [PMID: 27836724].
- [11] Shakhnovich EA *et al.* *Mol Microbiol.* 2007, **66**:6 [PMID: 17986190].
- [12] Minato Y *et al.* *PLoS One.* 2013, **8**:5 [PMID: 23675480].
- [13] Lowden MJ *et al.* *Proc Natl Acad Sci U S A.* 2010, **107**:7 [PMID: 20133655].
- [14] Irwin JJ & Shoichet BK. *J Chem Inf Model.* 2005, **45**:1 [PMID: 15667143].
- [15] Hess B *et al.* *J Chem Theory Comput.* 2008, **4**:3 [PMID: 26620784].
- [16] Schuttelkopf AW & van Aalten DM. *Acta Crystallogr D Biol Crystallogr.* 2004, **60**:8 [PMID: 15272157].
- [17] Kumari R *et al.* *J Chem Inf Model.* 2014, **54**:7 [PMID: 24850022].
- [18] Miteva MA *et al.* *Nucleic Acids Res.* 2006, **34** [PMID: 16845110].
- [19] Cheng F *et al.* *J Chem Inf Model.* 2012, **52**:11 [PMID: 23092397].

Edited by P Kanguane

Citation: Torktaz *et al.* *Bioinformation* 14(3): 101-105 (2018)

**License statement:** This is an Open Access article which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited. This is distributed under the terms of the Creative Commons Attribution License