Short Communication

Virtual screening studies to identify novel inhibitors for Sigma F protein of *Mycobacterium tuberculosis*

CrossMark

Kiran Kumar Mustyala [a], *Vasavi Malkhed* [a], *Venkataramana Reddy Chittireddy* [b], *Uma Vuruputuri* [a,*]

[a] *Molecular Modelling Research Laboratory, Department of Chemistry, University College of Science, Osmania University, Hyderabad 500 007, Telangana, India*
[b] *Department of Chemistry, JNTUH-CEH, Jawaharlal Nehru Technological University, Kukatpally, Hyderabad 500 085, Telangana, India*

ARTICLE INFO

ABSTRACT

Tuberculosis (TB) is one of the oldest threats to public health. TB is caused by the pathogen *Mycobacterium tuberculosis* (MTB). The Sigma factors are essential for the survival of MTB. The Sigma factor Sigma F (SigF) regulates genes expression under stress conditions. The SigF binds to RNA polymerase and forms a holoenzyme, which initiates the transcription of various genes. The Usfx, an anti-SigF protein, binds to SigF and alters the transcription initiation and gene expression. In the present work, virtual screening studies are taken up to identify the interactions between SigF and small molecular inhibitors which can inhibit the formation of holoenzyme. The studies reveal that ARG 104 and ARG 224 amino acid residues of SigF protein are forming important binding interactions with the ligands. The *in silico* ADME properties for the ligand data set are calculated to check the druggability of the molecules.

Introduction

The communicable disease tuberculosis (TB) is caused by the oldest human pathogen *Mycobacterium tuberculosis* (MTB). In the year 2013 nearly 1.5 million people died and 9.0 million new TB cases were reported [1]. MTB can survive in the host organism against the change in environmental conditions due to complex gene expression, which is controlled by specific Sigma factors [2,3]. The Sigma factors, a regulatory family of proteins, play a key role in the immunopathology of MTB [4,5]. The MTB encodes 13 Sigma factors, among which Sigma factor F (SigF) protein regulates the SigB and SigC factor protein expression, which are important in virulence [6–8]. The SigF is involved in direct and indirect regulation of many genes, which are essential for cell wall protein synthesis and survival of MTB in the host system [9,10]. Geiman et al. reported that 187 genes in stationary phase and 277 genes in late stationary phase show less expression in the SigF-deficient MTB [11]. The Usfx, an anti-Sigma factor, negatively regulates the activity of SigF, in response to a variety of

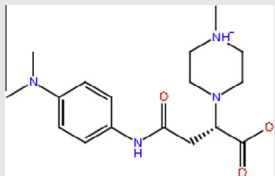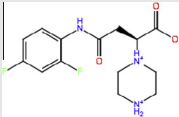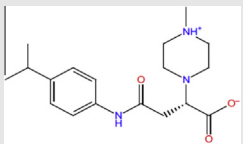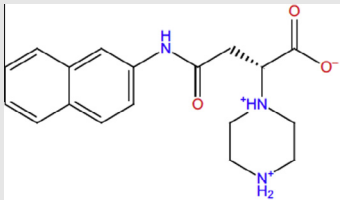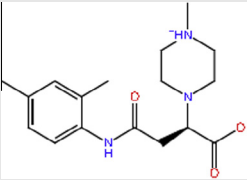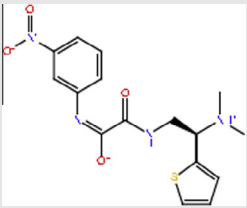**Table 1 – The structures and Hydrogen bonding interactions of Sigma lifechem databank ligand molecules with SigF protein, prioritized with best docking score and energy are represented.**

S. No	Structure	Glide score	Glide energy	Intermolecular interactions
M1	(2S)-4-{[4-(Dimethylamino)phenyl]amino}-2-(4-methyl-1-piperazinyl)-4-oxobutanoic acid	−9.39	−40.90	ILE54:O–M1:H50 ARG104:HE–M1:O12 ARG104:HH21–M1:O12 ARG104:HH21–M1:O13 ARG224:HH22–M1:O19
M2	(2S)-4-[(2,4-Difluorophenyl)amino]-4-oxo-2-(1-piperazinediiumyl)butanoate	−9.32	−38.46	ARG104:HE–M2:O12 ARG104:HH21–M2:N1 ARG104:HH21–M2:O12 ARG224:HH22–M2:O18
M3	(2S)-4-[(4-Isopropylphenyl)amino]-2-(4-methyl-1-piperazinyl)-4-oxobutanoic acid	−8.94	−36.34	ARG224:HH22–M3:O12 ARG224:HH12–M3:O12 ARG104:HH21–M3:O18 ARG57:HH12–M3:O11
M4	(2R)-4-(2-Naphthylamino)-4-oxo-2-(1-piperazinediiumyl)butanoate	−8.52	−44.35	ARG104:HH21–M4:O12 ARG104:HE–M4:O15 ARG57:HH12–M4:O11 GLU59:OE2–M4:H25 GLU59:OE1–M4:H25 Pi–cation M4–ARG57:NH2

Mol	Structure	Docking	MM-GBSA	Interactions
M5	(2R)-4-[(2,4-Dimethylphenyl)amino]-2-(4-methyl-1-piperazinyl)-4-oxobutanoic acid	−8.24	−33.89	GLU59:OE1–M5:H24 GLU59:OE2–M5:H24 ALA55:O–M5:H40 ARG104:HE–M5:O11 ARG104:HH21–M5:O11 ARG224:HH22–M5:O16 Pi–cation ARG224:NH1–M5 Pi–sigma PHE225:HE2–M5
M6	N-[2-(Dimethylamino)-2-(2-thienyl)ethyl]-N′-(3-nitrophenyl)ethanediamide	−8.08	−36.46	ARG224:HH22–M6:O16 ARG224:HH22–M6:N7 ARG224:HH12–M6:O16 ARG104:HH21–M6:O19 ASP174:OD1–M6:H43 Pi–cation ARG104:NH1–M6 ARG104:NH2–M6

The 12,000 molecules of Sigma lifechem database is used for screening process. The screening process carried out with HTVS, SP and XP docking modes, the output of 68 molecules are analyzed. The 6 molecules (M1 to M6) with the best docking score are represented with docking interactions in the table.

Table 2 – ADME properties.

C1 Mol	C2 Stars	C3 CNS	C4 Mol Weight	C5 SASA	C6 Volume	C7 DHB	C8 AHB	C9 QPlogP o/w	C10 QPP Caco	C11 QPlog BB	C12 Meta	C13 Human OralA	C14 % Human OralA	C15 N and O	C16 Rule of 5	C17 Rule of 3
M1	1	0	334.417	636.73	1116.938	2	9.5	−0.769	12.671	−0.148	6	2	42.182	7	0	1
M2	0	0	313.303	538.80	935.71	3	8	−1.087	9.197	0.097	3	2	37.828	6	0	1
M3	0	0	333.43	664.67	1154.632	2	8.5	−0.273	9.156	−0.351	5	2	42.558	6	0	1
M4	0	−1	327.382	615.62	1057.566	3	8	−0.754	5.476	−0.479	3	2	35.747	6	0	1
M5	0	0	319.403	626.1	1088.01	2	8.5	−0.545	14.135	−0.039	6	2	44.339	6	0	1
M6	0	−2	362.403	590.23	1049.348	2	6.5	1.687	59.566	−1.14	4	3	68.591	8	0	0

Optimum values for the parameter considered. 95% of available drugs fall in the range of stars (more number of stars indicate that the molecule is less drug like) [Stars]: 0–5, predicted central nervous system activity [CNS]: −2(inactive) + 2(active), molecular weight [Mol Weight]: (130–725), solvent accessible surface area using a probe with 1.4 Å radius [SASA]: 300.0–1000.0, total solvent accessible volume in cubic angstroms [Volume]: 500.0–2000.0, hydrogen bond donors [DHB]: (0.0–6.0), hydrogen bond acceptors [AHB]: (2.0–20.0), predicted octanol/water partition coefficient [QPlogP o/w]: (−2.0–6.5), Predicted apparent Caco-2 cell (model for gut blood barrier) permeability in nm/s [QPP Caco]: <25 poor, >500 great, predicted brain/blood barrier partition coefficient [QPlogBB]: (−3.0–1.2), number of likely metabolic reactions [Meta]: (1–8), human oral absorption [Human OralA]: 1 low, 2 medium, 3 high, % human oral absorption [% Human OralA]: >80% high, <25% low, number of nitrogens and oxygens [N and O]: 2–15, [Rule of 5] (4), [Rule of 3] (3). C = Column.
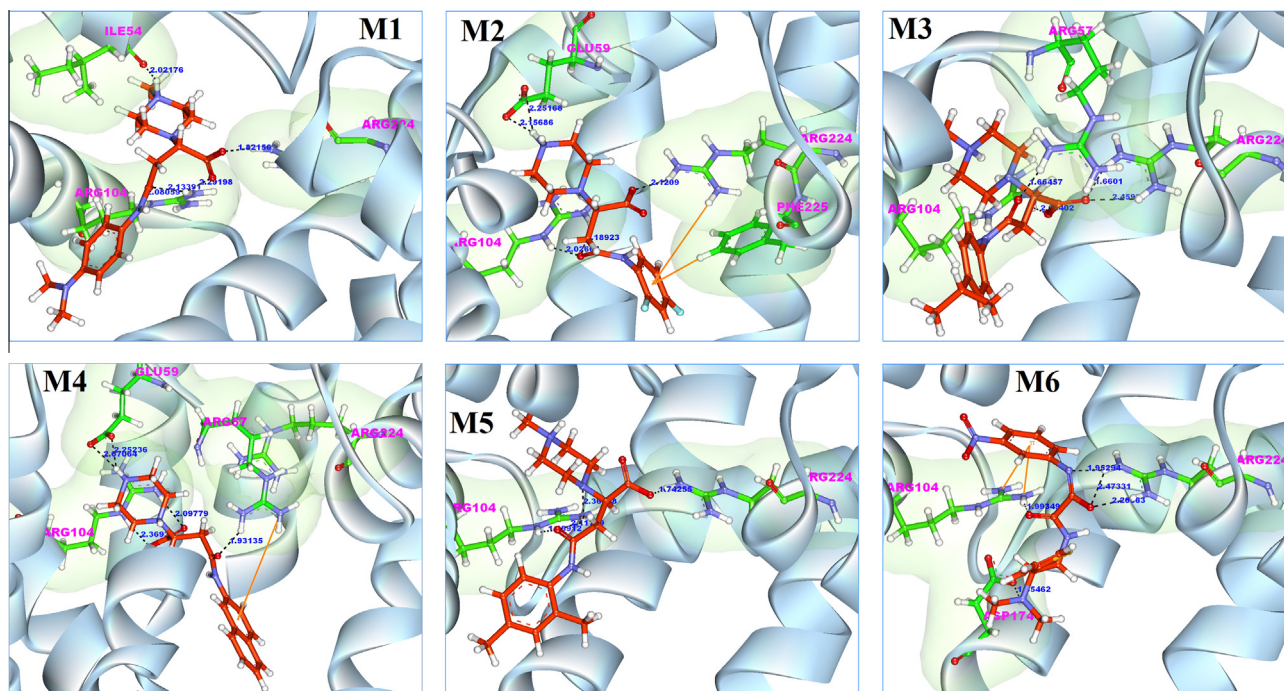
Fig. 1 – Binding interaction poses of ligands with SigF protein for the best docked molecules. The ligand molecules are represented in red ball and stick model, π–π stacking are shown in orange color, intermolecular Hydrogen Bonds are represented in black and the SigF protein is shown in light blue color.

physiological stress conditions [12,13]. In the present work the *in silico* screening has been taken up to identify small molecules, which can act as antagonists for the SigF protein.

Methodology

The SigF protein binds to its cognate anti-SigF (Usfx) in its regulatory circuit; in the absence of Usfx, SigF initiates transcription initiation and gene expression, which leads to protein synthesis and in turn helps in the survival of MTB. Finding inhibitors for the SigF protein at the Usfx binding site will arrest the survival of MTB. In the present work *in silico* screening is taken up to identify competitive inhibitors for Usfx.

Virtual screening

The structure of SigF was considered from an earlier work [13]. The SigF structure was energy minimized using the protein preparation wizard in Maestro 9.0.111 (Maestro v 9.0.111 Schrodinger LLC, New York, NY) applying OPLS 2001 (optimized potential for liquid simulations 2001) force field with default parameters [14]. The Virtual screening work flow of Schrodinger involves three consecutive steps: (a) receptor grid generation; (b) ligand preparation; and (c) Glide ligand docking [15]. The grid was generated using the Gridgen module of Schrodinger Suite at the active site amino acid residues [13,16]. The Sigma lifechem small molecule database was considered and retrieved in Sdf file format. The ligands were subjected to ligand preparation using the Ligprep 2.5 module of Schrodinger Suit [17] and during the process, tautomeric states and ionization states were generated using the epic module. The work flow utilizes the Glide module for Ligand

and Receptor docking. Glide filters the molecules using HTVS (high throughput virtual screening), SP (standard precision) and XP (extra precision) modes [16]. The OPLS 2001 force field [14,18] parameters were applied while performing docking calculations. The molecules with the best Glide score and Glide energy were visually inspected and considered for further analysis. The SASA (solvent accessible surface area) for the receptor and ligand complexes were calculated with the default parameters. The receptor–ligand complexes were analyzed using Accelrys Discovery Studio Visualizer (Accelrys Software Inc., 2007 Accelrys Discovery Studio Visualiser v 2.5.5. Accelrys Software Inc., San Diego).

ADME properties

The Absorption Distribution Metabolism and Elimination (ADME) properties were calculated using the QikProp [18] module of Schrodinger suite (QikProp, version 3.0, Schrodinger, LLC, New York, NY, 2010) for assessing the druggability and to filter the ligand molecules at an early stage of identifying the new antagonists.

Results and discussion

Virtual screening

The virtual screening studies are carried out with 12,000 small ligand molecules from the Sigma lifechem database. In the process, the grid box is generated with $75 \times 75 \times 75$ Å³ around the active site amino acids which were considered from an earlier study [13]. In the ligand preparation process using the epic program, 5 stereo isomers from 32 structures

Ligand M1
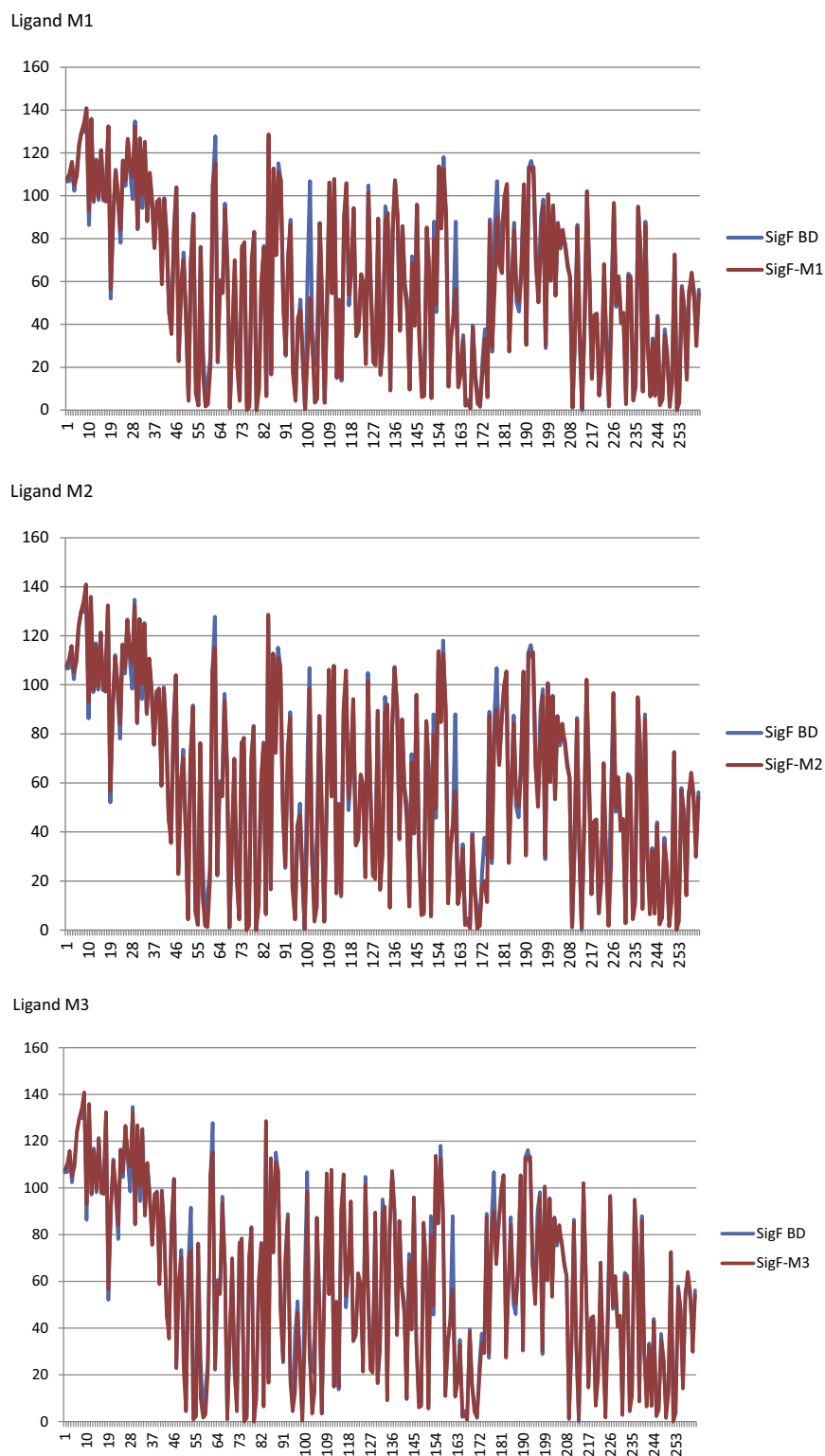
Ligand M2

Ligand M3

Fig. 2 – Solvent accessible surface area of SigF protein and the SigF protein–ligand complexes for the best docked molecules. The solvent accessible surface area (SASA) of SigF protein before docking (SigF BD) is represented in blue color line and the protein–ligand complex is represented with maroon color line for the best docked molecules. The amino acid numbers are represented on X-axis and SASA values are shown on Y-axis.

generated and one ring conformation generated for 5 and 6-membered rings with the least energy are considered; 20,646 molecular structures are generated in the ligand

preparation output file, which are used in the screening process. Among the 20,646 ligand molecules, 6790 molecules are docked in HTVS mode; the top 10% (679 of 6790) of the
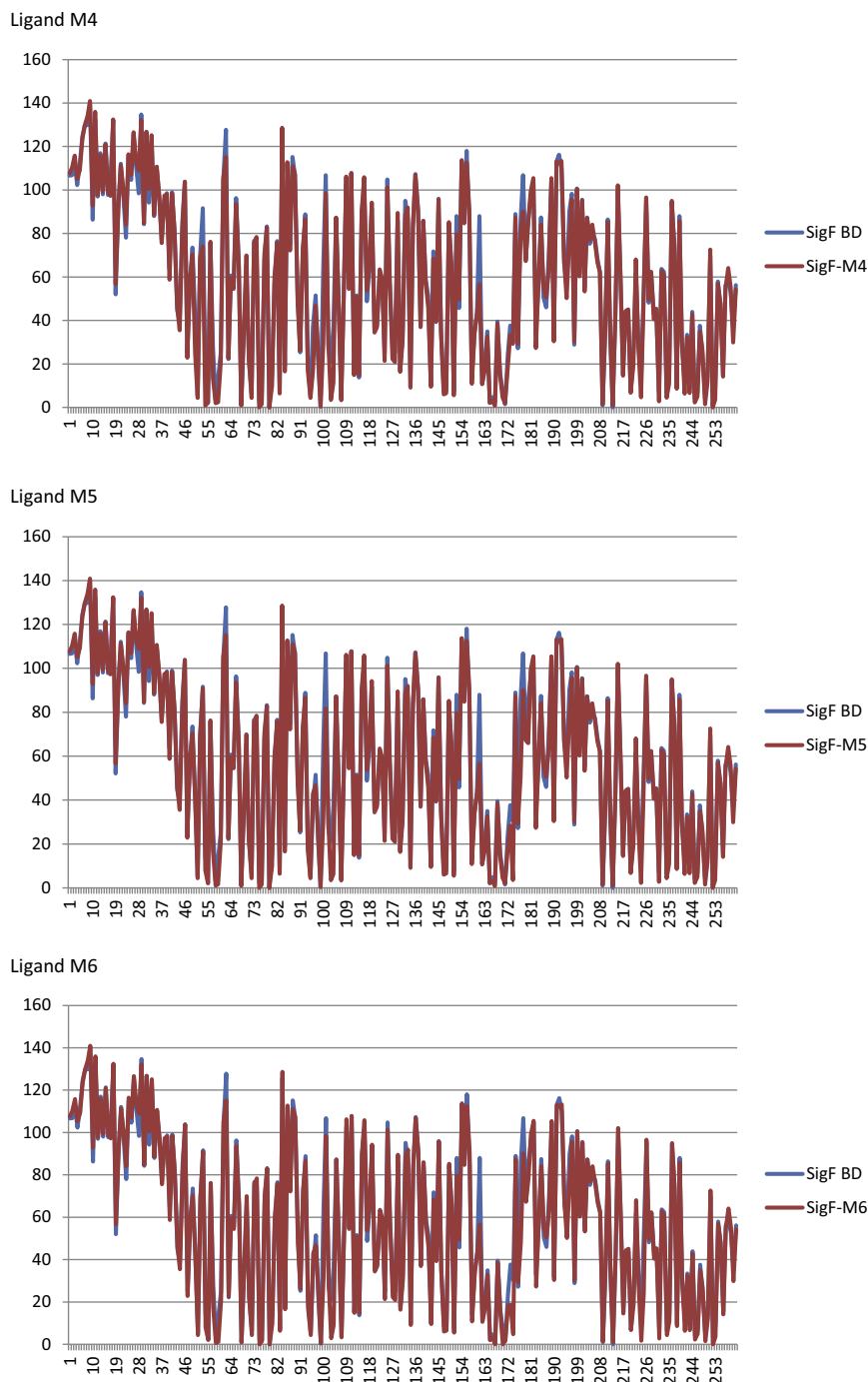
Ligand M4

Ligand M5

Ligand M6

Fig 2. *(continued)*

ligand molecules from the HTVS screening process are considered for SP docking. The 679 ligand molecules are docked and the top 10% of SP docked molecules are utilized for XP docking mode. Finally, 68 docked complexes are generated in the XP flexible docking mode. The docked complexes are analyzed, visually inspected and the data of a sample of 6 molecules (M1 to M6) with the corresponding docking properties generated namely docking score, docking energy, docking interactions and ADME properties, are presented in Tables 1 and 2. The hydrogen bond interactions and π cation interactions are depicted in Fig. 1. The virtual screening analysis

reveals that the amino acid residues ILE54, ARG57, GLU59, ARG104, ASP174 and ARG224 are involved in the hydrogen bond formation and π cation interactions with the M1 to M6 ligand molecules. The amide group oxygen in M6 molecule and the amide group oxygen in M1 to M5 molecules bound in the docked complex through hydrogen bonds with the amine hydrogen of ARG104 in SigF protein. The piperazine-1-yl acetic acid moiety present in the M1 to M5 molecules is consistently binding to the ARG224 amino acid of the SigF protein. The carboxyl oxygen forms a hydrogen bond with the amino group (hydrogen) of the ARG224. The docking result

analysis reveals the common presence of piperazine-l-yl acetic acid moiety and an amide group in all the ligands and is capable of binding effectively with ARG104 and ARG224 of the SigF protein. The SASA calculations are carried out for the SigF protein and the SigF–Ligand docked complexes and are represented in Fig. 2. The SASA values of the SigF protein for the amino acid residues which are involved in bond formation (ILE54, ARG57, GLU59, ARG104, ASP174 and ARG224) and spatially nearby residues in the binding site decreased after docking when compared with that before docking. The decrease in SASA values confirms that these amino acid residues are involved in the bond formation with the ligand molecules.

ADME properties

The ADME properties for the new ligands identified namely M1 to M6 are calculated and tabulated in Table 2. These molecules have properties within the limits projected as per the Lipinski rules of 5 and Jorgensen's rules of 3, with medium human oral absorption, which signifies that the ligand molecules have acceptable ADME properties.

Conclusion

The virtual screening studies performed using Sigma lifechem database against active site residues of SigF reveal ILE54, ARG57, GLU59, ARG104, ASP174 and ARG224 amino acid residues to be important for binding in the SigF protein. A sample of six ligands is presented in the present communication; several novel scaffolds are identified in the virtual screening studies. The piperazine-l-yl acetic acid moiety and an amide group in the ligands commonly exist and forms hydrogen bonds with ARG104 and ARG224 of the SigF protein. The ligand molecules show admissible ADME properties and are identified as novel antagonists for the SigF protein. Further work is in progress in the direction of identifying novel potent inhibitors for the SigF protein, which is important for virulence.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

REFERENCES

[1] A. Zumla, A. George, V. Sharma, R.H. Herbert, A. Oxley, M. Oliver, W.H.O. The, Global tuberculosis report – further to go, Lancet Glob Health 3 (2015) (2014) e10–12.

[2] J.D. Helmann, The extracytoplasmic function (ECF) sigma factors, Adv. Microb. Physiol. 46 (2002) 47–110.

[3] V. Malkhed, K. Mustyala, S. Potlapally, U. Vuruputuri, Modeling of alternate RNA polymerase Sigma D factor and identification of novel inhibitors by virtual screening, Cell. Mol. Bioeng. 5 (2012) 363–374.

[4] E.P. Williams, J.H. Lee, W.R. Bishai, C. Colantuoni, P.C. Karakousis, *Mycobacterium tuberculosis* SigF regulates genes encoding cell wall-associated proteins and directly regulates the transcriptional regulatory gene phoY1, J. Bacteriol. 189 (2007) 4234–4242.

[5] Y. Zhang, K. Post-Martens, S. Denkin, New drug candidates and therapeutic targets for tuberculosis therapy, Drug Discov. Today 11 (2006) 21–27.

[6] J.H. Lee, P.C. Karakousis, W.R. Bishai, Roles of SigB and SigF in the *Mycobacterium tuberculosis* sigma factor network, J. Bacteriol. 190 (2008) 699–707.

[7] V. Malkhed, B. Gudlur, B. Kondagari, R. Dulapalli, U. Vuruputuri, Study of interactions between *Mycobacterium tuberculosis* proteins: SigK and anti-SigK, J. Mol. Model. 17 (2011) 1109–1119.

[8] S. Rodrigue, R. Provvedi, P.E. Jacques, L. Gaudreau, R. Manganelli, The sigma factors of *Mycobacterium tuberculosis*, FEMS Microbiol. Rev. 30 (2006) 926–941.

[9] S. Gebhard, A. Humpel, A.D. McLellan, G.M. Cook, The alternative sigma factor SigF of *Mycobacterium smegmatis* is required for survival of heat shock, acidic pH and oxidative stress, Microbiology 154 (2008) 2786–2795.

[10] D.F. Browning, S.J. Busby, The regulation of bacterial transcription initiation, Nat. Rev. Microbiol. 2 (2004) 57–65.

[11] D.E. Geiman, D. Kaushal, C. Ko, S. Tyagi, Y.C. Manabe, B.G. Schroeder, et al, Attenuation of late-stage disease in mice infected by the *Mycobacterium tuberculosis* mutant lacking the SigF alternate sigma factor and identification of SigF-dependent genes by microarray analysis, Infect. Immun. 72 (2004) 1733–1745.

[12] J. Beaucher, S. Rodrigue, P.E. Jacques, I. Smith, R. Brzezinski, L. Gaudreau, Novel *Mycobacterium tuberculosis* anti-sigma factor antagonists control sigmaF activity by distinct mechanisms, Mol. Microbiol. 45 (2002) 1527–1540.

[13] K.K. Mustyala, V. Malkhed, S.R. Potlapally, V.R. Chittireddy, U. Vuruputuri, Macromolecular structure and interaction studies of SigF and Usfx in *Mycobacterium tuberculosis*, J. Recept. Signal Transduct. Res. 34 (2014) 162–173.

[14] W.L. Jorgensen, D.S. Maxwell, J. Tirado-Rives, Development and testing of the OPLS all-atom force field on conformational energetic and properties of organic liquids, J. Am. Chem. Soc. 118 (1996) 11225–11236.

[15] S. Kawatkar, H. Wang, R. Czerminski, D. Joseph-McCarthy, Virtual fragment screening: an exploration of various docking and scoring protocols for fragments using Glide, J. Comput. Aided Mol. Des. 23 (2009) 527–539.

[16] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, et al, Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for protein–ligand complexes, J. Med. Chem. 49 (2006) 6177–6196.

[17] I.J. Chen, N. Foloppe, Drug-like bioactive structures and conformational coverage with the LigPrep/ConfGen suite: comparison to programs MOE and catalyst, J. Chem. Inf. Model. 50 (2010) 822–839.

[18] L. Ioakimidis, L. Thoukydidis, A. Mirza, S. Naeem, J. Reynisson, Benchmarking the Reliability of QikProp. Correlation between Experimental and Predicted Values, QSAR Comb. Sci. 27 (2008) 445–456.