Design of LuxO based inhibitors to reverse engineer the genetic circuit of *Vibrio cholera-* an anti-virulent cholera therapy

Hema M.[1], Karthi S.[2] and Adline Princy S.*[1]

[1]*Quorum Sensing Laboratory, Centre for Research on Infectious Diseases (CRID), School of Chemical and Biotechnology, SASTRA University, Thanjavur, Tamil Nadu, India*
[2]*Department of Bioinformatics, School of Chemical and Biotechnology, SASTRA University, Thanjavur, Tamil Nadu, India*

_____

ABSTRACT

Quorum Sensing (QS) in Vibrio cholerae, ensures to co-ordinate its behavioural changes at the inter/intra species level to establish its infection. The LuxO, a well-known QS response regulator controls the global (three parallel) QS cascade of Vibrio cholerae. The structural/functional predictions of LuxO shows to have a receiver, central ATPase and a DNA binding domains and we have identified the ATP binding domain as our molecular target as its main function is to hydrolysis the ATP molecule (source of energy and phosphate group). In the present study, we have modelled ATP binding domain of LuxO based on the structure 3MOE (crystal structure of ATP-bound state of walker B mutant NtrC1 ATPase domain) as templates. High throughput virtual screening of 1 million compounds from the public database were done against our target, LuxO. Based on the Glide score, interaction sites and ADMET property, the top 10 best hits were identified. The compound, 3 [4-morpholine-anilio] carbonyl-2-pyrazine carboxylic acid (QSI^{Vc}) has a glide score of -11.873 with favourable interactions to the amino acids- G170, G172, K 173, R 357, and I 140 and Mg^{2+}. The conserved residues, R 357 that lies in the arginine finger of the ATPase domain and G 170 and G172 in the Walker A motif were functionally characterized to aid in the hydrolysis of ATP. So the proposed structure-function relationship of QSI^{Vc} and their exact interaction pattern to inhibit the ATP hydrolysis would be of significant interest to develop an effective cholera therapy.

Key words: *V. cholerae,* ADMET Property, Walker A motif, Glide Score, NtrC1

_____

INTRODUCTION

Vibrio cholerae, a causative agent of the disease cholera, develops a major health risks among the developing countries.The cell-density dependent process, Quorum Sensing (QS) dictate the behavioral pattern of the bacteria where the "Auto-inducers" act as a bacterial lingo. Bacteria sense their own population through QS [1]. The lesser bacterial loads (low cell density (LCD)) correspond to a low concentration of auto-inducers whereas the greater bacterial loads (High cell density (HCD)), correspond to a higher concentration of auto-inducers. Thus the concentration of auto-inducers is directly proportional to corresponding increase to the cell population. There are basically two systems of QS in *V. cholerae* that have been explored till date [2, 3]. The third system having an unknown cell signal remains to be unexplored whereas the system-I and system-II corresponds to anindependent intra and inter-species communication. The system I and II consists of a transmembrane receptor, CqsSand a periplasmic/transmembrane receptor LuxP/Q to specifically respond to the signals (CAI-1 and AI-2) synthesized by the enzymes CqsA and LuxS respectively [4, 5] The receptors are also well-characterized to have bi-functional roles

_____

to either behave as kinases (LCD) or Phosphatases (HCD). These signals at its lower concentration shows lower affinity towards their cognate receptors [6]. The receptors also serve as kinases to transfer the phosphate group from a protein called LuxU to a regulator protein, LuxO. Both system I and II converges at the regulator, LuxO [7-10] whereas, the system III integrates its signal to the same regulatory protein, LuxO on activation of the Vars/A pathway, but the mechanism is yet to be elucidated[7, 11]. Once the signal integrates the LuxO, it get phosphorylated, LuxO–P and synergistically work with an alternate σ54 to increase the transcript level of the 4 quorum regulatory RNAs (Qrrl 1-4). The Qrrl 1-4 in-turn interacts with the RNA chaperone Hfq, and antisense the HapR mRNA, a global repressor of virulence in *V. cholerae*[12].

Hap R, a negative regulator represses the transcriptional activation of the VpsT, and enhance the adherence of *V.cholerae* to the intestinal epithelial cells [5, 13]. Also down-regulates the production of the major virulence factors of *V. cholerae*, Cholera Toxin and Toxin co-regulated Pilus and hemolysin either through an direct/indirect mechanism [14]. In parallel, the HapR also act as an activator of HapA protease and RNA polymerase sigma factor (RpoS σ54). The functional significance of the HapA protease digest the GM-1 receptor and detaches the *V.cholerae* cells from the human epithelial cells [15] also the σ54 increases those bacterial cells to resist the extreme nutrition and oxidative conditions [16] . The *V.cholerae* cells are finely tuned or re-programmed to express virulence factors like cholera toxin, Toxin co-regulated pilus, hemolysin, biofilm under the LCD conditions whereas, the factors like HapA protease, σ54 are repressed on inhibiting the HapR expression at the post-transcriptional level. The fine tuning of the behavior of the cells solely depend on the autoinducer concentration and at HCD, the CAI-1 and AI-2 synergistically bind to their specific cognate receptors CqsS and LuxP/Q occurs to enable their binding switch from kinase to phosphatase. Consequently the inactivation of LuxO-P to LuxO occurs to depress the transcription of the qrrl 1-4 genes without any effect on the negative regulator, HapR. Thus the negative regulator, HapR inhibits the expression of the virulence factors and biofilm formation and paradoxically upregulates the production of HapA protease and stress response genes via, activating σ54. Therefore, we propose to reverse engineer the quorum sensing circuit to a HCD condition in a LCD state on targeting the quorum regulator, LuxO. This could probably provide an anti-virulent effect to overcome the emergence of multi-drug resistant strains [17]. Thus the outcome of this research would significantly prove that probing QS circuits would be a way to stifle *Vibrio cholera* pathogenesis.

EXPERIMENTAL SECTION

Sequence Alignment
The FASTA sequence of the ATP binding domain of LuxO (Q9KT 84) was retrieved from the database, Uniprot(www.uniprot.org) and subjected to BLAST search to identify the highly sequence similarity with a minimum E value [18]. The template structure (3M0E) was obtained from the protein structure database (http://www.rcsb.org). Further, the sequence alignment was performed between the target protein (Q9KT 84) and the template protein (3M0E) exploiting the web base tool named Clustral W[19]

Template based Modeling
The initial model of the ATP binding domain of LuxO was generated employing the molecular modeling software called Schrödinger (Prime). Crystal structure of 3M0E served as template input for the homology modelling [20]. Among the 10 focused models having distinctive geometrical conformation the best model was taken for the docking studies.

Quality assessment of Model
Various parameters including the Ramachandran plots [21], G- Factor, PSQS and Z-score were considered to score the best model. Structure validation program called ProSA (program to check the stereo chemical quality of the protein structures) [22] and PROCHECK [23] was used to analyses the structure of modelled protein. RMSD value of the modelled protein was obtained on superimposing the modelled protein along with the crystal structure protein using the PyMOLE software.

Molecular Docking
Prior to docking, the modelled protein was prepared using a tool called Prep Wize (Schrödinger) to optimize and minimize the validated protein. Grid was generated for the prepared protein using Schrödinger suite (Grid to generate tool). A high throughput virtual screening (HTVS) of millions of compounds from Maybridge (57000) and

_____

Zinc database (1 million compound) were on lead optimization with the docking interactions site followed by the standard precision (SP) and extra precision (XP).

Screening of Ligands
The parameters used for screening of ligand were glide scores, interaction patterns and the functional group of the compounds. The selected compounds were further subjected to clustering to rule out few of them having the same structure using Canvas module (Schrödinger). The 10 best hits were considered for characterizing the drug likeness property to optimize the lead molecules, LuxO based inhibitors using the ADMET and TOPKAT (Toxicity Prediction by Komputer Assisted Technology) [24]

RESULTS AND DISCUSSION

Sequence alignment
Since the NMR/Crystal structure of the ATPase domain of LuxO is unavailable, we have developed a target base model and the PDB protein, 3M0E was identified as a potential template on a BLAST search. Sequence alignment between the reference PDB protein and the target sequence plays a significant role in generating a perfect model. Hence the percentage sequence similarity between the reference (Q9KT84) and the target sequence (3M0E) was 39% (identity 42.4%). As the function of the proteins is the same; they are more likely to share a high degree of structural similarity. Bassler in the year 2013 has also shown a structural similarity of LuxO to the crystal structure of 3M0E on mutational studies [10]. In the present study, to refine the mode to a stable structure (Figure1) the residues from 1-111 and 378-455 were removed as it belongs to the receiver domain and the DNA binding domain
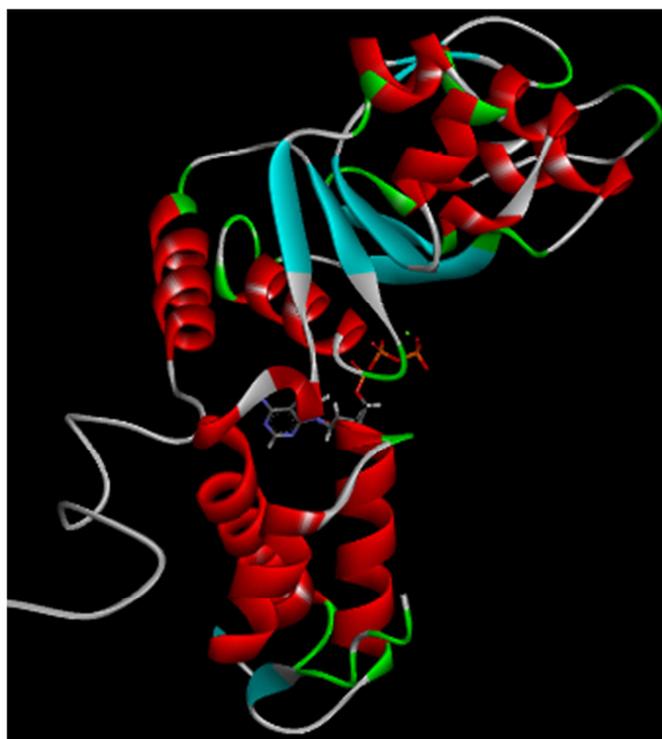

Figure 1: Ribbon Structure of modelled protein generated using Schrodinger suite

Structure Validation
Modelled protein was analyzed using ProSA-Z and PROCHECK. The Ramachandran plot analysis of the ATPase domain showed 11.1% in allowing region, 0.9% in generously allowed region, 87.26% inthe core region and 0.4% in the disallowed regions (Figure 2).
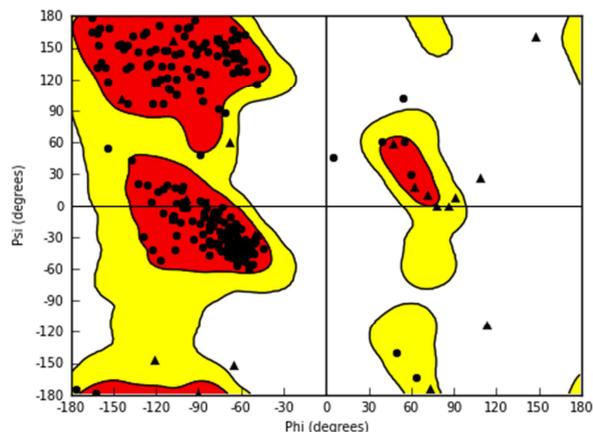
_____

Figure 2: Ramachandran Plot for modelled protein analyzed by PROCHECK

The observed G-factor for the modeled protein was -0.01Z- Score and PSQS are -6.61 and -0.3 and lies within the range (Figure 3 a and 3 b). The knowledge based energy graph also shows the position of the sequence that lie within the range.
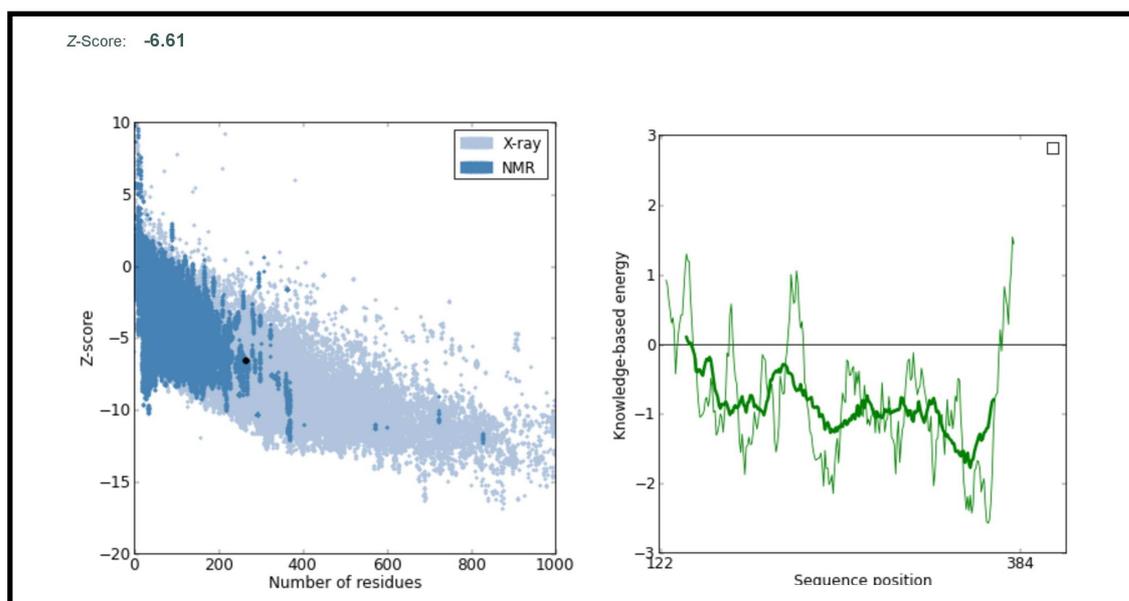

Figure 3 (a): ProsA- web Z score (b): Knowledge based energy graph generated by ProsA

ProSA energies enable us to know the stability of the modelled protein [25]. The graph elucidates the modeled protein is of fairly good quality. The RMSD value of the target protein was obtained on superimposing the parent protein with the modeled protein and the value was 2.2 Å suggesting a reliable 3D structure (Figure 4).
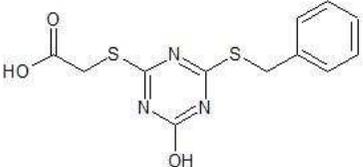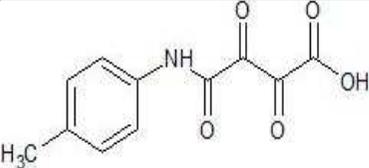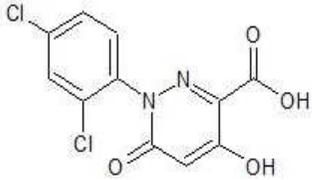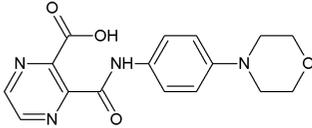
_____
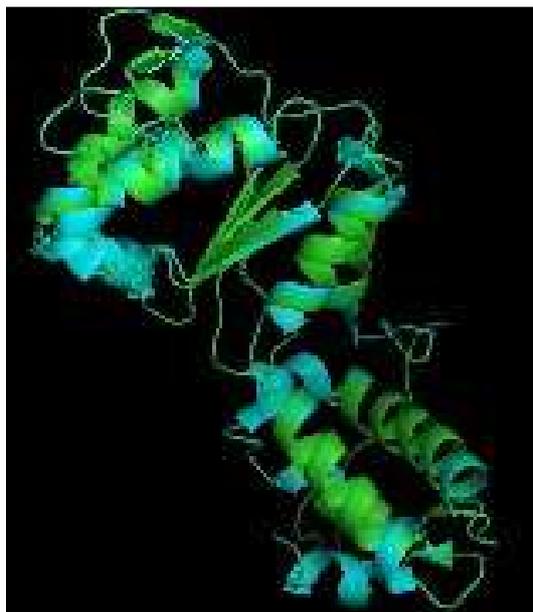
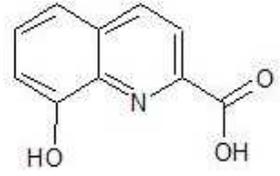**Figure 4: Superimposed image generated by Pym olé. Green colored ribbon shows modelled protein.**

Protein Ligand Docking

One of the most challenges in the computational chemistry is to understand the molecular interaction of the ligand with its active site of the target protein. The molecular docking analysis shows the mode and affinity of the binding of the small drug like molecules [26-29]. Initially, the target protein was docked with an ATP molecule where it was found to interact with the residues, S169, G170, G172, K173, R357 and the residues G 170 and G172 was used as reference residues to identify the inhibitors. Theidentified top-hit ligand, 3 [4-morpholine-anilio] carbonyl-2-pyrazine carboxylic acid (QSI[Vc]) also interact to those residues [30, 31]. Also, we score the ligand on the basis of the Glide score (G-Score) as it includes the negative sum of value of 16 parameters like protein–ligand hydrogen bond energy (external H-bond), protein–ligand van der waals energy (external vdw), ligand internal vdw energy (internal vdw), and ligand torsional strains energy (internal torsion) etc., [32, 33]. Also the drug- likeness properties of the top-scored ligands were analyzed using the software called TOPKAT and ADMET (Table 1) and the interaction pattern of QSI[Vc]is shown in the figure 5a and 5b.

Among the various interactions, the hydrogen bond shows specific and stable interaction of the protein, LuxO with the ligands forms the basis of molecular recognition [34]. QSI[Vc] has also shown a three hydrogen bond interaction with the three key amino acid residues (G 170 & 172 and I 140) that are involved in the hydrolysis of ATP molecules [35, 36]. Thus the LuxO selective inhibitor, QSI[Vc] would be a promising therapeutic candidate of the NtrC-family of proteins. This NtrC-like protein has a central ATP domain sandwiched between a receiver and a DNA binding domains with a unique characteristics of walkerA, walker B and arginine finger [35]. The walker A motif in the C1 region interacts with a phosphate group of an ATP molecule to form a P loop with a conserved region GXXGXGK (X represents any amino acid) and hydrolyze an ATP molecule [37]. The walker B of the C4 region has a conserved region, hhhhDE (h-hydrophobic amino acid) where the asparate residue co-ordinate with the Mg^{2+} to favor the hydrolysis of ATP molecule. Likewise, arginine finger was found to interact with the ɣ-phosphate of the ATP and to oligomerize the complex [38](Figure 6).Our studies had deduced the top 10 hit ligand to interact to those conserved residues (I 143, G170, G172, R 347 Mg^{2+}) of the ATPase domain that is involved in the hydrolysis of ATP.

Table: 1 Toxicity Prediction for the top 10 best hit LuxO selective ligand (QSI^{vc})

Sl.NO	Compound	Toxicity prediction			Molecular Property			
		QP log HERGE	QP log BB	QP log Khsa	Mol. Wt	Absorption	Log P	Rule of 5 and 3
1	G-score (-16.79) (G170, L173, R357,N280)	●	●	●	292.28	63.95%	1.871	0
2	G-score (-15.612) (G172,L173,N280)	●	●	●	309.35	74.02%	2.706	0
3	G-score (-13.117) (G170,L173)	●	●	●	259.64	64.07%	1.148	0
4	G-score (-13.487) (L173,A357)	●	●	●	301.08	69.84%	2.586	0
5	G-score (-11.873) (I143, G170, G172)	●	●	●	328.32	68.34%	1.412	0
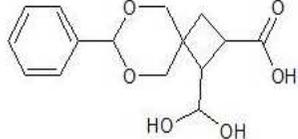
_____
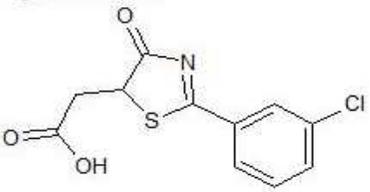
6	G-score (-12.471) (L173,N280)	🟢	🟢	🟢	269.72	69.19%	1.755	0
7	G-score (-12.426) (G170, R 357)	🟢	🟢	🟢	189.17	66.29%	1.109	0

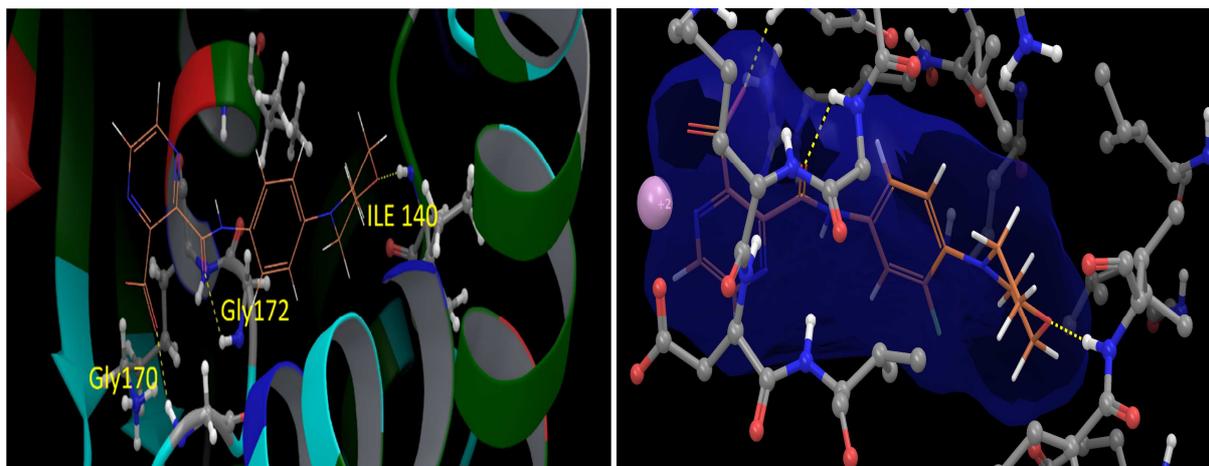*Toxic free behavior*

_____

Figure 5(a): Figure showing interaction of QSI^Vc with modelled protein. (b): Solid mesh image showing the binding pocket of inhibitor
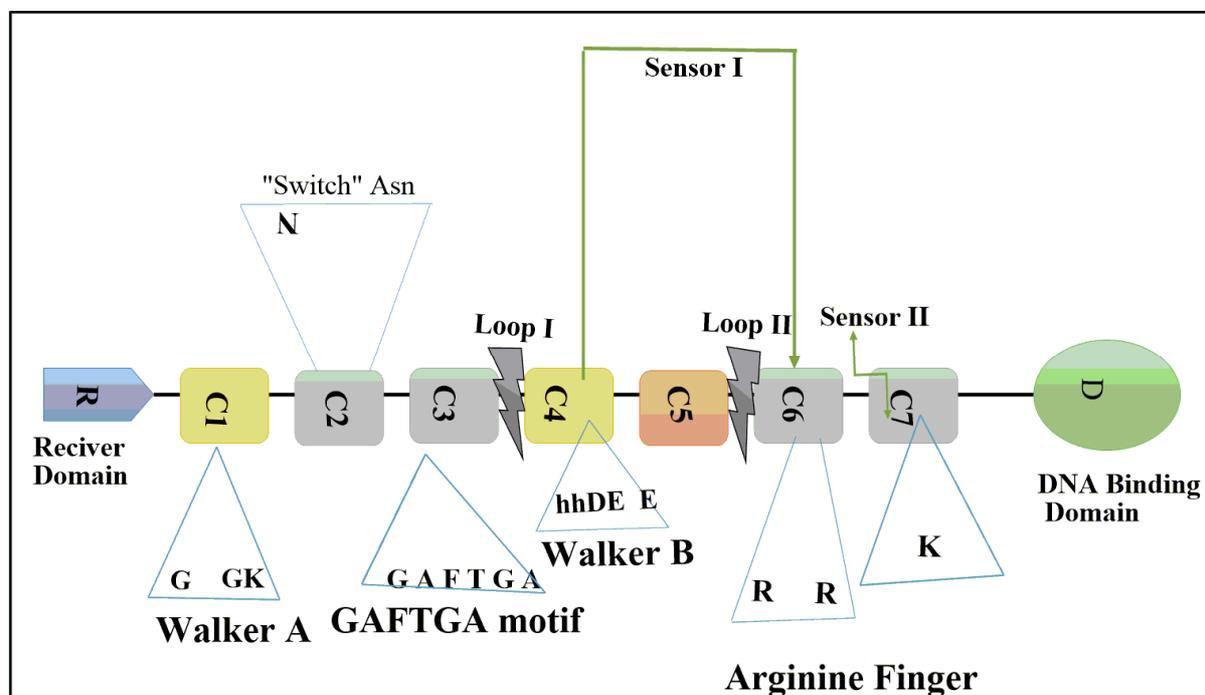

Figure 6: Genetic organization of the NtrC- type quorum regulator, LuxO

CONCLUSION

The study would result on the discovery of a potent novel LuxO selective anti-virulent drug (QSI^VC) for cholera treatment that could hinder the QS circuits I and II of *Vibrio cholerae* and stimulate HCD condition in a LCD state. This might probably halt the virulence of *Vibrio cholerae* and would efflux them from the human intestinal system, rendering the clinical condition harmless. Further the *invitro* and *invivo* biological evaluation will reveal the small molecule inhibitors to specifically target the ATPase domain of the NtrC-family regulator, LuxO. Since the NtrC-type of regulators are well-characterized in most of the bacteria to control virulence, nitrogen metabolism, motility and other vital processes would offer an promising future for anti-Two Component Signaling (TCS) drug development.

_____

REFERENCES

[1]Arya R ; Princy S A,*Future Microbiol.,***2013,**8(10), 1339-1353
[2]Adline Princy S; N S; Sajal Srivastava; Thomas Sabu,*Nova Publishers.,***2011,**123-133
[3]S. Adline Princy; V R; Rekha Arya; Sabu Thomas,*Nova Publishers.,***2011,**109-122
[4]Kelly R C; Bolitho M E; Higgins D A; Lu W; Ng W-L; Jeffrey P D; Rabinowitz J D; Semmelhack M F; Hughson F M ; Bassler B L,*Nature chemical biology.,***2009,**5(12), 891-895
[5]Hammer B K ; Bassler B L,*Molecular microbiology.,***2003,**50(1), 101-104
[6]Adline Princy S ; Praveen Krishna V, *Asia Pacific J Mol Biol Biotech.,***2013,**21(2), 67-74
[7]Miller M B, Skorupski K; Lenz D H; Taylor R K; Bassler B L,*Cell.,***2002,**110(3), 303-314
[8]Wei Y; Ng W L; Cong J ; Bassler B L,*Molecular microbiology.,***2012,**83(6), 1095-1108
[9]Ng W-L; Wei Y; Perez L J; Cong J; Long T; Koch M; Semmelhack M F; Wingreen N S; Bassler B L,*Proc Natl Acad Sci.,***2010,**107(12), 5575-5580
[10]Ng W-L; Perez L; Cong J; Semmelhack M F;Bassler B L,*PLoS Pathog.,***2012,**8(6), e1002767
[11]Lenz D H; Miller M B; Zhu J; Kulkarni R V; Bassler B L,*Mol microbiol.,***2005,**58(4), 1186-1202
[12]Lenz D H; Mok K C; Lilley B N; Kulkarni R V; Wingreen N S; Bassler B L,*Cell.,***2004,**118(1), 69-82
[13]Waters C M; Lu W; Rabinowitz J D; Bassler B L,*J bacteriol.,***2008,**190(7), 2527-2536
[14]Tsou A M ; Zhu J,*Infection and immunity.,***2010,**78(1), 461-467
[15]Finkelstein R A; Boesman-Finkelstein M; Chang Y and Häse C,*Infection and immunity.,***1992,**60(2), 472-478
[16]Joelsson A; Kan B; Zhu J, *Applenvironmental microbiol.,***2007,**73(11), 3742-3746
[17]Srikkanth Balasubramanian; Hema M, Santhosh RS; Karthi Shanmugam; Adline Princy S *Biotechnology-An Indian J.,***2014,**9(2),121-124
[18]Altschul SF G W; Miller W, Myers; EW & Lipman DJ,*J Mol biol.,***1990,**215, 403-410
[19]Thompson JD H D; Gibson TJ *Nucleic Acids res.,***1994,**22, 4673–4680
[20]B Webb;A Sali,*Springer.,* **2014**, 1-15
[21]RAMACHANDRAN GN R C; SASISEKHARAN V,*J Mol Biol.,***1963,**7(95),
[22]Wiederstein M M J; *Nucleic Acids Res 35.* **2007,**W407–W410
[23]Laskowski RA M M; Moss DS;Thornton JM,*J Appl Crystallogr.,***1993,**26, (283–291
[24]K Yadav D; Mudgal V; Agrawal J; K Maurya A; U Bawankule D; S Chanotiya C; Khan F ;T Thul S,*Curr Comput Aided.,***2013,**9(3), 360-370
[25]Wiederstein M; Sippl M J, *Nucleic Acids res.,***2007,**35(suppl 2), W407-W410
[26]Frimurer T M; Peters G H; Iversen L F; Andersen H S; Møller N P H; Olsen O H,*Biophys J.* **2003,**84(4), 2273-2281
[27]Ravichandiran V; Shanmugam K; Anupama K; Thomas S; Princy A,*Eurj of med chem.,***2012,**48(200-205
[28]Ravichandiran V; Shanmugam K, Princy Solomon A,*Med Chem.,***2013,**9(6), 819-827
[29]Vinothkannan Ravichandiran S K; Shrimathi R; Princy S A; *Asian J Chem.,***2013,**25(1), 95-100
[30]Fazil M H U T; Kumar S; Rao N S; Selvaraj C; Singh S K; Pandey H P; Singh D V,*JBiomolecu Stru Dyn.,***2012,**30(5), 574-584
[31]Faloon P; Weiner W S; Matharu D S; Neuenswander B; Porubsky P; Youngsaye W; Bennion M; Ng W-L; Hurley A; Mosher C M,*NIH Report.,***2014,**
[32]Jones G; Willett P; Glen R C,*J Mol Biol.,***1995,**245(1), 43-53
[33]Arya R ; Princy S A,*Med Chem Res.,***2013,**22(4), 1856-1865
[34]Hubbard R E; Kamran Haider M,*eLS.,***2010,**
[35]Lee S-Y; De La Torre A; Yan D; Kustu S; Nixon B T; Wemmer D E,*Genes Dev.,***2003,**17(20), 2552-2563
[36]Ogura T; Whiteheart S W ; Wilkinson A J,*J Struct Biol.,***2004,**146(1), 106-112
[37]Saraste M; Sibbald P R; Wittinghofer A,*Trends Biochem.,***1990,**15(11), 430-434
[38]Chen B; Sysoeva T A; Chowdhury S; Guo L; De Carlo S; Hanson J A; Yang H ; Nixon B T, *Structure.,***2010,**18(11), 1420-1430