Prediction of Novel Inhibitors against Exodeoxyribonuclease I of *H. influenzae* through *In Silico* Approach

Shaik Parveen, Natarajan Pradeep, Kanipakam Hema and Amineni Umamaheswari

Abstract — These *Haemophilus influenzae* is a Gram-negative bacterium which causes pneumonia in humans. Due to its multidrug resistance to penicillin, rifampin and polymyxin and adverse effects of existing treatments, the condition became an open challenge for many researchers to discover novel antagonists for the treatment of pneumonia caused by *H. influenzae*. 5 potential sRNA candidates were identified in *H. influenzae* using sRNA predict tool, among them3 were enzymes and 2 were non-enzymes. Among the three identified enzymes exodeoxyribonuclease I of *H. influenzae* was non-homologous to humans was selected as novel drug target in the present study. Exodeoxyribonuclease I is an enzyme involved in the mismatch repair mechanism of *H. influenzae*. The 3D structure of the exodeoxyribonuclease I was modeled based on the crystal structure of 4JRP using Modeler 9v13 and built model was validated using PROCHECK analysis, ProQ and ProSA. The existing eight inhibitors of exodeoxyribonuclease I were searched in 3D ligand database through shape screening against ASINEX database using Phasev3.2 module and structural analogs were docked with exodeoxyribonuclease I in Maestro v9.6 virtual screening workflow, that implements three stage Glide docking protocol. The docking results revealed that 11 leads were having better docking and ΔG scores when compared with the eight existing inhibitors among them lead 1 was having ΔG score of -68.78 kcal/mol. The major significance of the study revealed that lead 1 can obstruct the function of exodeoxyribonuclease I by blocking the replication mechanism in *H. influenzae* which stops the endurance of the bacterial growth. Hence, lead 1 was proposed as therapeutic antagonist in the treatment of pneumonia caused by pathogenesis of *H. influenzae.*

Index Terms— *Haemophilus influenzae*, Pneumonia, sRNA Predict, Exodeoxyribonuclease I, Glide, ΔG score.

1 Introduction

Haemophilus influenzae is a Gram-negative, coccobacillary, facultative anaerobic bacterium belonging to the Pasteurellaceae family [1]. *Haemophilus* "loves heme", more specifically it requires a precursor of heme to grow. Nutritionally, *H. influenzae* prefers a complex medium and requires preformed growth factors present in blood specifically X factor (hemin) and V factor (NAD or NADP). The bacterium grows best at 35-37^{o}C and has an optimal pH of 7.6. The bacterium was erroneously considered to be the cause of influenza until 1933, when the viral etiology of influenza became apparent and is still colloquially known as 'bacterial influenza'. *H. influenzae* is responsible for a wide range of localized and invasive infections [2]. *H. influenzae* was the first free-living organism to have its whole genome sequence. Naturally acquired diseases caused by *H. influenzae* seem to occur in humans like bacteremia, pneumonia, epiglottitis and acute bacterial meningitis [3]. In some cases, it also causes cellulitis, osteomyelitis and infectious arthritis [4]. However, statistics also revealed that this pathogen is also a major cause of lower respiratory tract infections in infants and children in developing countries [5], [2]. The patient records of Sri Venkateswara Institute of Medical Sciences (SVIMS), Tirupati (Rayalaseema region Andhra Pradesh, India) also reported it is one of the most predominant pathogen causing bacterial meningitis [6].

The bacterial gene expression was controlled by a regulator that interacts with either DNA or mRNA for the synthesis of proteins. When the target for regulation is DNA, the gene expression can be regulated in transcription process or when the target is RNA, the gene expression can be regulated in translation process. Moreover, in the both processes sRNAs are the regulators of the bacterial gene expression; hence identification of sRNAs in the whole genome sequence of pathogen was a novel approach for the identification of drug target. sRNAs are small and non-coding RNAs, that have been implicated in regulation of various cellular processes in living system, allowing them to adapt to changing environmental conditions. sRNAPredict tool was used to identify the sRNAs present in *H. influenzae* type b strain KW20Rd genome [6], [7]. The enzyme exodeoxyribonuclease I was selected as a drug target as it is non-homologous to Homo sapiens and also has key role in the bacterial transformation and mismatch repair mechanism of *H. influenzae* leading to cause pneumonia [8], [9], [10].

Eight antibiotics are being used for *H. influenzae* such as penicillin, ceftriaxone, chloramphenicol, rifampicin, ciprofloxacin, meropenem, celecoxibwere revealed from literature search. They are coupled with side effects like allergies and also exhibiting poor pharmacological properties [11], [12]. Because of the multidrug resistance of *H. influenzae* and adverse effects of existing treatment options, open up new challenges for researchers to discover novel lead molecules (inhibitors) for treatment of pneumonia [13], [3], [14]. In the present study, exodeoxyribonuclease I was selected as drug target and potent

Corresponding Author
Dr. Amineni Umamaheswari, Associate Professor& Coordinator of BIF, SVIMS Bioinformatics Centre, Department of Bioinformatics, SVIMS University, Tirupati, A.P., India – 517 507
Email: svims.btisnet@nic.in, Ph: +91- 877-2287727.

lead molecules were identified through systemic strategy of homology modeling, shape screening, molecular docking and free energy calculations.formatting an article or it will be returned for reformatting.

2 Materials and Methods

2.1 Hardware and Software

In the present study molecular modeling, docking, free energy calculations and ADME/T calculations were carried out using Schrödinger 2013 software in HPZ800 workstation [2].

2.2 sRNAPredict

The whole genome sequence of *H. influenzae* was retrieved from the National center for biotechnology information (NCBI). The genome was submitted to sRNAPredict tool. sRNAs are having pivotal role in regulating various cellular processes. sRNAPredict tool defines the supposed intergenic regions present within the coding regions of the genome, thus it reveals the possible sRNA candidates [15]. The retrieved sRNA candidates were checked for non-homology to Homo sapiens using BLASTP. Non human homologous sRNA candidate was taken for KEGG pathway analysis to concern the significance of the target in the organism.

2.3 Homology Modeling and Validation

Exodeoxyribonuclease I of *H. influenzae* sequence was retrieved from the UniProt and modeling was performed using Modeller9v13. Crystal structure of exodeoxyribonuclease I from *Escherichia coli* (4JRP) had 68% of identity was chosen as template. The template crystal structure was aligned with exodeoxyribonuclease I of *H. influenzae* and refined using ClustalX. Twenty models of exodeoxyribonuclease I of *H. influenzae* were generated with Modeller9v13 by spatial restraints of standard protocol. The reliable model with the compatible DOPE (discrete optimized protein energy) score was selected [16]. The structure was further validated using PROCHECK [17], ProSA [18] and ProQ analysis which are efficient tools for evaluating quality of the 3D models generated. PROCHECK was used to check valid stereochemistry of the model and ProSA optimized to find native structure compatibility. ProQ, a neural network based predictor that works on a number of structural features predicts the quality of a protein model [19].

2.4 Active Site Prediction

The template active site residues of exodeoxyribonuclease I of *E. coli* were defined through the residues interacting with 3'-end of the ssDNA. Further, conserved active site residues in template and target were conceived for the grid generation.

2.5 High Throughput Virtual Screening

The eight existing inhibitors of exodeoxyribonuclease I of *H. influenzae* were taken as query to search against ASINEX database through shape based similarity screening. Ligand preparation was done using LigPrep module of Schrödinger. A grid of 10 x 10 x 10 Å was defined around the active site residues of the receptor, exodeoxyribonuclease I. The Epik is a one of the module of Schrödinger suite was employed to enumerate tautomers (10,000 for each ligand) and protonation states [2]. Reactive filters and Lipinski's filters were applied to refine the generated tautomers using QikProp v3.6. All the retrieved hits were docked with exodeoxyribonuclease I of *H. influenzae* using Glide v6.0 [20]. Glide implements three tier docking protocol such as high throughput virtual screening (HTVS), standard precision (SP) and extra precision (XP) docking methods respectively [21], [22], [23]. HTVS followed by SP and further to XP from lower stringency to higher stringency for screening the compounds [24].

2.6 Prime MM-GBSA

As binding free energies were more accurate to define the binding affinity than the docking scores, binding free energies (ΔG) were calculated through Prime/MM-GBSA. ΔG, calculates the ligand binding energies and ligand strain energies for a ligand and the receptor alone. MM-GBSA is a method that combines OPLS-AA molecular mechanics energies (EMM), an SGB salvation model for polar solvation (GSGB), and a non-polar salvation term (GNP) composed of the non-polar solvent accessible surface area and van der Waals interactions [28], [25]. The binding energy is calculated by the following equation

$$\Delta G_{bind} = \Delta E + \Delta G_{solv} + \Delta G_{SA}$$

Where, ΔE is the minimized energies, ΔG_{solv} solvation free energies, ΔG_{SA} is the difference in surface area energy of the exodeoxyribonuclease I-lead 1 docking complex and sum of the surface energies of exodeoxyribonuclease I of *H. influenzae* and lead 1 respectively.

2.7 ADME/T Properties

QikProp v3.6, program was used to calculate ADME/T properties (absorption, distribution, metabolism, excretion and toxicity). Qikprop is quick, accurate tool predicts physically significant descriptors and pharmaceutically relevant properties of organic molecules [26], [27]. Ligprep minimized ligands were given as input for Qikprop module, which provides the ranges of both molecular properties and principal descriptor properties by comparing the molecules with those of 95% of known drugs (Schrödinger 2009).

3.0 Results And Discussion

3.1 Prediction of sRNAs in *H. influenzae*

The whole genome of *H. influenzae* was retrieved from the NCBI. The *H. influenzae* genome has size of about 2.18 MB with GC content of 51.18%. The *H. influenzae* contains 1765

coding genes that are coding for 1610 proteins. Small RNAs (sRNAs) are the non-coding genes which constitute large and heterogeneous bacterial gene expression regulators. sRNAPredict tool was used to define the sRNA candidates present in *H.influenzae* resulted five sRNA candidates. Among the five sRNA candidates, two are proteins and three were enzymes. The three enzymes were subjected to non-homology search against *Homo sapiens* using BLASTP. Exodeoxyribonuclease I was found to be non-homologous to *Homo sapiens* with >30 % identity and also plays major role in the DNA mismatch repair mechanism which is important for replication [28] survival and proliferation of pathogen. Thus exodeoxyribonuclease I was selected for the present study. Hence, designing a novel antagonist can inhibit the function of exodeoxyribonuclease I which further halts the replication in *H. influenzae*.

3.2 Homology Modeling

The generated homology model of exodeoxyribonuclease I of *H. influenzae* using template structure of exodeoxyribonuclease I of *E. coli* (4JRP) were validated using Ramachandran plot and Z-score. A set of 20 exodeoxyribonuclease I of *H. influenzae* models were generated and the model with least density optimization protein energy (DOPE) score of -55987.72 kcal/mol was selected for further validation (Fig. 1).

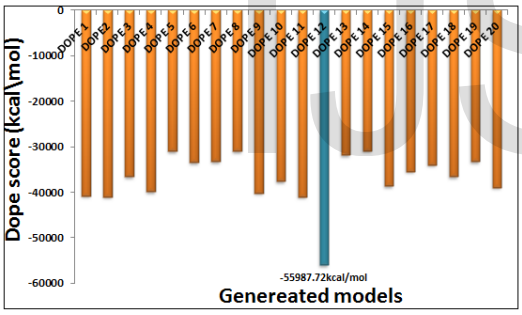

Fig. 1 DOPE scores of 20 homology models of exodeoxyribonuclease I.

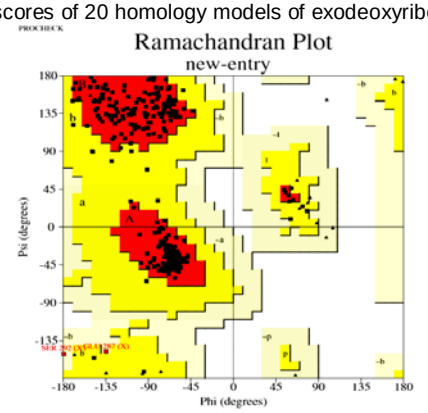

Fig. 2 Ramachandran plot of the best exodeoxyribonuclease I model of *H. influenzae*

The model was further considered for PROCHECK analysis. The PROCHECK results revealed that 95.6% of residues were in the most favorable regions of Ramachandran plot (Fig. 2). ProSA-web analysis had revealed a Z-score value of -10.4 which is found to be in the range of crystal structures native conformations. LG score >4.0 is expected for an extremely good quality model from ProQ model validation server. The exodeoxyribonuclease I model had LG score of 4.168 suggesting that the exodeoxyribonuclease I model was extremely good quality model (Fig. 3).

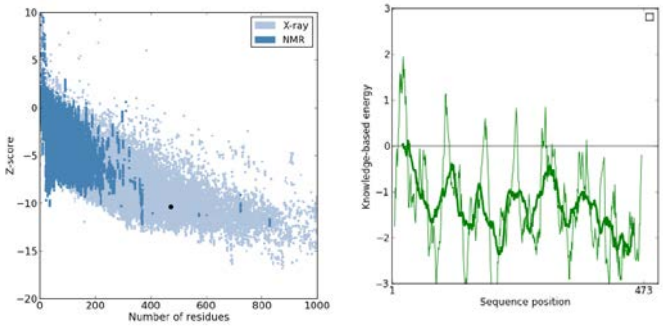

Fig. 3 ProSA result of eodeoxyribonuclease I

3.3 Prediction of Active Site Residues

Korada *et al.*, 2013 reported the template structure of 4JRP domain residues of Asp15, Glu17, Thr18, Thr21, Ala63, Ile66 and Asp186 were bound to 3′-end of the ssDNA of *E. coli*. The active site residues were conserved in the exodeoxyribonuclease I model of *H. influenzae* and correspond to Asp11, Ser14, Glu16, Val17, Ala59 and Asp104 were selected for grid generation.

3.4 Virtual Screening and Binding Free Energy Calculations

The eight known inhibitors of exodeoxyribonuclease I of *H. influenzae* were screened against ASINEX database using Phase v3.2 by shape based similarity screening toretrieve 2612 structurally similar compounds. By removing the high ionization/tautomer states using LigPrep and further 400 conformations were generated. 370 compounds were passed through the Lipinski's filter. After passing through the reactive filters 308 compounds were docked by targeting active site residues of exodeoxyribonuclease I of *H. influenzae* through HTVS docking results revealed 77 compounds. Thus the obtained 77 compounds were taken for SP docking. Further they were subjected to XP docking which revealed 38 best scoring lead molecules. The docking results revealed that 11 leads were having better docking and binding free energies (ΔG) by comparing to the eight existing inhibitors. Among 11 leads, lead 1 is having the lowest ΔG score of -68.78 kcal/mol (Fig. 4). Lead 1 showed the lowest ΔG binding free energy with two hydrogen bonds. The active site residues such as Tyr98, Asn104 are interacted with hydrogen bond formation with lead 1. Residues such as Glu8, Asn11, Tyr12, Glu13, Ser14, Gly16, Val17, Ala59, Val62, Asn99, Tyr103, and His177 of exodeoxyribonuclease I were found to be involved in the van der Waal's interactions.

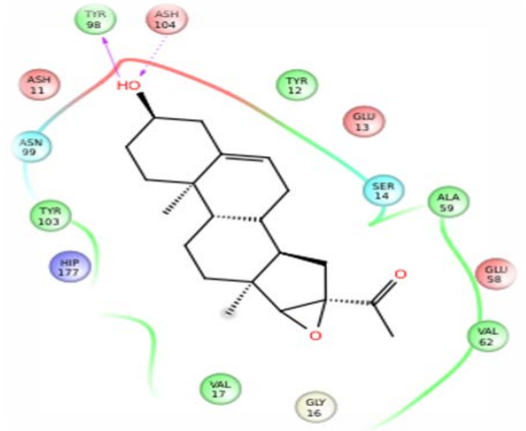

Fig. 4 Molecular interactions of exodeoxyribonuclease I with lead 1

The lead 1 interaction with exodeoxyribonuclease I of *H. influenzae* was well correlated with the active site residues of exodeoxyribonuclease I of *E. coli*. Interaction study of exodeoxyribonuclease I - lead 1 docking complex revealed that, lead 1 has the lowest XPG score and the lowest binding free energy (ΔG) with better binding affinity and enriched binding orientation towards exodeoxyribonuclease I of *H. influenzae*. Lead 1 had no violations for Lipinski's rule of 5, blood brain barrier (BBB) respectively and observed to possess beneficial pharmacological properties.

3.5 Pharmacokinetic Properties

The ADMET properties of lead 1 thrived within the normal ranges without any violations and biologically active without any toxic functional groups. Lead 1 had a good absorption prediction for metabolism. Molecular weight of 330.466 Daltons with 8 rotatable bonds (10 rotatable bonds are allowed). Log P for octanol / water 2.852, blood-brain barrier (BBB) -0.332, suggesting that the lead 1 can cross the BBB to show its activity. Lead 1 showed zero violations for Jorgensen rule of 3 and Lipinski's rule of 5.

Lead 1 showed good pharmacological properties with normal ranges without any violations than the known inhibitors, hence lead 1 might be a new scaffold molecule used to impede the activity of exodeoxyribonuclease I of *H. influenzae*. The present study revealed that lead 1 is having good drug like properties, binding affinity and binding orientation with exodeoxyribonuclease I of *H. influenzae* than the eight existing inhibitors. Lead 1 can act as potential inhibitor of exodeoxyribonuclease I of *H. influenzae* to inhibit its activity and would be useful for treating and developing new therapeutic *H. influenzae* infections.

4 Conclusion

Effective utilization of the various available Bioinformatics tools has enabled the successful identification of putative drug target against *H. influenzae*. The sRNAPredict tool was used in the present study to define the small RNAs present in *H. influenzae*. As the enzymes are important in the metabolic functions, enzymes were considered as drug targets. The non-homology towards *Homo sapiens* revealed exodeoxyribonuclease I as effective drug target against *H. influenzae*. Homology model was generated for the exodeoxyribonuclease I by 4JRP. The molecular docking analysis and binding free energy calculations suggests that lead 1 molecule would be a potent inhibitor against exodeoxyribonuclease I of *H. influenzae*. Eleven lead molecules were identified with better binding affinity when compared to the existing inhibitors of exodeoxyribonuclease I of *H. influenzae*. Among the obtained eleven best leads, lead 1 showed the better binding affinity, binding orientation with strong hydrogen bonds and good van der Waals interactions were observed with the active site residues. Good pharmacological properties were predicted as par with 95% of the existing drug molecule and no reactive functional groups were observed in lead 1. Thus, lead 1 might impede the exodeoxyribonuclease I enzyme activity and may provide a frame work for further development of drugs against exodeoxyribonuclease I of *H. influenzae* for treating pneumonia. The lead 1 molecule identified in the present study has to be further validated by *in vivo* and *in vitro* bioassays to support the findings of this study. These understandings toward the exodeoxyribonuclease I of *H. influenzae* will certainly enable better appreciation and aid in the design of better treatment strategies against the pathogen.

Acknowledgment

Authors are thankful to DBT, Ministry of Science and Technology, Govt. of India, for providing support to SVIMS Bioinformatics center through BIF program (No. BT/BI/25/001/2006).

References

[1] P. Kuhnert, H. Christensen & M. Caister, *Pasteurellaceae*: "Biology, Genomics and Molecular Aspects," *Academic Press,* vol. 24, pp. 0255-0857, 2008.

[2] J. Levy, G. Verhaegen, P. De. Mol, M. Couturier, D. Dekegel & J. P. Butzler, "Molecular characterization of resistance plasmids in epidemiologically unrelated strains of multiresistant *Haemophilus influenzae," J Infect Dis,* vol.168, pp.177-87, 1993.

[3] M. F. Hammer, A. E. Woerner, F. L. Mendez, J. C. Watkins, & J. D. Wall. "Genetic evidence for archaic admixture in Africa". *Clinical Infectious Diseases,* vol. 50: pp. 37-44, 2010.

[4] O.O. Soge & M. C. Roberts, "Characterization of macrolide resistance genes in *Haemophilus influenzae* isolated from children with cystic fibrosis," *J. Antimicrobial. Chemother,* vol. 66, pp. 100–4, 2011.

[5] M. Munikumar, I.V. Priyadarshini, D. Pradhan, A. Umamaheswari, B. Vengamma "Computational approaches to identify common subunit vaccine candidates against bacterial meningitis,. *Interdisciplinary Sciences: Computational Life Sciences*, vol. 5, pp. 155, 2013.

[6] T. Ando, D. A. Israel, K. Kusugami & M. J. Blaser**, "**a member of the DprA family, is involved in natural transformation in Helicobacter pylori," *J Bacterial,* vol**. 181**, pp. 5572-5580, 1999.

[7] M. M. Cox, "Motoring along with the bacterial RecA protein," *Nat Rev Mol Cell Bio,* vol. l8, pp. 127–138, 2007.

[8] **Schrödinger Release 2013-3**: Maestro, version 9.6, Schrödinger, LLC, New York, NY, 2013.

[9] K. Judith Christman Department of Biochemistry and Molecular Biology and UNMC/Eppley Cancer Center, University of Nebraska Medical Center, University Medical Center, Omaha, Nebraska, USA . *Academic Press,* vol. 21, pp. 5483-5495, 2013.

[10] J. I. Rearick, J. E. Sadler, J. C. Paulson & R.L. Hill, "Enzymatic characterization of beta D-galactoside alpha2 leads to 3 sialyltransferase from porcine submaxillary gland," *J. Biol. Chem,* vol. 254, pp. 4444-4451, 1979.

[11] D. B. Kitchen, H. Decornez, J. R. Furr & J. Bajorath. "Docking and scoring in virtual screening for drug discovery methods and application". *Nat Rev drug Discovery*, (), vol. 3, pp. 935-949, 2004.

[12] R. D. Fleischmann, M. D. Adams, O. White, R. A. Clayton, E. F. Kirkness, A. R. Kerlavage, C. J. Bult, J. Tomb, B. A. Dougherty, J. M. Merrick, K. McKenney & G. G. Sutton, "Whole-genome random sequencing and assembly of *Haemophilus influenzae," J. Biol. Chem,* vol.269, pp. 496-512, 1995.

[13] C. M. Chang, T. L. Lauderdale & H. C. Lee, "Colonization of fluoroquinolone-resistant *Haemophilus influenzae* among nursing home residents in southern Taiwan," *J. Hosp. Infect,* vol. 75, pp. 304–308, 2010.

[14] J. Livny, A. Brencic, S. Lory & M. K. Waldor, "Identification of 17 *Pseudomonas aeruginosa s*RNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2," *Nucleic Acids Res,* vol. 34, pp. 3484-3493, 2006.

[15] R.A. Laskowski, M.W. MacArthur, D.S. Moss, J.M. Thornton, "PROCHECK: a program to check the stereochemical quality of protein structures," *J Appl Cryst*, vol. 26, pp. 283–291, 1993.

[16] M. Wiederstein & M.J. Sippl, "ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins," *Nucleic Acids Res*, vol. 35, pp. W407-410, 2007.

[17] B. Wallner & A. Elofsson, "Can correct protein models be identified? Protein Science: A Publication of the Protein Society," vol. 12, pp. 1073-1086, 2003.

[18] A. Umamaheswari, D. Pradhan, I. Priyadarshini & G. Subramanyam, "*In silico* putative drug target in homology modeling of UDP-N-acetylglucosamine 1-carboxyvinyltransferase MurA," *Gen Med*, vol. 2, pp. 295, 2008.

[19] I. Shoichet & J. Bajorath, "Docking and scoring in virtual screening for drug discovery methods and application," *Nat Struct MolBiol*, vol. 3, pp. 935-949, 2004.

[20] M. Munikumar, I. Priyadarshini, D. Pradhan, S. Swargam, & A. Umamaheswari, "T-cell vaccine design for *S. pneumoniae in silico* approach," *Journal of Biomolecular Structure and Dynamics,* vol. 31, pp. 114-115, 2010.

[21] A. Umamaheswari, M. Munikumar, D. Pradhan & M. Hemanth kumar. "Docking studies towards exploring lead molecules aginst binding pocket of yellow fever virus envelop protein," *Interdiciplinary sciences: Computational Life Sciences*, vol. 3, pp. 64-77, 2011.

[22] Yagupsky, Seal Samar & F Richard. "Virtual screening for drug discovery methods and application," Discovery University of Montana, Missoula, Montana, *Nat Rev drug*. Vol.71, pp.1156-1162, 1993.

[23] H. M. Friesner, T. Tokino, V. E. Velculescu., D. B. Levy, R. Parsons, J. M. Trent, D. Lin, W. E. Mercer, K. W. Kinzler & B. Vogelstein. "A potential mediator of p53 tumor suppression," *Bio Cell,* vol. 75, pp. 817–825, 2004.

[24] J. Halgren, A. Brencic, S. Lory & M. K. Waldor, "Identification of 17 *Pseudomonas aeruginosa* sRNAs and prediction of sRNA-encoding genes in 10 diverse pathogens using the bioinformatic tool sRNAPredict2," *Nucleic Acids Res,* vol. 34, pp. 3484-93, 2004.

[25] J. Vivek & M. J. Blaser, "Comparative susceptibilities of ampicillin and chloramphenicol resistant *Haemophilus influenzae* to fifteen antibiotics," J *Anti microb Chemother,* vol. 19, pp. 297-301, 2008.

[26] P. Veal, J. P. Prendergast, M. Wagner, R. Nissen, D. E. Reeves, D. P. Pettijohn & Edwards, "The DNA-bending protein HMG-1 enhances progesterone receptor binding to its target DNA sequences," *Mol. Cell Biol*, vol. 5, pp. 3376–3391, 1998.