In silico Identification to Potential Drug siRNA Using Statistical Methods in *Helicobacter pylori*

Al Rikabi, Z. H.

Institute of Genetic Engineering and Biotechnology for Postgraduate Studies, University of Baghdad, Iraq

Key words: *In silico* study, siRNA, Target design, and DEG

Received (June), **Accepted** (December)

ABSTRACT

Helicobacter pylori a causative influence risk of carcinogenesis which has developed resistance to drugs. The present study was carried out to identify potential drug targets in *H. pylori* by design of potent siRNA. To reach the goal, the collection essential gene for *H. pylori* and host from DEG, then a variety of bioinformatics databases studying Blastp Database of essential genes, Blastp from NCBI and KEGG, for selection non-homologous genes, then the software OligoWalk to design siRNA. The study concluded to get seven (7) essential membrane protein of H. pylori to become drug target, then subjected to find siRNA which are shorter than 21 nt about 2423 siRNA, 110 siRNA (Effective siRNA rules), 56 siRNA (ratio GC (35%-60%)), 50 siRNA (Desired position), 30 siRNA (Probability range between (0.70-0.91)), 20 siRNA (Secondary structures), 8 siRNA (Drug target), this steps respectively. Conclusion: Eight (8) siRNA which could use as potential drug target for resistant of *Helicobacter pylori* 26695.

دراسه حاسوبيه لتشخيص التواليات الدوائية siRNA الاكفأ باستخدام طرق احصائية لبكتريا. *Helicobacter pylori*

زينب هاتف عباس

معهد الهندسة الوراثية والتقنيات الاحيائية للدراسات العليا – جامعه بغداد - العراق

الكلمات المفتاحية: دراسه حاسوبية، siRNA، تصميم الدواء،قاعده البيانات لجينات الاساسية

الخلاصة

بكتريا *Helicobacter pylori* من المسببات المرضيه الخطرة لاحداث التسرطن والتي تطورت لمقاومة العقاقير. وقد أجريت هذه الدراسة لتحديد أهداف دوائية لـ *H. pylori* بتصميم تواليات صغيره قويه التاثير من siRNA. للوصول للهدف، تم جمع الجينات الاساسية لـ *H. pylori* والمضيف (الإنسان) من قاعده بيانات ،وتم استخدام العديد من برامجيات المعلوماتية الحيوية blastp من قاعده البيانات DEG و NCBI وباستخدام Blastp كما استخدمت KEGG لاختيار الجينات غير المتجانسة. وتم استخدام برنامج OlioWalk لايجاد siRNA واستخلصت الدراسه للوصول الى سبعه من البروتينات الغشائية الاساسية لـ *H.pylori* لتصبح اهدافا دوائية، ثم اخضعت لايجاد siRNA والتي تكون أقصر من 21 نيوكلتيد والبالغ عددها siRNA 2423، siRNA110 (خاضعة قواعد تصميم siRNA الفعال)، 56 siRNA (ضمن حدود GC٪ المطلوبه (35%-60%))، و 50 siRNA (ضمن الموقع المطلوب)، 30 siRNA (ذات احتماليه عاليه تراوحت بين (0.70-0.91)، 20 siRNA (الهياكل الثانوية)، 8 siRNA (اهدافا دوائية) على التوالي. الاستنتاج: ثمانية تواليات صغيره من نيوكلتيدات (siRNA) تمثل اهدافا دوائية لمقاومه بكتريا *Helicobacter pylori 26695* .

1. INTRODUCTION

Helicobacter pylori is a microaerophilic, gram-negative bacteria that colonizes the gastric mucosa of approximately 50% of the world's population [1]. It is a primary factor in human diseases such as duodenal/gastric ulcer disease, gastritis, gastric adenocarcinoma, and

mucosa-associated tissue lymphoma and cell gastric [2, 3, 4]. When we reviewed the relationship between H. pylori and gastrointestinal tract adenocarcinomas. Many Reported that the complete sequence of the circular genome of H. pylori concluded that genetic effects influence the acquisition of H. pylori infection but if that sharing the same rearing environment also contributes to the familial tendency [1].

Although gastric adenocarcinoma is associated with the presence of H. pylori in the stomach, only a small fraction of colonized individuals develops this common malignancy. The H. pylori strain and host genotypes probably influence the risk of carcinogenesis [5]. H. pylori has been defined as a class carcinogenic factor by the world health organization (WHO) and its persistent colonization in the stomach leads to an increased risk of peptic ulcers and gastric adenocarcinoma [6]. In most cases the infection established during childhood and persists lifelong. This chronic infection leads to continuous low level gastric inflammation, which is asymptomatic in most cases. However, some people develop more pronounced inflammation, which can lead to peptic ulcer disease or gastric cancer [7]. The critical genes crucial for the survival of pathogen and absent in the host are also identified using the subtractive genomics approach. The critical genes required using Database of Essential Genes (DEG). To find the chances of cross reactivity and side effects to genes and their products which can be used as potential drug targets are also identified by analyzing these genes with the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database. Such approach will ensure that the drug target is available only in the pathogen and not in the human, by using successfully for various pathogens [8].

Long double stranded RNA molecules can be employed to induce RNA (RNAi) in lower eukaryotes, siRNAs being (21-19) nucleotides in length have to be used for gene silencing in mammalian cells in order to prevent the activation of an unspecific interferon response [9]. RNA interference (RNAi) is a naturally occurring phenomenon of RNA mediated gene silencing that is highly conserved among multicellular organisms [10]. Aim of study, designing better and effective drugs against resistant H. pylori by using small interfering RNA (siRNA).

2. MATERIALS AND METHODS

Samples Collection

Collection of amino acid sequence of human 118 essential protein and 323 essential protein for Helicobacter pylori 26695 using Database of Essential Genes (DEG) (http://tubic.tju.edu.cn/deg/)[11] was performed to identify the critical genes necessary for survival of the H. pylori 26695 and human.

Bioinformatics Programs

A: Compare DEG (Human Essential Genes)

Aliment 118 amino acid sequence of human, with 323 proteins of Helicobacter pylori 26695 using Blastp /DEG (http://tubic.tju.edu.cn/deg/blast.php?&db=e)[11] with expectation Value (E value) 10-5 [7].

B: NCBI Blast (National Center Biotechnology Information) Human Genome Compare 55 essential protein non-homologues of H.pylori gene from high step, by Blastp/ NCBI human genome (http://blast.ncbi.nlm.nih.gov/genome/seq/BlastGen.cgi?taxid=9606) with expectation value (E-value) 10-5. Determine non-homologues H. pylori essential gene against human genome [6].

C: KEGG (Kyoto Encyclopedia of Genes and Genomes) Human Metabolism Metabolic pathway analysis of essential proteins of H. pylori was done by Kyoto Encyclopedia of Genes

and Genomes (KEGG) pathway database[12] (Automatic Annotation Server) (www.genome.jp/tools/kass/). KAAS provides functional annotations of genes by Blast comparison against the manually curated KEGG genes database. The result contains KO (KEGG Orthology) assignments and automatically generated KEGG pathways [12]. To search the homologous proteins between H.pylori and human, with (E- value) 10-5.

D: Membrane proteins Screening of 30 proteins drug targets to identify the surface membrane proteins by using Trans Membranous Hidden Markov Model (TMHMM), (http://www.cbs.dtu.dk/services/TMHMM) to choice membrane proteins [13].

E: siRNA (small interfering RNA) Membrane proteins (That best drug target proteins) [14, 15] from *H.pylori*, can be use this proteins to design efficient siRNA by software OligoWalk (http://rna.urmc.rochester.edu/servers/oligowalk) from their RNA[16, 17]. The default siRNA candidate is an RNA oligonucleotide having 19 nucleotide.

G: Consensus Format for siRNA Sequences

The problem of choosing target siRNA sequences is to predict whether or not a candidate siRNA sequence for the target mRNA (typically 19 nucleotides) will result in effective gene silencing, so that can be transformed into the problem of finding the degree of gene silencing functionality of a given siRNA candidate, therefore reported eight criteria for select the optimal functional siRNAs to improving siRNA selection.

1. **Designing siRNA Sequences**[18, 19].

The earliest way in which the parts of target mRNA are arranged effectively by siRNA duplexes 21 nucleotides long: 19-nt base- paired sequences with 2-nt overhangs at the 3' ends [18]. Many siRNA design guidelines/rules have been reported since then, and we treats the following three rule (A, B and C) [19].

A: (1) at least 3 As or Us at positions 15–19, (2) absence of internal repeats, (3) an A at position 19, (4) an A at position 3, (5) a U at position 10, (6) a base other than G or C at position 19, (7) a base other than G at position 13.

B: (1) an A or U at position 19, (2) a G or C at position 1, (3) at least five U or A residues from positions 13 to 19, (4) no GC stretch more than 9 nt long.

C: (1) a U at position 19, (2) a C or G at position 11, (3) a G at position 16, (4) an A at position 13.

These guidelines are summarized in (Table 1).

Position	**1**	**3**	**6**	**10**	**11**	**13**	**16**	**19**
A		A		T		A/C/U		A/U
B	G/C							A/U
C					C/G	A	G	U

Table 1: Effective nucleotides specified in the individual guidelines [18, 19].

2. **GC Content 35-60% [20].**

Target sequences should have a G+C content between 35-60%.

3. Position

It has already been reported that potent siRNA sequences tend to have a less stable 5' end refer to the nucleotides in the antisense guide strand below with position in the sequence order in the 5' 3' direction on the guide strand [19, 21].

4. Probability [22, 23]

The great and important problem of selecting target siRNA sequences is to predict whether or not a candidate siRNA sequence for the target mRNA will be result in effective gene silencing, therefor it can be used to calculate the probability of individual nucleotide occurrence frequencies (typically 19 nucleotides X=X1, X2, ……, X19 where Xi is the i-th nucleotide) that candidate siRNAs will be effective this is because the effectiveness of siRNA sequences greatly depends on individual positions siRNAs of the sequences so that the evaluation of candidates can be based on the analyses of nucleotide presence features. The probability of individual nucleotide occurrences at positions from 5' to 3' in the RNA sites from 1 to 19 in the effective siRNA population is obtained, can be computed as follows [22, 23]:

$$\mathbf{f_p^N} = \frac{\sum_{i=1}^{I}\{A\ G\ C\ U\}}{I}$$

Where, N is the kind of nucleotide (A, G, C, or U), p is the position in the mRNA (1, 2, …,19 from 5' to 3'), and i is the number of the effective siRNA sequences.

Then the probability OF_i of each effective siRNA sequence is calculated in the following way[22, 23].

$$\mathbf{OF_{i=}}\ \prod_{f=1}^{19} f_{ip}^{N}$$

Where, i is the sequence identification number of the effective siRNAs (i.e., i=1, 2… I). The average sequence probability A_E for the effective siRNAs is therefore computed as follows[22, 23]:

$$\mathbf{A_E} = \frac{\sum_{i=1}^{I} OFi}{I}$$

A_E could be considered a criterion for candidate siRNA, that is, if the probability of the candidate sequence were greater than A_E, it would indicate a high likelihood of gene-silencing. On the other hand, if the probability of the candidate sequence were remarkably lower than A_E, it would indicate a low likelihood of effectiveness.

5. Secondary structure

The secondary structure for mRNA very important rule in medicine and treatment so that it's become one of beset rule to choosing target siRNA [24]. The folding energy of the siRNA sequence, calculated as in [25], is another factor, which seems to affect siRNA potency. For further clarification a flow chart shows complete methodology used for choosing of Potential siRNA molecules in this study (Figure 1).

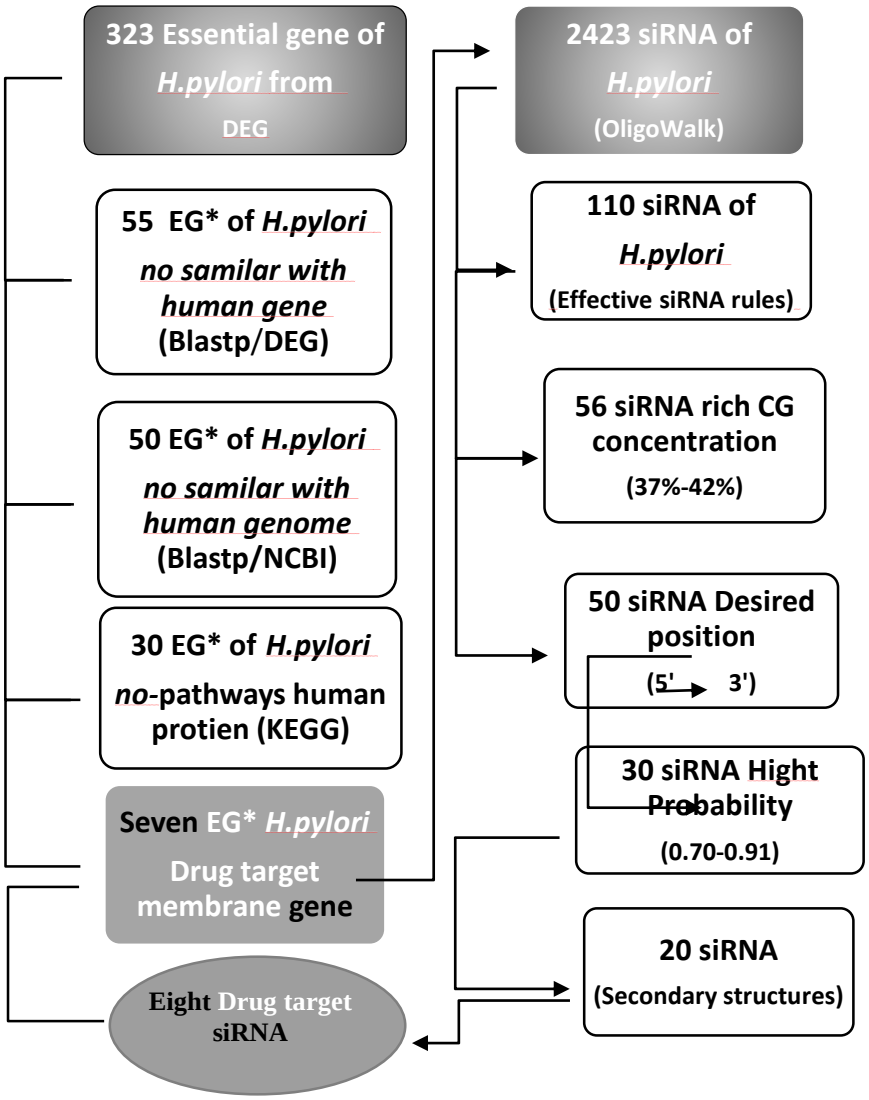


Figure 1: Flow chart showing complete methodology used for choosing of Potential siRNA in this study. *****EG (Essential Gene)

3. RESULTS AND DISCUSSION

In the present study human non-homology essential gene of H. pylori as well as their protein, products have been identified using drug target that strongly bind with the pathogen by using In Silico drug target identification. To choose the good, by using many bioinformatics programs, including Blastp of DEG (http://tubic.tju.edu.cn/deg/)[11], that compere are 118 essential protein with design 323 for H. pylori, appear 55 protein non-human homologous essential gene of H.pylori, the remaining 50 proteins were analyzed used Blast against human using NCBI Blastp human genome (http://blast.ncbi.nlm.nih.gov/genome /seq/BlastGen.cgi?taxid= 9606) [6] this proteins non-significant similarity with human protein. These proteins were analyzed using KEGG pathway database (http://www.genome.jp/kegg/pathway.html) from KEGG database (http://www.genome.jp/kegg/) [12] to secretion of the protein drug target by remove homologous essential gene metabolism with human and choice 30 non homologous essential

genes of H.pylori with human. Additionally the subcellular localization of the 30 protein of H.pylori by using TMHMM [13] to identify the membrane gene that (a good drug target) [14, 15] we get only seven essential genes that membrane cell membrane.

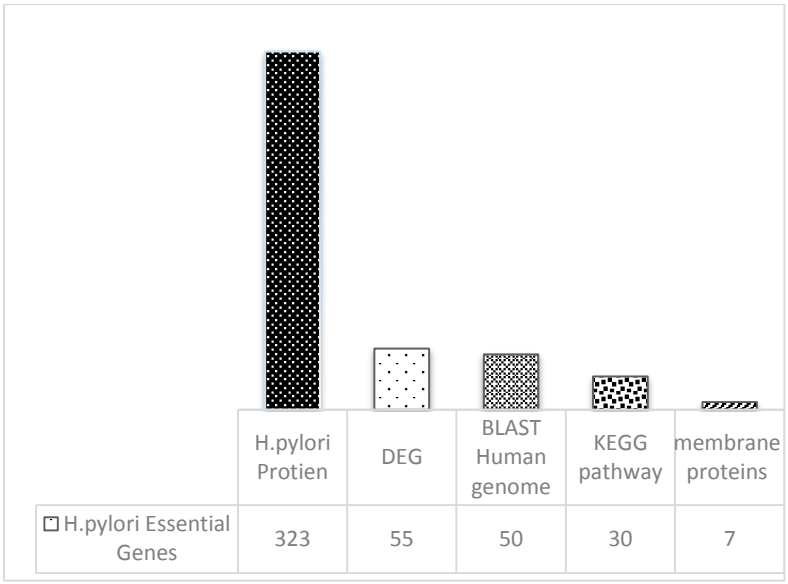


Figure 2: Summary of Target Identification, total genes 323 of *H.pylori*, 55 of DEG non homologous with human DEG, 50 of non homologous with human genes by blastp , 30 of KEGG pathway, 7 of membrane protein position represent best drug target.

We will review the genes that are considered good drug targets in this study with their names and accession number DEG are: (DEG10080047 (rfaC/waaC), DEG10080052 (dppA), DEG10080062 (ycf5), DEG10080124 (HP0746), DEG10080224 (HP1216), DEG10080247 (HP1289) and DEG10080323 (HP1580). And appear 2423 siRNA having 19 nucleotide when input the seven membrane gene of H.pylori to OligoWalk (http://rna.urmc.rochester.edu/servers/oligowalk) [16, 17] filtering inter 2423 siRNA to choose the best drug target, appear that about (110) siRNA sequence have effect design siRNA sequences rule [18, 19] Shawn in (Figure 3).

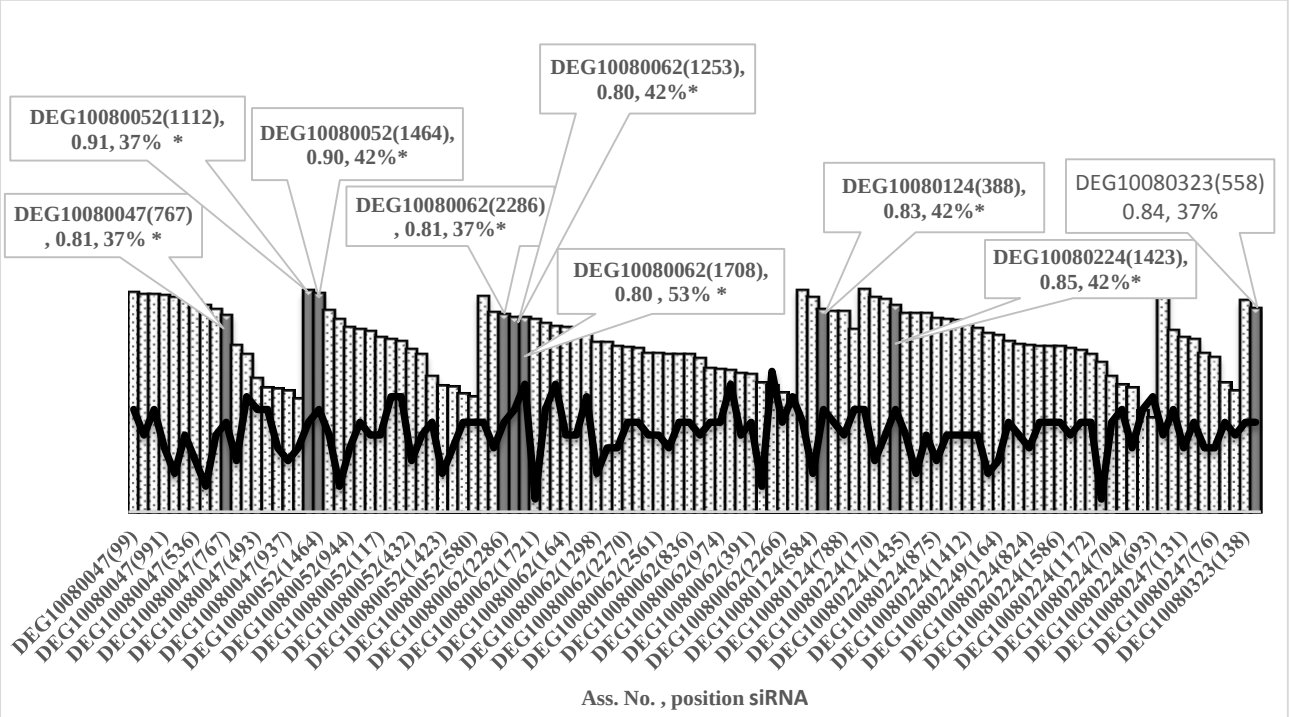
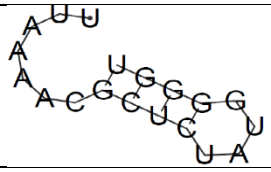
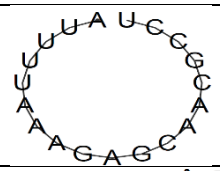
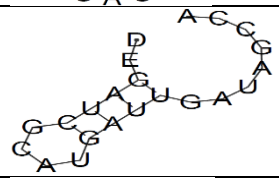
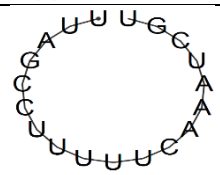
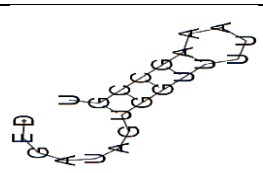
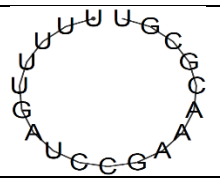
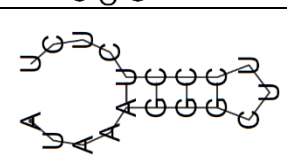
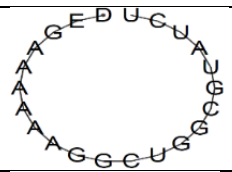
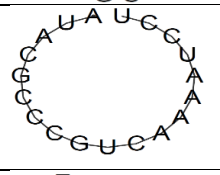
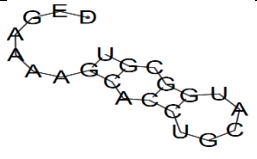
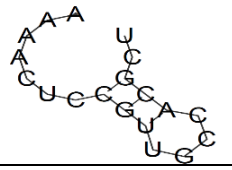


Figure (3): siRNA have effect design appear probability and GC% to siRNA for 110 siRNA, *black columns represent to target siRNA to *H.pylori*, other columns represent output of siRNA.
*Ass. No (DEG) gene, Starting target position, probability siRNA in seq., GC%

Next, putative target sequences are filtered based on GC-content, another strong determinant of siRNA potency. A range of 30–65% GC [20] is considered optimal for identifying effective siRNAs and is generally used among potency-based siRNA design algorithms. To improve our yield of siRNA with a potential for high specificity. Our evaluation of the siRNA database supports that siRNAs within this range exhibit a suitable potential for efficient silencing of >80%. The Deleting siRNA sequence that have low GC percentage 30 % [27, 28] appear the result about 56 siRNA rich CG concentration, than remove six sequence depended on location 150 near from ‘5, because the best siRNA that near from ‘3 [29] and output of 30 siRNA by their probabilities being efficient siRNA because the best of probability having silencing efficacy larger than 0.7 [18] and result about 20 siRNA. We obtained Refseq coding sequences for the 20 therapeutically relevant gene targets (Table 2). These sequences were used as input at the siRNA online web sites ViennaRNA (http://rna.tbi.univie.ac.at/cgi-bin/RNAfold.cgi)[30] to find secondary structures siRNA for deletion of hairloop include 19 nucleotide as shown in (Table 2),

No	siRNA Sequence(5'->3')	Ass. No. of gene (start Position on target)	Probability	GC %	secondary structure
1	UUAAAACGCUCUAUGGGGU	DEG10080047(608)*	0.89	42%	

2	AUUUUAAAGAGCAACGCCU	DEG10080047(767)*	0.81	37%	
3	AUCGCAUGAUUGAUAGCCA	DEG10080052(983)*	0.70	47%	
4	UUAGCCUUUUUCAAAUCGU	DEG10080052(1112)*	0.91	37%	
5	AUAGUGGUUUUAAAGCCGU	DEG10080052(1160)*	0.75	37%	
6	UUUUUGAUCCGAAACGCGU	DEG10080052(1464)*	0.90	42%	
7	AUAAAGGGCUUCCCUCUCU	DEG10080052(1482)*	0.71	47%	
8	AAAAAAGGCUGGCGUAUCU	DEG10080062(235)*	0.77	42%	
9	AUACGCCCGUCAAAAUCCU	DEG10080062(1253)*	0.80	42%	
10	AAAAGCACCUGCAUGGCGU	DEG10080062(1343)*	0.76	53%	
11	AAAACUCCGUUGCCACGCU	DEG10080062(1708)*	0.80	53%	

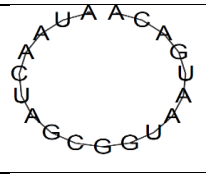
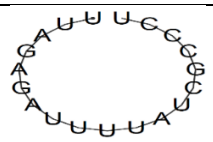
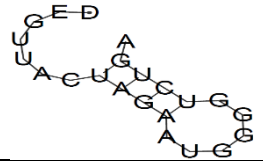
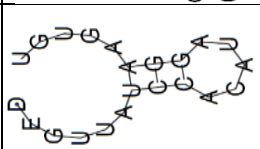
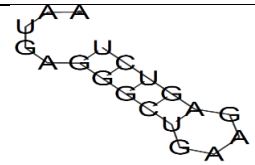
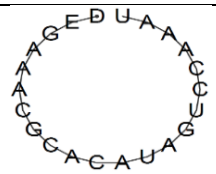
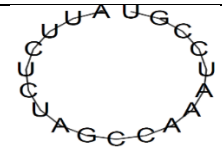
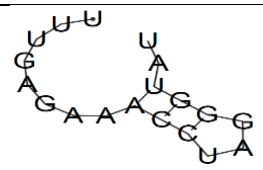
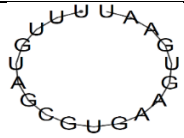
12	AUAACUAGCGGUAAUGACA	DEG10080062(2286)*	0.81	37%	
13	UUAGAGAUUUUAUCGCCCU	DEG10080124(365)*	0.83	37%	
14	UUACUAGAAUGGGGUCUGA	DEG10080124(388)*	0.83	42%	
15	UUAUCCACAUAGGAAGUGU	DEG10080124(857)*	0.91	37%	
16	AAUGAGGGCUGAAGAGUCU	DEG10080124(1020)*	0.75	42%	
17	AAACGCACAUAGUCCAAAU	DEG10080224(950)*	0.70	37%	
18	AUUCUCUAGCCAAAUCCGU	DEG10080224(1423)*	0.85	42%	
19	UUUGAGAAACCUAGGGUAU	DEG10080247(240)*	0.71	37%	
20	UUUGUAGCGUGAAGUGAAU	DEG10080323(558)*	0.84	37 %	

Table 2: List of high probability, GC and secondary structure of siRNA sequence (by ViennaRNA[30]) for *H.pylori*
* Position siRNA on target

After removed the siRNA include hairloop for example: siRNA starting at position 608 to gene DEG10080047, siRNA for DEG10080052 starting at 1482, DEG10080062 (1708), DEG10080124 (388), siRNA to DEG10080247 begin with 240 nt and other, these screening of siRNA to reach best drug targets, shown in
(Table 3).

Table (3): List of potential target siRNA (best probability and GC%).

No.	siRNA Sequence(5'-3')	Ass. No. of gene (Position on target)	No. Nucleotide for gene	Probability of being efficient siRNA	GC%
1	AUUUUAAAGAGCAACGCCU	DEG10080047(767)	1023 bp	0.81	37%
2	UUAGCCUUUUUCAAAUCGU	DEG10080052(1112)	1650 bp	0.91	37%
3	UUUUUGAUCCGAAACGCGU	DEG10080052(1464)	1650 bp	0.90	42%
4	AUACGCCCGUCAAAAUCCU	DEG10080062(1253)	2811 bp	0.80	42%
5	AUAACUAGCGGUAAUGACA	DEG10080062(2286)	2811 bp	0.81	37%
6	AAACGCACAUAGUCCAAAU	DEG10080224(950)	1983 bp	0.70	37%
7	AUUCUCUAGCCAAAUCCGU	DEG10080224(1423)	1983 bp	0.85	42%
8	UUUGUAGCGUGAAGUGAAU	DEG10080323(558)	663 bp	0.84	37%

We remove two of siRNA that have non-hairloop secondary structures one starting of them on gene (DEG10080062), where gene has 2811 bp and small interference RNA beging at position 235 near from ‘3 although it has 42% GC with 0.77 probability but we've raised it from our results and siRNA at position 365 to DEG10080124 the same thing show in Table (2).

5. CONCLUSION

The *In Silico* study based approach involves a series of screening, scoring and designing functional siRNA of essential gene of *H.pylori* that can be used as potential drug target the result of this study is designing of siRNAs per gene with (Ass. No. DEG10080047) starting at position 767nt, for gene DEG10080052 get tow siRNA one starting position at 1112 nt anther starting at 1464nt, the same case for gene DEG10080062 at (2286, 1253), and also for DEG10080224 ane at position 1423 anther one at 950 where finally we see the siRNA for gene DEG10080323 that starting at position (558) from 3’ prime for resistance of disease that cause *H. Pylori*.

ACKNOWLEDGEMENTS

Author is thankful Dr. Hassan Mahmood Abo Almaali to sharing his knowledge, and all who participates in this work.

REFERENCES

[1] Kumari, T. P.; Rani, P.; Amma, V.; Shamya, S. and Sundar, S.K. (2014). *In Silico* Identification drug targets of *Helicobacter pylori*. *International Journal of Research in Engineering and Bioscience*, 2: 1-13.
[2] Deng, W.; Ping, H.; Jeng, W.; Antone, R.; Chao, K.; Chen, W.; Sophie, S.; Angela, C.; Wen, H. and David Y. (2010). Sequential and Concomitant Therapy with Four Drugs Is Equally Effective for Eradication of H pylori Infection. *Clinical Gastroenterology and Hepatology*, 8:36–41.

[3] Patrick, P. B.; Krishna, R.; Tommy, P.; Niraj, J.; Nithya, K. and Molly, F. (2011). A Randomized Study Comparing Levofl oxacin, Omeprazole, Nitazoxanide, and Doxycycline versus Triple Therapy for the Eradication of Helicobacter pylori. *Am J. Gastroenterol* , 106(Nov.):1970–1975.
[4] Wenming, W.; Yunsheng, Y. and Gang, S. (2012). Recent Insights into Antibiotic Resistance in Helicobacter pylori Eradication. *Gastroenterology Research and Practice*,2012:1- 8
[5] Juozas, K.; Thomas, W.; Alexander, L.; Marcis, L.; Indre, B.; Ruta, S.; Simonas, J.; Ugne, G.; Audrius, I.; Guntis, A.; Vitalija, P.; Jurgita, S.; Limas, K. and Peter, M. (2014). Gene Polymorphisms of Micrornas in Helicobacter pylori -Induced. *PLOS ONE*, 9(1): e87467
[6] Dapeng, Z.; Zhifang, L.; Jian, D.; Wenjuan, Li. and Yundong, S. (2010). *Helicobacter pylori* CagA up regulation of CIP2A is dependent on the Src and MEK/ERK pathways. *Journal of Medical Microbiology*, 59: 259–265.
[7] Katherine, W. C; Darren, P. L. and Richard, J. M. (2014). CCL20/CCR6-mediated migration of regulatory T cells to the *Helicobacter pylori*-infected human gastric mucosa. *Gut*, 0:1–10.
[8] John, J.G. and Valentinia, U. (2011). *In silico* identification of putative drug targets in Klebsiella pneumonia MGH78578. *Indian Journal of Biotechnology*, 10: 432- 439.
[9] Jonathan, R. L.; Allison, K. W. and Ricarda, A. S. (2006). The Mutational Consequences of Plant. Transformation. *J. Biomed. Bioethanol.* , 2006: 1–7.
[10] Ryther, R.; Flynt, A.S.; Phillips, J.A. and Patton, J.G. (2005). siRNA therapeutics: big potential from small RNAs. *Gene Therapy*, 12(1): 5–11.
[11] Luo, H.; Lin, Y.; Gao, F.; Zhang, CT. and Zhang, R. (2014). DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic Acids Res., 42:D574-80.
[12] Altermann, E. and Klaenhammer, T.(2005). PathwayVoyager: pathway mapping using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database. *BMC Genomics* 2005, **6**:60.
[13] Krogh, A.; Larsson, B.; von Heijne, G. and Sonnhammer, EL. (2001). Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. J Mol Biol., 305(3):567-80
[14] Barh, D.; Kumar, A. and Misra, A. (2009).Genomic Target Database (GTD):A database of potential targets in human pathogenic bacteria. *Bioinformatics*, 4: 50-51.
[15] Dutta, A.; Singh, S.; Ghosh, P.; Mukherjee R.; Mitter S. and Bandyopadhyay D. (2006).In Silico identification of potential therapeutic targets in the human pathogen Helicobacter Pylori. *In Silico Biol*., 6: 43-47.
[16] Mathews, D. H. (2010). Using OligoWalk to identify efficient siRNA sequences. *Methods Mol Biol.*, 629: 109-21.
[17] Stanislav, B.; Jessica, S.; Matthew, G. and David, H. (2013). RNA structure: web servers for RNA secondary structure prediction and analysis. *Nucleic Acids Research*, 41: W471–W474.
[18] Jean-Philippe, V.; Nicolas, F. and Yves, V. (2006). An accurate and interpretable model for siRNA efficacy prediction. *BMC Bioinformatics*, 7 (520): 1471-2105.
[19] Shishir, K.; Bashir, A.; Mugdha, S. and Shailendra, K. (2010). A Novel Algorithm to Design an Efficient siRNA by combining the Pre Proposed Rules of siRNA Designing. *J Comput Sci Syst Biol*, 3(1): 005-009

[20] Matveeva, O.; Nechipurenko, Y.; Rossi, L.; Moore, B.; Saetrom, P.; Ogurtsov, A.Y.; Atkins, J.F and Shabalina, A.(2007). Comparison of approaches for rational siRNA design leading to a new efficient and transparent method. *Nucleic Acids Res*., 35(8):e63.
[21] Chia-Ying, C. and Tariq, M. R. (2008). Potent RNAi by short RNA triggers. *RNA*, 14 (9): 1714–1719.
[22] Junichi, Y.; Takayuki, M. and Haruhiko, F. (2011). siRNAs Induce Efficient RNAi Response in Bombyx mori Embryos. *PLoS One*, 6 (9): e25469.
[23] Reynolds, A.; Leake, D.; Boese, Q.; Scaringe, S. and Marshall, W.S. (2004). Rational siRNA design for RNA interference. *Nat Biotechnol*, 22: 326-330.
[24] Takasaki S. (2013). Methods for selecting effective siRNA target sequences using a variety of statistical and analytical techniques. *Methods Mol Biol*., 942:17.
[25] Jagla, B.; Aulner, N.; Kelly, P.; Song, D. and Volchuk, A. (2005). Sequence characteristics of functional siRNAs. *RNA*, 11: 864-872.
[26] Peek, A.S. (2007). Improving model predictions for RNA interference activities that use support vector machine regression by combining and filtering features. *BMC Bioinformatics*, 8: 182–201.
[27] Dieter, H.; Joerg, L.; Craig, M.; Jan, W. and Fred, A. (2005). Design of a genome-wide siRNA library using an artificial neural network. *Nature Biotechnology*, 23 (8): 995-1001.
[28] Jochen, W. K.; Loukas, M. and Chris, H. (2009). Approximate Bayesian feature selection on a large meta-dataset offers novel insights on factors that effect siRNA potency. *Bioinformatics*, 25 (13): 1594–1601.
[29] Olga, M.; Yury, N.; Leo, R; Barry, M; Pal, S.; Aleksey, Y. O.; John, F. A. and Svetlana, A. S.(2007). Comparison of approaches for rational siRNA design leading to a new efficient and transparent method. *Nucleic Acids Research*, 35: e63.
[30] Lorenz, R.; Bernhart, S.H.; Siederdissen, C.H.; Tafer, H.; Flamm, C.; Stadler ,P.F. and Hofacker, L. (2011) Vienna RNA Package 2.0. *Algorithms for Molecular Biology*, 6:26.