**Original Research Article**

DOI: 10.26479/2018.0406.23

**ANALYSIS OF *HELICOBACTER PYLORI* STRAIN 26695  
PROTEOME FOR PROBABLE TARGETS**Devo kanta pegu<sup>1\*</sup>, Madhusmita Bhagawati<sup>2</sup>, Rasana Paul<sup>1</sup>,Dipshikha Gogoi<sup>1</sup>, Tazid Ali<sup>3</sup>, Rajeev Sarmah<sup>4</sup>

1. Centre for Biotechnology and Bioinformatics, Dibrugarh University, Dibrugarh, Assam, India.
2. Department of Microbiology, Srimanta Sankardeva University of Health Sciences, Guwahati, Assam, India.
3. Department of Mathematics, Dibrugarh University, Dibrugarh, Assam, India.
4. Faculty of Science (Biotechnology), Assam Down Town University, Guwahati, Assam, India.

**ABSTRACT:** A comparative study of protein sequences of *Helicobacter pylori* with that of human and lactobacillus proteins was performed. The protein sequences of *Helicobacter pylori* which show no significant similarity with human and lactobacillus protein sequences were considered to be the probable targets. Out of 1590 protein sequences of *H. pylori* 1407 and 1409 had shown similarity at various levels to the proteins of Homo sapiens and lactobacillus species respectively. Two sets consisting of 183 and 181 dissimilar proteins were sorted. On union of these two sets 81 proteins were found to be in the intersection. However, of 81 sequences, excluding hypothetical protein sequences 18 protein sequences in this study were considered to be the candidate targets. A literature survey was conducted for the proteins considered as probable target and it was found that of the 18 proteins 13 were already considered as target. The rest five proteins may be considered as new probable targets these are Phospholipase A1, Flagellar basal body P-ring protein, NADH\_ubiquinone oxidoreductase subunit J, Siderophore-mediated iron transport protein (tonB) and Relaxase. Among these logical explanation illustrated in result section Phospholipase A1 can be considered as one of the potential targets.

**KEYWORDS:** *Helicobacter pylori*, Homo sapiens, lactobacillus, Phospholipase A1, Simulation.

**Corresponding Author: Dr. Devo kanta pegu\*** Ph.D.

Centre for Biotechnology and Bioinformatics, Dibrugarh University, Dibrugarh, Assam, India.

Email Address: devo.kanta.pegu@gmail.com

## 1. INTRODUCTION

*Helicobacter pylori* cause chronic persistent and atrophic gastritis in adults and children that often culminate the development of gastric and duodenal ulcers. Studies indicate that infected individuals have 2-6 fold increased risk of developing gastric cancer and mucosal associated lymphoid tissue lymphoma compared to their uninfected counterparts. *Helicobacter pylori* have a worldwide distribution and are associated with the pathogenesis of various diseases of the digestive system [1]. *H. pylori* is associated with diseases such as gastritis, peptic ulcers, and carcinoma[2]. Despite the fact that the optimal treatment for *Helicobacter pylori* infection has not been found, there are several combinations of drugs which can cure the infection in up to 80-90% of cases[3]. It is known that chronic gastritis caused by this infection may lead to atrophic gastritis, intestinal metaplasia and dysplasia, thus setting the role of *H. pylori* in gastric carcinogenesis cascade [4,5]. In fact, recent molecular studies showed high levels of *H. pylori* resistance to some antibiotics that are routinely used in eliminating schemes, especially clarithromycin and fluoroquinolones[5]. Molecular tests conducted from gastric biopsies have been used in molecular biology laboratories, showing high levels of sensitivity and specificity[6]. The main *H. pylori* resistance mechanism results from the acquisition of point mutations[7]. *H. pylori*'s resistance to the main drugs used in therapy have been frequently evaluated; however, in countries like Brazil, studies that have examined the association between the presence of point mutations and resistance to *H. pylori* are often restricted to clarithromycin[8]. In terms of the association between the presence of resistance to drugs and the gender of patients infected with *H. pylori*, studies have shown that women are more likely to have high rates of drug resistance to the treatments administered as part of *H. pylori* therapy[9]. This work is laid down to determine the protein 3D structure of the most vulnerable target proteins (structure not determined by experimental technique) so as to add 3D structure as well as to understand the structure-function relationship of the probable target proteins by molecular dynamics analysis. Presently, availability of genome scale sequenced data of microbes[10] and completion of human genome project has revolutionized the field of drug discovery[11]. Novel drug targets are required in order to defend pathogens which are readily becoming resistant to drugs. Comparative proteomics provide new opportunities for finding drug targets where in protein sequence and structure of the pathogen and the host may be compared in search of targets, besides exploring the biological process. In general a target should provide adequate selectivity; yielding a drug which is specific or highly selective against the pathogen with respect to the host[12]. Moreover the target should be essential for growth and viability of pathogen at least under the condition of infection. The entire process is built on the assumption that the potential target must play an essential role in the pathogen's survival and constitute a critical component in metabolic pathway[13]. At the same time the target should not have any well conserved homolog in the human host. This could preclude possibilities of unacceptable cross reactivity that might prove detrimental to the host[14]. The present

study used the database of *Helicobacter pylori* proteins and then applied a subtractive approach by comparing it with the human proteins as well as normal gut flora proteins to look for potential targets. Proteome of *Helicobacter pylori* has been known. The strain 26695 has a circular genome of 1,667,867 base pairs with approximately 1,590 predicted coding sequences. Sequence analyses indicate that bacterium has a well developed system of motility, iron scavenging, DNA restriction and modification. Most putative adhesion proteins have been identified indicating a rich surface topography and a potentially complex mechanism of host pathogen [15]. Phospholipase A 1 (PLA 1) hydrolyses the fatty acyl ester bond at the sn-1 position of the glycerol moiety. The action of PLA 1 results in the accumulation of free fatty acids and 2-acyl lysophospholipid or 1-acyl lysophospholipid, respectively [16]. The phospholipases A 1 in general constitute a large group of 1-acyl hydrolases, some of which also degrade neutral lipids (lipases). PLA 1 has been reported to play an important role as a virulence factor in some bacteria and fungi. An intact *H. pylori* outer membrane phospholipase A (OMPLA) has been suggested necessary for colonization of the human gastric ventricle [17, 18]. Only *H. pylori* with intact OMPLA will survive in acidic environments in vitro, whereas phase variants with truncated OMPLA can survive at neutral pH. Intact OMPLA is required for survival of *H. pylori* in acidic environments despite the fact that OMPLA is enzymatically inactive at low pH [17, 20]. Based on the literature review, Protein Model Assessment and Molecular Dynamic Simulation (MD) we can conclude that out of the 5 new probable target proteins Phospholipase A1 protein can be taken for further studies for drug likeliness.

## 2. MATERIALS AND METHODS

Proteome of *Helicobacter pylori* strain 26695 were retrieved from the NCBI ftp site (Index of/genomes/Bacteria/*Helicobacter pylori*\_26695\_uid178201/ and NC\_018939.faa). A total of 1590 protein sequences were available. These protein sequences were subjected for PSI-BLAST with Homo sapiens and Lactobacillus species protein sequence database, for maximum target sequence of 10 with expect threshold 10, i.e E value < 0.01. This comparison was made using scoring scheme **BLOSUM62**

The protein sequences of *Helicobacter pylori* showing no significant similarity were considered for further analysis. Therefore, after the similarity search two sets of no significant were obtained, first that is with Homo sapiens (taxid: 9606) protein database and second lactobacillus (taxid:1578) protein database. On union of these set the protein sequences in the intersection region were considered for the search of potential probable target.

### Protein Model Generation and Assessment

To generate the theoretical model the ab-initio based I-TASSER server (<http://zhanglab.ccmb.med.umich.edu/I-TASSER/>). Thousands of conformations called as decoys were generated through simulations. Based on pair wise similarity of structure, clustering of all the decoys were carried out for which I-Tasser uses SPICKER program reported the best model [20,21].

### Molecular Dynamic Simulation

The Molecular Dynamic (MD) simulations was conducted using GROMOS96 43A1 force field and the flexible SPC water model in GROMACS 4.5.1 package (Groningen Machine for Chemical Simulations) package running on a high performance Dell Precision Tower 7910 computer [22,23]. An octahedron water box was generated and the potential target protein solvation was completed within the box. Energy minimization of the solvated protein done through 50000 steps of the steepest descent minimization to generates a good starting structure for MD simulation [24,25]. Then the energy-minimized model was subjected to position-restrained MD under isothermal-isobaric condition where equilibration was carried out at 300 K and a pressure of 1 bar for 50, 000 steps of 100 ps. Finally the MD simulation was performed at 30 nano second (ns), maintaining the same temperature and pressure. The system stability and differences in the trajectories were analyzed using XMGRACE software.

### 3. RESULTS AND DISCUSSION

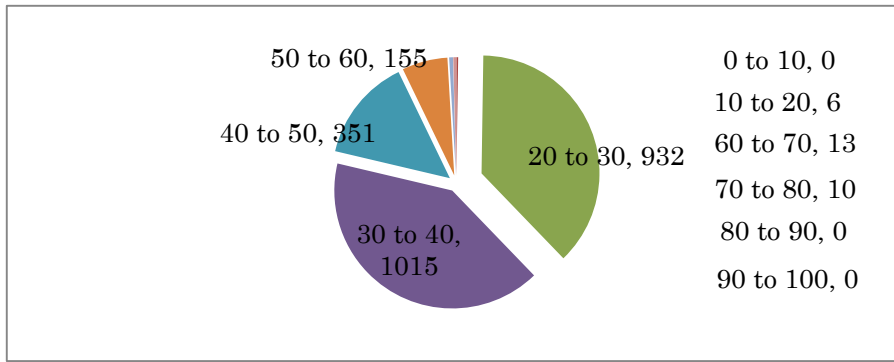
*Helicobacter pylori* strain 26695 has provided a wealth of information enabling us to carry out *in silico* studies to identify new molecular targets for pharmaceutical treatment. Out of 1590 predicted protein coding genes in *H.pylori*, experimentally determined structures are available for only 145 proteins in the PDB, therefore a wealth of structural information remains to be explored. In these work, PSI-BLAST methods is used to get structural as well as functional insight into the *H.pylori* proteome. Using PSI-BLAST, proteome of *H. pylori* strain 26695 were compared with the proteome of Human and Lactobacillus respectively by the above mentioned methodology using appropriate algorithm and certain parameter values to find the extent of similar sequences in the alignment by counting and measuring the identity percentage values separately for both the comparison.

Table 1: % IDENTITY Calculation of *H. pylori* and Lactobacillus Proteome

Class Interval of Percentage	Frequency of Identity	Percentage of pie diagram
0 to 10	0	0
10 to 20	6	0.87
20 to 30	932	135.18
30 to 40	1015	147.21
40 to 50	351	50.91
50 to 60	155	22.48
60 to 70	13	1.88
70 to 80	10	1.45
80 to 90	0	0
90 to 100	0	0
Sum	2482	360

**Percentage of pie diagram** =  $360/3338 * \text{Frequency of Identit}$

**Frequency of Identity** =  $\text{COUNTIF}(C2:C3338, ">=0") - \text{COUNTIF}(C2:C3338, ">=10")$



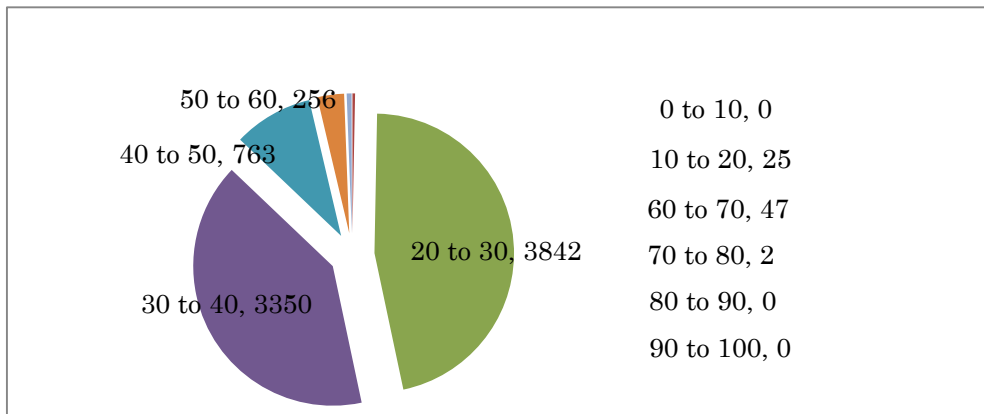
**Fig 1: pie-diagram showing the %identity calculation of *H. pylori* and Lactobacillus Proteome**

**Table 2: %IDENTITY Calculation of *H. pylori* and Human Proteome**

Class Interval of Percentage	Frequency of Identity	Percentage of pie diagram
0 to 10	0	0
10 to 20	25	1.08
20 to 30	3842	166.94
30 to 40	3350	145.56
40 to 50	763	33.15
50 to 60	256	11.12
60 to 70	47	2.04
70 to 80	2	0.08
80 to 90	0	0
90 to 100	0	0
Sum	8285	360

**Percentage of pie diagram** = 360/8296 \* percentage identity values

**Frequency of Identity** = COUNTIF (C2:C8296, ">=0")- COUNTIF (C2:C8296, ">=10")



**Fig 2: pie-diagram showing the %identity calculation of *H. pylori* and Human Proteome**

The table 1 and 2 above shows the pie-fraction of the percentage identity of the region separated by an interval of 10 units of percentage. Here, the sum total of identities for each fraction is used as divisor to the total angle count (i.e 360°). For each class interval their quotient is multiplied to the respective percentage identity. Class interval of the percentage denotes the class interval of the percentage of *H.pylori* compared with that of Human and Lactobacillus. Frequency of identity is the numbers of identity of *H.pylori* with respect to percentage class interval. Percentage of pie diagram denotes the percentage of *H.pylori* for drawing pie diagram for each class interval of percentage. After comparative study with *H .pylori* proteome a total number of 181 and 183 non

significant matches are found for Lactobacillus proteome and Homo sapiens proteome. Total significant matches and their extent of similarity are summarized by the above diagrammatic representations. From the above identity percentage analysis it shows a clear idea for these two proteome set as *H. pylori* 26695 protein sequences are much more evolutionary similar with Homo sapiens than Lactobacillus, i.e maximum numbers of identical sequences are found in Homo sapiens.

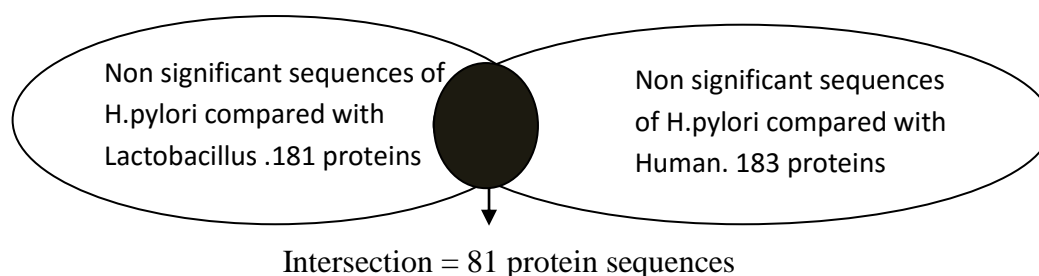


Fig3: Non significant sequences of Lactobacillus and Human proteome

**Table 3: List of Non significant sequences of *H. pylori* compared with *Lactobacillus* (181 proteins sequences)**

Sl no	Initial Reference	First Reference	Protein Name	Strain Name
1.	gi 410681694	YP_006934096.1	hypothetical protein C694_00860	Helicobacter pylori 26695
2.	gi 410681697	YP_006934099.1	fructose-bisphosphate aldolase	Helicobacter pylori 26695
3.	gi 410681698	YP_006934100.1	elongation factor P	Helicobacter pylori 26695
4.	gi 410681705	YP_006934107.1	hypothetical protein C694_00915	Helicobacter pylori 26695
5.	gi 410681706	YP_006934108.1	hypothetical protein C694_00920	Helicobacter pylori 26695
6.	gi 410681708	YP_006934110.1	hypothetical protein C694_00930	Helicobacter pylori 26695
7.	gi 410681710	YP_006934112.1	hypothetical protein C694_00940	Helicobacter pylori 26695
8.	gi 410681714	YP_006934116.1	fumarate reductase cytochrome b-556 subunit	Helicobacter pylori 26695
9.	gi 410681873	YP_006934275.1	hypothetical protein C694_01775	Helicobacter pylori 26695
10	gi 410682129	YP_006934531.1	membrane fusion protein (mtrC)	Helicobacter pylori 26695
11	gi 410681897	YP_006934299.1	hypothetical protein C694_01905	Helicobacter pylori 26695
12	gi 410682104	YP_006934506.1	flagellar motor switch protein	Helicobacter pylori 26695
13	gi 410682346	YP_006934748.1	hypothetical protein C694_04270	Helicobacter pylori 26695
14	gi 410682133	YP_006934535.1	ABC transporter permease	Helicobacter pylori 26695
15	gi 410682009	YP_006934411.1	hypothetical protein C694_02515	Helicobacter pylori 26695
16	gi 410682360	YP_006934762.1	type I restriction enzyme S protein (hsdS)	Helicobacter pylori 26695
17	gi 410681931	YP_006934333.1	hypothetical protein C694_02075	Helicobacter pylori 26695
18	gi 410681992	YP_006934394.1	molybdenum ABC transporter ModB	Helicobacter pylori 26695

19	gi 410682330	YP_006934732.1	osmoprotection protein (proWX)	Helicobacter pylori 26695
20	gi 410682006	YP_006934408.1	hypothetical protein C694_02500	Helicobacter pylori 26695
21	gi 410681824	YP_006934226.1	dipeptide ABC transporter permease (dppB)	Helicobacter pylori 26695
22	gi 410682510	YP_006934912.1	relaxase	Helicobacter pylori 26695
23	gi 410681948	YP_006934350.1	hypothetical protein C694_02160	Helicobacter pylori 26695
24	gi 410682117	YP_006934519.1	hypothetical protein C694_03075	Helicobacter pylori 26695
25	gi 410682529	YP_006934931.1	hypothetical protein C694_05255	Helicobacter pylori 26695
26	gi 410681959	YP_006934361.1	hypothetical protein C694_02230	Helicobacter pylori 26695
27	gi 410682354	YP_006934756.1	hypothetical protein C694_04315	Helicobacter pylori 26695
28	gi 410681814	YP_006934216.1	toxin-like outer membrane protein	Helicobacter pylori 26695
29	gi 410682359	YP_006934761.1	hypothetical protein C694_04340	Helicobacter pylori 26695
30	gi 410682011	YP_006934413.1	50S ribosomal protein L28	Helicobacter pylori 26695
31	gi 410682351	YP_006934753.1	outer membrane protein P1 (ompP1)	Helicobacter pylori 26695
32	gi 410682235	YP_006934637.1	hypothetical protein C694_03705	Helicobacter pylori 26695
33	gi 410682309	YP_006934711.1	flagellar sheath adhesin hpaA	Helicobacter pylori 26695
34	gi 410682182	YP_006934584.1	hypothetical protein C694_03410	Helicobacter pylori 26695
35	gi 410682488	YP_006934890.1	hypothetical protein C694_05020	Helicobacter pylori 26695
36	gi 410682340	YP_006934742.1	hypothetical protein C694_04240	Helicobacter pylori 26695
37	gi 410682254	YP_006934656.1	UDP-MurNac-pentapeptide synthetase (murF)	Helicobacter pylori 26695
38	gi 410681767	YP_006934169.1	hypothetical protein C694_01225	Helicobacter pylori 26695
39	gi 410682256	YP_006934658.1	hypothetical protein C694_03810	Helicobacter pylori 26695
40	gi 410681875	YP_006934277.1	flagellar motor switch protein G	Helicobacter pylori 26695
41	gi 410682012	YP_006934414.1	hypothetical protein C694_02530	Helicobacter pylori 26695
42	gi 410682119	YP_006934521.1	penicillin-binding protein 1A (PBP-1A)	Helicobacter pylori 26695
43	gi 410682461	YP_006934863.1	hypothetical protein C694_04875	Helicobacter pylori 26695
44	gi 410682355	YP_006934757.1	thiamine-phosphate pyrophosphorylase	Helicobacter pylori 26695
45	gi 410681876	YP_006934278.1	flagellar assembly protein H	Helicobacter pylori 26695
46	gi 410682369	YP_006934771.1	phosphoheptose isomerase	Helicobacter pylori 26695
47	gi 410682048	YP_006934450.1	cag pathogenicity island protein	Helicobacter pylori 26695
48	gi 410681942	YP_006934344.1	hypothetical protein C694_02130	Helicobacter pylori 26695
49	gi 410682373	YP_006934775.1	hypothetical protein C694_04410	Helicobacter pylori 26695
50	gi 410682435	YP_006934837.1	toxin-like outer membrane protein	Helicobacter pylori 26695
51	gi 410682480	YP_006934882.1	virulence associated protein D (vapD)	Helicobacter pylori 26695
52	gi 410682379	YP_006934781.1	lipid-A-disaccharide synthase	Helicobacter pylori 26695

53	gi 410682376	YP_006934778.1	hypothetical protein C694_04425	Helicobacter pylori 26695
54	gi 410682186	YP_006934588.1	chorismate synthase	Helicobacter pylori 26695
55	gi 410682091	YP_006934493.1	hypothetical protein C694_02945	Helicobacter pylori 26695
56	gi 410682050	YP_006934452.1	cag pathogenicity island protein (cag9)	Helicobacter pylori 26695
57	gi 410681952	YP_006934354.1	hypothetical protein C694_02180	Helicobacter pylori 26695
58	gi 410682093	YP_006934495.1	hypothetical protein C694_02955	Helicobacter pylori 26695
59	gi 410681778	YP_006934180.1	hypothetical protein C694_01280	Helicobacter pylori 26695
60	gi 410682488	YP_006934890.1	hypothetical protein C694_05020	Helicobacter pylori 26695
61	gi 410682126	YP_006934528.1	hypothetical protein C694_03120	Helicobacter pylori 26695
62	gi 410682049	YP_006934451.1	cag pathogenicity island protein (cag8)	Helicobacter pylori 26695
63	gi 410682103	YP_006934505.1	hypothetical protein C694_03005	Helicobacter pylori 26695
64	gi 410682545	YP_006934947.1	flagellar motor switch protein FlIM	Helicobacter pylori 26695
65	gi 410682178	YP_006934580.1	hypothetical protein C694_03390	Helicobacter pylori 26695
66	gi 410682479	YP_006934881.1	hypothetical protein C694_04975	Helicobacter pylori 26695
67	gi 410682123	YP_006934525.1	hypothetical protein C694_03105	Helicobacter pylori 26695
68	gi 410682041	YP_006934443.1	cag island protein	Helicobacter pylori 26695
69	gi 410682047	YP_006934449.1	hypothetical protein C694_02715	Helicobacter pylori 26695
70	gi 410682113	YP_006934515.1	type III restriction enzyme R protein (res)	Helicobacter pylori 26695
71	gi 410682115	YP_006934517.1	hypothetical protein C694_03065	Helicobacter pylori 26695
72	gi 410682026	YP_006934428.1	hypothetical protein C694_02610	Helicobacter pylori 26695
73	gi 410682456	YP_006934858.1	D-alanine glycine permease	Helicobacter pylori 26695
74	gi 410682176	YP_006934578.1	hypothetical protein C694_03380	Helicobacter pylori 26695
75	gi 410681881	YP_006934283.1	hypothetical protein C694_01815	Helicobacter pylori 26695
76	gi 410681839	YP_006934241.1	hypothetical protein C694_01585	Helicobacter pylori 26695
77	gi 410682547	YP_006934949.1	hypothetical protein C694_05345	Helicobacter pylori 26695
78	gi 410682083	YP_006934485.1	hypothetical protein C694_02905	Helicobacter pylori 26695
79	gi 410682099	YP_006934501.1	hypothetical protein C694_02985	Helicobacter pylori 26695
80	gi 410682032	YP_006934434.1	hypothetical protein C694_02640	Helicobacter pylori 26695
81	gi 410681997	YP_006934399.1	hypothetical protein C694_02455	Helicobacter pylori 26695
82	gi 410682040	YP_006934442.1	cag pathogenicity island protein Epsilon	Helicobacter pylori 26695
83	gi 410682007	YP_006934409.1	hypothetical protein C694_02505	Helicobacter pylori 26695
84	gi 410682100	YP_006934502.1	hypothetical protein C694_02990	Helicobacter pylori 26695
85	gi 410681838	YP_006934240.1	nitrite extrusion protein (narK)	Helicobacter pylori 26695
86	gi 410682082	YP_006934484.1	30S ribosomal protein S21	Helicobacter pylori 26695
87	gi 410682016	YP_006934418.1	hypothetical protein C694_02550	Helicobacter pylori 26695



88	gi 410681835	YP_006934237.1	hypothetical protein C694_01565	Helicobacter pylori 26695
89	gi 410682190	YP_006934592.1	hypothetical protein C694_03460	Helicobacter pylori 26695
90	gi 410681825	YP_006934227.1	dipeptide ABC transporter permease (dppC)	Helicobacter pylori 26695
91	gi 410682522	YP_006934924.1	hypothetical protein C694_05220	Helicobacter pylori 26695
92	gi 410682241	YP_006934643.1	hypothetical protein C694_03735	Helicobacter pylori 26695
93	gi 410681544	YP_006933946.1	hypothetical protein C694_00110	Helicobacter pylori 26695
94	gi 410681548	YP_006933950.1	hypothetical protein C694_00130	Helicobacter pylori 26695
95	gi 410681549	YP_006933951.1	dethiobiotinsynthetase	Helicobacter pylori 26695
96	gi 410681572	YP_006933974.1	hypothetical protein C694_00250	Helicobacter pylori 26695
97	gi 410681577	YP_006933979.1	Myosin-3	Helicobacter pylori 26695
98	gi 410681581	YP_006933983.1	hypothetical protein C694_00295	Helicobacter pylori 26695
99	gi 410681582	YP_006933984.1	hypothetical protein C694_00300	Helicobacter pylori 26695
10	gi 410681584	YP_006933986.1	hypothetical protein C694_00310	Helicobacter pylori 26695
10	gi 410681589	YP_006933991.1	urease accessory protein UreE	Helicobacter pylori 26695
10	gi 410681593	YP_006933995.1	urease subunit alpha	Helicobacter pylori 26695
10	gi 410681594	YP_006933996.1	lipoprotein signal peptidase	Helicobacter pylori 26695
10	gi 410681596	YP_006933998.1	30S ribosomal protein S20	Helicobacter pylori 26695
10	gi 410681599	YP_006934001.1	hypothetical protein C694_00385	Helicobacter pylori 26695
10	gi 410681600	YP_006934002.1	hypothetical protein C694_00390	Helicobacter pylori 26695
10	gi 410681636	YP_006934038.1	flagellin B	Helicobacter pylori 26695
10	gi 410681635	YP_006934037.1	hypothetical protein C694_00565	Helicobacter pylori 26695
10	gi 410681640	YP_006934042.1	hypothetical protein C694_00590	Helicobacter pylori 26695
11	gi 410681641	YP_006934043.1	hypothetical protein C694_00595	Helicobacter pylori 26695
11	gi 410681645	YP_006934047.1	translation initiation factor IF-3	Helicobacter pylori 26695
11	gi 410681648	YP_006934050.1	hypothetical protein C694_00630	Helicobacter pylori 26695
11	gi 410681650	YP_006934052.1	hypothetical protein C694_00640	Helicobacter pylori 26695
11	gi 410681651	YP_006934053.1	hypothetical protein C694_00645	Helicobacter pylori 26695
11	gi 410681654	YP_006934056.1	phospho-2-dehydro-3-deoxyheptonate aldolase	Helicobacter pylori 26695
11	gi 410681655	YP_006934057.1	hypothetical protein C694_00665	Helicobacter pylori 26695
11	gi 410681658	YP_006934060.1	iron-sulfur cluster binding protein	Helicobacter pylori 26695
11	gi 410682200	YP_006934602.1	hypothetical protein C694_03510	Helicobacter pylori 26695
11	gi 410681900	YP_006934302.1	cytochrome c biogenesis protein (ycf5)	Helicobacter pylori 26695
12	gi 410682022	YP_006934424.1	hypothetical protein C694_02590	Helicobacter pylori 26695
12	gi 410682542	YP_006934944.1	hypothetical protein C694_05320	Helicobacter pylori 26695

12	gi 410682019	YP_006934421.1	phospholipase A1	Helicobacter pylori 26695
12	gi 410682102	YP_006934504.1	hypothetical protein C694_03000	Helicobacter pylori 26695
12	gi 410682000	YP_006934402.1	type II adenine specific DNA methyltransferase	Helicobacter pylori 26695
12	gi 410682488	YP_006934890.1	hypothetical protein C694_05022	Helicobacter pylori 26695
12	gi 410682537	YP_006934939.1	hypothetical protein C694_05295	Helicobacter pylori 26695
12	gi 410682521	YP_006934923.1	integrase-recombinase protein	Helicobacter pylori 26695
12	gi 410682008	YP_006934410.1	hypothetical protein C694_02510	Helicobacter pylori 26695
12	gi 410681895	YP_006934297.1	hypothetical protein C694_01895	Helicobacter pylori 26695
13	gi 410681751	YP_006934153.1	hypothetical protein C694_01145	Helicobacter pylori 26695
13	gi 410682308	YP_006934710.1	hypothetical protein C694_04080	Helicobacter pylori 26695
13	gi 410682516	YP_006934918.1	hypothetical protein C694_05170	Helicobacter pylori 26695
13	gi 410682300	YP_006934702.1	hypothetical protein C694_04040	Helicobacter pylori 26695
13	gi 410682251	YP_006934653.1	hypothetical protein C694_03785	Helicobacter pylori 26695
13	gi 410681887	YP_006934289.1	spore coat polysaccharide biosynthesis protein C	Helicobacter pylori 26695
13	gi 410681769	YP_006934171.1	signal-transducing protein, histidine kinase (atoS)	Helicobacter pylori 26695
13	gi 410681874	YP_006934276.1	flagellar MS-ring protein	Helicobacter pylori 26695
13	gi 410681934	YP_006934336.1	hypothetical protein C694_02090	Helicobacter pylori 26695
13	gi 410682135	YP_006934537.1	hypothetical protein C694_03175	Helicobacter pylori 26695
14	gi 410682183	YP_006934585.1	hypothetical protein C694_03415	Helicobacter pylori 26695
14	gi 410682464	YP_006934866.1	rRNA large subunit methyltransferase	Helicobacter pylori 26695
14	gi 410682137	YP_006934539.1	chemotaxis protein (cheV)	Helicobacter pylori 26695
14	gi 410681891	YP_006934293.1	putative type II DNA modification enzyme/Methyl transferase	Helicobacter pylori 26695
14	gi 410682236	YP_006934638.1	hypothetical protein C694_03710	Helicobacter pylori 26695
14	gi 410682128	YP_006934530.1	hypothetical protein C694_03130	Helicobacter pylori 26695
14	gi 410682131	YP_006934533.1	hypothetical protein C694_03145	Helicobacter pylori 26695
14	gi 410682234	YP_006934636.1	hypothetical protein C694_03700	Helicobacter pylori 26695
14	gi 410682246	YP_006934648.1	hypothetical protein C694_03760	Helicobacter pylori 26695
14	gi 410681770	YP_006934172.1	hypothetical protein C694_01240	Helicobacter pylori 26695
15	gi 410682315	YP_006934717.1	hypothetical protein C694_04115	Helicobacter pylori 26695
15	gi 410682535	YP_006934937.1	hypothetical protein C694_05285	Helicobacter pylori 26695
15	gi 410682327	YP_006934729.1	flagellar motor protein MotA	Helicobacter pylori 26695

15	gi 410681939	YP_006934341.1	hypothetical protein C694_02115	Helicobacter pylori 26695
15	gi 410681884	YP_006934286.1	hypothetical protein C694_01835	Helicobacter pylori 26695
15	gi 410681777	YP_006934179.1	hypothetical protein C694_01275	Helicobacter pylori 26695
15	gi 410682124	YP_006934526.1	flagellin A	Helicobacter pylori 26695
15	gi 410682195	YP_006934597.1	integrase-recombinase protein	Helicobacter pylori 26695
15	gi 410682544	YP_006934946.1	flagellar motor switch protein FliY	Helicobacter pylori 26695
15	gi 410682291	YP_006934693.1	hypothetical protein C694_03995	Helicobacter pylori 26695
16	gi 410681935	YP_006934337.1	neuraminyllactose-binding hemagglutinin HpaA	Helicobacter pylori 26695
16	gi 410682470	YP_006934872.1	prolipoproteindiacylglyceryl transferase	Helicobacter pylori 26695
16	gi 410682112	YP_006934514.1	2-oxoglutarate-acceptor oxidoreductase subunit OorC	Helicobacter pylori 26695
16	gi 410681816	YP_006934218.1	hypothetical protein C694_01470	Helicobacter pylori 26695
16	gi 410682509	YP_006934911.1	integrase/recombinase (xerD)	Helicobacter pylori 26695
16	gi 410682368	YP_006934770.1	hypothetical protein C694_04385	Helicobacter pylori 26695
16	gi 410681890	YP_006934292.1	hypothetical protein C694_01870	Helicobacter pylori 26695
16	gi 410681753	YP_006934155.1	hypothetical protein C694_01155	Helicobacter pylori 26695
16	gi 410682318	YP_006934720.1	hypothetical protein C694_04130	Helicobacter pylori 26695
16	gi 410681813	YP_006934215.1	hypothetical protein C694_01455	Helicobacter pylori 26695
17	gi 410681771	YP_006934173.1	flagellar basal body P-ring protein	Helicobacter pylori 26695
17	gi 410682189	YP_006934591.1	glycerol-3-phosphate dehydrogenase	Helicobacter pylori 26695
17	gi 410682298	YP_006934700.1	preprotein translocase subunit SecA	Helicobacter pylori 26695
17	gi 410682307	YP_006934709.1	trigger factor	Helicobacter pylori 26695
17	gi 410682294	YP_006934696.1	hypothetical protein C694_04010	Helicobacter pylori 26695
17	gi 410682378	YP_006934780.1	transcription elongation factor GreA	Helicobacter pylori 26695
17	gi 410682380	YP_006934782.1	hypothetical protein C694_04445	Helicobacter pylori 26695
17	gi 410682451	YP_006934853.1	hypothetical protein C694_04825	Helicobacter pylori 26695
17	gi 410681937	YP_006934339.1	transposase-like protein, PS3IS	Helicobacter pylori 26695
17	gi 410682375	YP_006934777.1	hypothetical protein C694_04420	Helicobacter pylori 26695
18	gi 410681812	YP_006934214.1	hypothetical protein C694_01450	Helicobacter pylori 26695
18	gi 410682323	YP_006934725.1	hypothetical protein C694_04155	Helicobacter pylori 26695
18	gi 410682247	YP_006934649.1	hypothetical protein C694_03765	Helicobacter pylori 26695
18	gi 410682010	YP_006934412.1	potassium channel protein	Helicobacter pylori 26695

**Table 4: List of Non-significant sequences of *H. pylori* compared with Human (183 proteins sequences )**

Sl. no	Initial Reference	1 <sup>st</sup> Reference	Protein Name	Strain Name
1.	gi 410681694	YP_006934096.1	hypothetical protein C694_00860	Helicobacter pylori 26695
2.	gi 410681697	YP_006934099.1	fructose-bisphosphate aldolase	Helicobacter pylori 26695
3.	gi 410681698	YP_006934100.1	elongation factor P	Helicobacter pylori 26695
4.	gi 410681705	YP_006934107.1	hypothetical protein C694_00915	Helicobacter pylori 26695
5.	gi 410681706	YP_006934108.1	hypothetical protein C694_00920	Helicobacter pylori 26695
6.	gi 410681708	YP_006934110.1	hypothetical protein C694_00930	Helicobacter pylori 26695
7.	gi 410681710	YP_006934112.1	hypothetical protein C694_00940	Helicobacter pylori 26695
8.	gi 410681714	YP_006934116.1	fumarate reductase cytochrome b-556 subunit	Helicobacter pylori 26695
9.	gi 410681873	YP_006934275.1	hypothetical protein C694_01775	Helicobacter pylori 26695
10.	gi 410682129	YP_006934531.1	membrane fusion protein (mtrC)	Helicobacter pylori 26695
11.	gi 410681897	YP_006934299.1	hypothetical protein C694_01905	Helicobacter pylori 26695
12.	gi 410682104	YP_006934506.1	flagellar motor switch protein	Helicobacter pylori 26695
13.	gi 410682346	YP_006934748.1	hypothetical protein C694_04270	Helicobacter pylori 26695
14.	gi 410682133	YP_006934535.1	ABC transporter permease	Helicobacter pylori 26695
15.	gi 410682009	YP_006934411.1	hypothetical protein C694_02515	Helicobacter pylori 26695
16.	gi 410682360	YP_006934762.1	type I restriction enzyme S protein (hsdS)	Helicobacter pylori 26695
17.	gi 410681931	YP_006934333.1	hypothetical protein C694_02075	Helicobacter pylori 26695
18.	gi 410681992	YP_006934394.1	molybdenum ABC transporter ModB	Helicobacter pylori 26695
19.	gi 410682330	YP_006934732.1	osmoprotection protein (proWX)	Helicobacter pylori 26695
20.	gi 410682006	YP_006934408.1	hypothetical protein C694_02500	Helicobacter pylori 26695
21.	gi 410681824	YP_006934226.1	dipeptide ABC transporter permease (dppB)	Helicobacter pylori 26695
22.	gi 410682510	YP_006934912.1	relaxase	Helicobacter pylori 26695
23.	gi 410681948	YP_006934350.1	hypothetical protein C694_02160	Helicobacter pylori 26695
24.	gi 410682117	YP_006934519.1	hypothetical protein C694_03075	Helicobacter pylori 26695
25.	gi 410682529	YP_006934931.1	hypothetical protein C694_05255	Helicobacter pylori 26695
26.	gi 410681959	YP_006934361.1	hypothetical protein C694_02230	Helicobacter pylori 26695
27.	gi 410682354	YP_006934756.1	hypothetical protein C694_04315	Helicobacter pylori 26695
28.	gi 410681814	YP_006934216.1	toxin-like outer membrane protein	Helicobacter pylori 26695
29.	gi 410682359	YP_006934761.1	hypothetical protein C694_04340	Helicobacter pylori 26695
30.	gi 410682011	YP_006934413.1	50S ribosomal protein L28	Helicobacter pylori 26695
31.	gi 410682351	YP_006934753.1	outer membrane protein P1 (ompP1)	Helicobacter pylori 26695
32.	gi 410682235	YP_006934637.1	hypothetical protein C694_03705	Helicobacter pylori 26695

33.	gi 410682309	YP_006934711.1	flagellar sheath adhesin hpaA	Helicobacter pylori 26695
34.	gi 410682182	YP_006934584.1	hypothetical protein C694_03410	Helicobacter pylori 26695
35.	gi 410682488	YP_006934890.1	hypothetical protein C694_05020	Helicobacter pylori 26695
36.	gi 410682340	YP_006934742.1	hypothetical protein C694_04240	Helicobacter pylori 26695
37.	gi 410682254	YP_006934656.1	UDP-MurNac-pentapeptidepresynthetase (murF)	Helicobacter pylori 26695
38.	gi 410681767	YP_006934169.1	hypothetical protein C694_01225	Helicobacter pylori 26695
39.	gi 410682256	YP_006934658.1	hypothetical protein C694_03810	Helicobacter pylori 26695
40.	gi 410681875	YP_006934277.1	flagellar motor switch protein G	Helicobacter pylori 26695
41.	gi 410682012	YP_006934414.1	hypothetical protein C694_02530	Helicobacter pylori 26695
42.	gi 410682119	YP_006934521.1	penicillin-binding protein 1A (PBP-1A)	Helicobacter pylori 26695
43.	gi 410682461	YP_006934863.1	hypothetical protein C694_04875	Helicobacter pylori 26695
44.	gi 410682355	YP_006934757.1	thiamine-phosphate pyrophosphorylase	Helicobacter pylori 26695
45.	gi 410681876	YP_006934278.1	flagellar assembly protein H	Helicobacter pylori 26695
46.	gi 410682369	YP_006934771.1	phosphoheptose isomerase	Helicobacter pylori 26695
47.	gi 410682048	YP_006934450.1	cag pathogenicity island protein (cag7)	Helicobacter pylori 26695
48.	gi 410681942	YP_006934344.1	hypothetical protein C694_02130	Helicobacter pylori 26695
49.	gi 410682373	YP_006934775.1	hypothetical protein C694_04410	Helicobacter pylori 26695
50.	gi 410682435	YP_006934837.1	toxin-like outer membrane protein	Helicobacter pylori 26695
51.	gi 410682480	YP_006934882.1	virulence associated protein D (vapD)	Helicobacter pylori 26695
52.	gi 410682379	YP_006934781.1	lipid-A-disaccharide synthase	Helicobacter pylori 26695
53.	gi 410682376	YP_006934778.1	hypothetical protein C694_04425	Helicobacter pylori 26695
54.	gi 410682186	YP_006934588.1	chorismate synthase	Helicobacter pylori 26695
55.	gi 410682091	YP_006934493.1	hypothetical protein C694_02945	Helicobacter pylori 26695
56.	gi 410682050	YP_006934452.1	cag pathogenicity island protein (cag9)	Helicobacter pylori 26695
57.	gi 410681952	YP_006934354.1	hypothetical protein C694_02180	Helicobacter pylori 26695
58.	gi 410682093	YP_006934495.1	hypothetical protein C694_02955	Helicobacter pylori 26695
59.	gi 410681778	YP_006934180.1	hypothetical protein C694_01280	Helicobacter pylori 26695
60.	gi 410682488	YP_006934890.1	hypothetical protein C694_05020	Helicobacter pylori 26695
61.	gi 410682126	YP_006934528.1	hypothetical protein C694_03120	Helicobacter pylori 26695
62.	gi 410682049	YP_006934451.1	cag pathogenicity island protein (cag8)	Helicobacter pylori 26695
63.	gi 410682103	YP_006934505.1	hypothetical protein C694_03005	Helicobacter pylori 26695
64.	gi 410682545	YP_006934947.1	flagellar motor switch protein FliM	Helicobacter pylori 26695
65.	gi 410682178	YP_006934580.1	hypothetical protein C694_03390	Helicobacter pylori 26695
66.	gi 410682479	YP_006934881.1	hypothetical protein C694_04975	Helicobacter pylori 26695

67.	gi 410682123	YP_006934525.1	hypothetical protein C694_03105	Helicobacter pylori 26695
68.	gi 410682041	YP_006934443.1	cag island protein	Helicobacter pylori 26695
69.	gi 410682047	YP_006934449.1	hypothetical protein C694_02715	Helicobacter pylori 26695
70.	gi 410682113	YP_006934515.1	type III restriction enzyme R protein (res)	Helicobacter pylori 26695
71.	gi 410682115	YP_006934517.1	hypothetical protein C694_03065	Helicobacter pylori 26695
72.	gi 410682026	YP_006934428.1	hypothetical protein C694_02610	Helicobacter pylori 26695
73.	gi 410682456	YP_006934858.1	D-alanine glycine permease	Helicobacter pylori 26695
74.	gi 410682176	YP_006934578.1	hypothetical protein C694_03380	Helicobacter pylori 26695
75.	gi 410681881	YP_006934283.1	hypothetical protein C694_01815	Helicobacter pylori 26695
76.	gi 410681839	YP_006934241.1	hypothetical protein C694_01585	Helicobacter pylori 26695
77.	gi 410682547	YP_006934949.1	hypothetical protein C694_05345	Helicobacter pylori 26695
78.	gi 410682083	YP_006934485.1	hypothetical protein C694_02905	Helicobacter pylori 26695
79.	gi 410682099	YP_006934501.1	hypothetical protein C694_02985	Helicobacter pylori 26695
80.	gi 410682032	YP_006934434.1	hypothetical protein C694_02640	Helicobacter pylori 26695
81.	gi 410681997	YP_006934399.1	hypothetical protein C694_02455	Helicobacter pylori 26695
82.	gi 410682040	YP_006934442.1	cag pathogenicity island protein Epsilon	Helicobacter pylori 26695
83.	gi 410682007	YP_006934409.1	hypothetical protein C694_02505	Helicobacter pylori 26695
84.	gi 410682100	YP_006934502.1	hypothetical protein C694_02990	Helicobacter pylori 26695
85.	gi 410681838	YP_006934240.1	nitrite extrusion protein (narK)	Helicobacter pylori 26695
86.	gi 410682082	YP_006934484.1	30S ribosomal protein S21	Helicobacter pylori 26695
87.	gi 410682016	YP_006934418.1	hypothetical protein C694_02550	Helicobacter pylori 26695
88.	gi 410681835	YP_006934237.1	hypothetical protein C694_01565	Helicobacter pylori 26695
89.	gi 410682190	YP_006934592.1	hypothetical protein C694_03460	Helicobacter pylori 26695
90.	gi 410681825	YP_006934227.1	dipeptide ABC transporter permease (dppC)	Helicobacter pylori 26695
91.	gi 410682522	YP_006934924.1	hypothetical protein C694_05220	Helicobacter pylori 26695
92.	gi 410682241	YP_006934643.1	hypothetical protein C694_03735	Helicobacter pylori 26695
93.	gi 410681544	YP_006933946.1	hypothetical protein C694_00110	Helicobacter pylori 26695
94.	gi 410681548	YP_006933950.1	hypothetical protein C694_00130	Helicobacter pylori 26695
95.	gi 410681549	YP_006933951.1	dethiobiotinsynthetase	Helicobacter pylori 26695
96.	gi 410681572	YP_006933974.1	hypothetical protein C694_00250	Helicobacter pylori 26695
97.	gi 410681577	YP_006933979.1	Myosin-3	Helicobacter pylori 26695
98.	gi 410681581	YP_006933983.1	hypothetical protein C694_00295	Helicobacter pylori 26695
99.	gi 410681582	YP_006933984.1	hypothetical protein C694_00300	Helicobacter pylori 26695
100.	gi 410681584	YP_006933986.1	hypothetical protein C694_00310	Helicobacter pylori 26695
101.	gi 410681589	YP_006933991.1	urease accessory protein UreE	Helicobacter pylori 26695

102.	gi 410681593	YP_006933995.1	urease subunit alpha	Helicobacter pylori 26695
103.	gi 410681594	YP_006933996.1	lipoprotein signal peptidase	Helicobacter pylori 26695
104.	gi 410681596	YP_006933998.1	30S ribosomal protein S20	Helicobacter pylori 26695
105.	gi 410681599	YP_006934001.1	hypothetical protein C694_00385	Helicobacter pylori 26695
106.	gi 410681600	YP_006934002.1	hypothetical protein C694_00390	Helicobacter pylori 26695
107.	gi 410681636	YP_006934038.1	flagellin B	Helicobacter pylori 26695
108.	gi 410681635	YP_006934037.1	hypothetical protein C694_00565	Helicobacter pylori 26695
109.	gi 410681640	YP_006934042.1	hypothetical protein C694_00590	Helicobacter pylori 26695
110.	gi 410681641	YP_006934043.1	hypothetical protein C694_00595	Helicobacter pylori 26695
111.	gi 410681645	YP_006934047.1	translation initiation factor IF-3	Helicobacter pylori 26695
112.	gi 410681648	YP_006934050.1	hypothetical protein C694_00630	Helicobacter pylori 26695
113.	gi 410681650	YP_006934052.1	hypothetical protein C694_00640	Helicobacter pylori 26695
114.	gi 410681651	YP_006934053.1	hypothetical protein C694_00645	Helicobacter pylori 26695
115.	gi 410681654	YP_006934056.1	phospho-2-dehydro-3-deoxyheptonate aldolase	Helicobacter pylori 26695
116.	gi 410681655	YP_006934057.1	hypothetical protein C694_00665	Helicobacter pylori 26695
117.	gi 410681658	YP_006934060.1	iron-sulfur cluster binding protein	Helicobacter pylori 26695
118.	gi 410682200	YP_006934602.1	hypothetical protein C694_03510	Helicobacter pylori 26695
119.	gi 410681900	YP_006934302.1	cytochrome c biogenesis protein (ycf5)	Helicobacter pylori 26695
120.	gi 410682022	YP_006934424.1	hypothetical protein C694_02590	Helicobacter pylori 26695
121.	gi 410682542	YP_006934944.1	hypothetical protein C694_05320	Helicobacter pylori 26695
122.	gi 410682019	YP_006934421.1	phospholipase A1	Helicobacter pylori 26695
123.	gi 410682102	YP_006934504.1	hypothetical protein C694_03000	Helicobacter pylori 26695
124.	gi 410682000	YP_006934402.1	type II adenine specific DNA methyltransferase	Helicobacter pylori 26695
125.	gi 410682537	YP_006934939.1	hypothetical protein C694_05295	Helicobacter pylori 26695
126.	gi 410682521	YP_006934923.1	integrase-recombinase protein	Helicobacter pylori 26695
127.	gi 410682008	YP_006934410.1	hypothetical protein C694_02510	Helicobacter pylori 26695
128.	gi 410681895	YP_006934297.1	hypothetical protein C694_01895	Helicobacter pylori 26695
129.	gi 410681751	YP_006934153.1	hypothetical protein C694_01145	Helicobacter pylori 26695
130.	gi 410682308	YP_006934710.1	hypothetical protein C694_04080	Helicobacter pylori 26695
131.	gi 410682516	YP_006934918.1	hypothetical protein C694_05170	Helicobacter pylori 26695
132.	gi 410682300	YP_006934702.1	hypothetical protein C694_04040	Helicobacter pylori 26695
133.	gi 410682251	YP_006934653.1	hypothetical protein C694_03785	Helicobacter pylori 26695
134.	gi 410681887	YP_006934289.1	spore coat polysaccharide biosynthesis protein C	Helicobacter pylori 26695

135.	gi 410681769	YP_006934171.1	signal-transducing protein, histidine kinase (atoS)	Helicobacter pylori 26695
136.	gi 410681874	YP_006934276.1	flagellar MS-ring protein	Helicobacter pylori 26695
137.	gi 410681934	YP_006934336.1	hypothetical protein C694_02090	Helicobacter pylori 26695
138.	gi 410682135	YP_006934537.1	hypothetical protein C694_03175	Helicobacter pylori 26695
139.	gi 410682183	YP_006934585.1	hypothetical protein C694_03415	Helicobacter pylori 26695
140.	gi 410682464	YP_006934866.1	rRNA large subunit methyltransferase	Helicobacter pylori 26695
141.	gi 410682137	YP_006934539.1	chemotaxis protein (cheV)	Helicobacter pylori 26695
142.	gi 410681891	YP_006934293.1	putative type II DNA modification enzyme/Methyl transferase	Helicobacter pylori 26695
143.	gi 410682236	YP_006934638.1	hypothetical protein C694_03710	Helicobacter pylori 26695
144.	gi 410682128	YP_006934530.1	hypothetical protein C694_03130	Helicobacter pylori 26695
145.	gi 410682131	YP_006934533.1	hypothetical protein C694_03145	Helicobacter pylori 26695
146.	gi 410682234	YP_006934636.1	hypothetical protein C694_03700	Helicobacter pylori 26695
147.	gi 410682246	YP_006934648.1	hypothetical protein C694_03760	Helicobacter pylori 26695
148.	gi 410681770	YP_006934172.1	hypothetical protein C694_01240	Helicobacter pylori 26695
149.	gi 410682315	YP_006934717.1	hypothetical protein C694_04115	Helicobacter pylori 26695
150.	gi 410682535	YP_006934937.1	hypothetical protein C694_05285	Helicobacter pylori 26695
151.	gi 410682327	YP_006934729.1	flagellar motor protein MotA	Helicobacter pylori 26695
152.	gi 410681939	YP_006934341.1	hypothetical protein C694_02115	Helicobacter pylori 26695
153.	gi 410681884	YP_006934286.1	hypothetical protein C694_01835	Helicobacter pylori 26695
154.	gi 410681777	YP_006934179.1	hypothetical protein C694_01275	Helicobacter pylori 26695
155.	gi 410682124	YP_006934526.1	flagellin A	Helicobacter pylori 26695
156.	gi 410682195	YP_006934597.1	integrase-recombinase protein	Helicobacter pylori 26695
157.	gi 410682544	YP_006934946.1	flagellar motor switch protein FliY	Helicobacter pylori 26695
158.	gi 410682291	YP_006934693.1	hypothetical protein C694_03995	Helicobacter pylori 26695
159.	gi 410681935	YP_006934337.1	neuraminyllactose-binding hemagglutinin HpaA	Helicobacter pylori 26695
160.	gi 410682470	YP_006934872.1	prolipoproteindiacylglyceryl transferase	Helicobacter pylori 26695
161.	gi 410682112	YP_006934514.1	2-oxoglutarate-acceptor oxidoreductase subunit OorC	Helicobacter pylori 26695
162.	gi 410681816	YP_006934218.1	hypothetical protein C694_01470	Helicobacter pylori 26695
163.	gi 410682509	YP_006934911.1	integrase/recombinase (xerD)	Helicobacter pylori 26695
164.	gi 410682368	YP_006934770.1	hypothetical protein C694_04385	Helicobacter pylori 26695
165.	gi 410681890	YP_006934292.1	hypothetical protein C694_01870	Helicobacter pylori 26695



166.	gi 410681753	YP_006934155.1	hypothetical protein C694_01155	Helicobacter pylori 26695
167.	gi 410682318	YP_006934720.1	hypothetical protein C694_04130	Helicobacter pylori 26695
168.	gi 410681813	YP_006934215.1	hypothetical protein C694_01455	Helicobacter pylori 26695
169.	gi 410681771	YP_006934173.1	flagellar basal body P-ring protein	Helicobacter pylori 26695
170.	gi 410682189	YP_006934591.1	glycerol-3-phosphate dehydrogenase	Helicobacter pylori 26695
171.	gi 410682298	YP_006934700.1	preprotein translocase subunit SecA	Helicobacter pylori 26695
172.	gi 410682307	YP_006934709.1	trigger factor	Helicobacter pylori 26695
173.	gi 410682294	YP_006934696.1	hypothetical protein C694_04010	Helicobacter pylori 26695
174.	gi 410682378	YP_006934780.1	transcription elongation factor GreA	Helicobacter pylori 26695
175.	gi 410682380	YP_006934782.1	hypothetical protein C694_04445	Helicobacter pylori 26695
176.	gi 410682451	YP_006934853.1	hypothetical protein C694_04825	Helicobacter pylori 26695
177.	gi 410681937	YP_006934339.1	transposase-like protein, PS3IS	Helicobacter pylori 26695
178.	gi 410682375	YP_006934777.1	hypothetical protein C694_04420	Helicobacter pylori 26695
179.	gi 410681812	YP_006934214.1	hypothetical protein C694_01450	Helicobacter pylori 26695
180.	gi 410682323	YP_006934725.1	hypothetical protein C694_04155	Helicobacter pylori 26695
181.	gi 410682247	YP_006934649.1	hypothetical protein C694_03765	Helicobacter pylori 26695
182.	gi 410682010	YP_006934412.1	potassium channel protein	Helicobacter pylori 26695
183.	gi 410682012	YP_006934224.1	hypothetical protein C694_02580	Helicobacter pylori 26695

**Intersection of 81 protein sequences ID of Non significant Sequences of *H.pylori* compared with that of Human and Lactobacillus (81 protein sequences)**

YP\_006935354.1, YP\_006935044.1, YP\_006934379.1, YP\_006934821.1,  
 YP\_006934096.1,YP\_006934108.1, YP\_006934299.1, YP\_006934408.1, YP\_006934912.1,  
 YP\_006934350.1,YP\_006934519.1, YP\_006934216.1, YP\_006934584.1, YP\_006934890.1,  
 YP\_006934742.1,YP\_006934658.1, YP\_006934278.1, YP\_006934450.1, YP\_006934775.1,  
 YP\_006934837.1,YP\_006934452.1, YP\_006934890.1, YP\_006934451.1, YP\_006934525.1,  
 YP\_006934449.1,YP\_006934428.1, YP\_006934501.1, YP\_006934409.1, YP\_006934924.1,  
 YP\_006934643.1,YP\_006933946.1, YP\_006934001.1, YP\_006934002.1, YP\_006934052.1,  
 YP\_006934053.1,YP\_006934057.1, YP\_006934602.1, YP\_006934302.1, YP\_006934421.1,  
 YP\_006934504.1,YP\_006934653.1, YP\_006934638.1, YP\_006934179.1, YP\_006934155.1,  
 YP\_006934215.1,YP\_006934173.1, YP\_006934072.1, YP\_006935105.1, YP\_006935401.1,  
 YP\_006935305.1,YP\_006935314.1, YP\_006935107.1, YP\_006934250.1, YP\_006935072.1,  
 YP\_006935170.1,YP\_006935160.1, YP\_006934605.1, YP\_006934461.1, YP\_006934464.1,  
 YP\_006935262.1,YP\_006934560.1, YP\_006935112.1, YP\_006935033.1, YP\_006935363.1,  
 YP\_006935108.1,YP\_006935335.1, YP\_006935361.1, YP\_006935041.1, YP\_006934795.1,  
 YP\_006935254.1,YP\_006934797.1, YP\_006935190.1, YP\_006934460.1, YP\_006934003.1,

YP\_006935062.1, YP\_006935161.1, YP\_006935348.1, YP\_006935073.1, YP\_006934245.1, YP\_006934603.1, YP\_006935005.1

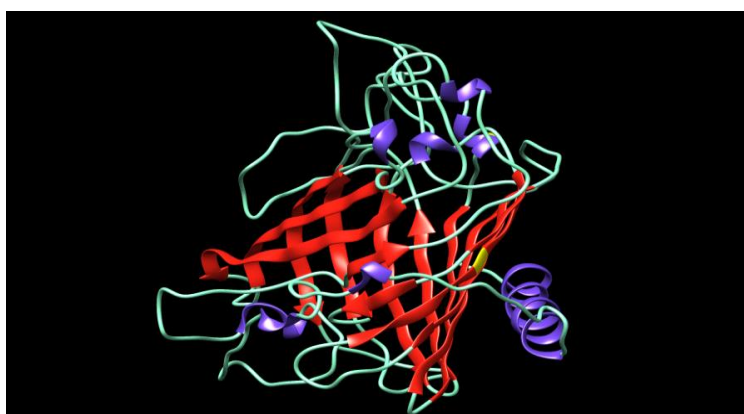
Out of the total number of 81 protein sequences from the intersection of non significant sequences of *Lactobacillus* and Human proteome, some of the proteins found in database search are hypothetical in nature, those were not considered for further analysis as eliminated out. By literature survey and pathway studies to functional analysis for all the 18 probable targets protein sequences it was found that some of the protein sequences are already known targets and the remaining can be used as a new potential probable target protein of *H. pylori* 26695 proteome. 5(five) numbers of knew probable targets which can be used for further analysis are as shown below Table 5.

**Table 5: List of probable target proteins names, accession no, GI no and amino acid length**

Sl. No.	Protein Names	Accession No.	GI NO.	Amino Acid Length
1	Phospholipase A1	AFV42658	409894600	338aa
2	Flagellar basal body P-ring protein	AFV41470	409893412	342aa
3	NADH_ubiquinone oxidoreductase subunit J	AFV42487	409894429	182aa
4	Siderophore-mediated iron transport protein(tonB)	AFV42559	409893489	505aa
5	Relaxase	AFV42209	40989151	608aa

#### Model Generation and Assessment

Protein model was generated by I-TASSER server. Structural analysis and verification results of the generated protein model revealed that the quality of the protein model generated by I-TASSER server was reliable (Figure:6). The Ramachandran plot obtained from the PROCHECK analysis reveal the stereo chemical properties of the generated model. The Exp.TM score and Exp RMSD of the protein model are 0.57+-0.14 and 9.2+-4.6. The Phospholipase A1 quality result of model showing c-score of -2.07 with cluster Density 0.0277.



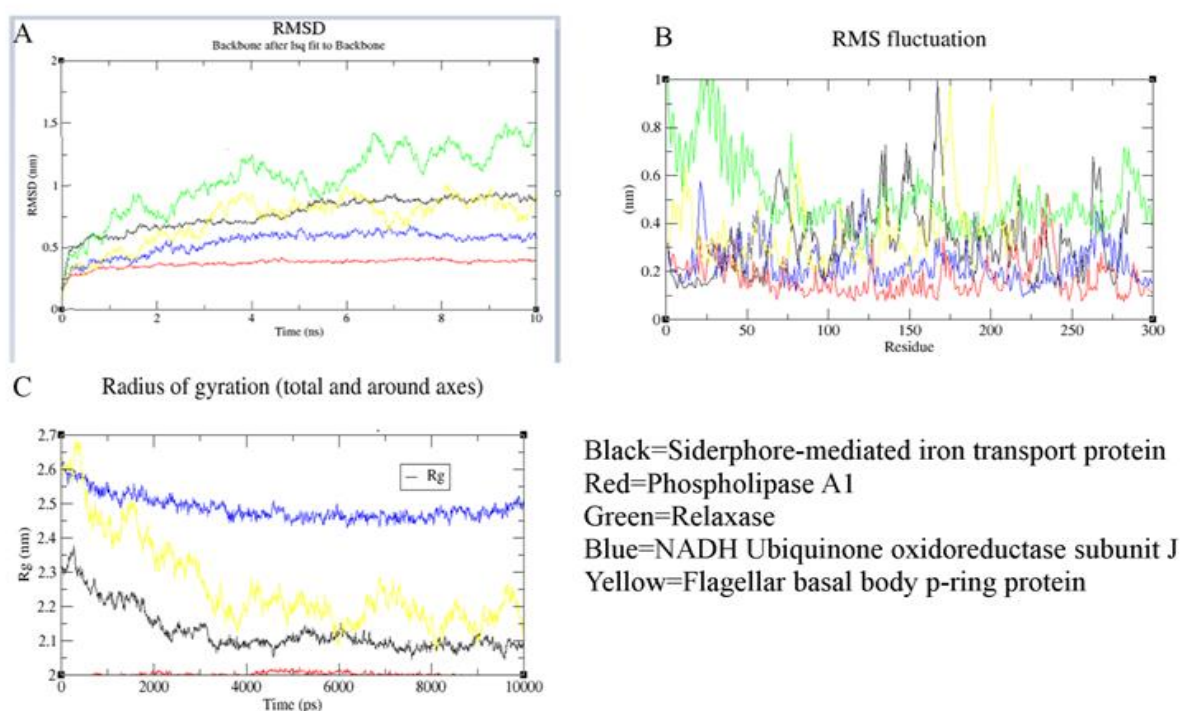
**Fig 4: Secondary structure of Phospholipase A1 protein. In this structure Red, Violet and Green colour indicates beta sheets, alpha helix and coil. The yellow colour indicates the Active site region of the protein.**

**Table 6: Phospholipase A1 generated models assessment score comparison**

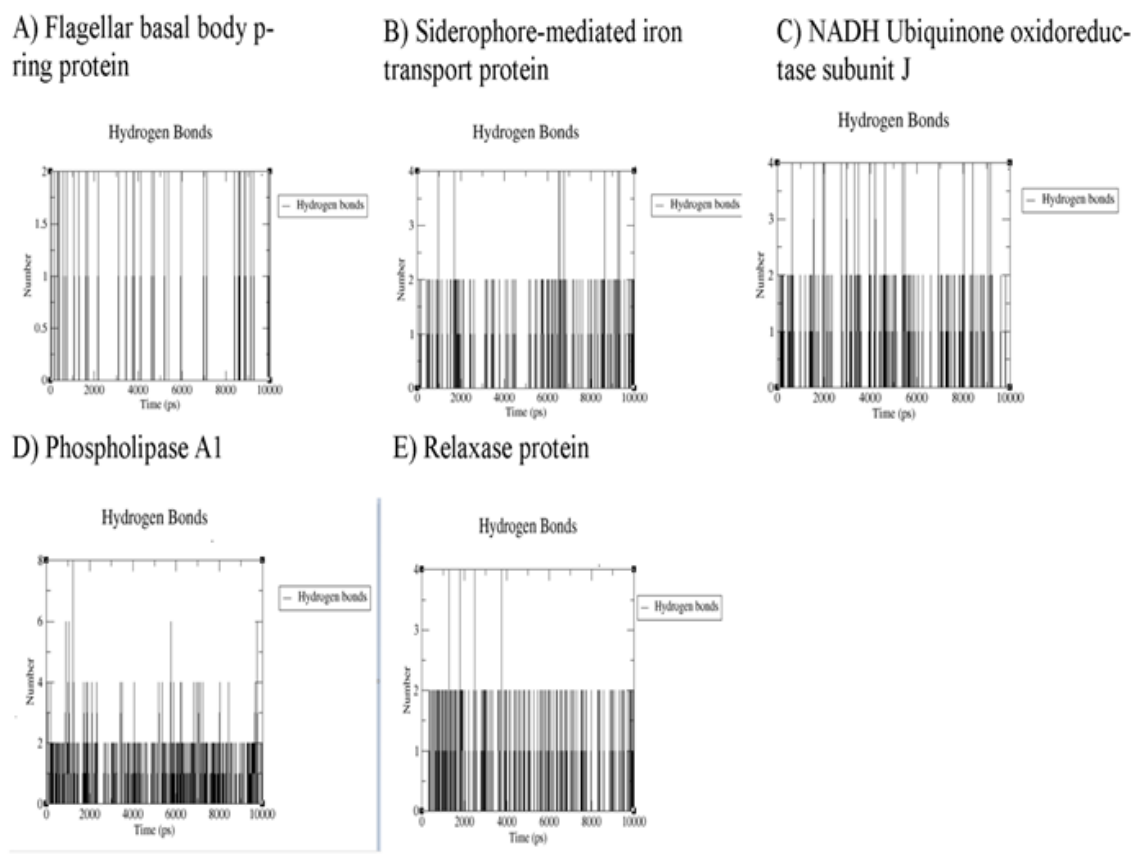
Protein	FavouredReg (%)	Allowed Reg(%)	Disallowed Reg(%)	C-Score	Exp.TM Score	Exp RMSD
I-TASSER	69.4	17.7	12.9	-2.07	0.57+-0.14	9.2+-4.6

### Molecular Dynamic (MD) Simulation

The molecular dynamic (MD) simulation study was done with modeled Phospholipase A1 protein in explicit water. A constant variation was observed in the RMSD curve after 10 ns. The radius of gyration, RMSD plot and RMS fluctuation graph of the protein model were generated (Supplementary figure, Figure: 5) showing flexibility of the modeled protein with gradual increase of time.



**Fig 5: Trajectory Analysis of Wild type Protein; A) RMSD Plot represents Phospholipase A1 is more stable; B) Represents highly fluctuating residues in NADH Ubiquinone oxidoreductase subunit J protein; C) The Radius of gyration is more stable in Phospholipase A1.**



**Fig 6: Hydrogen bond Interaction; 8 hydrogen bonds are found in Phospholipase A1 protein which is the higher value in compare to the others**

## DISCUSSION

*Helicobacter pylori* cause chronic persistent and atrophic gastritis in adults and children that often culminate the development of gastric and duodenal ulcers. From literature survey and molecular dynamics simulation of the protein *H.pylori*, out of 1590 protein sequences 5 proteins can be taken as a probable target protein in order to eradicate the disease caused by *H.pylori*. In the present study, out of the 5 proteins, Phospholipase A1 shows high rate of drug likeliness and molecular dynamics stability. The protein can be forwarded for further drug development study of gastric and duodenal ulcers.

## 4. CONCLUSION

The present study mainly aim to identification of new probable target proteins from *H. pylori* strain 26695 proteome comparison which can be further used as a novel drug targets without focusing on putative targets through database search, and activated signaling pathway studies for functional analysis of proteins by differential proteome approach. Out of 1590 protein sequences of *H. pylori* 1407 and 1409 had shown similarity at various levels to the proteins of Homo sapiens and lactobacillus species respectively. Two set consisting of 183 and 181 dissimilar proteins were sorted. On union of these two sets 81 proteins were found to be in the intersection. Out of the 18 protein sequences 18 were considered to be the candidate targets, excluding hypothetical protein. From the

18 protein sequences 13 were already known and 5 numbers of proteins was studied for further analysis. The identity percentage analysis shows a clear idea for these two proteome set as *H. pylori* 26695 protein sequences are much more evolutionary similar with Homo sapiens than Lactobacillus, i.e maximum numbers of identical sequences are found in Homo sapiens. Protein model was generated by I-TASSER server. Structural analysis and verification results of the generated protein model revealed that the quality of the protein model generated by I-TASSER server was reliable. The molecular dynamic (MD) simulation study was done with modeled Phospholipase A1, Flagellar basal body P-ring protein, NADH\_ubiquinone oxidoreductase subunit J, Siderophore-mediated iron transport protein (tonB) and Relaxase protein in explicit water. A constant variation was observed in the RMSD curve after 10 ns. The radius of gyration, RMSD plot and RMS fluctuation graph of the protein model were generated showing flexibility of the modeled protein with gradual increase of time. The Root mean square deviation, Root mean square fluctuation and Radius of gyration using g\_rmsd, g\_rmsf and g\_gyrate respectively which is available in Gromacs was analyzed for the following 5 proteins. Based on the literature review, Protein Model Assessment and Molecular Dynamic Simulation (MD) it can be concluded that out of the 5 new potential probable target proteins Phospholipase A1 protein can be one of the potential probable targeted protein which can be taken for further studies for drug likeliness.

#### **ACKNOWLEDGEMENT**

We acknowledge DBT BIF, Centre for Biotechnology and Bioinformatic, Dibrugarh University for the infrastructural facility for the research work. We also acknowledge the DELCON facility of Dibrugarh University for Literature Search and UGC for providing the National Fellowship for Higher Education of ST student.

#### **CONFLICT OF INTEREST**

The Authors have no conflict of interest.

#### **REFERENCES**

1. S Diaconu, Predescu A, Moldoveanu A, Pop C.S, and Fierbințeanu-Braticevici C. "Helicobacter pylori infection: old and new." Journal of medicine and life 10, no. 2 (2017): 112.
2. Kuipers E. J. "Helicobacter pylori and the risk and management of associated diseases: gastritis, ulcer disease, atrophic gastritis and gastric cancer." Alimentary pharmacology & therapeutics 11, no. S1 (1997): 71-88.
3. RWM Van der Hulst, JJ Keller, EAJ Rauws et al. "Treatment of Helicobacter pylori infection: a review of the world Literature". Helicobacter 1996; 1: 6-19.
4. P Correa, MB. Piazzuelo, "Natural history of Helicobacter pylori infection. Dig Liver Dis". 2008;40:490-6.
5. Graham DY. "Helicobacter pylori Update: Gastric Cancer, Reliable Therapy, and Possible

- Benefits” *Gastroenterology*. 2015;148:719-31.
6. Mégraud F, Floch P, Labenz J, Lehours P. “Diagnostic of *Helicobacter pylori* infection.” *Helicobacter*. 2016;21(Suppl 1):8-13.
  7. James Versalovic, S. Michael Osato, Klaudia Spakovsky, Dore M. Pina, Reddy Rita, Stone Gregory G, et al. “Point mutations in the 23S rRNA gene of *Helicobacter pylori* associated with different levels of clarithromycin resistance.” *The Journal of antimicrobial chemotherapy* 40, no. 2 (1997): 283-286.
  8. Megraud, F. “H pylori antibiotic resistance: prevalence, importance, and advances in testing.” *Gut* 53, no. 9 (2004): 1374-1384.
  9. Li, Bao-Zhu, Diane Erin Threapleton, Wang Ji-Yao, Xu Jian-Ming, Yuan Jin-Qiu, Zhang Chao, Li Peng et al. “Comparative effectiveness and tolerance of treatments for *Helicobacter pylori*: systematic review and network meta-analysis.” *bmj* 351 (2015): h4052.
  10. De Groot, Rudolf S., Matthew A. Wilson, and Roel of MJ Boumans. “A typology for the classification, description and valuation of ecosystem functions, goods and services.” *Ecological economics* 41, no. 3 (2002): 393-408.
  11. Kushwaha, Sandeep K., and Madhvi Shakya. “Protein interaction network analysis—approach for potential drug target identification in *Mycobacterium tuberculosis*.” *Journal of theoretical biology* 262, no. 2 (2010): 284-294.
  12. Moir, Donald T., Karen J. Shaw, Roberta S. Hare, and Gerald F. Vovis. “Genomics and antimicrobial drug discovery.” *Antimicrobial agents and chemotherapy* 43, no. 3 (1999): 439-446.
  13. Shanmugam Anusuya, and Jeyakumar Natarajan. “Computational genome analyses of metabolic enzymes in *Mycobacterium leprae* for drug target identification.” *Bioinformatics* 4, no. 9 (2010): 392.
  14. Dutta Anirban, Singh Shashi Kr, Ghosh Payel, Mukherjee Runni, Mitter Sayak, and Bandyopadhyay Debashis. “In silico identification of potential therapeutic targets in the human pathogen *Helicobacter pylori*.” *In silico biology* 6, no. 1, 2 (2006): 43-47.
  15. Yuri Churin, Al-Ghoul Laila, Kepp Oliver, Meyer Thomas F., Birchmeier Walter, and Naumann Michael. “*Helicobacter pylori* CagA protein targets the c-Met receptor and enhances the motogenic response.” *The Journal of cell biology* 161, no. 2 (2003): 249-255.
  16. TS1 Istivan, PJ Coloe. “Phospholipase A in Gram-negative bacteria and its role in pathogenesis”. *Microbiology Society*, 01 May 2006: 1263-1274.
  17. N Dorrell, MC Martino, RA Stabler, Ward SJ, Zhang ZW, McColm AA, Farthing MJ, Wren BW. “Characterization of *Helicobacter pylori* PIdA, a phospholipase with a role in colonization of the gastric mucosa.” *Gastroenterology*. 1999;117:1098–104.
  18. Istivan TS, Coloe PJ. “Phospholipase A in Gram-negative bacteria and its role in pathogenesis.”

- Microbiology. 2006;152:1263–74.
19. Tannaes T, Dekker N, Bukholm G, Bijlsma JJ, Appelmelk BJ.” Phase variation in the *Helicobacter pylori* phospholipase A gene and its role in acid adaptation.” *Infect Immun.* 2001;69:7334–40.
  20. Roy A, Kucukural A, Zhang Y , “I-TASSER: a unified platform for automated protein structure and function prediction.” *Nature Protocols* (2010);5(4): 725-738.
  21. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993),” PROCHECK: A program to check the stereochemical quality of protein structures.” *J ApplCryst* 26: 283-291.
  22. Scott WRP, Hünenberger PH, Tironi IG, Mark AE, Billeter SR, et al. “The GROMOS Biomolecular Simulation Program Package.” *J PhysChem A* (1999) ;103: 3596-3607.
  23. Lüthy R, Bowie JU, Eisenberg D “Assessment of protein models with three-dimensional profiles.” *Nature* (1992) ; 356(6364): 83-85.
  24. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, et al. "GROMACS: fast, flexible, and free.” *J ComputChem* (2005) 26(16): 1701-1718.
  25. Hess B, Bekker H, Berendsen HJC, Fraaije J, “GEM LINC: A linear constraint solver for molecular simulations.” *J ComputChem* (1997) 18(12): 1463-1472.