IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

Journal home page: www.ijreb.org

IN SILICO IDENTIFICATION OF DRUG TARGETS OF *HELICOBACTER PYLORI*

T.P.Kumari Pushpa Rani[1], Vijayalakshmi Amma[1], S.Shamya[2] and S.K.Sundar[3]

[1]Department of chemistry, Noorul Islam University, Kumaracoil

[2]Department of Biotechnology, Noorul Islam College of Arts & Science, Kumaracoil

[3]Department of Microbiology, M.R.Government Arts College, Mannargudi

Corresponding author mail id: sksundar@yahoo.com

ABSTRACT

Helicobacter pylori is a microaerophilic, gram-negative bacterium that colonizes the gastric mucosa of approximately 50% of the world's population. The HopQ protein of *H. pylori* has been found to be involved in adherence and hence in pathogenesis. In the present study, analysis of the HopQ sequence revealed that the protein is a stable one and possesses a domain region (HP OMP) as analysed using CDD and Pfam which includes four motif regions. Another result to be of importance was that the HopQ protein showed transmembrane location. The outer membrane protein has also been found to be specific to *H. pylori.* Hence the protein would be a better choice of target in the pathogen. All these results show that HopQ is a stable protein having conserved regions in its sequence. The organism seems to be resistant to many antibiotics. Considering these factors, designing a suitable cost-effective drug to counter the infection of this bacterium is the need of the hour.

KEYWORDS: *H. pylori*, Gastric mucosa drug targets, motif

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

INTRODUCTION

H. pylori is a primary pathogenic factor in benign and malignant gastroduodenal disease (Warren and Marshall 1983; Blaser and Parsonnet, 1994). Tomb *et al*. (1997) reported the complete sequence of the circular genome of *H. pylori*. The 1,667,867-bp genome contains 1,590 predicted coding sequences (genes). Sequence analysis of these genes indicated that the organism has systems for motility, for scavenging iron, and for DNA restriction and modification. Its survival in acid conditions depends, in part, on its ability to establish a positive inside-membrane potential in low pH.

Malaty *et al.* (1994) determined the *H. pylori* status in monozygotic and dizygotic twins from the Swedish Twin Registry: 36 MZ twin pairs reared apart, 64 MZ twin pairs reared together, 88 DZ twin pairs reared apart, and 81 DZ twin pairs reared together. The *H. pylori* status was determined by testing for anti-*H. pylori* IgG. The concordance rate for infection was higher in monozygotic twin pairs (81%) than in dizygotic twin pairs (63%). For 124 pairs of twins reared apart, the concordance rates were 82% and 66% for MZ and DZ twins, respectively. The correlation coefficient was 0.66 for monozygotic twins reared apart. Malaty *et al*. (1994) concluded that genetic effects influence the acquisition of *H. pylori* infection but that sharing the same rearing environment also contributes to the familial tendency.

Mendall and Northfield (1995) stated that most studies of *H. pylori* transmission have shown an increased rate of infection in the families of seropositive children, but there have been no controlled studies for variation in socioeconomic circumstances of the families. Hence, the findings may merely represent greater environmental exposure of the index positive children. In a large study involving 277 couples in a fertility clinic, Perez-Perez *et al.* (1991) found no increased rate of infection among the spouses of seropositive index cases. Mendall and Northfield (1995) noted that the study by Perez-Perez *et al*. (1991) was the only such study with sufficient power to detect modest effects and the only one to control for socioeconomic circumstances. Mendall and Northfield (1995) stated that it is unlikely that *H. pylori* could multiply in the

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

environment, suggesting that humans were probably the only source of *H. pylori* infection.

Because *H. pylori* is rarely found in deeper portions of the gastric mucosa, where O-glycans are expressed that have terminal alpha-1,4-linked N-acetylglucosamine, Kawakubo *et al.* (2004) tested whether these O-glycans might affect H. pylori growth. Kawakubo *et al.* (2004) reported that these O-glycans have antimicrobial activity against H. pylori, inhibiting its biosynthesis of cholesteryl-alpha-D-glucopyranoside, a major cell wall component. Thus, the unique O-glycans in gastric mucin appeared to function as a natural antibiotic, protecting the host from *H. pylori* infection.

Peek and Blaser (2002) reviewed the relationship between *H. pylori* and gastrointestinal tract adenocarcinomas. Although gastric adenocarcinoma is associated with the presence of *H. pylori* in the stomach, only a small fraction of colonized individuals develop this common malignancy. The authors suggested that *H. pylori* strain and host genotypes probably influence the risk of carcinogenesis by differentially affecting host inflammatory responses and epithelial cell physiology.

Kwok *et al*. (2007) found that the *H. pylori* adhesin protein CagL was targeted to the bacterial type IV secretion pilus surface, where it bound and activated the ITGA5 (135620)/ITGB1 (135630) receptor on gastric epithelial cells through its arg-gly-asp motif. CagL interaction with the integrin receptor triggered delivery of the *H. pylori* oncoprotein CagA into target cells and activation of FAK (PTK2; 600758) and SRC (190090) tyrosine kinases. Kwok *et al.* (2007) suggested that CagL may be used as a molecular tool to better understand integrin signaling and the mechanism by which *H. pylori* causes gastric ulcer and cancer.

Helicobacter pylori genomes contain about 30 different hop genes, which encode outer membrane proteins. Several Hop proteins mediate adherence of *H. pylori* to gastric epithelial cells. One outer membrane protein, HopQ (omp27), is of interest because two highly divergent families of *hopQ* alleles have been identified (Cao & Cover, 2002). HopQ is

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

known to be localized to the surface of *H. pylori* (Sabarth *et al.*, 2005).

The present study aims to study the outer membrane protein of *Helicobacter pylori* using bioinformatics tools as a target against which drugs could be designed in the future.

MATERIALS AND METHODS

The tools used in the present study were computing techniques to analyse DNA and amino acid sequences in a biologically meaningful manner. The query sequence was retrieved from Genbank were computed using Prot Param (Gasteiger *et al.*, 2005). Fasta and Blast are tools applied for sequence alignment. BlastP was used in the present study to find homologous sequences (Altschul *et al*, 1990) to the query sequence and the sequences were then compared with the query sequences and arranged in the ascending values of E.

The conserved regions were detected using PROSITE and Pfam tools to establish functional domains and regular expression patterns (Finn *et al.*, 2008). The CDD is used to identify the conserved protein domain (Marchler-Bauer *et al.*, 2009). Finger print Analysis was done to build diagnostic signatures of the protein family membership and the fingerprints thus created were used to identify the distant relative of the protein in Prints database (Attwood *et al.*, 1998).

The secondary structure for the query protein was predicted using Jpred. The steric properties of amino acids were considered for the prediction of secondary structure. The numbers of proline and glycine residues were taken into consideration as they exhibit reduced and complete torsional freedom respectively. Prediction of transmembrane regions and orientation of the query protein was predicted using TMpred (Kroug *et al.*, 2001). The hydrophobic analysis was first being carried out by computing the percentage of non-polar amino acids present in the query protein.

RESULTS AND DISCUSSION

The outer membrane protein (HopQ) sequence of *Helicobacter pylori* with accession number HQ343310 was retrieved from GenBank .The query sequence which was a linear protein of 631 amino acid residues (Table 1) was subjected to compositional analysis using various bioinformatics tools. The individual amino acid composition, aliphatic and aromatic

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

composition and hydrophobicity of the query protein were computed using ProtParam and the results were given in Table 2. The protein possesses both acidic and basic amino acid residues. The amino acids which provide stable secondary structures such as glutamic acid and leucine were present with an instability index of 22.40 and therefore it was classified as a stable protein. The molecular weight of the protein was found to be 68732.7 Daltons. The composition of aliphatic residues was more (80.92%) than that of aromatic residues in the query protein.

The query protein was subjected to BLASTP analysis & the alignments of similar sequences were computed on the basis of expectation (E) values. The sequences in the database showed high level of similarity with the query sequences as denoted by the e-values (E^{-100} to E^{-50}). They included sequences of the outer membrane protein of *H. pylori* strains isolated from different geographical locations.

The Conserved Domain Search of the query sequence revealed an HP OMP domain at the region between 458 to 631. Gram-negative bacteria outer membranes constitute a semi-permeable, size- dependent permeability barrier, for example to hydrolytic enzymes, detergents, dyes and hydrophobic anti-microbials. The outer membrane protein (OMP) profile of *Helicobacter pylori* differs from that of other Gram-negative bacteria, where the highly non-selective porins are absent and a number of less abundant protein species are observed [PUBMED:9252185]. OMPs from *H. pylori* have been identified as porins, gastric epithelial cell adhesins and Lewis B binding adhesins [PUBMED:9430586]. Extensive C-terminal sequence similarity between these OMPs has been used to define a much larger paralogous family.

H. pylori is the causative agent of gastritis and peptic ulceration in humans. Numerous subtypes of OMPs have also been identified. Attempts have been made to construct recombinant vectors that are able to express these OMPs in order to develop a vaccine protecting against Hp infection and a diagnostic reagent kit to quickly detect infection. OMPs were chosen as possible targets of vaccine development as they are *H. pylori* specific, surface exposed and highly antigenic. The Pfam analysis also

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

Journal home page: www.ijreb.org

showed the HP OMP domain between the residues 304 and 472 (Table 3).

The query protein was then subjected to Motif analysis using PRINTS39 and Matrix Blos62. The results showed ten fingerprints, of which the HP OMP family fingerprint had four motif regions (Table 4). The motif 1 (447-469) lies within the region encoded by Pfam pattern HP OMP. The other three motif regions, motif 2 (473-494), motif 3 (494-516) and motif 4 (591-613) lie in the domain region encoded by CDD (, Figure 1).

In the present study secondary structure of the query protein was predicted using Jpred analysis. The query protein mainly exhibited α- helix conformation (Table 5). The amino acids like Tryptophan and Methionine which have propensity for α helix were dominant in the query protein (1.1%, 0.7% respectively).

Since the query protein is a surface protein, computational techniques (TMpred) were used to predict the transmembrane helices.

The results are presented in the (Table 6). Four helices of inside to outside orientation (6 to 32, 111 to 130, 491 to 513 and 527 to 543) were predicted. Similarly three helices of outside to inside orientation were predicted for the query protein. Of them three of the orientations from inside to outside and one orientation form outside to inside showed more preference. The presence of transmembrane helices indicates that a protein has the membrane bound location (Krogh *et al.*, 2001). This confirms that the protein lies on the cell membrane of the organism. The observations made in the present study regarding the HP OMP domain and motifs will definitely throw light in designing a new drug candidate to suitably dock these regions or in constructing a recombinant vaccine using these motifs as epitopes which will be a great leap in the treatment of dreadful diseases caused by *H. pylori*.

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

IJREB

Journal home page: www.ijreb.org

Table: 1 Query sequence (hq343310) retrieved from Genbank

```
      1 mkktkktill sltlassllh aedngvflsv gyqigeavqk vknadkvqkl sdayenlnki
     61 lanhdhsnpe aintnsatai nqaignlnan tqnlidktdn spayqatlla lkstvglwns
    121 iayavicggy tdkpnhnite tfynqpgqns nsitcgsngl gtlpagknsh lsieqfatln
    181 kayqiiqaal kqglpalsdt kktvevtikt atnaqninvn nnnnnaadat ietkntyind
    241 aqnlltqaqt iintlqdncp mlkgkssgt ngantpswqt sanqnscsvf gtefsaisdm
    301 isnaqnivqe tqqlnttplk siaqphnfnl nspnsvalaq smlknaqsqa avlklanqvg
    361 ndfnristgv lknyieecna nassesvsnn twgkgcagvk qtltslessn asfssqtpqi
    421 nqaetianti vqelghnpfk rvgiissqtn ngamnglgvq vgykqffgek krwglryygf
    481 fdynhayiks sffnsasdvw tygvgsdllf nfindkntnf lgknnqisfg lfggialagt
    541 swlnsqfvnl ktisnvysak vntanfqflf nlglrtnlar pkkkdshhaa qhgmelgvki
    601 ptintnyysf ldtkleyrrl ysvylnyvfa y
```

Table: 2 Protparam results of the query sequence

```
Number of amino acids                                  : 631
Molecular weight                                       : 68732.7
Theoretical pI                                         : 9.07
Total number of negatively charged residues (Asp + Glu)  : 38
Total number of positively charged residues (Arg + Lys)  : 49
Ext. coefficient                                       : 67645
Abs 0.1% (=1 g/l)                                      : 0.984

  assuming all pairs of Cys residues form cystines

Ext. coefficient                                       : 67270

Abs 0.1% (=1 g/l)                                      : 0.979

   assuming all Cys residues are reduced

The N-terminal of the sequence considered              : M (Met)
The estimated half-life is                             : 30 hours

                          (mammalian reticulocytes, in vitro)

                           >20 hours (yeast, in vivo)

                           >10 hours (Escherichia coli, in vivo)

Instability index (II)                                 : 22.40
                             (This classifies the protein as stable)

Aliphatic index                                        : 80.92
Grand average of hydropathicity (GRAVY)                : -0.374
```

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

Table: 3 Pfam domain analysis result

Source	Domain	Start	End
sig_p		1	44
low_complexity		23	34
low_complexity		82	101
low_complexity		106	208
low_complexity		153	175
Pfam A	HP OMP	304	472

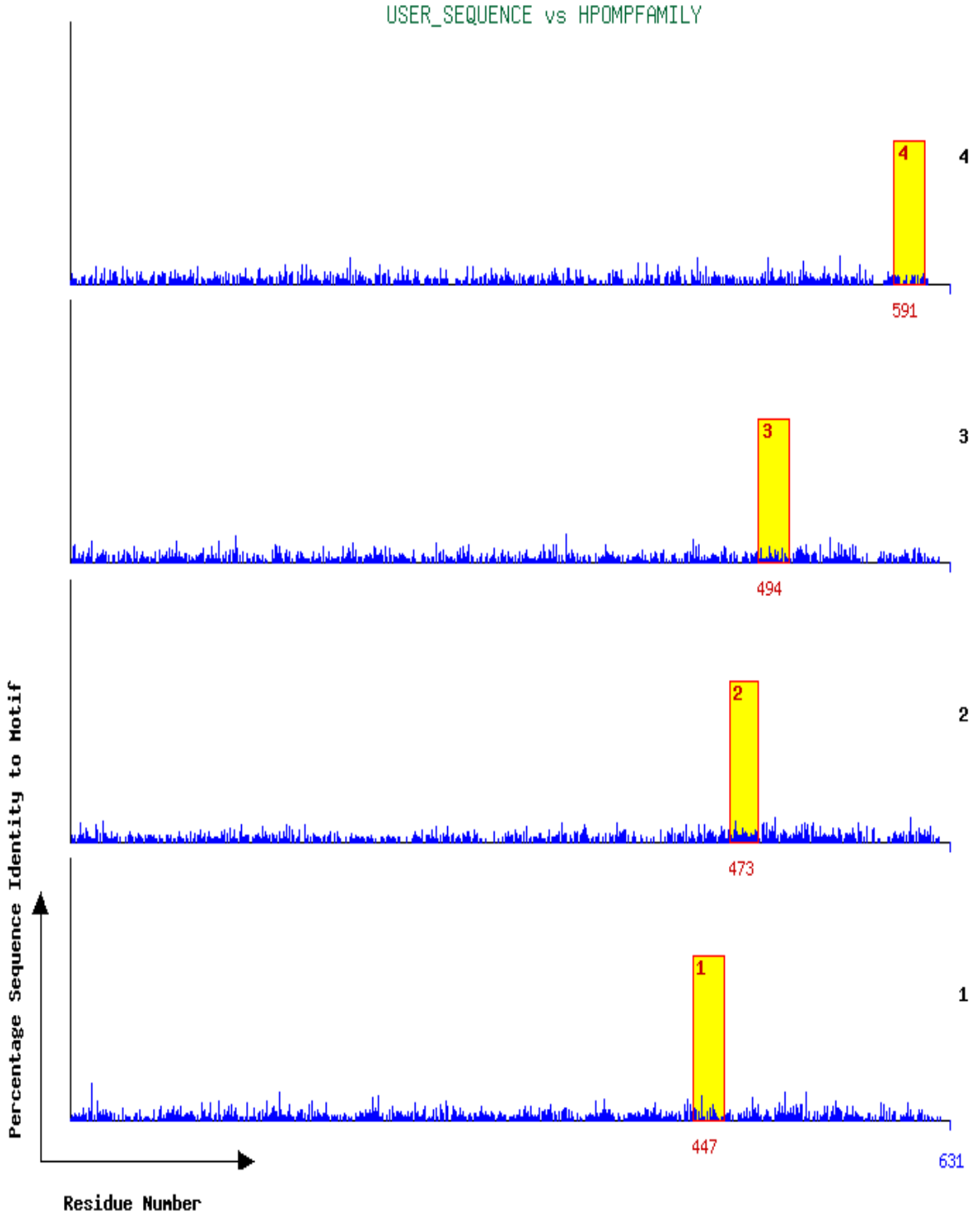
Table: 4 Fingerprint analysis of query sequence

Ten top scoring fingerprints for your query. Detailed by motif

FingerPrint Name	Motif Number	IdScore	PfScore	Pval	Sequence	Length	low	Pos	high
HPOMPFAMILY	1 of 4	63.03	737	3.89e-15	SQTNNGAMNGLGVQVGYKQFFG	22	0	447	0
	2 of 4	61.12	710	7.25e-14	WGLRYYGFFDYNHAYIKSSFF	21	0	473	0
	3 of 4	54.61	632	1.05e-12	NSASDVWTYGVGSDLLFNFIND	22	0	494	0
	4 of 4	54.55	640	5.06e-13	QHGMELGVKIPTINTNYYSFLD	22	0	591	0
ALPHATUBULIN	4 of 13	44.63	254	4.81e-03	DKVQKLSDA	9	0	45	0
	9 of 13	40.30	249	3.59e-03	INVNNNNNNAADATIETKNTYI	22	0	217	0
VACCYTOTOXIN	13 of 15	32.55	236	1.39e-02	PEAINTNSATAINQAIGNLNANT	23	0	69	0
	15 of 15	33.33	280	1.54e-03	FGGIALAGTSWLNSQFVNLKT	21	0	532	0
CD97PROTEIN	2 of 13	20.00	199	8.96e-02	AQNLLTQAQTIINTLQDNCP	20	0	241	0
	5 of 13	29.63	207	1.46e-02	QTLTSLESSNASFSSQTP	18	0	401	0
	9 of 13	24.56	185	3.93e-02	LLFNFINDKNTNFLGKNNQ	19	0	508	0
STAPHEXOTOXN	1 of 4	27.45	156	1.11e-02	TGVLKNYIEECNANA	15	0	368	0
	4 of 4	30.15	151	2.08e-02	NTNYYSFLDTKLEYRR	16	0	604	0
POAALLERGEN	7 of 8	31.58	210	7.53e-03	ATLNKAYQIIQAALKQGLP	19	0	177	0
	8 of 8	27.41	130	6.01e-02	TKKTVEVTIKTATNA	15	0	200	0
INTEGRINB	10 of 14	18.96	177	1.88e-02	TAINQAIGNLNANTQNLIDKTDNSP	25	0	78	0
	12 of 14	29.42	144	2.42e-02	GKNNQISFGLFGGIALAG	18	0	522	0
OCTOPAMINER	2 of 7	31.11	206	4.38e-02	VICGGYTDKPNHNITETF	18	0	125	0
	6 of 7	41.43	218	3.31e-02	YKQFFGEKKRWGLR	14	0	463	0
P2X4RECEPTOR	3 of 10	28.89	187	3.00e-02	TLNKAYQII	9	0	178	0
	6 of 10	44.44	184	6.37e-02	RISTGVLKN	9	0	365	0
THIKCHANNEL	1 of 7	37.50	175	4.50e-02	HAEDNGVFLSVGYQIG	16	0	20	0
	7 of 7	34.62	185	4.74e-02	SVALAQSMLKNAQ	13	0	335	0

ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

Figure: 1 Graphical representation of OMP MOTIF regions



ISSN 2321-743X
International Journal of Research in
Engineering and Bioscience
Volume 2 Issue 1
IJREB
Journal home page: www.ijreb.org

Table: 5 Secondary structure prediction

```
QuerySeq : 1---------11--------21--------31--------41--------51--------61--------71--------81
         : MKKTKKTILLSLTLASSLLHAEDNGVFLSVGYQIGEAVQKVKNADKVQKLSDAYENLNKILANHDHSNPEAINTNSATAINQ
           --------91--------101-------111-------121-------131-------141-------151-------161-
           AIGNLNANTQNLIDKTDNSPAYQATLLALKSTVGLWNSIAYAVICGGYTDKPNHNITETFYNQPGQNSNSITCGSNGLGTLP
           ------171-------181-------191-------201-------211-------221-------231-------241---
           AGKNSHLSIEQFATLNKAYQIIQAALKQGLPALSDTKKTVEVTIKTATNAQNINVNNNNNNAADATIETKNTYINDAQNLLT
           ----251-------261-------271-------281-------291-------301-------311-------321-----
           QAQTIINTLQDNCPMLKGKSSSGTNGANTPSWQTSANQNSCSVFGTEFSAISDMISNAQNIVQETQQLNTTPLKSIAQPHNF
           --331-------341-------351-------361-------371-------381-------391-------401-------
           NLNSPNSVALAQSMLKNAQSQAAVLKLANQVGNDFNRISTGVLKNYIEECNANASSESVSNNTWGKGCAGVKQTLTSLESSN
           411-------421-------431-------441-------451-------461-------471-------481-------49
           ASFSSQTPQINQAETIANTIVQELGHNPFKRVGIISSQTNNGAMNGLGVQVGYKQFFGEKKRWGLRYYGFFDYNHAYIKSSF

           1-------501-------511-------521-------531-------541-------551-------561-------571-

           FNSASDVWTYGVGSDLLFNFINDKNTNFLGKNNQISFGLFGGIALAGTSWLNSQFVNLKTISNVYSAKVNTANFQFLFNLGL

           ------581-------591-------601-------611-------621--------: QuerySeq
           RTNLARPKKKDSHHAAQHGMELGVKIPTINTNYYSFLDTKLEYRRLYSVYLNYVFAY: QuerySeq

           1---------11--------21--------31--------41--------51--------61--------71--------81
Jnet     : ---HHHHHHHHHHHHHHH------EEEEEEEEHHHHHHH----HHHHHHHHHHHHHHHHHHHHHHHHHHH---HHHHHHH
jhmm     : --HHHHHHHHHHHHHHHH------EEEEEEEEEHHHHHHE--HHHHHHHHHHHHHHHHHHHHHHHHHHHH---HHHHHHH
jpssm    : -----HHHHHHHHHHHHH-------EEEEHHHHHHHHHH------HHHHHHHHHHHHH-----------EE----HHHHH

           --------91--------101-------111-------121-------131-------141-------151-------161-
Jnet     : HHHHHHHHHHH---------HHHHHHHHHHHHHHHHHHHHEEEEE-------------------------EEE---------
jhmm     : HHHHHHHHHHH--------HHHHHHHHHHHHHHHHHHHHEEEEEE-------------------------EEE--------
jpssm    : HHHHHHHHHHHH----------HHHHHHHHHHHHHHHHHHHEEEE-------------E----------EEE---------
           ------171-------181-------191-------201-------211-------221-------231-------241---

Jnet     : --------HHHHHHHHHHHHHHHHHH------------EEEEEEEEEE----------------------HHHHHHHHH
jhmm     : --------HHHHHHHHHHHHHHHHHH------------EEEEEEEEEE----------------------HHHHHHHHH
jpssm    : --------HHHHHHHHHHHHHHHHHH-----------EEEEEEEEE---------------E--------E---HHHHHHH
           ----251-------261-------271-------281-------291-------301-------311-------321-----
Jnet     : HHHHHHHHHHHHHHHHE--------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHH---------------
jhmm     : HHHHHHHHHHHHHHHHH-------------------HHHHHHHHHHHHHHHHHHHHHHHHHHHHH---------------
jpssm    : HHHHHHHHHHHHH---EEE-------------------------HHHHHHHHHHHHHHHHHHHHHHH---------------

           --331-------341-------351-------361-------371-------381-------391-------401-------
Jnet     : -------HHHHHHHHHHH-HHHHHHHHHHHHHHH------HHHHHHHHHH----------------EEHHHHHHHHHHHH
```

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

IJREB

Journal home page: www.ijreb.org

```
jhmm   : -------HHHHHHHHHHHHHHHHHHHHHHHHHHHH------HHHHHHHHHH----------------EEEEHHHHHHHHHHH
jpssm  : -------HHHHHHHHHH----HHHHHHHHHHHHHH------HHHHHHHHHH----------------HHHHHHHH-HHH--

          411-------421-------431-------441-------451-------461-------471-------481-------49
Jnet   : H------HHHHHHHHHHHHHHHHHH------EEEE----------EEEEEEEE--------EEEEEEEEE------EEE---
jhmm   : HH-----HHHHHHHHHHHHHHHHHH-----EEEEEE----------EEEEEEE-------EEEEEEEEEE------EEE---
jpssm  : ------HHHHHHHHHHHHHHHHHH-----------------EEEEEEEEEEEE-----EEEEEEEEEE-----EE----

          1-------501-------511-------521-------531-------541-------551-------561-------571-

Jnet   : -------H--HHHHHHHHEEE-----------EEEEEE-HHHHHH-------HEE----------------HHHHHHHHH
jhmm   : -----------HHHHHHHEEE-----------EEEEEE-HHHHHH------HHHH-HH-------EEEE---HHHHHHHHH
jpssm  : ---EEEEHHHHHHHHHHEEEEE----------EEEEEEHHHHHHHHH------EE--------------HHHHHHHHHH-E

          ------581-------591-------601-------611-------621--------

Jnet   : HE--E-----------H--EEEEEEE------EE-----HHEEEEEHHHHHHHHH--
jhmm   : H---HH------HHHHHH--EEEEEE------HHHH---HHHHHHHHHHHHHHHH--
jpssm  : EEEEE-------------EEEEEEEEEEEEEEEE-----EEEEEEEEEEEEEEEE--
```

H = helix, E = strand, - =no prediction

Jnet - Final secondary structure prediction for query

jalign - Jnet alignment prediction

jhmm - Jnet hmm profile prediction

jpssm - Jnet PSIBLAST pssm profile predict

Table: 6 - Transmembrane helices prediction

1) Possible transmembrane helices

```
Inside to outside helices :   4 found
      from         to     score center
   6 (   6)  32 (  25)    434     17
 111 ( 111) 130 ( 130)   1157    120
 491 ( 495) 513 ( 511)     96    503
 527 ( 527) 543 ( 543)   1571    535

Outside to inside helices :   3 found
      from         to     score center
 106 ( 110) 131 ( 126)    913    118
 442 ( 442) 468 ( 463)    180    453
 527 ( 527) 543 ( 543)   1465    535
```

2) Table of correspondences

Helices shown in brackets are considered insignificant.
A "+"-symbol indicates a preference of this orientation.
A "++"-symbol indicates a strong preference of this orientation.

```
          inside->outside | outside->inside
(    6-  32 (27)  434 ++ ) |
  111- 130 (20) 1157 ++    |   106- 131 (26)  913
                           |(  442- 468 (27)  180 ++ )
(  491- 513 (23)   96 ++ ) |
  527- 543 (17) 1571  +    |   527- 543 (17) 1465
```

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

Journal home page: www.ijreb.org

3) Suggested models for transmembrane topology

```
2 possible models considered, only significant TM-segments used

-----> STRONGLY prefered model: N-terminus inside
 2 strong transmembrane helices, total score : 2622
 # from   to length score orientation
 1  111  130 (20)     1157 i-o
 2  527  543 (17)     1465 o-i

------> alternative model
 2 strong transmembrane helices, total score : 2484
 # from   to length score orientation
 1  106  131 (26)      913 o-i
 2  527  543 (17)     1571 i-o
```

REFERENCES

1. Altschul, Stephen F., Warren Gish, Webb Miller, Eugene W. Myers, and David J. Lipman (1990). Basic local alignment search tool. *J. Mol. Biol*. 215:403-10.

2. Attwood, T. K., Beck, M. E., Flower, D. R., Scordis, P. & Selley, J. (1998). The PRINTS protein fingerprint database in its fifth year. *Nucl. Acids Res*. 26(1), 304-308.

3. Blaser, M. J., Parsonnet, J. Parasitism by the 'slow' bacterium Helicobacter pylori leads to altered gastric homeostasis and neoplasia. J. Clin. Invest. 94: 4-8, 1994. [PubMed: 8040281].

4. Cao P, Cover TL. Two different families of *hopQ* alleles in *Helicobacter pylori*. J Clin Microbiol. 2002;40:4504–4511.

5. Finn R.D. , J. Tate, J. Mistry, P.C. Coggill, J.S. Sammut, H.R. Hotz, G. Ceric, K. Forslund, S.R. Eddy, E.L. Sonnhammer and A. Bateman, (2008). *The Pfam protein families database. Nucleic Acids Research*, Database Issue 36:D281-D288.

6. Gasteiger E., Hoogland C., Gattiker A., Duvaud S., Wilkins M.R., Appel R.D., Bairoch A. *Protein Identification and Analysis Tools on the ExPASy Server;* (In) John M. Walker (ed): The Proteomics Protocols Handbook, Humana Press (2005). pp. 571-607.

7. Kawakubo, M., Ito, Y., Okimura, Y., Kobayashi, M., Sakura, K., Kasama, S., Fukuda, M. N., Fukuda, M., Katsuyama, T., Nakayama, J. Natural antibiotic function of a human gastric mucin against Helicobacter pylori infection. Science 305: 1003-1006, 2004. [PubMed: 15310903].

8. Kroug, A., Larsson, B., Von Heijine, G. and Sonnhammer, E.L.L. (2001). Predicting transmembrane protein topology with a hidden Markov model: Application to complete

IJREB

ISSN 2321-743X

International Journal of Research in

Engineering and Bioscience

Volume 2 Issue 1

Journal home page: www.ijreb.org

genomes. *Journal of molecular biology*, 305(3): 567-580.

9. Kwok, T., Zabler, D., Urman, S., Rohde, M., Hartig, R., Wessler, S., Misselwitz, R., Berger, J., Sewald, N., Konig, W., Backert, S. Helicobacter exploits integrin for type IV secretion and kinase activation. Nature 449: 862-866, 2007. [PubMed: 17943123].

10. Marchler-Bauer A et al. (2009), "CDD: specific functional annotation with the Conserved Domain Database.", *Nucleic Acids Res*. 37(D) 205-10.

11. Mendall, M. A., Northfield, T. C. Transmission of Helicobacter pylori infection. Gut 37: 1-3, 1995. [PubMed: 7672655].

12. Peek, R. M., Jr., Blaser, M. J. Helicobacter pylori and gastrointestinal tract adenocarcinomas. Nature Rev. Cancer 2: 28-37, 2002. [PubMed: 11902583].

13. Perez-Perez, G. I., Witkin, S. S., Decker, M. D., Blaser, M. J. Seroprevalence of Helicobacter pylori infection in couples. J. Clin. Microbiol. 29: 642-644, 1991. [PubMed: 2037687].

14. Sabarth N, Hurvitz R, Schmidt M, Zimny-Arndt U, Jungblut PR, Meyer TF, Bumann D. Identification of *Helicobacter pylori* surface proteins by selective proteinase K digestion and antibody phage display. J Microbiol Methods. 2005;62:345–349.

15. Tomb, J.-F., White, O., Kerlavage, A. R., Clayton, R. A., Sutton, G. G., Fleischmann, R. D., Ketchum, K. A., Klenk, H. P., Gill, S., Dougherty, B. A., Nelson, K., Quackenbush, J., and 30 others The complete genome sequence of the gastric pathogen Helicobacter pylori. Nature 388: 539-547, 1997. [PubMed: 9252185].

16. Warren, J. R., Marshall, B. Unidentified curved bacilli on gastric epithelium in active chronic gastritis. Lancet 321: 1273-1275, 1983. (Note: Originally Volume I.) [PubMed: 6134060].