A bioinformatics assay on Catalase epitopes of *Helicobacter pylori*

Niloufar Rashidi [1,*], Hajieh Gh. Safaei[2]

[1]Department of medical laboratory sciences, Faculty of paraMedicine, Ahvaz jundishapour University of Medical Sciences, Ahvaz, Iran

[2]Department of Microbiology, Faculty of Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

* Presenter & Corresponding author

Abstract:

Introduction:

T cell responses are stimulated by the presence of short peptides bound on the surface of antigen presenting cells. These epitopes are derived from antigens of the pathogen by a series of steps that include the proteolytic processing of antigenic proteins by the antigen presenting cells and binding of the resulting peptides to major histocompatibility complex (MHC) molecules.

Protective immune responses may be generated based on a single immunodominant epitope or more epitopes.

Immunoinformatics tools permit us to decide on epitopes or sub-sequences of proteins that interface with the T cells of the host and predict the MHC binder. Such tools let the scanning of genome-derived protein sequences for T cell epitopes, the 8-10 mer peptides that bind to MHC and interact with the Tcell receptor, stimulating Tcell response. These epitope prediction tools are proved to be very useful, since they significantly reduce the time and effort implicated in screening probable epitopes, mainly for pathogens which there is no vaccines available now. Prediction of catalase epitopes *insilico* would be a valuable tool for developing useful immuoprophylatic strategy against *H. pylori* infection. In addition, it could be helpful for the analysis of other *H.pylori* antigens and other pathogens and provide a successful progress for the design of epitope-based vaccines against many pathogens.

Methods:

Specific sequences for *H.pylori* (26695) catalase in comparison with human catalase using blast identified. We selected epitopes located in regions which have shown a very low degree of sequence similarity with the human enzyme. These regions can be introduced as specific regions for *H. pylori* protein.

Sequence of *Helicobacter pylori* 26695 catalase has 98% to 100% similarity with catalase of other strains of *H.pylori* using blast. We used Propred software for prediction promiscuous T-cell epitopes (www.imtech.res.in/raghava). The propred web interface allows users to predict MHC class II binding regions in antigen sequence. Fifty- one MHC class II epitope-mapping matrices were developed at Propred.

Results:

According to blast with human catalase, our result showed the entirely specific regions and specific regions. In this study we selected seven epitopes from entirely specific and specific regions. The epitopes were chosen from these regions: the first epitope from 1-10, the second epitope from 32-44, the third epitope from 267- 288, the fourth epitope from 210-233, the fifth epitope from 210-233, the sixth epitope from 420-438 and the seventh epitope from 420-438 regions of *H.pylori* catalase.

Discussion:

Prevalence of *Helicobacter pylori* infection is from 25% in developed countries to more than 90% in developing areas. *H.pylori* is a major cause of gastric carcinoma and lymphoma. Catalase is highly conserved, important for survival of the organism. Also it is expressed in high level and is exposed on surface of bacteria; therefore it is a suitable antigen candidate for epitope vaccine assays. The suggested epitope should be rather a composition of selected seven epitopes: MVNKDVKQTTKKVLLQSTWFLKKFHPFDVTKIKKWVKFHFHTMQKKVKHLTNEEAKKYRADDSDYYKK YYRSLPADEK. This epitope involved 43 out of 51 alleles of MHC class II and stimulated T-cell responses. Lysines (KK) that located between epitopes have some advantages: a) cathepsine B, a protease that has role in processing antigens for presentations by MHC II, considers KK as a target. b) It is possible that a new epitope forms from linkage of some epitopes; therefore, lysines (KK) are an avoiding factor for this undesired event.

Selected epitopes had the high score for reactivating MHC out of 51 MHC class II in Propred software in comparison with other epitopes. The first epitope (MVNKDVKQTT) contain 26, the second (VLLQSTWFL) contain 21, the third (FHPFDVTKI) contain 16, the fourth (WVKFHFHTMQ) contain 22, the fifth (VKHLTNEEA) contain 7, the sixth(YRADDSDYY) contain 10 and the seventh(YYRSLPADEK) contain 11 alleles of 51 MHCII. Therefore, since immunoinformatics tools can develop effective immuoprophylatic strategy against *H. pylori* infection, we consider those as suitable plans. We registered these epitopes in NCBI; GenBank accession number: Cat JQ361787

Keywords: *Helicobacter pylori*, catalase, epitopes

References:

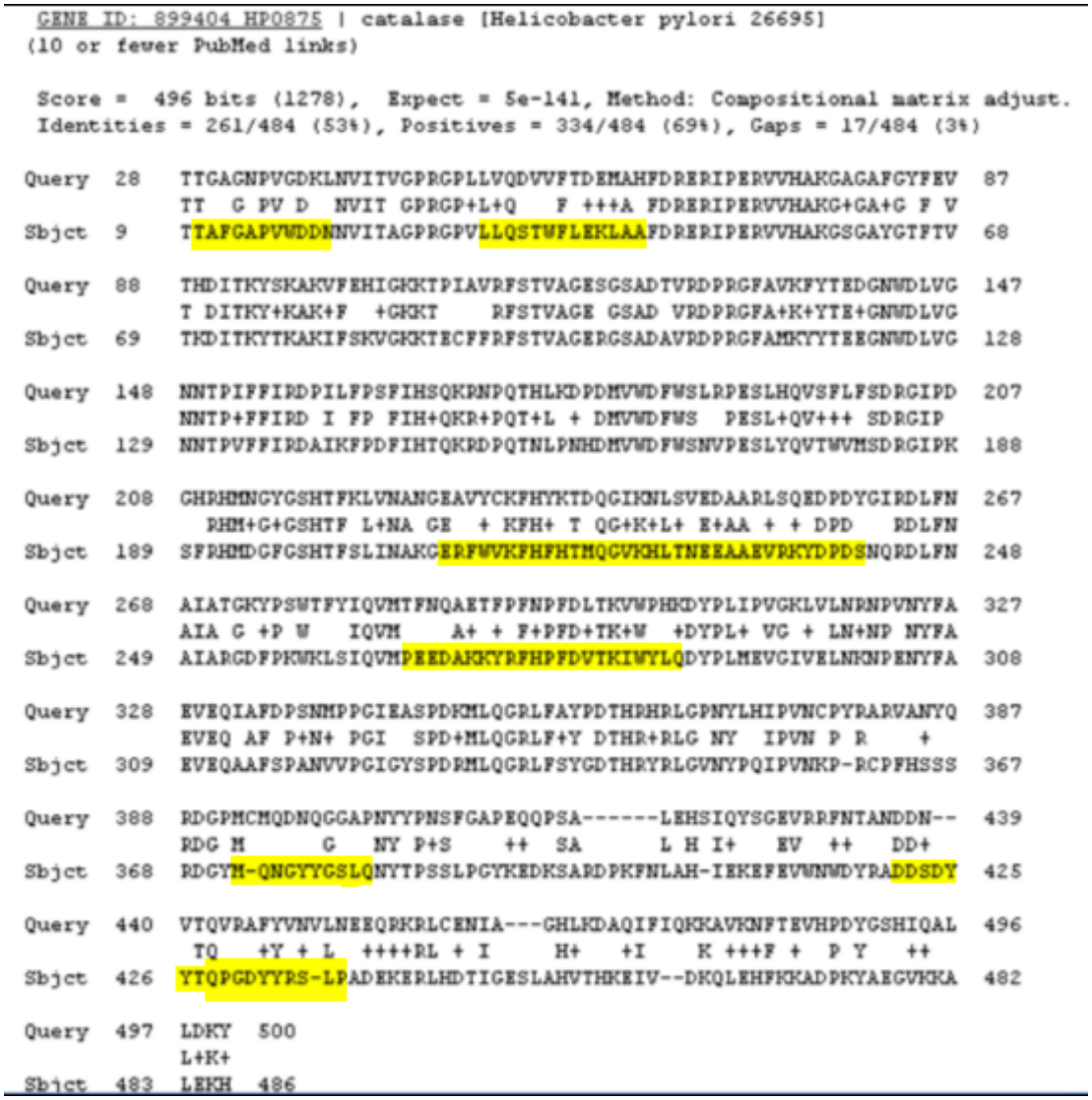
1. Moise, L., et al., Identification of genome-derived vaccine candidates conserved between human and mouse-adapted strains of H. pylori. Human vaccines, 2007. ***4****(3).*

2. Zhou, W.Y., et al., Therapeutic efficacy of a multi-epitope vaccine against Helicobacter pylori infection in BALB/c mice model. Vaccine, 2009. ***27****(36): p. 5013-5019.*

3. De Groot, A.S. and J.A. Berzofsky, From genome to vaccine—new immunoinformatics tools for vaccine design. Methods, 2004. ***34****(4): p. 425-428.*

4. Khan, A.M., et al., A systematic bioinformatics approach for selection of epitope-based vaccine targets. Cellular immunology, 2006. ***244****(2): p. 141-147.*

5. Guy, B., et al., Do Th1 or Th2 sequence motifs exist in proteins? Identification of amphipatic immunomodulatory domains in Helicobacter pylori catalase. Immunology letters, 2005. ***96****(2): p. 261-275.*



```
 GENE ID: 899404 HP0875 | catalase [Helicobacter pylori 26695]
(10 or fewer PubMed links)

 Score =  496 bits (1278),  Expect = 5e-141, Method: Compositional matrix adjust.
 Identities = 261/484 (53%), Positives = 334/484 (69%), Gaps = 17/484 (3%)

Query  28   TTGAGNPVGDKLNVITVGPRGPLLVQDVVFTDEMAHFDRERIPERVVHAKGAGAFGYFEV  87
            TT  G PV D  NVIT GPRGP+L+Q    F +++A FDRERIPERVVHAKG+GA+G F V
Sbjct  9    TTAFGAPVWDDNNVITAGPRGPVLLQSTWFLEKLAAFDRERIPERVVHAKGSGAYGTFTV  68

Query  88   THDITKYSKAKVFEHIGKKTPIAVRFSTVAGESGSADTVRDPRGFAVKFYTEDGNWDLVG  147
            T DITKY+KAK+F  +GKKT    RFSTVAGE GSAD VRDPRGFA+K+YTE+GNWDLVG
Sbjct  69   TKDITKYTKAKIFSKVGKKTECFFRFSTVAGERGSADAVRDPRGFAMKYYTEEGNWDLVG  128

Query  148  NNTPIFFIRDPILFPSFIHSQKRNPQTHLKDPDMVWDFWSLRPESLHQVSFLFSDRGIPD  207
            NNTP+FFIRD I FP FIH+QKR+PQT+L + DMVWDFWS  PESL+QV+++ SDRGIP
Sbjct  129  NNTPVFFIRDAIKFPDFIHTQKRDPQTNLPNHDMVWDFWSNVPESLYQVTWVMSDRGIPK  188

Query  208  GHRHMNGYGSHTFKLVNANGEAVYCKFHYKTDQGIKNLSVEDAARLSQEDPDYGIRDLFN  267
              RHM+G+GSHTF L+NA GE  + KFH+ T QG+K+L+ E+AA + + DPD   RDLFN
Sbjct  189  SFRHMDGFGSHTFSLINAKGERFWVKFHFHTMQGVKHLTNEEAAEVRKYDPDSNQRDLFN  248

Query  268  AIATGKYPSWTFYIQVMTFNQAETFPFNPFDLTKVWPHKDYPLIPVGKLVLNRNPVNYFA  327
            AIA G +P W   IQVM    A+ + F+PFD+TK+W  +DYPL+ VG + LN+NP NYFA
Sbjct  249  AIARGDFPKWKLSIQVMPEEDAKKYRFHPFDVTKIWYLQDYPLMEVGIVELNKNPENYFA  308

Query  328  EVEQIAFDPSNMPPGIEASPDKMLQGRLFAYPDTHRHRLGPNYLHIPVNCPYRARVANYQ  387
            EVEQ AF P+N+ PGI  SPD+MLQGRLF+Y DTHR+RLG NY  IPVN P R    +
Sbjct  309  EVEQAAFSPANVVPGIGYSPDRMLQGRLFSYGDTHRYRLGVNYPQIPVNKP-RCPFHSSS  367

Query  388  RDGPMCMQDNQGGAPNYYPNSFGAPEQQPSA------LEHSIQYSGEVRRFNTANDDN--  439
            RDG M       G   NY P+S    ++  SA      L H I+   EV  ++   DD+
Sbjct  368  RDGYM-QNGYYGSLQNYTPSSLPGYKEDKSARDPKFNLAH-IEKEFEVWNWDYRADDSDY  425

Query  440  VTQVRAFYVNVLNEEQRKRLCENIA---GHLKDAQIFIQKKAVKNFTEVHPDYGSHIQAL  496
             TQ   +Y + L  ++++RL + I     H+   +I    K +++F +  P Y   ++
Sbjct  426  YTQPGDYYRS-LPADEKERLHDTIGESLAHVTHKEIV--DKQLEHFKKADPKYAEGVKKA  482

Query  497  LDKY  500
            L+K+
Sbjct  483  LEKH  486
```

Fig 1: result of blast search of specific sequences of *H.pylori* catalase (26695) and Human catalase.
Query: Human catalase enzyme sequence, Subject: *H.pylori* catalase enzyme sequence

Table1: Number of reactivating MHC class II for selected epitopes according to Propred software

Epitopes	Site of amino acid	Reactivating MHC
MVNKDVKQTT	**1-10**	**26**
VLLQSTWFL	**31-39**	**21**
FHPFDVTKI	**276-284**	**16**
WVKFHFHTMQ	**212-221**	**22**
VKHLTNEEA	**223-231**	**7**
YRADDSDYY	**418-426**	**10**
YYRSLPADEK	**432-441**	**11**