BIOINFORMATION Discovery at the interface of physical and biological sciences

open access

www.bioinformation.net Volume 10(8)

Hypothesis

Identification and validation of T-cell epitopes in outer membrane protein (OMP) of *Salmonella typhi*

Arifur Rahman Tanu, Mohammad Arif Ashraf, Md Faruk Hossain, Md Ismail & Hossain Uddin Shekhar*

Department of Biochemistry & Molecular Biology, University of Dhaka, Dhaka-1000, Bangladesh; Hossain Uddin Shekhar – Email: shekhardu@hotmail.com; Phone: 880-2-9661900/7654; *Corresponding author

Received July 15, 2014; Accepted August 04, 2014; Published August 30, 2014

Abstract:

This study aims to design epitope-based peptides for the utility of vaccine development by targeting outer membrane protein F (Omp F), because two available licensed vaccines, live oral Ty21a and injectable polysaccharide, are 50% to 80% protective with a higher rate of side effects. Conventional vaccines take longer time for development and have less differentiation power between vaccinated and infected cells. On the other hand, Peptide-based vaccines present few advantages over other vaccines, such as stability of peptide, ease to manufacture, better storage, avoidance of infectious agents during manufacture, and different molecules can be linked with peptides to enhance their immunogenicity. Omp F is highly conserved and facilitates attachment and fusion of *Salmonella typhi* with host cells. Using various databases and tools, immune parameters of conserved sequences from Omp F of different isolates of *Salmonella typhi* were tested to predict probable epitopes. Binding analysis of the peptides with MHC molecules, epitopes conservancy, population coverage, and linear B cell epitope prediction were analyzed. Among all those predicted peptides, ESYTDMAPY epitope interacted with six MHC alleles and it shows highest amount of interaction compared to others. The cumulative population coverage for these epitopes as vaccine candidates was approximately 70%. Structural analysis suggested that epitope ESYTDMAPY fitted well into the epitope-binding groove of HLA-C*12:03, as this HLA molecule was common which interact with each and every predicted epitopes. So, this potential epitope may be linked with other molecules to enhance its immunogenicity and used for vaccine development.

Background:

Typhoid fever is caused by *Salmonella typhi*, a Gram-negative bacterium, which is an obligatory human pathogen **[1]**. It has three outer membrane proteins- OmpC, OmpF and OmpA, those are good as vaccine candidate **[1]**. Among these three, OmpF, a major general diffusion porin of *Salmonella typhi*, has more conserved sequences and higher antigenicity than others **[1, 2]**. In spite of serious health complications, the scientific knowledge of the epidemiology and ecology of this virus is limited **[3, 4, 5]** to design an universal vaccine. Bioinformatics, especially immunoinformatics, is an emerging field in vaccine design. The combination experimental and *in silico* methods are crucial to solve complex problems such as revealing immune responses and vaccine design **[6]**. Available bioinformatics tools provide the searching option to scan for

probable epitope candidate from huge sets of protein antigens which are encoded by complete viral genomes. This computational vaccine design approach has proven very effective in fighting against few diseases such as multiple sclerosis [7], malaria [8], and tumors [9]. The most crucial job in this experiment is the identification of HLA ligands and T-cell epitopes [10]. Through T-cell epitope prediction tools to identify allele-specific binding peptides, it is possible to reduce the number of potential peptides considered as vaccine candidates. Along with these tools, different types of methods have been developed for the identification of proteosomal peptide cleavage sites, major histocompatibility complex (MHC) binding peptides and transporters associated with antigen presentation (TAP) molecules [6, 11-14]. T lymphocytes only recognize processed peptides or antigens and usually these peptides are presented by antigen presenting cells in association with HLA molecules. As a result, an epitope will only able to elicit immune response in human individual if he or she expressing this particular HLA and capable of binding with it efficiently [15]. In addition, over thousands of different HLA allelic variants have been identified so far [16]. And specific HLA alleles are prevalent in different ethnic group. Therefore, binding of predicted epitopes with different HLA results increased population coverage. It has been found that as much as 90% population coverage with different ethnic groups can be obtained by targeting different HLA molecules [17]. Due to the tedious and expensive nature of experimental screening procedures, computational compound screening has been pursued extensively in recent years [18], which make ligand-protein interaction more vivid in knowledge before any expensive wet lab trial.

Methodology:

Retrieving the protein sequences

The sequences of outer membrane protein F (OmpF) of *Salmonella typhi* were retrieved from the NCBI protein database (http://www.ncbi.nlm.nih.gov/protein/) from different isolates.

Multiple sequence alignment (MSA)

Retrieved sequences were used as a platform to generate multiple sequence alignments, which were the basis of finding the conserved regions of the protein sequences. The MEGA5 software package (www.megasoftware.net) was used to build the alignment.

T cell epitopes processing prediction

To evaluate the immunogenicity of the conserved peptide, different bio-computational tools were employed. In order to discover the immunogenicity of the conserved peptide, firstly a reverse immunogenic approach was employed for the selection of candidate epitopes. The T-cell epitopes from the conserved peptide were predicted using the NetCTL (http://tools.immuneepitope.org

/stools/netchop/netchop.do) prediction method, which is based on neural network architecture and predicts the epitope candidates based on the processing of the peptides *in vivo*. The threshold was set at 0.50 during analysis. A combined algorithm integrating MHC class I binding, transporter of antigenic peptides (TAP) transport efficiency and proteosomal cleavage prediction was involved to predict a total/overall score. Based on this overall score the best epitope candidates were selected for further analysis.

The Immuno Epitope Database (IEDB) MHC-1 binding prediction tool made the option of using a different prediction method available. Stabilized Matrix Method-based (SMM) prediction was used to calculate IC50 values for peptides binding to specific MHC molecules. Note that all the alleles were selected and length was set at 9.0 prior to the prediction. For the selected epitopes a web based tool 'Proteosomal cleavage/TAP transport/MHC class I combined predictor' was used to predict proteosomal scores, TAP score, processing score and MHC-1 binding score using the SMM based prediction method for each individual peptide. The parameters for immunogenicity detection (TAP score, proteosomal score,

IC50 values) were normalized on a scale of 0 to 1 and were given a weighted score to prioritize the parameters in order to determine immunogenicity. The binding energies for HLAepitope interactions were predicted using the MHC-1 peptide binding energy predictor tool.



Figure 1: Docking to predict the binding of predicted epitopes to MHC class I molecule, HLA-C*12:03. Binding of "ESYTDMAPY" to the binding grooves of the retrieved structure of HLA-C*12:03 (binding energy: -5.4Kcal/mol). The yellow colored portion and green portion in both figure (A & B) represent HLA-C*12:03 molecule and ESYTDMAPY epitope, respectively.

Epitope conservancy analysis

Epitope conservancy analysis was done with Epitope Conservancy Analysis of Immuno Epitope Database (http://tools.immuneepitope.org/tools/conservancy/iedb_inp ut) by comparing the conservancy of predicted epitope with the sequence of Omp F.

Retrieving 3D structure of HLA

The 3D structure of HLA-C*12:03 (PDB ID: 1EFX) was downloaded from the Protein Data Bank Database (www.rcsb. org/) and visualized in the PYMOL molecular graphics system. Prior to docking, all the water molecules were removed from the 3D structure of epitope free HLA-C*12:03.

Designing the 3D epitope structure

For the docking study, the ESYTDMAPY epitope was chosen because it showed the maximum interactions with different HLAs. The 3D structure of ESYTDMAPY was designed using the PEP-FOLD server: De novo peptide structure prediction (http://bioserv.rpbs.univ-paris-diderot.fr/PEP-FOLD/).

HLA-epitope binding prediction

The AutoDock tool from the MGL software package (version 1.5.6) was employed for docking purpose. Both the protein [HLA-C*12:03] and ligand (epitope ESYTDMAPY) files were firstly converted to PDBQT format to use them for the docking study. The grid space for center was set manually to facilitate the binding of epitope at the binding groove of HLA-C*12:03.

Population coverage

Population coverage for individual epitope was predicted by the population coverage tool (http://tools.immuneepitope.org / tools/population) from the IEDB analysis resource. The allele frequency of the interacting alleles was used to measure the population coverage for the corresponding epitope.



Figure 2: Population coverage by MHC Class I + II restricted epitopes from outer membrane protein F (OmpF) of Salmonella typhi. In case of epitopes from Omp F, the maximum coverage was found for the population of Europe, South Africa and East Asia.

B-cell epitope prediction

Linear B-cell epitopes were predicted from the given protein sequence using the B-cell epitope prediction tool from the IEDB analysis resource (http://tools.immuneepitope.org/ tools/bcell/iedb_input). The Kolaskar and Tongaonkar antigenicity prediction method was employed for prediction in which, predictions are based on a table that reflects the occurrence of amino acid residues in experimentally known segmental epitopes. This method can predict antigenic peptides with approximately 75% accuracy.



Figure 3: Linear B-cell epitope prediction using Kolaskar and Tongaonkar prediction tool of IEDB. The pick shows above the threshold corresponding to those sequence position which can act as B-cell epitope to elicit humoral immune response.

Results:

Identification of evolutionary conserved proteins/regions From the output of multiple sequence alignment for Omp F protein, it has been found that the sequence is conserved over all of the sequences. But, it's not true for Omp A and Omp C. That's why we exclude those proteins from out study.

Immunogenicity of conserved peptides

The NetCTL prediction tool predicted 355 overlapping potential epitopes from the given sequence, but only 16 potential T-cell epitope candidates were chosen on the basis of a total score (above 0.5) that was combined result of different parameters Table 1 (see supplementary material). SMM based IEDB MHC-1 binding prediction tool retrieved 784 possible MHC-1 allele interactions with the 16 different T-cell epitope those were selected before. The MHC-1 alleles for which the epitopes showed higher affinity (IC50<100) were selected for further analysis Table 2 (see supplementary material). Among the 16 T-cell epitopes, one 9-mer epitope (ESYTDMAPY) was found those interact with six HLA molecules. Here, ESYTDMAPY interacts with HLA-C*03:03, HLA-C*12:03, HLA-C*14:02, HLA-A*25:01, HLA-A*26:01, HLA-B*35:01. No other T-cell epitope interacted with as many MHC-1 alleles with as great an affinity as this epitope Table 3 (see supplementary material).

Epitope conservancy analysis

Epitope conservancy analysis with all of the Omp F protein sequences from different strains revealed that all 14 out of 16 peptides have a 100% protein sequence match. The results are summarized in **Table 4 (see supplementary material)**.

Molecular docking study of HLA-epitope interaction

AutoDock Vina predicted nine possible binding models. On the basis of higher binding energy with HLA-C*12:03, the best output model for ESYTDMAPY epitope was found to have a

ISSN 0973-2063 (online) 0973-8894 (print) Bioinformation 10(8): 480-486 (2014)

binding energy of -5.4 kcal/mol. The interacting and binding of HLA-epitope is illustrated in **Figure 1**.

Population coverage

The population coverage for different population is summarized in **Table 5 (see supplementary material).** And the graph shows that Europe, South Africa and East Asia have highest population coverage and it's over 60% (**Figure 2**).

Prediction of B-cell epitopes

The antigenic determinant plot was presented in **Figure 3**; the x-axis shows sequence position and the y-axis shows antigenic propensity. Average antigenic propensity of this protein is 1.003. There are 13 antigenic determinants in the sequence and shown in **Table 6 (see supplementary material).** The average for the whole protein is above 1.0; all residues above 1.0 are potentially antigenic.

Conclusion:

In vitro and *in vivo* studies are required to determining the actual effectiveness of the peptide for mounting an immune response. Binding chip assay for HLA and peptide would also be useful to determine the binding affinity of the peptide as a whole. The best predicted T cell epitope (ESYTDMAPY) is nonamer and it covers 126 to 134 positions of amino acids. At the same time, there is a potential B-cell (YGIVYDVESY) epitope which covers 119 to 128 positions of amino acids. If we take peptide from 119 to 134 amino acid positions and add potential molecules to enhance immunogenicity, then it would be possible to design a universal epitope based vaccine against *Salmonella typhi*.

Acknowledgement:

The authors express sincerest thanks to University Grants Commission, Agargaon, Dhaka-1207, Bangladesh to support this project during the fiscal year 2014-15.

References:

- [1] Toobak H et al. Biologicals. 2013 41: 224 [PMID: 23796754]
- [2] Balasubramaniam D et al. J Struct Biol. 2012 178: 233
 [PMID: 22525817]
- [3] Field H et al. Microbes Infect. 2001 3: 307 [PMID: 11334748]
- [4] Field H et al. Arch Virol Suppl. 2004 18: 113 [PMID: 15119766]
- [5] Plowright RK et al. Proc Biol Sci. 2008 275: 861 [PMID: 18198149]
- [6] Korber B et al. PLoS Comput Biol. 2006 2: e71 [PMID: 16846250]
- [7] Bourdette DN et al. Mult Scler. 2005 11: 552 [PMID: 16193893]
- [8] Lopez JA et al. Eur J Immunol. 2001 31: 1989 [PMID: 11449351]
- [9] Knutson KL et al. J Clin Invest. 2001 107: 477 [PMID: 11181647]
- [10] Petrovsky N & Brusic V, Immunol Cell Biol. 2002 80: 248 [PMID: 12067412]
- [11] Brusic V et al. Methods. 2004 34: 436 [PMID: 15542369]
- [12] Peters B et al. J Immunol. 2003 171: 1741 [PMID: 12902473]
- [13] Bhasin M & Raghava GP, Protein Sci. 2004 13: 596 [PMID: 14978300]
- [14] Nielsen M et al. Immunogenetics. 2005 57: 33 [PMID: 15744535]
- [15] Patronov A & Doytchinova I, Open Biol. 2013 3: 120139 [PMID: 23303307]
- [16] Ovsyannikova IG et al. Expert Rev Vaccines. 2006 5: 33 [PMID: 16451106]
- [17] Longmate J et al. Immunogenetics. 2001 52: 165 [PMID: 11220618]
- [18] Vennila S *et al. Bioinformation.* 2014 10: 263 [PMID: 24966532]

Edited by P Kangueane

Citation: Tanu et al. Bioinformation 10(8): 480-486 (2014)

License statement: This is an open-access article, which permits unrestricted use, distribution, and reproduction in any medium, for non-commercial purposes, provided the original author and source are credited

Supplementary material:

Table 1: Selected epitopes on the basis of total score (above 0.5) predicted by NetCTL.

Number	Peptides	Overall score		
1	DSKNADQTY	2.8047		
2	ATANAAEIY	2.8045		
3	YAEVGSIDY	2.6898		
4	LTGFGQWEY	2.5606		
5	YTDNYMTSR	1.6814		
6	NSDFFGLVD	1.3308		
7	YTDMAPYFS	1.3069		
8	YIQAGATYY	1.2632		
9	ESWAVGAKY	1.2326		
10	GSIDYGRNY	1.18		
11	SGETWGGAY	1.1778		
12	NTDLTGFGQ	1.1334		
13	ESYTDMAPY	1.1054		
14	RLAFAGLKY	1.9252		
15	WVDYRFNLL	0.9241		
16	NSNLVRLAF	0.8646		

Table 2: Allele selection for different peptide sequences. The proteasome complex contains enzymes that cleave peptide bonds, converting proteins into peptides. The antigenic peptides from proteasome cleavage associate with class I MHC molecules, and the peptide-MHC complexes are then transported to the cell membrane. TAP transports the peptide to the endoplasmic reticulum.

Allele	Peptide	Proteasome score	Tap	MHCscore	Processing	Total	MHC
			score		score	score	ic50
HLA-C*12:03	DSKNADQTY	1.47	1.21	-0.64	2.68	2.05	4.34
HLA-C*03:03	ESYTDMAPY	0.98	1.28	-0.7	2.26	1.56	5
HLA-C*12:03	NSDFFGLVD	1.14	-0.8	-0.75	0.35	-0.41	5.65
HLA-C*05:01	NSDFFGLVD	1.14	-0.8	-0.79	0.35	-0.45	6.19
HLA-C*03:03	YAEVGSIDY	1.15	1.19	-0.81	2.34	1.53	6.44
HLA-C*12:03	YAEVGSIDY	1.15	1.19	-0.9	2.34	1.44	7.98
HLA-C*05:01	NTDLTGFGQ	0.88	-0.07	-0.9	0.81	-0.1	7.99
HLA-C*12:03	YTDNYMTSR	1.03	0.54	-0.91	1.57	0.66	8.06
HLA-C*12:03	NTDLTGFGQ	0.88	-0.07	-0.91	0.81	-0.11	8.19
HLA-C*12:03	YTDMAPYFS	1.05	-1.08	-0.94	-0.04	-0.97	8.61
HLA-C*05:01	WVDYRFNLL	1.42	0.33	-1	1.75	0.76	9.9
HLA-C*05:01	YTDMAPYFS	1.05	-1.08	-1.18	-0.04	-1.22	15.16
HLA-C*12:03	YIQAGATYY	1.23	1.25	-1.24	2.48	1.24	17.19
HLA-C*12:03	ESYTDMAPY	0.98	1.28	-1.26	2.26	1	18.37
HLA-C*03:03	YIQAGATYY	1.23	1.25	-1.31	2.48	1.16	20.64
HLA-C*12:03	LTGFGQWEY	1.31	1.21	-1.32	2.52	1.21	20.71
HLA-C*12:03	SGETWGGAY	1.27	1.17	-1.32	2.45	1.12	20.95
HLA-C*15:02	NSDFFGLVD	1.14	-0.8	-1.35	0.35	-1	22.15
HLA-C*12:03	GSIDYGRNY	1.45	1.29	-1.38	2.74	1.36	24
HLA-A*29:02	YIQAGATYY	1.23	1.25	-1.4	2.48	1.08	24.99
HLA-A*29:02	RLAFAGLKY	1.35	1.32	-1.46	2.67	1.21	28.89
HLA-C*14:02	YTDNYMTSR	1.03	0.54	-1.5	1.57	0.07	31.52
HLA-C*12:03	ESWAVGAKY	1.28	1.28	-1.6	2.56	0.95	40.11
HLA-C*05:01	YTDNYMTSR	1.03	0.54	-1.61	1.57	-0.04	41.08
HLA-C*05:01	YAEVGSIDY	1.15	1.19	-1.62	2.34	0.72	41.66
HLA-C*03:03	ATANAAEIY	1.42	1.34	-1.62	2.76	1.13	42.03
HLA-B*15:02	YIQAGATYY	1.23	1.25	-1.69	2.48	0.79	49.34
HLA-C*14:02	ESYTDMAPY	0.98	1.28	-1.71	2.26	0.55	51.36
HLA-A*25:01	ESYTDMAPY	0.98	1.28	-1.72	2.26	0.54	52.25
HLA-C*05:01	NSNLVRLAF	1.38	1.13	-1.72	2.51	0.78	52.8
HLA-C*12:03	ATANAAEIY	1.42	1.34	-1.72	2.76	1.04	52.87
HLA-C*14:02	YAEVGSIDY	1.15	1.19	-1.73	2.34	0.61	53.66
HLA-C*03:03	GSIDYGRNY	1.45	1.29	-1.73	2.74	1.01	53.78
HLA-B*15:02	SGETWGGAY	1.27	1.17	-1.73	2.45	0.71	53.86
HLA-C*14:02	YIQAGATYY	1.23	1.25	-1.74	2.48	0.74	55.03
ISSN 0973-2063 (online) 0973-8894 (print)							
Bioinformation 10	(8): 480-486 (2014)			484			

HLA-C*03:03	LTGFGQWEY	1.31	1.21	-1.78	2.52	0.75	59.65
HLA-A*29:02	LTGFGQWEY	1.31	1.21	-1.78	2.52	0.74	60.64
HLA-C*12:03	RLAFAGLKY	1.35	1.32	-1.8	2.67	0.88	62.84
HLA-B*15:02	DSKNADQTY	1.47	1.21	-1.82	2.68	0.87	65.5
HLA-A*68:01	YTDNYMTSR	1.03	0.54	-1.85	1.57	-0.28	71.37
HLA-A*26:01	ESYTDMAPY	0.98	1.28	-1.87	2.26	0.39	73.72
HLA-C*07:01	NSDFFGLVD	1.14	-0.8	-1.87	0.35	-1.52	73.82
HLA-A*02:06	WVDYRFNLL	1.42	0.33	-1.87	1.75	-0.12	74.91
HLA-B*35:01	ESYTDMAPY	0.98	1.28	-1.88	2.26	0.38	76.66
HLA-B*15:02	NSNLVRLAF	1.38	1.13	-1.93	2.51	0.58	84.38
HLA-C*07:01	YAEVGSIDY	1.15	1.19	-1.96	2.34	0.38	90.82
HLA-C*12:03	NSNLVRLAF	1.38	1.13	-1.96	2.51	0.54	92.09
HLA-C*14:02	SGETWGGAY	1.27	1.17	-1.96	2.45	0.48	92.18
HLA-C*03:03	SGETWGGAY	1.27	1.17	-1.98	2.45	0.46	96.3

Table 3: Selected epitopes and the list of HLA molecules with those they interact. The highlighted epitope interacts with six HLA molecules and no other epitope interact with so many HLA molecules.

Number	Epitope	HLA to interact with
1	ATANAAEIY	HLA-C*03:03, HLA-C*12:03
2	DSKNADQTY	HLA-C*12:03, HLA-B*15:02
3	ESWAVGAKY	HLA-C*12:03
4	ESYTDMAPY	HLA-C*03:03, HLA-C*12:03, HLA-C*14:02, HLA-A*25:01, HLA-A*26:01, HLA-B*35:01
5	GSIDYGRNY	HLA-C*03:03, HLA-C*12:03
6	LTGFGQWEY	HLA-C*03:03, HLA-C*12:03, HLA-A*29:02
7	NSDFFGLVD	HLA-C*12:03, HLA-C*05:01, HLA-C*15:02, HLA-C*07:01
8	NSNLVRLAF	HLA-C*05:01, HLA-C*15:02, HLA-C*12:03
9	NTDLTGFGQ	HLA-C*05:01, HLA-C*12:03
10	RLAFAGLKY	HLA-A*29:02, HLA-C*12:03
11	SGETWGGAY	HLA-C*12:03, HLA-B*15:02, HLA-C*14:02, HLA-C*03:03
12	WVDYRFNLL	HLA-C*05:01, HLA-C*12:03, HLA-A*02:06
13	YAEVGSIDY	HLA-C*03:03, HLA-C*12:03, HLA-C*05:01, HLA-C*14:02, HLA-C*07:01
14	YIQAGATYY	HLA-C*03:03, HLA-C*12:03, HLA-C*14:02, HLA-A*29:02, HLA-B*15:02
15	YTDMAPYFS	HLA-C*12:03, HLA-C*05:01
16	YTDNYMTSR	HLA-C*12:03, HLA-C*05:01, HLA-C*14:02, HLA-A*68:01

Table 4: Epitope conservancy analysis through IEDB tools.

Epitope No.	Epitope Sequence	Epitope Length	Minimum Identity	Maximum identity
1	DSKNADQTY	9	100.00%	100.00%
2	ATANAAEIY	9	100.00%	100.00%
3	YAEVGSIDY	9	100.00%	100.00%
4	LTGFGQWEY	9	100.00%	100.00%
5	YTDNYMTSR	9	100.00%	100.00%
6	NSDFFGLVD	9	100.00%	100.00%
7	YTDMAPYFS	9	100.00%	100.00%
8	YIQAGATYY	9	100.00%	100.00%
9	ESWAVGAKY	9	100.00%	100.00%
10	GSIDYGRNY	9	100.00%	100.00%
11	SGETWGGAY	9	22.22%	33.33%
12	NTDLTGFGQ	9	22.22%	33.33%
13	ESYTDMAPY	9	100.00%	100.00%
14	RLAFAGLKY	9	100.00%	100.00%
15	WVDYRFNLL	9	100.00%	100.00%
16	NSNLVRLAF	9	100.00%	100.00%

Table 5: Population coverage for best predicted epitope is summarized here.

Population/Area Class I+II Coverage

1 /	
Europe	70.39%
South Africa	67.84%
East Asia	66.82%
North America	59.60%

North Africa	59.49%
South Asia	59.14%
West Africa	55.76%
Central Africa	50.65%
Southwest Asia	50.48%
South America	46.86%
Oceania	46.77%
Northeast Asia	46.59%
Southeast Asia	42.73%
East Africa	42.37%
West Indies	31.86%

Table 6: Potential linear peptides predicted to be antigenic determinants

Start position	End position	Peptide	Peptide length
5	20	KILAAVIPALLAAATA	16
33	43	LDLYGKAVGRH	11
95	113	SNLVRLAFAGLKYAEVGSI	19
119	128	YGIVYDVESY	10
154	159	GGLLTY	6
165	177	FGLVDGLSFGIQY	13
205	211	GFGVTAA	7
233	239	SWAVGAK	7
243	253	NNVYLAAVYAE	11
276	300	NLEVVAQYQFDFGLRPAISYVQSKG	25
311	323	DLAKYIQAGATYY	13
331	337	WVDYRFN	7
345	351	SSSYVGT	7