RESEARCH ARTICLE

In silico Design of Discontinuous Peptides Representative of B and T-cell Epitopes from HER2-ECD as Potential Novel Cancer Peptide Vaccines

Mahdavi Manijeh[1], Keyhanfar Mehrnaz[2], Moreau Violaine[3], Mohabatkar Hassan[2], Jafarian Abbas[1]*, Rabbani Mohammad[1]

Abstract

At present, the most common cause of cancer-related death in women is breast cancer. In a large proportion of breast cancers, there is the overexpression of human epidermal growth factor receptor 2 (HER2). This receptor is a 185 KDa growth factor glycoprotein, also known as the first tumor-associated antigen for different types of breast cancers. Moreover, HER2 is an appropriate cell-surface specific antigen for passive immunotherapy, which relies on the repeated application of monoclonal antibodies that are transferred to the patient. However, vaccination is preferable because it would stimulate a patient's own immune system to actively respond to a disease. In the current study, several bioinformatics tools were used for designing synthetic peptide vaccines. PEPOP was used to predict peptides from HER2 ECD subdomain III in the form of discontinuous-continuous B-cell epitopes. Then, T-cell epitope prediction web servers MHCPred, SYFPEITHI, HLA peptide motif search, Propred, and SVMHC were used to identify class-I and II MHC peptides. In this way, PEPOP selected 12 discontinuous peptides from the 3D structure of the HER2 ECD subdomain III. Furthermore, T-cell epitope prediction analyses identified four peptides containing the segments 77 (384-391) and 99 (495-503) for both B and T-cell epitopes. This work is the only study to our knowledge focusing on design of *in silico* potential novel cancer peptide vaccines of the HER2 ECD subdomain III that contain epitopes for both B and T-cells. These findings based on bioinformatics analyses may be used in vaccine design and cancer therapy; saving time and minimizing the number of tests needed to select the best possible epitopes.

Keywords: HER2 receptor - discontinuous B cell epitope - T-cell epitope - bioinformatics - peptide vaccine

Asian Pac J Cancer Prev, **14 (10)**, 5973-5981

Introduction

Worldwide, breast cancer represents 22.9% of all cancers in women and 16% of all female cancers. In 2008 breast cancer caused 13.7% of cancer deaths in women and the number of cases worldwide has significantly increased since the 1970s. There are two different types of breast cancers. In the most common breast cancers, infiltrating ductal carcinoma, Human Epidermal growth factor Receptor 2 (HER2) is overexpressed due to multiple copies of the HER2 gene instead of two copies in the normal cells (Jacot et al., 2013) Its overexpression usually results in malignant transformation of cells and aberrant cell growth accounting for 25% of all breast cancer cases (Tai et al., 2010).

HER2 is a 185 KDa growth factor receptor transmembrane glycoprotein and is described as the first tumor-associated antigen (TAA) for breast cancers (Wallecha et al., 2012). In addition, HER2 has key roles in all processes of cell development, regulating cell growth, differentiation, and survival. The overexpression of its genes is also related to many other cancers such as ovarian, prostate, lung, and gastric cancer (Siyi et al., 2008). Structurally, HER2 is comprised of three domains: N-terminal extracellular domain (ECD), a single transmembrane helix (TM), and an intracellular tyrosine kinase domain, which its largest part (ECD) divided into four subdomains (I-IV) (Lax, 1998). Moreover, HER2 is an ideal antigen for passive immunotherapy, in which effector cells or monoclonal antibodies (mAbs) are transferred to the patient (Disis et al., 2001). Currently, in treatments for some types of breast cancers, humanized mAbs, Pertuzumab, and Herceptin, against subdomain II and IV of HER2 ECD are used (Cho et al., 2003; Franklin et al., 2004; Awada et al., 2012). Although the mAbs have shown promising results in clinical settings, several concerns such as repeated treatments and related costs, limited duration of effective therapy, undesired immunogenicity, and possible tolerance are remained (Garrett et al., 2007; Calabrich et al., 2008).

*[1]Department of Pharmaceutical Biotechnology, School of Pharmacy and Pharmaceutical Sciences, Isfahan University of Medical Sciences, [2]Department of Biotechnology, Faculty of Advanced Sciences and Technologies, University of Isfahan, Isfahan, Iran, [3]UMR 3145 SysDiag CNRS/Bio-Rad, Parc Euromédecine, Montpellier, France *For correspondence: Jafarian@pharm.mui.ac.ir*

To overcome the disadvantages of passive immunotherapy, active immunization or vaccination is used that involves a specific TAA eliciting an endogenous immune or antitumor response. For instance, HER2 is considered as TAA for vaccine development and primary prevention of cancers. Strategies for active immunization include the use of whole cell vaccines, protein, and DNA vaccines, as well as peptide vaccines (Wang et al., 2010). However, such active immunization, resulting in an autoimmune response against HER2, may produce adverse affects that are difficult to reverse, and they might be more severe and prolonged compared with passive immunotherapy. As a result, immunization needs a vaccine design that targets a portion of the protein instead of its whole domains for circumventing tolerance to self tumor antigens (Garrett et al., 2007). Therefore, vaccination with a synthetic peptide instead of using whole protein increases the probability of generating an adaptive immune response and eliciting protein reactive high affinity mAbs (Miyako et al., 2011). Briefly, peptides are attractive anticancer vaccines that they are safe, free of pathogens and oncogenic materials, easily synthesized and result in sustained immune responses and memory. There are advantages using subunit peptide-based immunogens to develop vaccines but it has some limitations such as the fact that unmodified peptides are rarely immunogenic (Dakappagari et al., 2000; Arensa et al., 2013).

In the design of a synthetic peptide vaccine, it is necessary that a B-cell epitope be presented along an appropriate T-cell epitope, because the interaction between B and T-cells recognizing the same antigen is essential for B-cell proliferation and differentiation and the generation of high-affinity antibodies of the IgG isotype (Garrett et al., 2007). Therefore, in several studies, it has been shown that synthetic peptide vaccines containing a targeted B-cell epitope and a T-cell epitope generated substantial immune responses (Axelsena et al., 2011). In other words, recognition of antigenic epitopes either small T-cell epitopes or large conformational epitopes by B-cells is the key event of immune responses (Doytchinova and Flower, 2002; Singh et al., 2011). Moreover, up to 90% of B-cell epitopes are constituted by discontinuous epitopes on the surface of proteins. To this end, in the present study, we decided to identify discontinuous epitopes and T-cell epitopes of HER2 ECD by *in silico* methods and select peptides that have both B and T-cell epitopes.

Epitope identification using bioinformatics tools saves time and it is cost and labor effective. Recently, immunoinformatics, a developed branch of bioinformatics tools, is used for selecting epitopes from immunological target proteins (Bian et al., 2003; Li et al., 2005). In this work, PEPOP (Moreau et al., 2008) for prediction of peptides representative of discontinuous epitopes (in short "discontinuous peptides") and SYFPEITHI (Rammensee et al., 1999), MHCPred (Guan et al., 2003), Propred (Singh and Raghava, 2001), SVMHC (Dönnes and Elofsson, 2002), and HLA peptide motif search (Parker et al., 1994) for prediction of T-cell epitope peptides were used.

Since the anti-HER2 antibodies targeting distinct epitopes have different biological functions on cancer cells, it is needed to explore further new epitopes of HER2 ECD and use them as peptide vaccines. As subdomain III of HER2 ECD plays an important role in the formation of ligand binding site of the receptor and also there is no mAb and peptide vaccine against it, we attempted to design peptide vaccines from subdomain III of HER2 ECD by bioinformatics methods.

Materials and Methods

Prediction of "discontinuous" B-cell epitope peptides

3D structure selection: two PDB files (1N8Z, 1S78) describe the HER2 ECD. The chains in PDB files and the resolution of their structures were identified. Furthermore, sequence alignments were carried on extracted sequences from PDB files by ClustalW (Thompson et al., 1994) to choose the one for using in PEPOP.

Peptide design methods: PEPOP is a bioinformatics tool for prediction of antigenic or immunogenic peptides using 3D structure of a protein (Moreau et al., 2008). PEPOP identifies segments comprised of accessible and sequence continuous amino acids. These segments are clustered according to their spatial distances and the peptides are designed by finding a path among a set of close segments. The set of close segments are segments included in either a 10Å radius patch or a cluster defined by PEPOP. The path among the segments is found by a method of extension according to different criteria. In our study, the most appropriate methods of extension were selected among the 34 methods of PEPOP predicting "discontinuous" peptides. Optimized Nearest Neighbor (ONN), Optimized Flanking Nearest Neighbor (OFN), Optimized Patched segments Path (OPP), Traveling Salesman Problem (TSP), and Shortest Path (SHP) methods were thought to be the most appropriate, because these methods optimize the path among the segments composing peptides. In ONN method, from a reference segment, the nearest neighbor segment from the C-terminal of the reference segment was collected until a defined peptide length was achieved. Also, in OFN method, from a reference segment, the nearest neighbor segment from the C-terminal and N-terminal of the reference segment was collected until a defined peptide length was achieved. In addition, in OPP method, the collected segments were inside a 10Å radius sphere from a reference segment. For each of these three methods, ONN, ONF, and OPP, an optimized path was searched among the collected segments. Furthermore, in SHP-based method, from a set (either in a cluster or inside a 10Å radius sphere) of segments, a peptide was the shortest path between two segments, which included more amino acids. In TSP-based method, from a set (either in a cluster or inside a 10Å radius sphere) of segments, the algorithm for the traveling salesman problem was used to find the optimal path (the shortest distance) among all of the segments.

Peptide prediction and selection: in all clusters, each segment was considered as a reference segment. To design a peptide, PEPOP started from a reference segment and extended its sequence with the sequence of neighbor segments to yield a peptide with suitable amino acid length by a method of extension in a specified "area of extension". Thus, each peptide sequence has been generally comprised

of several segments. In this way, all the possible peptides were predicted. Then, for reducing the great number of produced peptides (915 peptides), a process was applied to select a small number of peptides progressively. At first, the peptides of less than 12 amino acids and more than 20 amino acids were removed. Also, peptides comprised of only one segment (S38 and S53) were removed because they targeted continuous epitopes. In this step, 701 peptides were predicted. In the next step, the predicted peptides from a reference segment including targeted region of the subdomain III from residues 318-508 (i.e. segments 67-103) and those having at least three amino acids length were kept. Besides, all the peptides including segments from 1-317 and from 509-577, were removed. Then, duplicated sequences or sequences including the shortest ones were removed. In addition, among similar peptides only one of them was kept. Finally, 12 peptides remained.

Prediction of T-cell epitope peptides

The sequence of the targeted region of HER2 ECD was submitted to the web servers MHCpred, SVMHC, SYFPEITHI, Propred, and HLA peptide motif search. Next, the desired alleles of MHC molecules were selected and the program was run. The results page produced subsequently a sorted list of nine amino acid substrings of the submitted antigen sequence to calculate affinities. Default settings were applied to all the tools used. These tools are matrix based, which rank nine and ten amino acid long segments of a protein that overlaps by eight and nine amino acids, respectively, based on the estimated probability of binding to a selected MHC molecule. In this work, MHC class-I and class-II binding peptides of HER2 ECD were identified by inhibitory concentration (IC_{50}) value based on the quantitative prediction method, MHCpred. Also, online T-cell epitope prediction tools SYFPEITHI (Rammensee et al., 1999), Propred (Singh and Raghava, 2001), HLA peptide motif search (Parker et al., 1994), and SVMHC (Dönnes and Elofsson, 2002), which employ binding status scoring qualitative prediction methods were used.

3D structure predictions of peptides

3D structure of the B-cell epitope peptides were predicted by de novo peptide structure prediction approach as built at the PEP-FOLD server (Maupetit et al., 2009). Visualization of all the models was performed in Pymol V1. Furthermore, selection of the best model for each peptide was carried out by GROMOS96 force field application in SPDBV that calculated to minimize the energy of the models. Besides, turn extraction of protein structure models was performed by extractTurn server at RPBS Mobyle Portal.

Theoretical physicochemical properties of the peptides

The theoretical physicochemical properties of the synthetic peptides such as the ionic status, calculated as the isoelectric point, and the hydrophobicity, measured as the grand average of hydropathicity (GRAVY) index were analysed using the Prot Param algorithm. The GRAVY index indicated the hydrophobicity of the peptide and was calculated as the sum of the hydropathy values (Kyte and Doolittle parameters) of the composing amino acids, divided by the number of residues in the sequence. Peptides with positive GRAVY index are hydrophilic whereas peptides with negative GRAVY index are hydrophobic (Lebreton et al., 2011).

Results

Prediction of discontinuous B-cell epitope peptides

Sequence alignments of 1S78 and 1N8Z indicated that a great part of chain C of 1N8Z and a great part of chain B of 1S78 were in common. Finally, chain B of 1S78, that was a little longer compared to the chain A was chosen because 1N8Z contained a short region in the targeted region (340-530). It should be noted that the numbering in PDB file has shifted: the fragment 340-530 from HER2 ECD subdomain III in the sequence corresponds to 318-508 in the PDB file.

From 3D structure of HER2 ECD, PEPOP identified 116 segments gathered in three clusters according to their spatial distances. PEPOP methods were run on the whole protein, systematically, with the requested length of 16 amino acids. Most of the predicted peptides in ONN, ONF, and OPP methods were in the form of segments but in TSP and SHP methods were in the form of segments and clusters. In this study, we wanted to predict peptides from a specific region of the protein, the region 340-530 on HER2 ECD subdomain III. It corresponds to the segments S67 to S103. Table 1 listed these segments and their position in the sequence of HER2 ECD. Predicted

Table 1. Segments of HER2 ECD Subdomain Iii and their Position in the Sequence

Segment number	Length	Residue	Position
S67	9	SKPCARVCY	325-343
S68	4	GMEH	346-349
S69	4	REVR	351-354
S70	2	SA	358-359
S71	2	QE	362-363
S72	2	AG	365-366
S73	2	KK	368-369
S74	1	F	371
S75	2	PE	378-379
S76	2	FD	381-382
S77	8	DPASNTAP	384-391
S78	4	QPEQ	393-396
S79	1	Q	398
S80	2	ET	401-402
S81	2	EE	404-405
S82	1	Y	409
S83	1	Y	411
S84	4	DSLP	417-420
S85	1	S	423
S87	1	Q	429
S88	2	QG	435-438
S89	2	SW	446-447
S90	1	Q	451-452
S93	1	H	470
S94	2	TH	472-473
S99	9	HTANRPEDE	495-503
S100	9	VGEGLACHQ	505-513

Table 2. Peptides with Rather Long Segments (>7 aa) Predicted by PEPOP

ID	Sequence	Length	Method	Peptide composition	Mw	pI	Gravy score
626	DPASNTAPSAFDPE	14	TSPnat1	S77, S70, S76, S75	1418.4	3.49	-0.921
1402	FDPEDPASNTAPQPEQQ	17	SHPnat	S76, S75, S77, S78, S79	1870.9	3.43	-1.735
123	HTANRPEDEVGEGLACHQ	18	ONF, ONN, SHPnat, TSPnat1,2,3,4	S99, S100	1963	4.8	-1.161
729	QGHTHSWHTANRPEDE	16	TSPnat4	S89, S93, S94, S90, S99	1901.9	5.75	-2.181
106	SKPCARVCYGMEHREVR	17	ONF	S67, S68, S69	2021.3	8.88	-0.859

Table 3. Discontinuous Peptides Predicted by PEPOP

ID	Sequence	Length	Method	Peptide composition	Mw	pI	Gravy score
591	ILHNYYGMEHREVRPE	16	TSPnat2, 4	S88, S83, S82, S68, S69, S75	2043.2	6.02	-1.225
656	QGTHSWDSLPFDQPEQ	16	TSPnat1, 2, 4	S89, S94, S90, S84, S76 ,S78	1871.9	4.02	-1.525
1412	QPEQQETKKEEQ	12	SHPnat	S78, S79, S80, S73, S81, S87	1501.5	4.49	-3.175

Table 4. Predicted Peptides by PEPOP Which the Path between the Segments on the Protein Seems to be Less Natural

ID	Sequence	Length	Method	Peptide composition	Mw	pI	Gravy score
632	DSLPPEFDQPEQSAQE	16	TSPnat1	S84, S75, S76, S78, S70, S71	1816.8	3.39	-1.625
658	FDQPEQQSQGTHSWDSLP	18	TSPnat3	S76, S78, S79, S85, S89, S94, S90, S84	2087.1	4.02	-1.594
590	GMEHREVRPESAQEFILHN	19	TSPnat1	S68, S69, S75, S70, S71, S74(371), S88	2279.5	5.4	-1.063
633	PEFDQPEQSAQEAGET	16	TSPnat2, 3, 4	S75, S76, S78, S70, S71, S72, S80	1762.7	3.45	-1.669

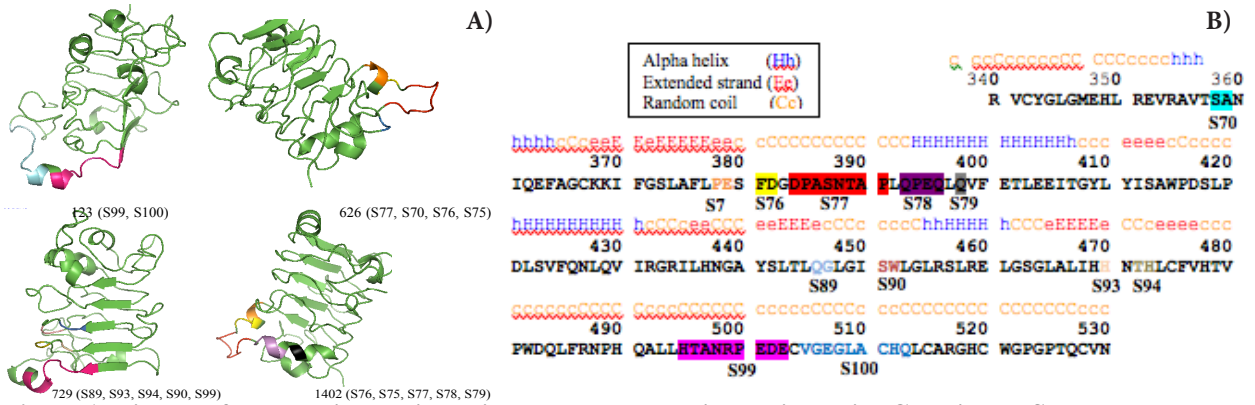

Figure 1. Display of the Predicted Discontinuous Immunogenic Peptides with Constituent Segments. A) the 3D structure of HER2 ECD subdomain III: peptides of table 1 with ID Number: 123, 1402, 729 and 626; **B)** the sequence of HER2 ECD subdomain III: the corresponding segments in relation with the sequence and secondary structure are displayed in colored residues

peptides, which did not relate to this region (340-530) were removed and at last 12 peptides were selected. By inspecting the peptides on 3D structure of the protein, three tables were made, from the most relevant to the least (Table 2-4). Figure 1 displays how the segments of predicted discontinuous-continuous peptides are located on 3D structure of HER2 ECD and how they are related to the primary and secondary structure of the protein. In Table 2 the peptides having a rather long segment (>7 amino acid) are listed. It is supposed that these peptides should more easily lead to the production of antibodies. Peptides were shown in Table 3 are more "discontinuous" than the other predicted peptides. In Table 4 peptides that the path among their segments on the protein seems to be less natural are listed.

Prediction of T-cell epitope peptides

Five publicly available tools trained on different data sets were employed for determination of MHC class I and II binders of HER2 ECD subdomain III. The peptides with the best predicted binding affinities (IC_{50} values>10(nM)) and frequently recognized by three and four tools used in this work are listed in Table 5. Among all predicted alleles, only the peptide of HLA-A*0201 allele in MHC class I molecule was predicted by SYFPEITHI, HLA peptide motif search, MHCpred, and SVMHC. Also the peptides of DRB1_0401 and DRB5_0101 alleles in MHC class II molecules were identified by Propred, SVMHC, and MHCpred. As shown in Table 5, P2 and P3 in MHC Class I alleles were the same as P5 and P10 in MHC class II alleles.

Linking B- and T-cell peptide epitope predictions

Comparing the results of B and T-cell epitopes indicated that segments S77: 384-DPASNTAP-391 and S99: 495-HTANRPEDE-503 were epitopes for both B and T-cells. S77 (384-391) was anchoring peptide motif in H2 Db, HLA B702, and HLA B5101 alleles of MHC class-I molecules and DRB 0101 allele of MHC class-II molecule. Furthermore, S99 was identified as H2 Db and HLA-A*0201 alleles of MHC class I molecules and DRB1_1304 allele of MHC Class II molecule by T-cell epitope prediction tools. From these findings, it was concluded that the peptides 626:

Table 5. Predicted T-cell Epitope Peptides of HER2 ECD Subdomain III. Only High Score MHC is Shown

ID	Peptides	Position	Tools*
MHC class I alleles			
P1 HLA-A*0201	KIFGSLAFL	369-377	1, 2, 3, 4
P2	ALIHHNTHL	466-474	1, 2, 3, 4
P3	GLGISWLGL	447-455	1, 2, 3, 4
P4	QLFRNPHQA	484-492	1, 2, 3, 4
MHC class II alleles			
P5 DRB1_0401	LIHHNTHL	467-484	3, 4, 5
P6	FQNLQVIRGRIL	425-436	3, 4, 5
P7	WLGLRSLRE	452-460	3, 4, 5
P8 DRB5_0101	FVHTVPWDQ	476-484	3, 4, 5
P9	VRAVTSANI	353-361	3, 4, 5
P10	LGISWLGLR	448-456	3, 4, 5
P11	YLYISAWPD	409-417	3, 4, 5

*Tools that recognized the corresponding T cell epitopes or part of the epitope are numbered as follows: SYFPEITHI (1), HLA peptide motif search (2), MHCpred (3), SVMHC (4), Propred (5)

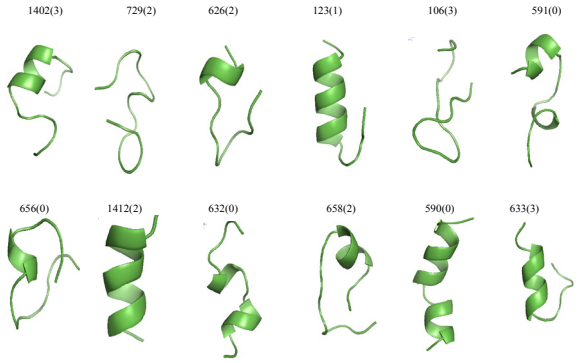

Figure 2. Predicted Structures of B-cell Epitope Peptides and The Number of Their Turn as Obtained From Pepfold Server and Extractturn Server, Respectively

DPASNTAPSAFDPE, 1402: FDPEDPASNTAPQPEQQ, 123: HTANRPEDEVGEGLACHQ, and 729: QGHTHSWHTANRPEDE (Table 2) contained anchoring motifs for both classes of MHC molecules.

3D structure predictions of peptides

The conformation of the peptides plays an important role in epitope recognition by antibodies and in their presentation by helper T-cells to MHC molecules. Therefore, 3D structure of selected peptides listed in table 2-4, was predicted using the PEP-FOLD server. In each case, the most stable predicted structure having minimum total energy was considered as the best prediction result. In addition, beta turn analyses using extractTurn server indicated that the peptides 590, 591, 632, and 656 had no turn in their 3D structure and the peptides 1402, 633, and 106 had the maximum number of turns (three) in their 3D structures (Figure 2).

Discussion

Generation of an effective immune response to foreign proteins needs the participation of antigen presenting cells (APC), T and B lymphocytes (Axelsena et al., 2011). While B cells and antibodies generally recognize conformational epitopes from surface proteins, T cells recognize linear epitopes derived from proteins that are processed by APCs (Weber et al., 2009). In addition, the presentation of T-cell epitopes in the context of MHC is critical to the recognition of self vs. foreign proteins and to the development of both cellular and humoral immune responses (De Groot et al., 2010). Furthermore, the ability of proteins to interface with the humoral (B-cell) and cellular (T-cell) immune systems, influence their potential immunogenicity. Therefore, an immunogen construct should contain both B and T-cell epitopes (Wang et al., 2010). It could be either large proteins that have many B and T-cell epitopes or synthetic peptides containing only the target B-cell epitope or T-cell epitope for generating specific antibodies (Garrett et al., 2007; Axelsena et al., 2011).

Synthetic peptide vaccines containing B-cell epitopes are poor immunogens (Gangwara et al., 2012). Also, clinical responses of vaccination with a peptide containing only T-cell epitope, which is considered to be vital for antitumor activity, are poor (Rosenberg, 2001). For instance, the HER2/neu peptide vaccine, HER2 369-377, that binds to MHC class I molecules which has been used in most published clinical trials, did not show any effectiveness for treatments (Zaks and Rosenberg, 1998; Noguchi and Sasada, 2013). Also, the cytotoxicity of HER2/neu peptides corresponding to the extracellular domain of the protein (HER2 42-56, HER2 98-114, and HER2 328-345), or peptides corresponding to the intracellular domain of the protein (HER2 776-790, HER2 927-941, and HER2 1166-1180) has not been reported in eight treated patients with stage III or IV breast or ovarian cancer. But immunization with the peptides HER2 369-384, HER2 688-703, and HER2 971-984, each containing a putative HLA-A2 binding motif elicited proliferative responses (Correa and Plunkett, 2001). Among the peptides used in different studies, only the peptide HER2 369-384 that corresponded to the subdomain III of HER2 ECD was similar to the peptide P1: KIFGSLAFL 369-377 of our study (Table 5) for HLA-A*0201 allele in MHC class I molecules. This sequence had already been reported by Firat et al. (1999) as a T-cell epitope for the same HLA allele. Also, Kastenmuller et al., 2007 identified the peptide ILHNGAYSL (435-443) as an epitope with the same HLA allele. In addition, Gritzapis et al., in 2008 and 2010, represented HER-2/neu (657-665) and (85-94) as immunogenic epitopes of HER-2/neu oncoprotein with potent antitumor properties. In the current study predictions failed to identify these mentioned peptides because the purpose was to predict peptides from subdomain III of HER2 ECD and one of them that was frequently selected by three and four tools used in this work.

To act as an effective immunogen, B and T-cell epitopes should be incorporated as a single chimeric peptide. In this way, Dakappagari et al. (2003) incorporated identified linear HER2 ECD B-cell epitopes (316-339 and 485-503) based on computer-aided analyses to measles virus fusion (MVF) 'promiscuous' T cell epitope via a four-residue linker sequence. They induced high-titer Abs inhibiting the growth of human breast cancer cells. Furthermore, Garrett et al. (2007) reported phase I clinical trial using

the first generation peptide vaccines of HER2 ECD: MVF 316-339 and MVF 628-647. They also used peptide constructs MVF 613-626, MVF 563-598, and MVF 597-626. All epitopes were immunogenic in FVB/n mice and the 597-626 epitope significantly reduced cancer growth in transgenic BALB-neuT mice. These epitopes were linear B-cell epitope and from subdomain II and IV of HER2 ECD. Thus, the findings of our work are different from those predicted epitopes or peptides of the previous mentioned studies. In our study the predicted peptides were discontinuous and were from subdomain III of HER2 ECD. Kumar et al. (2011) reported the peptide p185 (378-394) (PESFDGDPASNTAPLQPE), which encompasses several segments (S75, S76, S77, and S78) of HER2 ECD subdomain III shown in Table 1, as an immunodominant auto antigen mimotope possibly serving as an antigenic stimulus to trigger systemic lupus erythematosus in a subset of patients. Thus, it presents a potential safety issue for prospective vaccine development of constructs containing the said segments. Additionally, Fisher et al., 2010 reported several novel discontinuous B-cell epitopes of HER2 subdomain III that Fab37 binds to it with high affinity. These epitopes are primarily on subdomain III and IV, remote from the binding sites for therapeutic anti-HER2 antibodies pertuzumab and trastuzumab on subdomain II or IV, respectively. The Fab37 binding site on subdomain III overlaps with some sequence segments (S80, S84, S85, S90, and S94) of HER2 ECD shown in Table 1 and the binding site almost opposite the region of subdomain III contacted by subdomain I.

Active immunization against HER2 has yet to be experimentally evaluated in any form. For example, the CHP-HER2 vaccine, comprising truncated 146HER2 protein 127-146, combined with nanogels of Colesteryl Pullulan (CHP), is used as a novel protein antigen vaccine that elicits 146HER2-specific CD8(+) and CD4(+) T-cell immune responses in patients with HER2-expressing tumors. Also, this vaccine can induce HER2-specific humoral responses in patients with HER2-expressing tumors (Kageyama et al., 2008).

Another way for generating effective immune responses against peptide vaccines or construction of a functional or effective immunogen is using discontinuous peptides containing epitopes for both B and T-cells. In the present study, our aim was *in silico* designing of novel discontinuous peptide vaccines of HER2 ECD subdomain III, containing segments as epitopes for both B and T-cells, by immunoinformatics tools. Similar to our work, Miyako et al. (2011) identified a new epitope N: 163-182 of HER2 ECD containing epitopes for both B and helper T-cells. It was immunogenic and caused a high titer of peptide specific IgG antibodies in immunized BALB/c mice. In contrast to the present study, this epitope was located on subdomain I of HER2 ECD. Moreover, Senpuku et al. (1997) specified a unique peptide, containing both B and T-cell epitopes for the induction of cross-reacting antibodies to the surface protein antigen of *Streptococcus mutans*. Their results provided useful information for the design of an effective anticaries peptide vaccine.

Both B and T-cell epitope predictions were essential for epitope-based vaccine design (Chen et al., 2011). In recent years for prediction of B and T-cell epitopes of proteins with various origins, many immunoinformatics tools are available (Singh et al., 2011). In the current study SYFPEITHI, MHCPred, Propred, SVMHC, and HLA peptide motif search for prediction of T-cell epitope peptides were employed. Different studies have used distinctive methods of T-cell epitope prediction. Similar to our work, Gritzapis et al. (2008; 2010), used SYFPEITHI algorithm for T-cell epitope prediction of HER2 ECD. Based on that study they were able to introduce potent antitumor peptides. Furthermore, Singh et al. (2011) and Wiwanitkit et al. (2007) performed computational analysis of human papilloma virus and Lig A of *Leptospira interrogans* sequences to find T-cell epitopes using the bioinformatics tool MHCpred. In addition, Nair et al. (2011), reported the identification of epitopes of Alcohol dehydrogenase from *Curvularia lunata*. They predicted HLA class II binding peptides of this antigen using *in silico* methods MHCpred, SVRMHC, and SMM-Align and also, the online tools Propred, Multipred, SVMHC, and RANKPEP.

Despite of majority of B-cell epitopes being conformational; most of the B-cell epitope prediction tools are based on primary sequence for prediction of linear B-cell epitopes (Vita et al., 2010). However, a few recent studies propose bioinformatics tools based on 3D structure to predict epitopes such as SUPERFICIAL (Goede et al., 2010), Disco tope (Haste Andersen et al., 2006), Conformational Epitope Predictor (CEP) (Kulkarni-Kale et al., 2005), and PEPOP (Moreau et al., 2008). Although, computational methods have different performance; identification and prediction of suitable epitope peptides for vaccine design using these methods are cheaper and quicker than that of experimental screening. For example, CEP correctly identifies 76% of the amino acid residues known to describe the selected epitopes in data sets consisting of 63 antigen-antibody complexes (Kulkarni-Kale et al., 2005) but Disco Tope detects 15.5% (sensitivity) of residues located in discontinuous epitopes with a specificity of 95%. So, CEP can predict more accurately than Disco tope (Roggen, 2008). Furthermore, Moreau et al. (2008) explained that the performance of PEPOP is better than CEP and Disco tope. Therefore, in the present study we used newly developed immunoinformatics tool, PEPOP, for prediction of discontinuous B-cell epitope peptides. This is due to the fact that most of B-cell epitopes (~90%) are conformational epitopes (Van Regenmortel, 1996) and it is known that most of antibody interacting residues are discontinuous, hydrophobic, and exposed surface parts of antigens (Ansari and Raghav, 2010). In our previous study, a combination of linear B cell epitope prediction web servers such as ABCpred, BCPREDs, Bepired, Bcepred, Elliprro, Discotope, CBtope, and SUPERFICIAL software was used for prediction of B-cell epitope peptides of HER2 ECD subdomain III (Mahdavi et al., 2012) and the predicted peptides were conformational B cell epitopes P_1C: 378-393 (PESFDGDPASNTAPLQ) and P_2C: 500-510 (PEDECVGEGLA) by the integrated strategy and P4: PESFDGD-X-TAPLQ, P5: PESFDGDP X TAPLQ, P6: ESFDGDP X NTAPLQP, P7: PESFDGDP-X-NTAPLQ,

P8: ESFDG-XX-TAPLQPEQL, and P9: ESFDGDP-X-NTAPLQP by SUPERFICIAL software. P_1C and P_2C overlap with S75-S78 and S99-S100, respectively. In SUPERFICIAL software, predicted conformational peptides were linear B-cell epitopes that were linked by the linkers. With this software, the selection of an appropriate linker was difficult and the methodology employed was not clear. Similar to SUPERFICIAL software, PEPOP uses 3D structure of proteins to identify surface accessible segments but it does not use the linkers. Actually, PEPOP defines regions (clusters and patches) on the protein to predict discontinuous B-cell epitope peptides that can be used as immunogens for generating site specific Abs. Lebreton et al. (2011) predicted discontinuous epitopes of the C2 domain of Factor VIII (FVIII) using the bioinformatics algorithm PEPOP for epitope mapping. Also, Alvarenga et al. (2009), *in silico* designed 24 discontinuous-continuous peptides of scorpion toxin using PEPOP for selection of ones that can be recognized by anti Amm VIII antibodies. In our study, optimized peptide design methods of PEPOP selected 12 "discontinuous-continuous peptides" (Tables 2-4) from 3D structure of subdomain III of HER2 ECD (PDB: 1S78). T-cell epitope prediction results indicated peptides containing epitopes for B-cells and both of MHC class molecules: 626: DPASNTAPSAFDPE, 1402: FDPEDPASNTAPQPEQQ, 123: HTANRPEDEVGEGLACHQ, and 729: QGHTHSWHTANRPEDE (Table 2). The peptides defined in this work may induce not only specific antibodies but also generate antitumor effects due to cytotoxic T-cells (CTL) because both helper T-cells and CTLs in the context of MHC class II and I have critical roles in antitumor responses (Kobayashi et al., 2000). Helper T-cells, which are activated by peptides presented on MHC class II molecules, are critical to the supply of activated dendritic cells (DC) and for the generation and persistence of CTL (Kennedy and Celis, 2008). Moreover, helper T-cell activity is essential for IgG antibody production with high affinity (Kobayashi and Celis, 2008). Therefore, vaccination with the peptides predicted in the present study could be useful for inducing effective antibody mediated antitumor therapy. However, limitations of this study should be considered. The results from this research are only predicted results and further confirmation studies are required. *In vitro* synthesis of the identified peptides and *in vivo* experimental screening to test the efficacy are necessary for vaccine development. Another potential difficulty that might be encountered during subsequent attempts at experimental validation is conformational differences between cognate peptide and protein sequences. This means, even if the predicted peptides were to be used for immunization; the resulting antipeptide antibodies would likely fail to cross react with native HER2; especially if the peptides tended to adopt conformations different from those of the corresponding sequences in native HER2. According to Chen et al. (2009) linear peptides could be poor structural mimic of discontinuous epitopes. However, the peptides used by the community are generally continuous fragments of more complex epitopes whereas, in PEPOP, the proposed peptides gather several fragments of the same region of a protein. At first, the residues of these fragments are all surface exposed contrary to the 20-30% of the amino acid found to be buried in continuous epitopes by Chen et al. (2009). Secondly, the PEPOP methods attempt to find an optimal arrangement of these fragments in the peptide that reflect the most possible conformation of the corresponding region in the cognate protein. Moreover, the conformation is not the solely criteria influencing the binding of the Abs. The amino acid composition is also important. However, many of the potential epitopes have a natural preference for beta turn conformation (Ripoll, 1992; Pellequer et al., 1994). For these reasons in this work the conformation of the selected peptides and the number of beta turn in their 3D structures were identified by PEP-FOLD server (Figure 2). The occurrence of beta turns in 3D structure of the predicted peptides for further studies is important.

In conclusion, this study is the first that reports prediction of novel discontinuous B-cell epitope peptide vaccine of HER2 ECD subdomain III containing both B and T-cells epitopes. Our findings in the current study using immunoinformatics tools could be used in diagnostic tools, cancer therapy, vaccine design and also reducing the time and minimizing the total needed tests to find possible proper epitopes.

References

Alvarenga L, Moreau V, Felicori L, et al (2010). Design of antibody-reactive peptides from discontinuous parts of scorpion toxins. *Vaccine*, **28**, 970-80.

Ansari HR, Raghava GP (2010). Identification of conformational B-cell Epitopes in an antigen from its primary sequence. *Immunol Res*, **6**, 6.

Arensa R, Hallb TV, van der Burgb SH (2013). Prospects of combinatorial synthetic peptide vaccine-based immunotherapy against cancer. *Sem Immunol*, **25**, 182-90.

Awada A, Bozovic-Spasojevic I, Chow L (2012). New therapies in HER2-positive breast cancer: a major step towards a cure of the disease? *Cancer Treatment Reviews*, **38**, 494-504.

Axelsena TV, Holma A, Christiansena G, Birkelund S (2011). Identification of the shortest AB-peptide generating highly specific antibodies against the C-terminal end of amyloid-B42. *Vaccine*, **29**, 3260-9.

Bian H, Reidhaar-Olson JF, Hammer J (2003). The use of bioinformatics for identifying class II-restricted T-cell epitopes. *Meth*, **29**, 299-309.

Calabrich A, Fernandes GS, Katz A (2008). Trastuzumab: mechanisms of resistance and therapeutic opportunities. *Oncol*, **22**, 1250-8.

Chen P, Rayner S, Hu KH (2011). Advances of bioinformatics tools applied in virus epitopes prediction. *Virologica Sinica*, **26**, 1.

Chen SW, Van Regenmortel MH, Pellequer JL (2009). Structure-activity relationships in peptide-antibody complexes: implications for epitope prediction and development of synthetic peptide vaccines. *Curr Med Chem*, **16**, 953.

Cho HS, Mason K, Ramyar KX, et al (2003). Structure of the extracellular region of HER2 alone and in complex with the herceptin fab. *Nature*, **421**, 756-60.

Correa I, Plunkett T (2001). Update on HER-2 as a target for cancer therapy HER2/neu peptides as tumour vaccines for T cell recognition. *Breast Cancer Res*, **3**, 399-403.

Dakappagari NK, Douglas DB, Triozzi PL, et al (2000). Prevention of mammary tumors with a chimeric Her-2 B-cell

epitope peptide vaccine. *Cancer Res*, **60**, 3782-9.
Dakappagari NK, Pyles J, Parihar R, et al (2003). Multi-human epidermal growth factor receptor-2 B cell epitope peptide vaccine mediates superior antitumor responses. *J Immunol*, **170**, 4242-53.
De Groot AS, Cohen T, Ardito M, et al (2010). Use of bioinformatics to predict MHC ligands and T-cell epitopes: application to epitope-driven vaccine design. *Methods Microbiol*, **37**, 35-66.
Disis M, Schiffman K (2001). Cancer vaccines targeting the HERZ/neu oncogenic protein. *Sem Oncol*, **28**, 12-20.
Dönnes P, Elofsson A (2002). Prediction of MHC class I binding peptides, using SVMHC. *BMC Bioinformatics*, **3**, 25.
Doytchinova IA, Flower DR (2002). Quantitative approaches to computational vaccinology. *Immunol Cell Biol*, **80**, 270-9.
Firat H, Garcia-Pons F, Tourdot S, et al (1999). H-2 class I knockout, HLA-A2.1-transgenic mice: a versatile animal model for preclinical evaluation of antitumor immunotherapeutic strategies. *Eur J Immunol*, **29**, 3112.
Fisher RD, Ultsch M, Lingel A, et al (2010). Structure of the complex between HER2 and an antibody paratope formed by side chains from tryptophan and serine. *J Mol Biol*, **402**, 217.
Franklin MC, Carey KD, Vajdos FF, et al (2004). Insights into ErbB signaling from the structure of the ErbB2-pertuzumab complex. *Cancer Cell*, **5**, 317-28.
Gangwara RS, Shil P, Sapkala GN, Khanc SA, Gore MM (2012). Induction of virus-specific neutralizing immune response against West Nile and Japanese encephalitis viruses by chimeric peptides representing T-helper and B-cell epitopes. *Vir Res*, **163**, 40-50.
Garrett JT, Kaumaya P, Pei D, Dalbey R, Magliery T (2007). Peptide based B cell epitope vaccines targeting HER-2/Neu. PhD thesis. Ohio State University, Ohio, USA.
Goede A, Jaeger IS, Preissner R (2005). SUPERFICIAL-Surface mapping of proteins via structure-based peptide library design. *BMC Bioinformatics*, **6**, 223.
Gritzapis AD, Fridmanb A, Pereza SA, et al (2010). HER-2/neu (657-665) represents an immunogenic epitope of HER-2/neu oncoprotein with potent antitumor properties. *Vaccine*, **28**, 162-70.
Gritzapis AD, Voutsas IF, Lekka E, et al (2008). Identification of a novel immunogenic HLA-A*0201-binding epitope of HER-2/neu with potent antitumor properties. *J Immunol*, **181**, 146-54.
Guan P, Doytchinova IA, Zygouri C, Flower DR (2003). MHCPred: a server for quantitative prediction of peptide-MHC binding. *Nucleic Acids Res*, **31**, 3621-4.
Haste Andersen P, Nielsen M, Lund O (2006). Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Sci*, **15**, 2558-67.
Jacot W, Fiche M, Zaman K (2013). The HER2 amplicon in breast cancer: topoisomerase IIA and beyond. *Biochim Biophys Acta (BBA)*, **1836**, 146-57.
Kageyama S, Kitano S, Hirayama M, et al (2008). Humoral immune responses in patients vaccinated with 1-146 HER2 protein complexed with cholesteryl pullulan nanogel. *Cancer Sci*, **99**, 601.
Kastenmuller W, Gasteiger G, Gronau JH, et al (2007). Cross-competition of CD8+T cells shapes the immunodominance hierarchy during boost vaccination. *J Exp Med*, **204**, 2187.
Kennedy R, Celis E (2008). Multiple roles for CD4+ T cells in antitumor immune responses, *Immunol Rev*, **222**, 129.
Kobayashi H, Celis E (2008). Peptide epitope identification for tumorreactive CD4 T cells. *Cur Opin Immunol*, **20**, 221.
Kobayashi H, Wood M, Song Y, Appella E, Celis E (2000). Defining promiscuous MHC class II helper T-cell epitopes for the HER2/neu tumor antigen. *Cancer Res*, **60**, 5228-36.
Kulkarni-Kale U, Bhosle S, Kolaskar AS (2005). CEP: a conformational epitope prediction server. *Nucleic Acids Res*, **33**, 168-71.
Kumar S, Hinks JA, Maman JC, et al (2011). p185, an immunodominant epitope, is an autoantigen mimotope. *J Bio Chem*, **286**, 26220.
Lax I (1988). Localization of a major receptor-binding domain for epidermal growth factor by affinity labeling. *Mol Cell Biol*, **8**, 1831-4.
Lebreton A, Moreau V, Lapalud P, et al (2011). Discontinuous epitopes on the C2 domain of coagulation Factor VIII mapped by computer-designed synthetic peptides. *Bri J Haematol*, **155**, 487-97.
Li GF, Wang Y, Zhang ZS (2005). Identification of immunodominant Th1-type T cell epitopes from Schistosoma japonicum 28 kDa glutathione-Stransferase, a vaccine candidate. *Acta Biochim Biophys Sin*, **37**, 751-8.
Mahdavi M, Mohabatkar H, Keyhanfar M, Jafarian Dehkordi A, Rabbani M (2012). Linear and conformational B cell epitope prediction of HER 2 ECD-subdomain III by *in silico* methods. *Asian Pac J Cancer Prev*, **13**, 3053.
Maupetit J, Derreumaux P, Tuffery P (2009). PEP-FOLD: an online resource for de novo peptide structure prediction. *Nucleic Acids Res*, **37**, 498-503.
Miyako H, Kametani Y, Katano I, et al (2011). Antitumor effect of new HER2 peptide vaccination based on B cell epitope. *Anticancer Res*, **31**, 3361-8.
Moreau V, Fleury C, Piquer D, et al (2008). PEPOP: computational design of immunogenic peptides. *BMC Bioinformatics*, **9**, 71.
Nair S, Kukreja N, Singh BP, Arora N (2011). Identification of B cell epitopes of alcohol dehydrogenase allergen of *Curvularia lunata*. *PLoS ONE*, **6**, 20020.
Noguchi M, Sasada T (2013). Personalized peptide vaccination: a new approach for advanced cancer as therapeutic cancer vaccine. *Cancer Immunol Immunother*, **62**, 919-29.
Parker KC, Bednarek MA, Coligan JE (1994). Scheme for ranking potential HLA-A2 binding peptides based on independent binding of individual peptide side-chains. *J Immunol*, **152**, 163.
Pellequer JL, Westhof E, Van Regenmortel MH (1994). Epitope prediction from primary structure of proteins. In Peptide Antigens: A Practical Approach. GB Wisdow ed. Oxford, IRL Press.
Rammensee HG, Bachmann J, Emmerich NN, Bachor OA, Stevanovic S (1999). SYFPEITHI: database for MHC ligands and peptide motifs. *Immunogenetics*, **50**, 213-9.
Ripoll DR (1992). Conformational study of a peptide epitope shows large preferences for beta-turn conformations. *Int J Pept Protein Res*, **40**, 575-81.
Roggen EL (2008). B-cell epitope engineering: a matter of recognizing protein features and motives. *Drug Discovery Today: Technologies*, **5**, 49-55.
Rosenberg S (2001). Progress in human tumor immunology and immunotherapy. *Nature*, **411**, 380-4.
Senpuku H, Kato H, Takeuchi H, Noda A, Nisizawa T (1997). Identification of core B cell epitope in the synthetic peptide inducing cross inhibiting antibodies to surface protein antigen of streptococcus mutans. *Immunol invest*, **26**, 531-48.
Singh AK, Rathb SK, Misraa K (2011). Identification of epitopes in Indian human papilloma virus 16 E6: A bioinformatics approach. *J Vir Methods*, **177**, 26-30.
Singh H, Raghava GPS (2001). ProPred: Prediction of HLA-DR binding sites. *Bioinformatics*, **17**, 1236-7.
Siyi H, Zhiqiang Z, Liangwei L, et al (2008). Epitope mapping and structural analysis of an anti-ErbB2 antibody A21: Molecular basis for tumor inhibitory mechanism. *Proteins*,

70, 938-49.

Tai W, Mahato R, Cheng K (2010). The role of HER2 in cancer therapy and targeted drug delivery. *J Control Rel*, **146**, 264-75.

Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res*, **22**, 4673.

Van Regenmortel MH (1996). Mapping epitope structure and activity: from onedimensional prediction to four-dimensional description of antigenic specificity. *Meth*, **9**, 465.

Vita R, Zarebski L, Greenbaum JA, et al (2010). The immune epitope database 2.0. *Nucleic Acids Res*, **38**, 854-62.

Wallecha A, Ramos K, Malinina I, Singh R (2012). Listeria monocytogenes-based bivalent Lm-LLO immunotherapy for the treatment of HER2/neu positive and triple negative breast cancer and its impact on immunosuppression. *Cancer Research*, **72**, 4-5.

Wang B, Kaumaya PTP, Cohn DE (2010). Immunization with synthetic VEGF peptides in ovarian cancer. *Gyn Oncol*, **119**, 564-70.

Weber CA, Mehta PJ, Ardito M, et al (2009). T cell epitope: friend or foe? Immunogenicity of biologics in context. *Adv Drug Delivery Rev*, **61**, 965-76.

Wiwanitkit V (2007). Predicted epitopes of Lig A of *Leptospira interrogans* by bioinformatics method: a clue for further vaccine development. *Vaccine*, **25**, 2768-70.

Zaks TZ, Rosenberg SA (1998). Immunization with a peptide epitope from HER-2/neu leads to peptide-specific cytotoxic T lymphocytes that fail to recognize HER-2/neu+ tumors. *Cancer Res*, **58**, 4902-8.