Journal of Pharmaceutical Research International

33(22A): 1-13, 2021; Article no.JPRI.66974
ISSN: 2456-9119
(Past name: British Journal of Pharmaceutical Research, Past ISSN: 2231-2919, NLM ID: 101631759)

Screening of Human Epidermal Growth Factor Receptor 2 (HER2) Extracellular Domain for Potential Epitopes by Using Immuno-informatics Tools

Hasanain Abdulhameed Odhar[1*], Salam Waheed Ahjel[1], Zanan Abdulhameed Odhar[2], Suhad Sami Humadi[1], Ali Mahmood Rayshan[1] and Ahmed Fadhil Hashim[1]

[1]Department of Pharmacy, Al-Zahrawi University College, Karbala, Iraq.
[2]Ministry of Health, Baghdad, Iraq.

Authors' contributions

This work was carried out in collaboration among all authors. Author HAO conceived project idea, carried out computational work and wrote first draft of manuscript. Authors SWA and ZAO helped in results interpretation and project supervision. Authors SSH, AFH and AMR managed the literatures search. All the authors read and approved the final manuscript.

Article Information

DOI: 10.9734/JPRI/2021/v33i22A31383
Editor(s):
(1) Dr. Vasudevan Mani, Qassim University, Kingdom of Saudi Arabia.
Reviewers:
(1) Sayeh Lavasani, City of Hope National Cancer Institute, USA.
(2) Rong Xiu Li, Shanghai Jiao Tong University, China.
Complete Peer review History: http://www.sdiarticle4.com/review-history/66974

Original Research Article

Received 25 January 2021
Accepted 01 April 2021
Published 09 April 2021

ABSTRACT

The human epidermal growth factor receptor 2 (HER2) is a well-studied oncoprotein that is overexpressed in a considerable proportion of breast cancer patients. The increased expression of this tyrosine kinase receptor is usually associated with poor clinical prognosis in female patients with breast cancer. In these patients, specific response of immune system against HER2 had been observed. This suggests that immunotherapy approaches can be employed for enhancing the response of tumor infiltrating lymphocytes against HER2 in susceptible tumor microenvironment. In this regard, peptide vaccines are considered one of the most affordable immunotherapy modalities due to their low production cost and long-term effect. For this purpose, we have screened the extracellular domain of HER2 crystal for potential B-cells and T-cells epitopes by using different

**Corresponding author: E-mail: hodhar3@gmail.com;*

immuno-informatics tools. The output peptides were then refined and filtered according to their antigenicity, allergenicity and vulnerability to selected proteases. Here, we present multiple B-cells and T-cells epitope candidates against HER2 extracellular domain with high antigenicity, low allergenicity and good resistance for selected proteolytic enzymes. These filtered epitopes can be used for design and construction of anti-HER2 peptide vaccine for potential use in HER2 positive breast cancer patients. Additionally, the sequence of linear B-cells epitopes can be used for the design of monoclonal antibody variable region against HER2 extracellular domain.

Keywords: Breast cancer; HER2; immunotherapy; epitope; peptide vaccine.

1. BACKGROUND

Breast cancer is considered the second main cause of cancer death in female patients with an estimated mortality of 2.6% [1]. Therapeutic options available for management of breast cancer involves surgery, chemotherapy, radiotherapy, hormonal therapy and immunotherapy. A significant advancement in cancer immunotherapy has been accomplished due to better understanding of immune cells regulatory roles in tumor microenvironment (TME) [2]. Cancer immunotherapy does include different modalities like vaccination, adoptive T-cells therapy and chimeric antigen receptor (CAR) T-cells therapy [3]. These forms of immunotherapy are designed to enhance the capacity of tumor infiltrating lymphocytes (TILs) to recognize tumor associated antigen (TAA) and hence halt tumor progression [4]. The human epidermal growth factor receptor 2 (HER2) also known as ErbB2 (Erythroblastosis homolog B2) is a well-known oncoprotein. HER2 is a receptor tyrosine kinase that is overexpressed in 20% to 30% of invasive breast cancer cases [5]. The higher expression of HER2 on the surface of tumor cells is usually associated with poor clinical outcome and more tumor invasiveness. Immune system specific activities against HER2 had been observed in HER2 positive breast cancer patients. Thus, stimulating immune cells to target HER2 can be considered as a potential therapeutic tool in HER2 positive breast cancer patients [4].

Unlike other cancer immunotherapy modalities, vaccination represents a cost effective method with ability to induce long term memory effect [2]. Peptides derived from different parts of HER2 molecule had been used to generate several anti-HER2 vaccine candidates. In this regard, one of the most effective and promising vaccine candidates is E75 with ability to induce cytotoxic T-lymphocytes response against HER2 molecules as observed in clinical trials. This experimental peptide vaccine was derived from the extracellular domain of HER2 molecule with a sequence of “KIFGSLAFL” and a position located at 369–377 [6].

In the current study, we have screened the extracellular domain of HER2 crystal with several immuno-informatics tools to identify potential linear epitopes for B-cells and T-cells. The predicted epitopes were then filtered according to their antigenicity, allergenicity and susceptibility to selected proteases. The final filtered epitope candidates can be used to design novel anti-HER2 peptide vaccine or even monoclonal antibody variable region when considering B-cells epitopes.

2. METHODOLOGY

2.1 Setting up Screening Study Plan

The general framework for this study is similar to our previously published works [7,8]. A flowchart summary for screening and filtration steps to identify potential epitopes can be seen in Fig. 1.

2.2 Prediction of Physicochemical Characteristics for the Extracellular Domain of HER2 Crystal

ProtParam online tool was used to predict various physical and chemical features for the extracellular domain of HER2 [9]. For this purpose, FASTA sequence of HER2 extracellular domain crystal with PDB code 6OGE was submitted to the prediction tool. We have reported different physicochemical characteristics for the submitted sequence like molecular weight, isoelectric point and instability index. Additionally, both allergenicity score and antigenicity potential were predicted for extracellular domain of HER2 sequence by using AllerTOP v. 2.0 and VaxiJen v. 2.0 respectively [10,11]. For prediction of antigenicity, a threshold value of 0.5 was used.

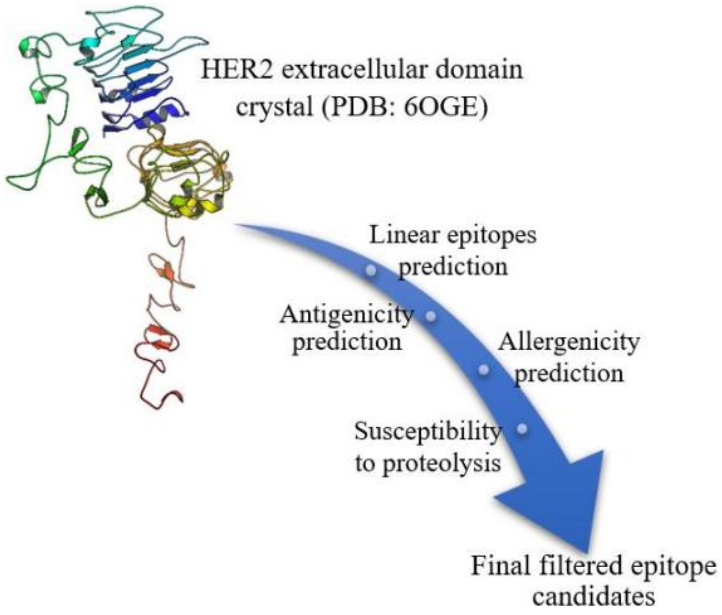

Fig. 1. A concise illustration for study plan steps

2.3 Prediction of Linear B-Cells Epitopes

Antigen sequence properties online tool was employed to screen HER2 extracellular domain sequence for continuous B-cells epitopes. This virtual screening tool was accessed through The Immune Epitope Database (IEDB) [12]. Three prediction methods were used to screen the FASTA sequence of the submitted crystal and these are: Emini surface accessibility scale [13], BepiPred-2.0 [14] and Kolaskar and Tongaonkar antigenicity scale [15]. A default threshold was used for screening by these three prediction methods. Then, antigenicity score was calculated for each generated epitope by using VaxiJen v. 2.0 [11]. We have reported only those epitopes with antigenicity score greater than the threshold value of 0.5.

2.4 Prediction of T-Cells Epitopes Presented by Major Histocompatibility Complex Class I (MHC-I)

The sequence of HER2 extracellular domain was submitted in FASTA format to a combined predictor tool accessible through IEDB [12]. This online tool predicts the potential of a peptide to become a T-cells epitope by calculating peptide ability for processing by proteasomes, transporter associated with antigen processing (TAP) and also MHC-I molecules. For this tool, we have used NetMHCpan version 3.0 method [16] and a selected panel of 51 human leukocyte antigen (HLA) as seen in Table 1. The length of the generated T-cells epitopes was specified to 9-mer. Finally, we have presented only those epitopes with VaxiJen score greater than 0.5.

2.5 Prediction of T-Cells Epitopes Presented by Major Histocompatibility Complex Class II (MHC-II)

We have used Tepitool, accessible through IEDB website, to predict T-cells peptides that can be presented through MHC-II pathway [12,17]. This online tool provides a flexible interface of six steps that facilitates the screening of submitted FASTA sequence for prediction of peptides that can bind either MHC-I or MHC-II molecules. Here, the extracellular domain of HER2 crystal (PDB: 6OGE) was submitted in FASTA format. Then, a panel of pre-selected MHC-II restricted alleles was used for screening the sequence as seen in Table 1. A default setting was applied to generate moderate number of potential epitopes with a length of 15-mer. We have used NetMHCIIpan-3.0 method to predict peptides with potential capacity of MHC-II binding [18]. The generated 15-mer peptides were sorted according to their binding affinity percentile rank with a cutoff value of 2.5. Again, we have only reported those peptides with VaxiJen score more than 0.5.

Table 1. List of MHC restricted alleles employed for screening HER2 extracellular domain crystal for T-cells epitopes

MHC-I restricted alleles	MHC-II restricted alleles
A*01:01, A*02:01, A*02:06, A*03:01, A*11:01, A*23:01, A*24:02, A*25:01, A*26:01, A*29:02, A*30:01, A*30:02, A*31:01, A*32:01, A*33:03, A*68:01, A*68:02, A*74:01, B*07:02, B*08:01, B*13:01, B*13:02, B*14:02, B*15:01, B*15:25, B*18:01, B*27:02, B*27:05, B*35:01, B*35:03, B*37:01, B*38:01, B*39:01, B*40:01, B*40:02, B*44:02, B*44:03, B*46:01, B*48:01, B*49:01, B*50:01, C*01:02, C*02:02, C*03:02, C*04:01, C*05:01, C*06:02, C*07:01, C*08:01, E*01:01, G*01:01	DRB1*01:01, DRB1*03:01, DRB1*04:01, DRB1*04:05, DRB1*07:01, DRB1*08:02, DRB1*09:01, DRB1*11:01, DRB1*12:01, DRB1*13:02, DRB1*15:01, DRB3*01:01, DRB3*02:02, DRB4*01:01, DRB5*01:01, DPA1*01, DPA1*01:03, DPA1*02:01, DPA1*03:01, DPB1*01:01, DPB1*02:01, DPB1*04:01, DPB1*04:02, DPB1*05:01, DQA1*01:01, DQA1*01:02, DQA1*03:01, DQA1*04:01, DQA1*05:01, DQB1*02:01, DQB1*03:01, DQB1*03:02, DQB1*04:02, DQB1*05:01, DQB1*06:02

2.6 Prediction of Allergenicity and Proteolysis Susceptibility for the Generated B-Cells and T-Cells Epitopes

AllerTOP v. 2.0 web-based tool was employed to filtrate and refine the generated epitopes according to their predicted potential to induce allergic reaction [10]. Only those epitopes that are probably non-allergenic were then submitted for proteolysis susceptibility prediction by PeptideCutter tool [19]. The submitted one letter sequence for each epitope was evaluated for degradation vulnerability by Arg-C proteinase, Asp-N endopeptidase, Caspase-1, Neutrophil elastase and Trypsin. Only those epitopes that are resistant to degradation by ≥ 3 enzymes were then subjected for further consideration.

2.7 Evaluating Surface Accessibility of Final Filtered B-Cells Epitopes

Efficient B-cells epitopes must be located in a solvent accessible region of the antigen under evaluation. Surface accessibility of the epitope is essential for successful recognition by B-cells receptors, these receptors are actually membrane bound immunoglobulins [20,21]. We have used PyMOL version 2.3 to visualize the position of filtered linear B-cells epitopes within HER2 extracellular domain crystal [22].

2.8 Molecular Docking of Filtered T-Cells Epitope Candidates against MHC-I Molecule

The tertiary structure for each sequence of T-cells peptides with MHC-I binding capacity was modelled by using PEP-FOLD 2.0 server [23]. The generated PDB file for each epitope was then docked against HLA-A*02:01 crystal (PDB: 5SWQ) by using PatchDock server [24]. For docking process, the receptor binding site was defined with the number of the following residues in chain A of HLA-A*02:01 crystal: 63, 66, 77, 99, 146, 147 and 171. Docking results were then further refined by using FireDock server [25]. The interaction between each 9-mer T-cells epitope and HLA-A*02:01 molecule was then visualized by using LigPlot+ v.1.4.5 [26] and for the first ranked complex only.

2.9 Population Coverage of Final Filtered T-Cells Epitopes

The sequence for each T-cells epitope presented by MHC-I or MHC-II molecules was submitted to population coverage prediction tool via IEDB server [12]. This tool can calculate population response to specific T-cells epitope in various locations of the world by using HLA genotypic frequencies and also collected data about MHC binding and/or T-cells restriction [27]. Here, class I and II combined calculation option was employed to predict population coverage for T-cells epitopes presented by MHC-I or MHC-II pathways. We have used a large panel of MHC restricted alleles as can be seen in Table 1 in order to make sure that challenges like MHC polymorphism and difference in MHC expression frequency among various populations can be minimized.

3. RESULTS AND DISCUSSION

The prediction of physicochemical properties for HER2 extracellular domain crystal, as summarized in Table 2, indicates that the whole crystal cannot be used as anti-HER2 vaccine candidate. This is because the extracellular domain of HER2 seems to be unstable as the instability index is greater than 40, also the crystal is probably a non-antigenic protein with

antigenicity potential less than 0.5 [28]. Therefore, we have screened HER2 extracellular domain for potential B-cells and T-cells epitopes as alternative strategy. It is worth to mention that the extracellular domain of HER2 looks to have a net negative charge as the number of negatively charged residues is greater than those with positive charge, also the predicted isoelectric point is less than 7 [29].

Twenty-two linear B-cells epitopes were reported in HER2 extracellular domain crystal, as seen in Table 3, and by using three prediction methods. These continuous epitopes have variable length and position, all have antigenicity score greater than the threshold value of 0.5.

Regarding the prediction of T-cells epitopes that are presented by MHC-I pathway, 18 peptides were reported in Table 4. All these T-cells epitopes have 9-mer length with antigenicity score more than 0.5. These epitopes were ranked according to their total score, this score represents a cumulative measure for peptide processing by proteasome, TAP and MHC-I. In general, higher total score reflects more efficient presentation of a peptide by MHC-I pathway.

For T-cells epitopes with potential capacity for MHC-II binding, 27 peptide candidates were predicted in Table 5. All these epitopes have 15-mer length and VaxiJen score greater than 0.5. These T-cells epitopes were sorted based on their predicted percentile rank, a lower percentile rank value is usually associated with better peptide binding to MHC-II molecules [12].

Table 2. Predicted physicochemical properties for the extracellular domain crystal of human epidermal growth factor receptor 2 (HER2)

Property	Predicted value
Number of amino acid residues	622
Molecular weight	68465.94 kDa
Theoretical isoelectric point (PI)	5.80
Number of negatively charged residues	63
Number of positively charged residues	46
Instability index (II)	52.08
Antigenicity score	0.4639
Allergenicity potential	Probable non-allergen

Table 3. Continuous B-cells epitopes predicted on the extracellular domain crystal of HER2

No.	Position	Length	Epitope sequence	Antigenicity	Prediction method
1	7-12	6	TDMKLR	1.412	Emini surface
2	150-155	6	IFHKNN	1.015	accessibility
3	296-301	6	HNQEVT	0.895	
4	303-309	7	EDGTQRC	1.443	
5	326-331	6	EHLREV	1.859	
6	461-468	8	DQLFRNPH	0.525	
7	474-479	6	TANRPE	1.683	
8	6-23	18	GTDMKLRLPASPETHLDM	0.624	BepiPred-2.0
9	98-115	18	GDPLNNTTPVTGASPGGL	0.710	
10	179-213	35	GSRCWGESSEDCQSLTRT VCAGGCARCKGPLPTDC	0.535	
11	294-347	54	PLHNQEVTAEDGTQRCEK CSKPCARVCYGLGMEHLR EVRAVTSANIQEFAGCKK	0.648	
12	390-404	15	ISAWPDSLPDLSVFQ	1.105	
13	581-607	27	GVKPDLSYMPIWKFPDEE GACQPCPIN	0.773	
14	169-176	8	ACHPCSPM	1.064	Kolaskar and
15	199-208	10	AGGCARCKGP	1.077	Tongaonkar
16	241-251	11	ICELHCPALVT	1.255	antigenicity
17	270-280	11	GASCVTACPYN	0.504	
18	288-297	10	SCTLVCPLHN	0.935	
19	372-378	7	PEQLQVF	1.013	
20	504-512	9	TQCVNCSQF	0.592	
21	571-581	11	PPFCVARCPSG	0.624	
22	601-615	15	CQPCPINCTHSCVDL	0.876	

Table 4. T-cells epitopes predicted on extracellular domain of HER2 crystal, all these antigenic peptides are 9-mer long and mainly presented by MHC-I molecules

No.	Position	Sequence	Proteasome score	TAP score	Processing score	MHC score	Total score	Antigenicity
1	133-141	IQRNPQLCY	1.54	1.35	2.89	-1.37	1.53	0.643
2	379-387	ETLEEITGY	1.10	1.13	2.24	-0.77	1.47	0.835
3	20-28	HLDMLRHLY	1.19	1.17	2.35	-1.27	1.08	1.323
4	464-472	FRNPHQALL	1.48	0.45	1.93	-0.92	1.01	0.866
5	524-532	VLQGLPREY	1.38	1.35	2.73	-1.84	0.89	0.736
6	457-465	TVPWDQLFR	1.21	0.68	1.89	-1.03	0.86	1.128
7	249-257	LVTYNTDTF	1.28	1.21	2.49	-1.66	0.84	0.680
8	273-281	CVTACPYNY	1.17	1.32	2.49	-1.70	0.80	0.736
9	456-464	HTVPWDQLF	1.34	1.10	2.45	-1.73	0.72	0.887
10	334-342	VTSANIQEF	1.46	1.18	2.65	-1.99	0.66	0.746
11	271-279	ASCVTACPY	1.14	1.34	2.48	-1.83	0.65	0.686
12	413-421	ILHNGAYSL	1.53	0.51	2.04	-1.39	0.65	0.536
13	274-282	VTACPYNYL	1.54	0.44	1.98	-1.41	0.57	1.268
14	151-159	FHKNNQLAL	1.53	0.38	1.91	-1.42	0.49	0.709
15	586-594	LSYMPIWKF	1.41	1.14	2.55	-2.07	0.48	0.597
16	370-378	LQPEQLQVF	1.55	1.06	2.62	-2.17	0.45	0.716
17	389-397	YISAWPDSL	1.57	0.48	2.05	-1.62	0.43	1.677
18	395-403	DSLPDLSVF	1.42	0.92	2.35	-2.21	0.14	1.108

TAP: Transporter associated with antigen processing; MHC: major histocompatibility complex

Table 5. T-cells epitopes predicted on extracellular domain of HER2 crystal and presented by MHC-II pathway. All these predicted peptides are 15-mer long

No.	Position	Sequence	Percentile rank	Antigenicity score
1	409-423	IRGRILHNGAYSLTL	0.08	0.503
2	410-424	RGRILHNGAYSLTLQ	0.09	0.728
3	411-425	GRILHNGAYSLTLQG	0.14	0.639
4	83-97	TQLFEDNYALAVLDN	0.28	0.699
5	61-75	YVLIAHNQVRQVPLQ	0.40	0.521
6	67-81	NQVRQVPLQRLRIVR	0.43	0.518
7	84-98	QLFEDNYALAVLDNG	0.46	0.794
8	147-161	WKDIFHKNNQLALTL	0.54	0.667
9	85-99	LFEDNYALAVLDNGD	0.54	0.935
10	386-400	GYLYISAWPDSLPDL	0.61	0.935
11	387-401	YLYISAWPDSLPDLS	0.71	1.225
12	328-342	LREVRAVTSANIQEF	0.73	0.841
13	385-399	TGYLYISAWPDSLPD	0.79	0.563
14	383-397	EITGYLYISAWPDSL	0.79	0.606
15	326-340	EHLREVRAVTSANIQ	0.87	0.831
16	329-343	REVRAVTSANIQEFA	0.96	0.770
17	146-160	LWKDIFHKNNQLALT	0.99	0.668
18	327-341	HLREVRAVTSANIQE	1.10	0.785
19	325-339	MEHLREVRAVTSANI	1.30	0.808
20	45-59	TNASLSFLQDIQEVQ	1.50	0.808
21	44-58	PTNASLSFLQDIQEV	1.90	0.811
22	46-60	NASLSFLQDIQEVQG	2.10	0.759
23	87-101	EDNYALAVLDNGDPL	2.20	1.238
24	330-344	EVRAVTSANIQEFAG	2.20	0.671
25	371-385	QPEQLQVFETLEEIT	2.40	1.143
26	373-387	EQLQVFETLEEITGY	2.50	0.852
27	47-61	ASLSFLQDIQEVQGY	2.50	0.645

Then, the antigenic B-cells and T-cells epitopes were further filtered and refined based on their potential to induce allergic reaction as reported in Table 6. Only those peptides that are probably non-allergenic were then assessed for their vulnerability to proteolytic degradation by five selected enzymes as seen in Table 7. B-cells and T-cells epitopes that are probably non-allergenic and resistant to degradation by ≥ 3 enzymes were then considered for further

analysis. The sequence along with length and position of these final filtered epitopes are presented in Table 8 as potential candidates.

The position of each potential B-cells epitope, as listed in Table 8, was then visually assessed by PyMOL for surface accessibility. According to Fig. 2, the location of these four linear B-cells epitopes is accessible by solvent. This may facilitate the interaction between these surface peptides in HER2 extracellular domain and membrane bound immunoglobulins in B-cells.

Docking results for interaction between filtered T-cells epitopes and HLA-A*02:01 molecule is summarized in Table 9. For these six T-cells epitopes, we have reported the global energy of binding to MHC-I molecules. A lower global binding energy reflects better interaction between T-cells epitope and MHC-I binding groove. Table 9 also reports the contribution of attractive (VdW) Van der Waals forces energy, (ACE) Atomic contact energy and energy of hydrogen bonds towards global energy. Finally, the table also shows residues in MHC-I molecule that may be involved in hydrogen bond interaction with each T-cells epitope. Fig. 3 represents a three-dimensional illustration for interaction between each T-cells epitope and MHC-I molecule. Modelling of interaction between T-cells epitopes and MHC-I molecule indicates that these six peptides are potential binders.

Table 6. Filtration of the generated linear epitopes according to their predicted allergenicity

No.	Epitope sequence	Allergenicity potential
	Continuous B-cells epitopes	
1	TDMKLR	Probable allergen
2	IFHKNN	Probable allergen
3	HNQEVT	Probable non-allergen
4	EDGTQRC	Probable non-allergen
5	EHLREV	Probable allergen
6	DQLFRNPH	Probable allergen
7	TANRPE	Probable allergen
8	GTDMKLRLPASPETHLDM	Probable allergen
9	GDPLNNTTPVTGASPGGL	Probable non-allergen
10	GSRCWGESSEDCQSLTRTVCAGGCARCK GPLPTDC	Probable allergen
11	PLHNQEVTAEDGTQRCEKCSKPCARVCY GLGMEHLREVRAVTSANIQEFAGCKK	Probable non-allergen
12	ISAWPDSLPDLSVFQ	Probable allergen
13	GVKPDLSYMPIWKFPDEEGACQPCPIN	Probable non-allergen
14	ACHPCSPM	Probable allergen
15	AGGCARCKGP	Probable non-allergen
16	ICELHCPALVT	Probable allergen
17	GASCVTACPYN	Probable non-allergen
18	SCTLVCPLHN	Probable allergen
19	PEQLQVF	Probable allergen
20	TQCVNCSQF	Probable non-allergen
21	PPFCVARCPSG	Probable allergen
22	CQPCPINCTHSCVDL	Probable allergen
	T-cells epitopes (MHC-I)	
1	IQRNPQLCY	Probable allergen
2	ETLEEITGY	Probable allergen
3	HLDMLRHLY	Probable allergen
4	FRNPHQALL	Probable allergen
5	VLQGLPREY	Probable non-allergen
6	TVPWDQLFR	Probable non-allergen
7	LVTYNTDTF	Probable allergen
8	CVTACPYNY	Probable non-allergen
9	HTVPWDQLF	Probable non-allergen
10	VTSANIQEF	Probable non-allergen
11	ASCVTACPY	Probable allergen
12	ILHNGAYSL	Probable non-allergen
13	VTACPYNYL	Probable allergen
14	FHKNNQLAL	Probable allergen
15	LSYMPIWKF	Probable non-allergen

No.	Epitope sequence	Allergenicity potential
16	LQPEQLQVF	Probable allergen
17	YISAWPDSL	Probable allergen
18	DSLPDLSVF	Probable non-allergen
	T-cells epitopes (MHC-II)	
1	IRGRILHNGAYSLTL	Probable allergen
2	RGRILHNGAYSLTLQ	Probable allergen
3	GRILHNGAYSLTLQG	Probable allergen
4	TQLFEDNYALAVLDN	Probable non-allergen
5	YVLIAHNQVRQVPLQ	Probable allergen
6	NQVRQVPLQRLRIVR	Probable non-allergen
7	QLFEDNYALAVLDNG	Probable allergen
8	WKDIFHKNNQLALTL	Probable non-allergen
9	LFEDNYALAVLDNGD	Probable allergen
10	GYLYISAWPDSLPDL	Probable allergen
11	YLYISAWPDSLPDLS	Probable allergen
12	LREVRAVTSANIQEF	Probable non-allergen
13	TGYLYISAWPDSLPD	Probable non-allergen
14	EITGYLYISAWPDSL	Probable allergen
15	EHLREVRAVTSANIQ	Probable non-allergen
16	REVRAVTSANIQEFA	Probable non-allergen
17	LWKDIFHKNNQLALT	Probable allergen
18	HLREVRAVTSANIQE	Probable non-allergen
19	MEHLREVRAVTSANI	Probable non-allergen
20	TNASLSFLQDIQEVQ	Probable allergen
21	PTNASLSFLQDIQEV	Probable allergen
22	NASLSFLQDIQEVQG	Probable allergen
23	EDNYALAVLDNGDPL	Probable allergen
24	EVRAVTSANIQEFAG	Probable non-allergen
25	QPEQLQVFETLEEIT	Probable non-allergen
26	EQLQVFETLEEITGY	Probable allergen
27	ASLSFLQDIQEVQGY	Probable allergen

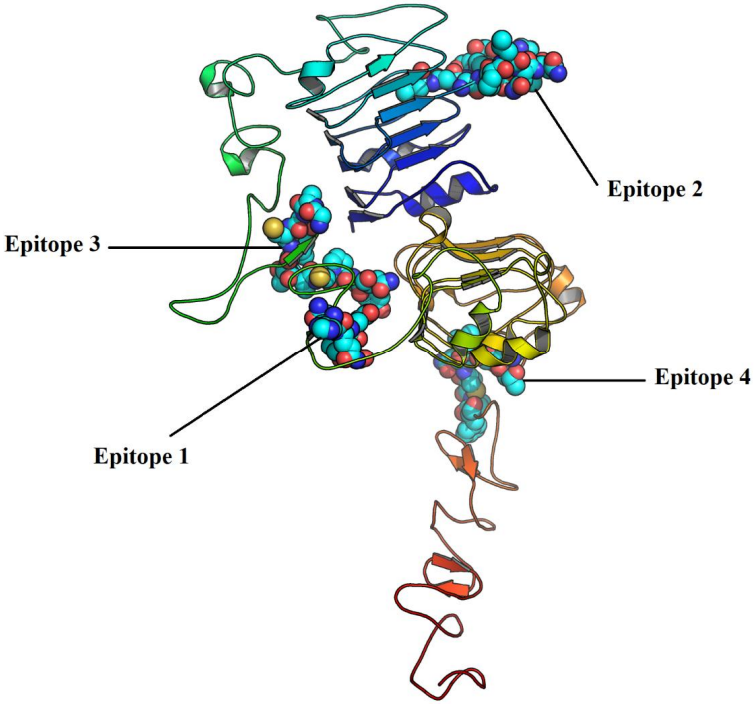

Fig. 2. Locations of final filtered B-cells epitopes are highlighted within HER2 extracellular domain crystal

Table 7. Prediction of the filtered epitopes susceptibility for degradation by selected enzymes

No	Epitope sequence	Susceptibility for digesting enzymes				
		Arg-C proteinase	Asp-N endopeptidase	Caspase-1	Neutrophil elastase	Trypsin
		Continuous B-cells epitopes				
1	HNQEVT	No	No	No	Yes	No
2	EDGTQRC	Yes	Yes	No	No	Yes
3	GDPLNNTTPVTGASPGGL	No	Yes	No	Yes	No
4	PLHNQEVTAEDGTQRCEKCSKP CARVCYGLGMEHLREVRAVTS ANIQEFAGCKK	Yes	Yes	No	Yes	Yes
5	GVKPDLSYMPIWKFPDEEGAC QPCPIN	No	Yes	No	Yes	Yes
6	AGGCARCKGP	Yes	No	No	Yes	Yes
7	GASCVTACPYN	No	No	No	Yes	No
8	TQCVNCSQF	No	No	No	Yes	No
		T-cells epitopes (MHC-I)				
1	VLQGLPREY	Yes	No	No	Yes	Yes
2	TVPWDQLFR	Yes	Yes	No	Yes	Yes
3	CVTACPYNY	No	No	No	Yes	No
4	HTVPWDQLF	No	Yes	No	Yes	No
5	VTSANIQEF	No	No	No	Yes	No
6	ILHNGAYSL	No	No	No	Yes	No
7	LSYMPIWKF	No	No	No	No	Yes
8	DSLPDLSVF	No	Yes	No	Yes	No
		T-cells epitopes (MHC-II)				
1	TQLFEDNYALAVLDN	No	Yes	No	Yes	No
2	NQVRQVPLQRLRIVR	Yes	No	No	Yes	Yes
3	WKDIFHKNNQLALTL	No	Yes	No	Yes	Yes
4	LREVRAVTSANIQEF	Yes	No	No	Yes	Yes
5	TGYLYISAWPDSLPD	No	Yes	No	Yes	No
6	EHLREVRAVTSANIQ	Yes	No	No	Yes	Yes
7	REVRAVTSANIQEFA	Yes	No	No	Yes	Yes
8	HLREVRAVTSANIQE	Yes	No	No	Yes	Yes
9	MEHLREVRAVTSANI	Yes	No	No	Yes	Yes
10	EVRAVTSANIQEFAG	Yes	No	No	Yes	Yes
11	QPEQLQVFETLEEIT	No	No	No	Yes	No

Table 8. List of final filtered epitope candidates predicted on extracellular domain of HER2 crystal

No.	Epitope sequence	Position	Length	Epitope type
1	HNQEVT	296-301	6	Linear B-cells
2	GDPLNNTTPVTGASPGGL	98-115	18	
3	GASCVTACPYN	270-280	11	
4	TQCVNCSQF	504-512	9	
5	CVTACPYNY	273-281	9	T-cells (MHC-I)
6	HTVPWDQLF	456-464	9	
7	VTSANIQEF	334-342	9	
8	ILHNGAYSL	413-421	9	
9	LSYMPIWKF	586-594	9	
10	DSLPDLSVF	395-403	9	
11	TQLFEDNYALAVLDN	83-97	15	T-cells (MHC-II)
12	TGYLYISAWPDSLPD	385-399	15	
13	QPEQLQVFETLEEIT	371-385	15	

Table 9. Results of molecular docking for 9-mer T-cells epitopes against MHC-I molecule

No.	Epitope sequence	Global energy	Attractive VdW	ACE	Hydrogen bonds	Interacting MHC-I residues
1	CVTACPYNY	-44.93	-19.80	-7.00	-1.31	Thr73
2	HTVPWDQLF	-43.73	-26.39	-4.38	-3.83	Lys146
3	VTSANIQEF	-40.97	-19.32	-2.72	-2.63	His114, Gln155
4	ILHNGAYSL	-49.06	-26.63	-5.66	-4.42	Glu63, Thr163
5	LSYMPIWKF	-43.46	-27.38	-8.77	-1.63	None
6	DSLPDLSVF	-55.72	-21.86	-5.64	-1.63	None

VdW: Van der Waals forces; ACE: Atomic contact energy

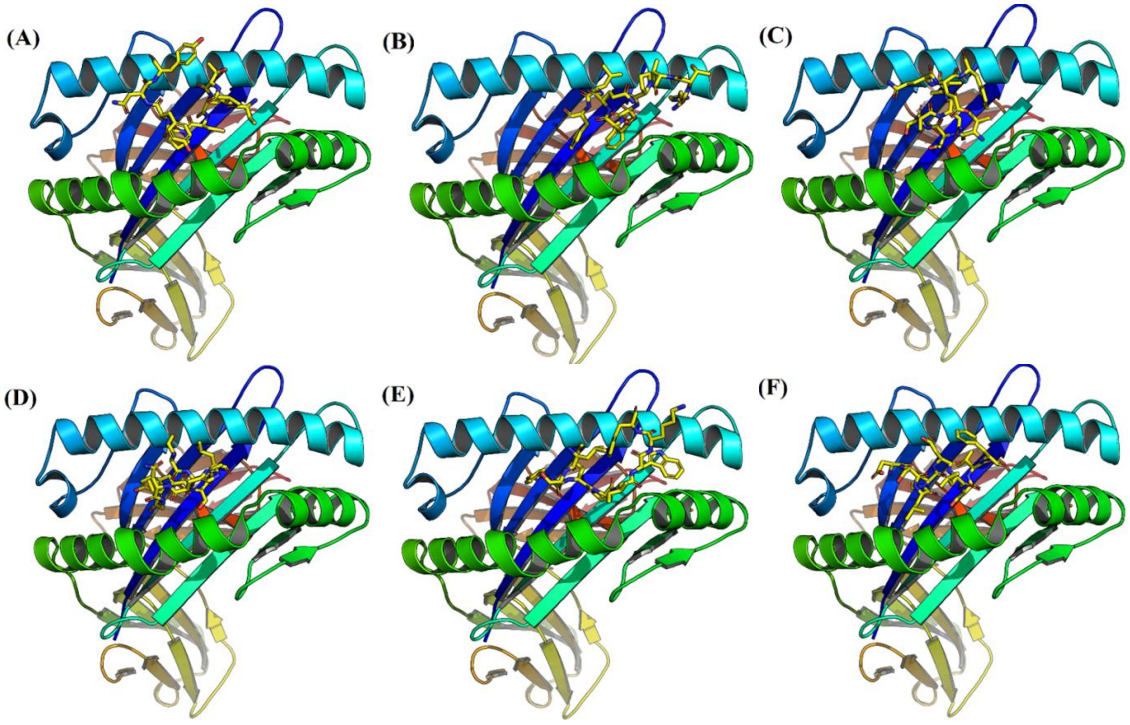

Fig. 3. A three-dimensional illustration for interaction between HLA-A*02:01 crystal and filtered T-cells epitopes with the following sequence: (A) CVTACPYNY, (B) HTVPWDQLF, (C) VTSANIQEF, (D) ILHNGAYSL, (E) LSYMPIWKF, (F) DSLPDLSVF

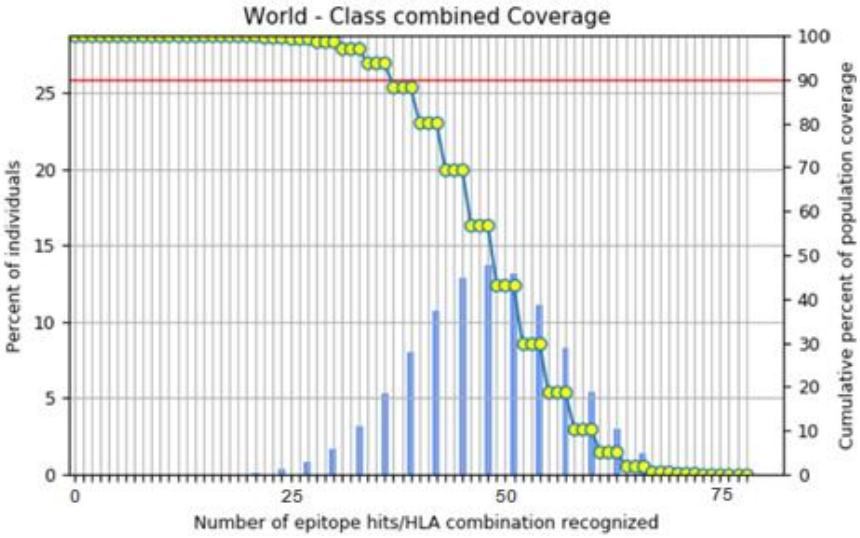

Fig. 4. Worldwide population coverage analysis for filtered T-cells epitopes presented by MHC-I or MHC-II molecules

Finally, the world population coverage analysis of T-cells epitopes presented by either MHC-I or MHC-II pathways shows that these nine peptides have excellent coverage against MHC restricted alleles employed. According to Fig. 4, the combination of these epitopes resulted in a projected worldwide coverage of 100% and 47.81 as average number of epitope hits, while the minimum number of epitope hits was 38.08 as recognized by 90% of the population.

4. CONCLUSION

Here, we report multiple B-cells and T-cells epitopes by screening HER2 extracellular domain crystal with various immuno-informatics tools. The final refined epitopes are predicted to be antigenic, non-allergenic with good resistance against selected proteolytic enzymes. The location of linear B-cells epitopes seems to be solvent accessible; these peptides can be used for the design of antibody variable regions against HER2. On the other hand, T-cells epitopes are believed to be good binders to MHC-I or MHC-II molecules with excellent population coverage. These filtered B-cells and T-cells epitopes can be used for the construction of anti-HER2 peptide vaccine candidate for potential use against HER2 positive breast cancer

CONSENT

It is not applicable.

ETHICAL APPROVAL

According to international standards, ethical approval has been collected and maintained by the authors.

COMPETING INTERESTS

Authors have declared that no competing interests exist.

REFERENCES

1. How Common is Breast Cancer? Breast Cancer Statistics.
Accessed 4 Aug 2020.
Available:https://www.cancer.org/cancer/breast-cancer/about/how-common-is-breast-cancer.html
2. Arab A, Yazdian-Robati R, Behravan J. HER2-Positive Breast Cancer Immunotherapy: A Focus on Vaccine Development. Arch Immunol Ther Exp (Warsz). 2020;68(1).
DOI: 10.1007/s00005-019-00566-1
3. Liu M, Guo F. Recent updates on cancer immunotherapy. Precis Clin Med. 2018; 1(2):65–74.
DOI: 10.1093/pcmedi/pby011
4. Nocera NF, Lee MC, De La Cruz LM, Rosemblit C, Czerniecki BJ. Restoring lost anti-HER-2 Th1 immunity in breast cancer: A crucial role for Th1 cytokines in therapy and prevention. Front Pharmacol. 2016; 7(OCT):356.

DOI: 10.3389/fphar.2016.00356

5. Witton CJ, Reeves JR, Going JJ, Cooke TG, Barlett JMS. Expression of the HER1-4 family of receptor tyrosine kinases in breast cancer. J Pathol. 2003;200(3):290–297.
DOI: 10.1002/path.1370
6. Patil R, Clifton GT, Holmes JP, Amin A, Carmichael MG, Gates JD, et al. Clinical and Immunologic Responses of HLA-A3+ Breast Cancer Patients Vaccinated with the HER2/neu-Derived Peptide Vaccine, E75, in a Phase I/II Clinical Trial. J Am Coll Surg. 2010;210(2):140–147.
DOI: 10.1016/j.jamcollsurg.2009.10.022
7. Odhar H, Ahjel S, Humadi S. Towards the design of epitope candidates for Coronavirus 2. Bioinformation. 2020;16(5): 375–386.
DOI: 10.6026/97320630016351
8. Odhar HA, Ahjel SW, Humadi SS. Towards the design of multiepitope-based peptide vaccine candidate against SARS-CoV-2. bioRxiv. 2020;2020.07.07.186122.
DOI: 10.1101/2020.07.07.186122
9. ExPASy - ProtParam tool.
Accessed 5 Aug 2020.
Available:https://web.expasy.org/protparam/
10. Bioinformatics Tool for Allergenicity Prediction.
Accessed 5 Aug 2020.
Available:https://www.ddg-pharmfac.net/AllerTOP/
11. Doytchinova IA, Flower DR. VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines. BMC Bioinformatics. 2007;8.
DOI: 10.1186/1471-2105-8-4
12. IEDB.org: Free epitope database and prediction resource.
Accessed 5 Aug 2020.
Available:https://www.iedb.org/
13. Emini EA, Hughes J V, Perlow DS, Boger J. Induction of hepatitis A virus-neutralizing antibody by a virus-specific synthetic peptide. J Virol. 1985;55(3):836–839.
DOI: 10.1128/jvi.55.3.836-839.1985
14. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45(W1):W24–29.
DOI: 10.1093/nar/gkx346
15. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276(1–2):172–174.
DOI: 10.1016/0014-5793(90)80535-Q
16. Hoof I, Peters B, Sidney J, Pedersen LE, Sette A, Lund O, et al. NetMHCpan, a method for MHC class i binding prediction beyond humans. Immunogenetics. 2009; 61(1):1–13.
DOI: 10.1007/s00251-008-0341-z
17. Paul S, Sidney J, Sette A, Peters B. TepiTool: A pipeline for computational prediction of T cell epitope candidates. Curr Protoc Immunol. 2016;2016:18.19.1-18.19.24.
DOI: 10.1002/cpim.12
18. Karosiene E, Rasmussen M, Blicher T, Lund O, Buus S, Nielsen M. NetMHCIIpan-3.0, a common pan-specific MHC class II prediction method including all three human MHC class II isotypes, HLA-DR, HLA-DP and HLA-DQ. Immunogenetics. 2013;65(10):711–724.
DOI: 10.1007/s00251-013-0720-y
19. PeptideCutter.
Accessed 6 Aug 2020.
Available:https://web.expasy.org/peptide_cutter/
20. Sanchez-Trincado JL, Gomez-Perosanz M, Reche PA. Fundamentals and Methods for T- and B-Cell Epitope Prediction. Selvan SR, editor. J Immunol Res. 2017;2017:2680160.
DOI: 10.1155/2017/2680160
21. Odhar HA, Ahjel SW. Potential Trends for COVID-19 Fighting: An Immuno-informatics Overview. OSF Prepr; 2020.
DOI: 10.31219/osf.io/d63mp
22. PyMOL.
Accessed 6 Aug 2020.
Available:https://pymol.org/2/
23. Shen Y, Maupetit J, Derreumaux P, Tufféry P. Improved PEP-FOLD approach for peptide and miniprotein structure prediction. J Chem Theory Comput. 2014;10(10):4745–4758.
DOI: 10.1021/ct500592m
24. Schneidman-Duhovny D, Inbar Y, Nussinov R, Wolfson HJ. Patch Dock and SymmDock: servers for rigid and symmetric docking. Nucleic Acids Res. 2005;33(suppl_2):W363–367.
DOI: 10.1093/nar/gki481
25. Andrusier N, Nussinov R, Wolfson HJ. FireDock: Fast interaction refinement in

molecular docking. Proteins Struct Funct Genet. 2007;69(1):139–159.
DOI: 10.1002/prot.21495

26. Laskowski RA, Swindells MB. LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery. J Chem Inf Model. 2011;51(10):2778–2786.
DOI: 10.1021/ci200227u

27. Bui HH, Sidney J, Dinh K, Southwood S, Newman MJ, Sette A. Predicting population coverage of T-cell epitope-based diagnostics and vaccines. BMC Bioinformatics. 2006;7.
DOI: 10.1186/1471-2105-7-153

28. Gamage DG, Gunaratne A, Periyannan GR, Russell TG. Applicability of Instability Index for *In-vitro* Protein Stability Prediction. Protein Pept Lett. 2019;26(5): 339–347.
DOI: 10.2174/0929866526666190228144219

29. Bunkute E, Cummins C, Crofts FJ, Bunce G, Nabney IT, Flower DR. PIP-DB: the Protein Isoelectric Point database. Bioinformatics. 2014;31(2):295–296.
DOI: 10.1093/bioinformatics/btu637

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sdiarticle4.com/review-history/66974