

Developing cancer vaccine with carcinoembryonic antigen and IGF-1R as immunostimulants using immunoinformatics approach

Louis Odinakaose Ezediuno¹, Michael Asebake Ockiya², Luqman Oluwaseun Awoniyi³, Adeola Oyepeju Sangodare^{4,*}, Kehinde Busuyi David⁵, Faith Owabhel Robert⁶

¹Department of Microbiology, University of Benin, Benin City, Nigeria

²Department of Animal Science, Faculty of Agriculture, Niger Delta University, Amassoma, Nigeria

³Department of Medicine, University of Turku, Turku, Finland

⁴Department of Microbiology, University of Ilorin, Ilorin, Nigeria

⁵Department of Physiology, Ladoke Akintola University of Technology, Ogbomoso, Nigeria

⁶Department of Medical Biochemistry, Rivers State University, Port Harcourt, Nigeria

Purpose: Colorectal cancer (CRC) remains a significant global health burden, necessitating innovative approaches for prevention and treatment. This study proposes a multipeptide vaccine targeting carcinoembryonic antigen (CEA) and insulin-like growth factor-1 receptor (IGF-1R), two prominent biomarkers associated with CRC progression.

Methods: Sequences of CEA and IGF-1R proteins were retrieved from NCBI databank, the sequences were aligned on the MEGA5 tool to identify conserved regions. Immunological and structural predictive analysis which include antigenic potential prediction, cytotoxic T-lymphocytes (CTLs), helper-T lymphocytes (HTLs), B-cell epitopes predictions, and prediction of the vaccine secondary and tertiary structure were performed. The vaccine was evaluated to validate its physiochemical and immunological properties. To determine the binding energy and domain, the tertiary structure of the vaccine was docked to Toll-like receptor 4, and viewed on PyMOL and LigPlot+ tools.

Results: CEA and IGF-1R were revealed to be highly antigenic, and non-allergens demonstrating the capacity to elicit robust immune responses, which include CTLs, HTLs, and B cells activation. The secondary structure revealed a conformation closely resembling native protein, with alpha helices, beta sheets, and coils, indicative of favorable interactions. Tertiary structure prediction predicted five models, model 0 was selected and validated due its highest confidence, and validation revealed that 87.5% of residues were within favored regions, with a z-score of 4.03. Molecular docking predicted strong binding complex with low binding energy.

Conclusion: Based on our analysis, the proposed multipeptide vaccine holds promise as an effective preventive measure against colorectal cancer development.

Keywords: Colorectal neoplasms, Multipeptide vaccine, Carcinoembryonic antigen, Insulin-like growth receptor

INTRODUCTION

The World Health Organization in 2023 reported that colorectal cancer ranks third as the world's most common cancer type. Regardless of existing effective technologies for early detection and treatment aimed at lowering mortality rate, over two million new cases are diagnosed yearly, resulting in about one million deaths annually. This places colorectal cancer as the second leading cause of cancer-associated deaths worldwide [1].

Cancer therapy has evolved tremendously over the years from rudimentary surgeries to advanced, personalized, and target-specific therapies. Formerly, cancer was treated through methods like surgery and radiotherapy. These methods were effective for treat-

Received: Dec 4, 2024 **Revised:** Jan 8, 2025 **Accepted:** Jan 14, 2025

Correspondence to: Louis Odinakaose Ezediuno

Department of Microbiology, University of Benin, 1154, P.M.B, Ugbowo Lagos Rd, Benin City 300103, Nigeria

Tel: +234-8067364550

E-mail: ezediunolouis@gmail.com

*Current affiliation: Department of Science Laboratory Technology, Ladoke Akintola University of Technology, Ogbomoso, Nigeria

© 2025 Korean Society of Surgical Oncology

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

ing localized cancer; they are, however, limited in addressing complications like metastasized cancer, they also often cause harm to the surrounding tissues. By the mid-20th century, advancement in technology and breakthrough in chemotherapy provided more treatment options with better success rates and solutions for the challenges [2]. Today, cancer treatments are increasingly target specific and personalized, these strategies include biomarker testing for cancer, chemotherapy, hormone therapy, photodynamic therapy, immunotherapy (CAR T-cell therapy, cancer vaccines, cytokines, immunomodulators, immune checkpoint inhibitors, monoclonal antibodies), targeted therapy, surgery, stem cell transplant, radiation therapy, hyperthermia [3,4].

Cancer vaccines are a state of the art immunotherapeutic approach developed to prevent and treat cancer. There are two main types of cancer vaccines: preventive and therapeutic cancer vaccines. While preventive cancer vaccine targets infectious agents such as viruses that can lead to cancer, therapeutic vaccines are designed to treat existing cancers. Therapeutic cancer vaccines in general are cancer-associated antigens that are introduced into the body to stimulate immunogenic recognition, attack and destruction of cancer cells by the immune system. In addition to treating fully developed malignancies, cancer vaccines reduce tumor size, block or slow the proliferation of cancer cells, prevent recurrence, and eradicate cancer cells that evade other therapies [5].

Carcinoembryonic antigen (CEA) is an oncofetal antigen that is overexpressed in human adenocarcinomas, especially in gastrointestinal cancer. It is a promising target for gastrointestinal cancer-specific immunotherapy [6,7]. Their antigenic property makes them a good potential target for immunotherapy, this was reported by Larocca and Schlom [8] and Meraviglia-Crivelli et al. [9], who stated that CEA can induce a stronger CD8 T-cell cytotoxic response when animals are immunized with it. In another study, a recombinant vaccinia virus vaccine that encodes complete or internally deleted cDNAs for human CEA induced CEA-specific autoantibodies in seven out of the 32 participants involved in the study. It was reported that the participants previously had no CEA antibodies before vaccination [10].

Insulin-like growth factor-1 receptor (IGF-1R) is another biomarker implicated in the development of several human malignancies [11], including breast, pancreatic, and colorectal cancers [12]. It is reported that IGF-1R is significantly involved in cellular growth, spread, alteration, and apoptotic inhibition [13]. IGF-1R signaling causes tumor cells to be resistant to treatments such as chemotherapy and anti-hormonal therapy.

Many cancer types affecting humans have been found to overexpress IGF-1R, an instance of such observation was reported by Hakam et al. [14], who also wrote in their findings that cases of

pancreatic cancer show an overexpression of IGF-1R, this claim was also reported by Xu et al. [15], who reported an increased level of IGF-1R in blood plasma for every advanced stage of pancreatic cancer and a total absence of same in the surrounding non-cancerous tissues.

Owing to the role of IGF-1R in cancer cell proliferation and growth, resistance to treatment, its overexpression in specific malignant cells and lack of it in surrounding non-malignant cells, it may be suggested that immunotherapeutic approach involving the epitope IGF-1R could be a better anticancer therapy compared to the traditional methods [13].

There is an increasing amount of data supporting the fact that epitope vaccines can induce immunological response against malignant cells by activating helper T cells (Th), cytotoxic T cells, and B cells. As regards these, IGF-1R has become a prominent molecular target for anticancer therapeutic approaches.

Immunological responses are critically important in fighting cancer; adapting this response in order to treat cancer may be a vital approach in combating the recent rise in cancer morbidity and mortality. To achieve this, antigenic epitopes are introduced to the host immune system through the process referred to as vaccination [16].

An epitope with antigenic properties is a unit from the pathogen or in this case the tumor cell that can induce innate, humoral or cellular immune response [16]. Combining antigenic epitopes to develop a vaccine results into a multiepitope vaccine which is one composed of a series of overlapping chain of amino acids with the potential to elicit cytotoxic T-lymphocytes (CTLs), helper T cells (Th), B cells and also induce efficient immune responses against the targeted tumor.

Multiepitope approach for cancer therapy is gaining recognition because compared to other techniques, it is more convenient, time-saving and cost-effective [17]. Based on the above-mentioned factors, this research was focused on designing a multiepitope vaccine developed by combining CEA and IGF-1R epitope to treat and prevent colorectal cancer.

METHODS

Retrieval of sequences

Amino acid sequences for CEA (> AAA51963.1, > AAB59513.1, > AAA51967.1, > AAA62835.1) and IGF-1R (> AAI43722.1, > NP_001278787.1, > NP_000866.1, > XP_047288401.1, > XP_047288400.1, > XP_047288399.1, > XP_047288398.1, > XP_016877626.1, > BAG11657.2, > EAX02222.1, > EAX02220.1, > AAI13611.1, > AAI13613.1) were downloaded from the NCBI database (<https://ncbi.nlm.nih.gov/>).

Determination of conserved regions and the assessment of their antigenic and allergenic score

Conserved region determination for the two protein types (CEA and IGF-1R) was carried out on the MEGA5 tool. Using ANTI-GENpro online tool on the <https://scratch.proteomics.ics.uci.edu/>, AllergenFP v1.0 server (<https://ddg-pharmfac.net/AllergenFP/>) and AllerCatPro v1.7 server (<https://allercatpro.bii.a-star.edu.sg/>), the antigenic and allergenic properties of the selected conserved subsequences were determined.

Prediction of immunogenic epitopes

Stimulation of cellular immunity requires proper interaction between antigens and major histocompatibility complex (MHC) molecules, therefore, to determine the potential epitopes that will properly bind to MHC molecules, the NetCTL (<https://services.healthtech.dtu.dk/>) and IEDB (<http://www.iedb.org/>) web servers were used. These web servers predict both CTL and helper-T lymphocytes (HTL) epitopes, respectively. Humoral immunity stimulating epitopes were determined on the bcpred web server (<http://ailab-projects2.ist.psu.edu/bcpred/predict.html>); this server predicts B-cell binding epitopes.

Construction of primary structure of the vaccine candidate

Primary vaccine construction involves linking CTL, HTL, and B-cell epitopes together by linkers, adjuvants are usually attached to vaccines to improve their antigenicity. The candidate vaccine's primary structure was assembled by linking the predicted CTL, HTL, and B-cell epitopes together with appropriate linkers, 50S RIBOSOMAL PROTEIN L7/L12 (LOCUS RL7_MYCTU) was also attached to the N-terminal of the vaccine as adjuvant to improve the vaccine's immunogenicity [18].

Screening for primary vaccine antigenic, allergenic, toxicity and physicochemical properties

Screening of the candidate vaccine primary construct was performed to determine its antigenicity, allergenicity, and toxicity to prospective hosts. ANTIGENpro server was used to analyze for antigenicity of a proposed candidate vaccine. The allergenic and toxicity potential of a vaccine is also analyzed on the AllergenFP v1.0 server, AllerCatPro v1.7 server, and the ToxIBTL server (<https://github.com/WLYLab/ToxBTL>), respectively; these are re-

ported in Table 1 in the result section of our study. Protparam was used in analyzing the physicochemical properties of the vaccine to predict its biochemical properties.

Secondary structure prediction for candidate vaccine

Proper folding of protein and the formation of their tertiary structure depends significantly on their secondary structure. Hence, this research involved analyzing the secondary structure features such as β -sheet, α -helix and turns of the candidate vaccine using the SOPMA online tool (https://npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?page=/NPSA/npsa_sopma.htm).

Prediction of 3D structure

Proper folding is crucial for the functionality of a polypeptide, predicting the tertiary structure of the candidate vaccine was essential to understanding its potential immune response. To accomplish this, the Swiss Model server available <https://swissmodel.expasy.org> was utilized as outlined in the study of [19].

Epitopes mapping of conformational Linear and discontinuous B cell

B-cell epitopes can be categorized into two types vis linear (continuous) and conformational (discontinuous) epitopes. Linear B-cell epitopes are sequential and account for a small percentage of B cells. While conformational B-cell epitopes consist of fragments of non-sequential epitopes that are exposed to solvent, they make the larger percentage of B cells. Mapping both linear and conformational epitopes is vital because antibody that can recognise linear B-cell epitopes may not necessarily recognise conformational B-cell epitopes. To map B-cell epitope for vaccine constructs, the Ellipro suite available on <http://tools.iedb.org/ellipro> was employed. The suite utilizes a minimum threshold of 0.5 and maximum distance of 6 to map epitopes. Additionally, it utilizes both the MOD-ELLER program and Jmol viewer to evaluate and visualize the three-dimensional (3D) structure of B-cell epitopes within the vaccine construct [20].

Refinement of predicted tertiary structure

Refinement is done by subjecting the vaccine candidate to Galaxy refinement tool GalaxyWEB. Galaxy refine improves the tertiary structure of the vaccine candidate by increasing the number of res-

Table 1. Selected conserved epitopes

Origin of epitope	Protein name	Sequence length	% of conserved region	Antigenic score	AllergenFP	AllerCatPro
IGF-1R conserved epitopes	Insulin-like growth 1	1,337	95.97	0.938057	Nonallergenic	Nonallergenic
CEA conserved epitopes	Carcinoembryonic antigen	695	99.0	0.931601	Nonallergenic	Nonallergenic

Table 2. Predicted cytotoxic T-lymphocytes epitopes

No	Origin of epitope	Epitopes	Comb score	Length	IEDB immunogenicity score	Selected/not-selected
1	Conserved IGF-1R epitopes	YLTVDWSL	0.8494	9	0.9706	Selected
		STVDWSLIL	0.8171	9	0.9691	Selected
		ILDAVSNNY	0.9723	9	2.8337	Selected
		EKTINNEY	0.7992	9	0.9682	Selected
		TTINNEYNY	1.5449	9	0.7815	Selected
		CVACRHYYY	1.9877	9	0.9255	Selected
		SSDSEGFVI	1.1752	9	0.4463	Selected
		QLEGNYSFY	2.3829	9	0.8761	Selected
		LTIKAGKMY	1.3355	9	0.5259	Selected
		RIIITWHRY	0.8094	9	0.9621	Selected
		DLJSFTVYY	0.8402	9	0.9775	Selected
		SLPNGNLSY	1.4413	9	0.9754	Selected
		VTENPKTEV	0.7709	9	0.9758	Selected
		QAEKEEAAY	1.5985	9	0.7213	Selected
		ITDPEELET	0.9770	9	0.1454	Selected
		LILMYEIKY	0.7848	9	0.9176	Selected
		VSRQEYRKY	0.8278	9	0.5791	Selected
		WTDPVFFVY	2.5053	9	0.9402	Selected
		LYASVNPEY	0.8375	9	0.9653	Selected
		YASVNPEYF	0.7985	9	0.7771	Selected
		GQGSFGMVY	0.7752	9	0.9625	Selected
		TRDIYETDY	1.2961	9	0.9760	Selected
		LKDGVFITY	1.4427	9	0.9744	Selected
		TTYSDVWSF	0.8151	9	0.8020	Selected
		YSDVWSFGV	2.3969	9	0.8291	Selected
		LSNEQVLR	1.2156	9	0.6324	Selected
		ELMRMCWQY	0.9689	9	0.9188	Selected
		ASFDERQPY	1.0046	9	0.8960	Selected
2	Conserved CEA epitopes	QQATPGPAY	0.7687	9	0.9468	Selected
		IIQNDTGFY	1.2680	9	0.5266	Selected
		DTGFYTLHV	0.8121	9	0.7876	Selected
		VTRNDTASY	1.2734	9	0.9776	Selected
		RSDSVILNV	1.6723	9	0.9622	Selected
		DSVILNVLY	1.9542	9	0.9541	Selected
		TISPLNTSY	1.9334	9	0.9779	Selected
		YSWFVNGTF	0.7728	9	0.9725	Selected
		ITVNNSGSY	1.8189	9	0.9594	Selected
		TTVTITVY	2.2342	9	1.9329	Selected
		VTRNDVGPY	1.2770	9	0.9473	Selected
		HSDPVILNV	2.1280	9	0.9763	Selected
		PTISPSYTY	1.1070	9	0.9769	Selected
		TISPSYTTY	2.1177	9	0.9683	Selected
		AASNPPAQY	0.9809	9	0.9684	Selected
		ITEKNSGLY	3.5590	9	0.9000	Selected
		VTRNDARAY	0.9685	9	0.4370	Selected
		RSDPVTLDV	1.9754	9	0.9773	Selected
		IISPPDSSY	0.8897	9	0.9737	Selected
		YLSGANLNL	0.7711	9	0.9774	Selected
		SASNPSPQY	1.0145	9	0.9759	Selected
		ITPNNNGTY	2.5351	9	0.9612	Selected

IGF-1R, insulin-like growth factor-1 receptor; CEA, carcinoembryonic antigen.

idues positioned within the favored region, hence optimizing the structure for improved efficiency and to achieve a more favorable conformation. Galaxy refine tool achieved this by molecular dynamics and rearrangement of side-chains of the polypeptides [21].

Model stability and validity evaluation

Validating the refined tertiary structure of the candidate vaccine was done on ProSA server (<https://prosa.services.came.sbg.ac.at/prosa.php>), PROCHEK and ERRAT server (<https://saves.mbi.ucla.edu/>). ProSA server identifies potential error by comparing the predicted structure with experimental structures, thereby providing a z-score of the predicted structure, two graphical representations of the validation are also provided. The PROCHEK and ERRAT servers validate protein tertiary structures by using the Ramachandran plot and statistical evaluation of nonbonding interactions between different atoms by comparing them with databases [22].

Disulfide engineering for vaccine stability

Disulfide bond are covalent bonds interaction that stabilizes molecular structures by confirming specific geometric conformations to the biological 3D structures. Disulfide bonds for our vaccine model was done on the Disulfide by Design 2.0 server (<http://cptweb.cpt.wayne.edu/DbD2/>), the server search for residue pairs that can form disulfide bonds on the bases of their X3 value, energy and B-factor score. Two residue pairs were accepted to form the mutant, the PyMOL visualizing tool was utilized in viewing the binding site [23].

Receptor ligand molecular docking analysis

To determine the interactions between the ligands and receptors of a protein, molecular docking analysis should be carried out. The vaccine candidate which is the ligand in our research was docked against TLR4, a known Toll-like receptor (TLR) reported to be overexpressed in colorectal cancer and is also suggested to be directly correlated with patient's survival [24]. Docking was achieved on the PatchDock online docking tool (<https://bioinfo3d.cs.tau.ac.il/PatchDock/>). Visualization of the bound complex was carried out on PyMOL software, the interacting ligand-receptor complex

between the TLR4 and vaccine residue was viewed with LigPlot+ v.2.2.5 tool.

RESULTS

Screening for CTL, HTL, and B cells from the conserved epitopes

Seventeen 9mers high comb scoring CLTs epitopes were selected from the predicted epitopes from the conserved Chikungunya virus structural epitopes, this are listed on Table 2. Thirty-two 15mers helper T-lymphocytes (HTC) epitopes were also selected from the total predicted HTC epitopes on the bases of their low percentile rank score (Table 3) and 16 B-cell epitopes were chosen from the total B-cell epitopes predicted from the virus's structural protein based on their high score (Table 4). The presence of conformational epitopes in our research (Table 5) indicates the availability of B-cell epitopes since B-cell epitopes which are approximately 90% conformational epitopes. Conformational epitopes are crucial in peptide-based vaccines thus making their availability an advantage for the candidate vaccine.

Construction of primary structure of vaccine candidate

A total of 1,236 amino acid residues comprising the selected CTLs, HTC, and B cells linked together by suitable linkers were used to construct the primary vaccine candidate. Linking an adjuvant to the N-terminal of the candidate vaccine improves its antigenicity. Fig. 1 shows a graphical representation of the primary vaccine structure.

Antigenicity, allergenicity, toxicity, and physicochemical screening of the constructed primary vaccine candidate

The vaccine candidate was predicted to be nontoxic, nonallergenic and have an antigenicity score of 0.873038. The physicochemical assessment of the primary structure revealed that the sequence has a molecular weight of 240493.64, a total number of negative charge residue (Asp+Glu) score of 141, and a score of 193 as the total number of positive charge (Arg+Lys). The sequence has a 30-hour half-life in mammalian reticulocyte, > 20 hours in yeast and > 10

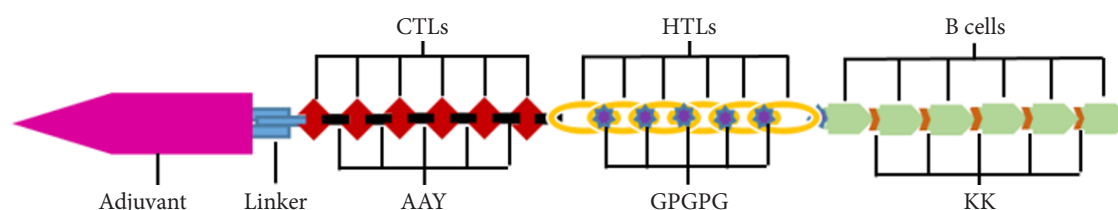


Fig. 1. Graphical representation of the primary vaccine construct. Antigenicity, allergenicity, toxicity, and physicochemical screening of the constructed primary vaccine candidate. CTLs, cytotoxic T-lymphocytes; HTLs, helper-T lymphocytes.

Table 3. Predicted helper-T lymphocytes epitopes

Allele	Start	End	Length	Method	Peptide	Percentile_rank	IC50
H2-IEd	831	845	15	Consensus (simm/nn)	VIMLYVFHRKRNNNSR	0.3	293.3
H2-IEd	832	846	15	Consensus (simm/nn)	IMLYVFHRKRNNNSRL	0.33	258.3
H2-IEd	830	844	15	Consensus (simm/nn)	LVIMLYVFHRKRNNNS	0.49	412.8
H2-IEd	833	847	15	Consensus (simm/nn)	MLYVFHRKRNNNSRLG	0.54	295.4
H2-IAb	696	710	15	Consensus (simm/nn)	ASNFVFARTMPAEGA	0.65	69.6
H2-IAb	695	709	15	Consensus (simm/nn)	SASNFVFARTMPAEG	0.73	74
H2-IAb	697	711	15	Consensus (simm/nn)	SNPVFARTMPAEGAD	0.76	80
H2-IAb	694	708	15	Consensus (simm/nn)	CSASNFVFARTMPAE	0.77	87.7
H2-IEd	829	843	15	Consensus (simm/nn)	GLVIMLYVFHRKRNN	0.8	570.6
H2-IAb	693	707	15	Consensus (simm/nn)	GCSASNFVFARTMPA	0.95	125.8
H2-IAb	977	991	15	Consensus (simm/nn)	MENNPVLAPPSLSKM	1.02	192.2
H2-IAb	976	990	15	Consensus (simm/nn)	EMENNPVLAPPSLSK	1.18	224.6
H2-IEd	607	621	15	Consensus (simm/nn)	NSIFVPRPERKRRDV	1.3	547.9
H2-IEd	608	622	15	Consensus (simm/nn)	SIFVPRPERKRRDVM	1.3	487
H2-IAb	975	989	15	Consensus (simm/nn)	PEMENNPVLAPPSLS	1.34	234.8
H2-IAb	978	992	15	Consensus (simm/nn)	ENNPVLAPPSLSKMI	1.5	195.1
H2-IAb	478	492	15	Consensus (simm/nn)	SEILYRTNASVPSI	1.5	113.5
H2-IAb	266	280	15	Consensus (simm/nn)	YVKIRHSHALVSLF	1.5	147.6
H2-IAb	479	493	15	Consensus (simm/nn)	EILYRTNASVPSIP	1.6	109
H2-IEd	606	620	15	Consensus (simm/nn)	HNSIFVPRPERKRRD	1.65	653.1
H2-IAb	979	993	15	Consensus (simm/nn)	NNPVLAPPSLSKMIQ	1.7	231
H2-IAb	265	279	15	Consensus (simm/nn)	GYVKIRHSHALVSL	1.75	148.2
H2-IEd	609	623	15	Consensus (simm/nn)	IPVPRPERKRRDVMQ	1.8	581.7
H2-IAb	480	494	15	Consensus (simm/nn)	ILYRTNASVPSIPL	1.8	138.6
H2-IAb	779	793	15	Consensus (simm/nn)	NYTARIQATSLSGNG	1.95	106
H2-IEd	834	848	15	Consensus (simm/nn)	LYVFHRKRNNNSRLGN	1.99	407.8
H2-IAb	778	792	15	Consensus (simm/nn)	GNYTARIQATSLSGN	2	109.8
H2-IAb	418	432	15	Consensus (simm/nn)	SPSYTYRPGVNLSL	0.38	80.1
H2-IAb	419	433	15	Consensus (simm/nn)	PSYTYRPGVNLSSL	0.4	89.3
H2-IAb	417	431	15	Consensus (simm/nn)	ISPSYTYRPGVNLS	0.48	104.7
H2-IAb	23	37	15	Consensus (simm/nn)	SLTLFWNPPTAKLT	0.59	65.1
H2-IAb	22	36	15	Consensus (simm/nn)	ASLLTFWNPPTAKL	0.6	65.8
H2-IAb	24	38	15	Consensus (simm/nn)	LLTFWNPPTAKLTI	0.65	73.3
H2-IAb	231	245	15	Consensus (simm/nn)	LNVLGPDAPTISPL	0.97	212.9
H2-IAb	420	434	15	Consensus (simm/nn)	SYTYRPGVNLSLSC	1	105.3
H2-IAb	256	270	15	Consensus (simm/nn)	NLSCHAASNPPAQYS	1.07	216.3
H2-IAb	232	246	15	Consensus (simm/nn)	NVLGPDAPTISPLN	1.07	230.9
H2-IAb	431	445	15	Consensus (simm/nn)	SLSCHAASNPPAQYS	1.07	211.1
H2-IAb	233	247	15	Consensus (simm/nn)	VLYGPDAPTISPLNT	1.07	208.5
H2-IAb	230	244	15	Consensus (simm/nn)	ILNVLYGPDAPTISP	1.17	253.4
H2-IAb	416	430	15	Consensus (simm/nn)	ITSPSYTYRPGVNL	1.22	203.9
H2-IAb	229	243	15	Consensus (simm/nn)	VILNVLYGPDAPTIS	1.22	267.6
H2-IAb	257	271	15	Consensus (simm/nn)	LSCHAASNPPAQYSW	1.25	237.1
H2-IAb	25	39	15	Consensus (simm/nn)	LTFWNPPTAKLTIE	1.36	128.5
H2-IAb	21	35	15	Consensus (simm/nn)	TASLLTFWNPPTAK	1.39	133.2
H2-IAb	430	444	15	Consensus (simm/nn)	LSLSCHAASNPPAQY	1.4	264.1
H2-IAb	255	269	15	Consensus (simm/nn)	LNLSCHAASNPPAQY	1.45	275.3
H2-IAb	429	443	15	Consensus (simm/nn)	NLSLSCHAASNPPAQ	1.45	278.3
H2-IAb	421	435	15	Consensus (simm/nn)	YTYRPGVNLSSLCH	1.5	212.3
H2-IAb	415	429	15	Consensus (simm/nn)	PTISPSYTYRPGVN	1.56	277.1
H2-IAb	254	268	15	Consensus (simm/nn)	NLNLSCHAASNPPAQ	1.59	299
H2-IAb	584	598	15	Consensus (simm/nn)	LDVLYGPDPTIISPP	1.6	200.1
H2-IAb	583	597	15	Consensus (simm/nn)	TLDVLYGPDPTIISP	1.65	225.2
H2-IAb	258	272	15	Consensus (simm/nn)	SCHAASNPPAQYSWF	1.75	336.7
H2-IAb	433	447	15	Consensus (simm/nn)	SCHAASNPPAQYSWL	1.75	332
H2-IAb	582	596	15	Consensus (simm/nn)	VTLDVLYGPDPTIIS	1.75	243.3
H2-IAb	20	34	15	Consensus (simm/nn)	LTASLLTFWNPPTTA	1.95	214.6
H2-IAb	585	599	15	Consensus (simm/nn)	DVLYGPDPTIISPPD	2	223.8

hours *Escherichia coli* *in vivo*. The aliphatic index was 66.73, instability index was 40.35, which classify the protein as unstable and grand average of hydropathicity is -0.354 .

Candidate vaccine secondary structure prediction

SOPMA online prediction tool provides both the secondary structure of the candidate vaccine and the alpha helix, extended strand, random coil, and beta turns for the vaccine construct as 15.80%, 27.05%, 50.82%, and 6.33%, respectively.

Prediction of 3D structure

Swiss Model 3D structure predicting server predicted 50 models through the homology modeling technique, model 1 (Fig. 2) was

Table 4. Predicted B cells

Rank	Sequence	Start position	Score
Conserved IGF-1R epitopes			
1	GVVKDEPETRVAIKTV	899	0.95
1	GGLVIMLYVFHRKRNN	828	0.95
2	NGSWTDPVFFYVQAKT	792	0.94
2	CRHYYYAGVCVPACPP	130	0.94
3	TCGKRACTENNECCHP	97	0.93
3	TARIQATSLSGNGSWT	781	0.93
4	HNSIFVPRPERKRRDV	606	0.91
4	DGTIDIEEV TENPKTE	554	0.91
4	YIVGNKPPKECGDLCP	47	0.91
4	SFTVYYKEAPFKNVTE	401	0.91
4	NGSQSMYCIPCEGPCP	193	0.91
4	TRDIYETDYRKGGKG	1036	0.91
5	VSRQYRKYGGAALNR	759	0.90
5	PVTWEPRPENSIFLKW	716	0.90
5	ACPPNTYRFEGWRCVD	142	0.90
5	RCMCWQYNPKMRPSFLE	1125	0.90
Conserved CEA epitopes			
1	GREIYPNASLLIQNI	97	0.96
2	PETQDATYLLWWNNQS	169	0.93
3	ISPPDSSYLSGANLNL	595	0.91
3	TVYAEPPKPFITSNNS	317	0.91
4	NRQIIIGYVIGTQQATP	76	0.90

IGF-1R, insulin-like growth factor-1 receptor; CEA, carcinoembryonic antigen.

selected on the bases of the resolution of the structure 1.55 Å, quaternary structure quality estimation of 0.00 which should range between 0 and 1, which, for a model constructed using a given alignment and template, is a number between 0 and 1 that represents the expected accuracy of the inter-chine contact with sequence similarity of 0.26, and coverage of 0.16.

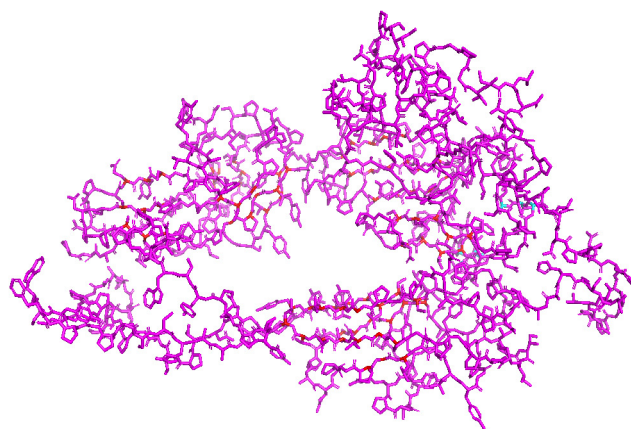


Fig. 2. Tertiary structure of candidate vaccine as predicted by Swiss Model.

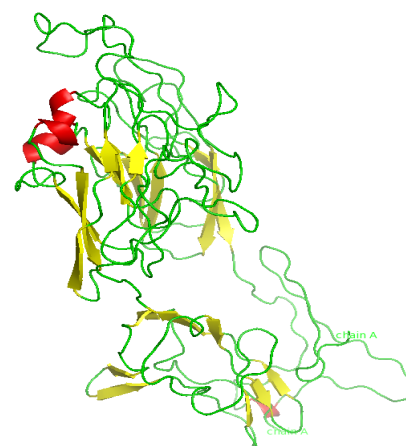


Fig. 3. Final model of the tertiary structure of the multiepitope vaccine after refinement.

Table 5. List of discontinuous B-cell conformational epitopes

No.	Residues	No. of residues	Score
1	TPIISPGPGPGSCHAASNPPAQYSWFGPGPGSCHAASPPA	40	0.837
2	PDAPTISPLPPGVNLSLSGPGPGNLS2CHAASNPPAQYGPDAPTISPLNAASNPPAAQYSWLGPGPVGT	69	0.683
3	IPLNTGPGPGILNVLYGPDAPTISPGPGPGTISP NLGPVILP	42	0.668
4	IACHAAANPPAQYGPGPGLNLSLSCHAASNPPGPTISPSYTY YRPGVNGPGPGPLDVLGPDTPISPPGPG	71	0.653
5	PPAQGP	6	0.587
6	NVLYGP	6	0.568
7	GPGLKGP GPGLS	12	0.560

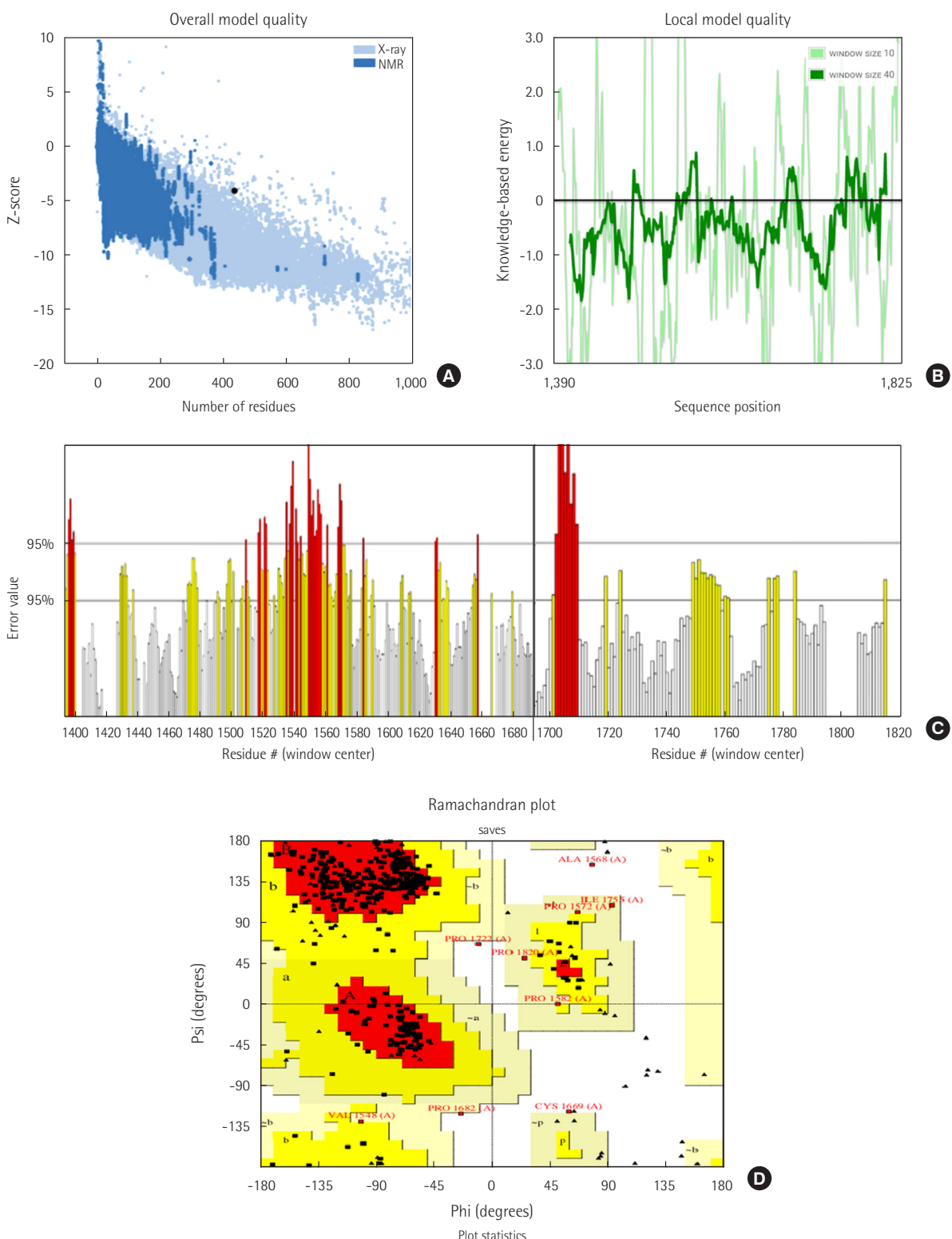


Fig. 4. (A) The figure shows a Z-score plot of the tertiary structure of the candidate vaccine, the vaccine is located in the X-ray crystallography region. (B) The local model quality plot revealed the quality of the candidate vaccine based on the position and energy function of the vaccine amino acids sequences. (C) The ERRAT plot reveals normal and faulty regions of the vaccine peptide sequence, these are seen as white, yellow, and red bars, respectively. (D) The Ramachandran plot indicating the percentages of the amino acid in the favored, allowed and disallowed regions as 87.5%, 10.9%, and 0.8%, respectively.

Refinement of predicted tertiary structure

The vaccine tertiary structure refinement provided five models, the model 1 (Fig. 3) was selected on the criteria as listed in Table 6.

Model stability and validation evaluation

The result obtained for 3D structure validation on ProSA calculated a z-score of -4.03 while ERRAT predicted an overall quality score of 63.9037%, and the graph of the vaccine local model quality revealed some potential low-quality regions within the sequence. VERIFY, a 3D validation tool on PROCHECK, predicted an overall G-factor score of -0.26, the Ramachandran plot analysis predicted an 87.5% favored region, 10.9% allowed region, and a 0.8% disallowed region (Fig. 4).

Vaccine disulfide engineering

To stabilize the structure of the candidate vaccine, disulfide engineering was performed. Forty-one pairs of amino acid residues were predicted to be useable for disulfide engineering, evaluating the residue pairs on the threshold of energy, Chi3 value and high B-factor, only two pairs of amino acid residues were finalized (GLY1561-GLY1701, TYR1769-ASN1810), this was carried out on Disulfide by Design v2.0 server (<http://cptweb.cpt.wayne.edu/DbD2/index.php>) (Fig. 5).

DISCUSSION

In this study, we sought to design novel multiepitope vaccine biomolecule that targets colorectal cancer using bioinformatic and

immunoformatic approaches to linking CEA and IGF-1R epitopes. The prime rationale in this approach is based on CEA and IGF-1R roles in resistance to targeted therapies, involvement in growth of malignancy and their oncological biomarker status [25,26]. Multiepitope vaccines as a therapeutic tool is becoming progressively accepted because of their efficiency and nontoxic property when paralleled to other treatment modalities [27].

To achieve this, CEA and IGF-1R sequences were selected and retrieved from the NCBI database. The selection was based on their antigenic potential which was predicted on the Scratch database. To develop the vaccine, conserve regions of the retrieved CEA and IGF-1R sequences were screened for and selected to ensure that the vaccine would be active for large variant of the cancer type and would still be active even if mutation should occur [28,29]. Analysis of the conserved sequence for antigenicity revealed that the sequences maintained the antigenic criterion of selection as the original CEA and IGF-1R sequences, hence possessing the potential of an immune stimulant.

To determine their immune-stimulating properties the conserved sequences were screened for CTLs, HTLs, and B cells inducing epitopes as a vaccine should be able of inducing these immune components to initiate humoral and cellular immune response. Epitopes that can initiate the production of these immune components were linked together using suitable linkers, adjuvant was also linked with the epitopes at the N-terminal to enhance the antigenic potential [18]. The complex now formed makes up the *in silico* primary vaccine structure.

It was also necessary to subject the vaccine's primary structure to antigenic and allergenic screen as was carried out using the Scratch online tool. Screening result predicted that the vaccine's primary structure preserved the antigenic potential while remaining nonallergenic. Physicochemical properties prediction also showed that the vaccine properties fell within the accepted threshold [30].

Vaccine structural stability is essential as the vaccine should be able to withstand biological, biochemical and metabolic pressure of the host system, we evaluate the vaccine structural stability by predicting the secondary structure of the vaccine, result obtained showed that the structure fall within accepted threshold, this was on the premise of the alpha helix and beta sheet percentages.

Table 6. Structural Information for refined tertiary structure of the multiepitope candidate vaccine

Model	GDT-HA	RMSD	MolProbity	Clash score	Poor rotamers	Rama favored
Initial	1.0000	0.000	2.361	13.0	0.6	79.7
Model 1	0.9484	0.415	2.348	17.1	0.9	87.3

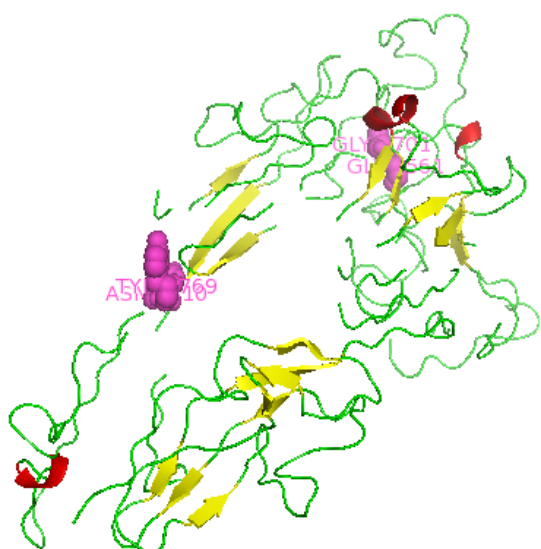


Fig. 5. Protein stability by disulfide engineering represented in magenta.

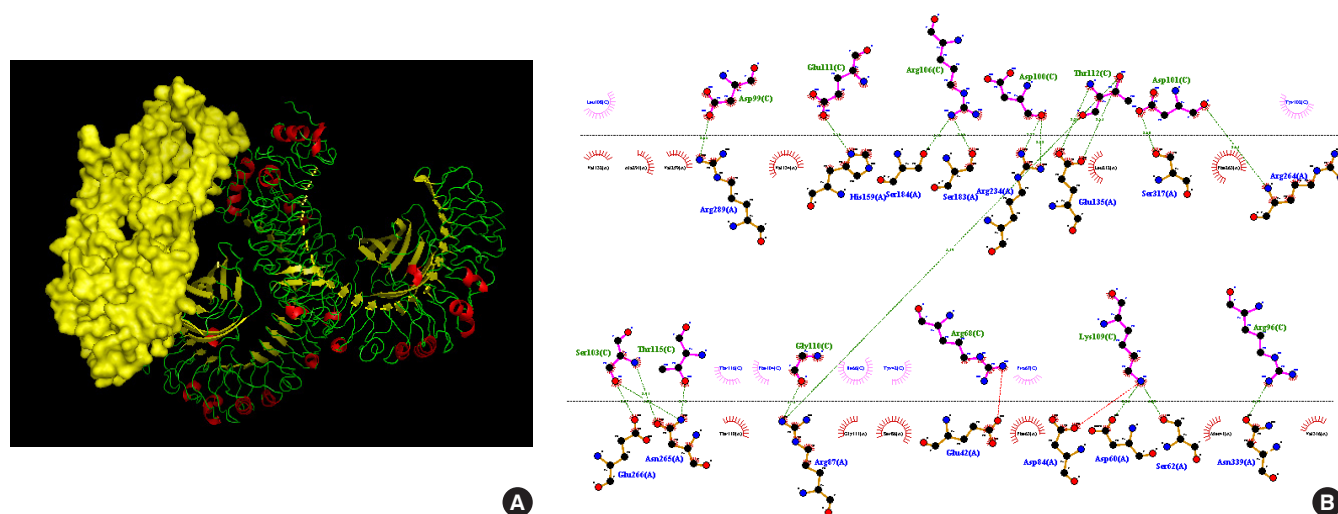


Fig. 6. (A) Model 1 was selected based on the Docking score (-317.92). This was the prediction with the highest negative value indicating its higher chances of docking. Other parameters include Confidence score of 0.9664 and Ligand RMSD (\AA) of 139.72. (B) Interaction between Toll-like receptor 4 (TLR4) and the vaccine candidate viewed using LigPlot+ v2.2.5 app, TLR4 amino acid residues are indicated in blue while the candidate vaccine amino acid residues are indicated in green.

The tertiary structure of a polypeptide plays a major role in its function. We predicted the vaccine tertiary structure by homology modeling. This method predicts the tertiary structure of an unknown polypeptide by comparing it with known structures on the database based on the similarity of their sequences.

In order to validate the predicted tertiary structure, the PDB file was subjected to statistical analysis, a z-score of -4.03 , and a Ramachandran plot with 87.5% in the most favored region was obtained. This result is satisfactory as it indicates a peptide that conforms to native protein.

Docking the vaccine peptide with the TLR4 molecule was done to evaluate the vaccine's ability to bind to TLR4. Binding an antigen with a TLR4 activates inflammatory response and innate immune system. The vaccine polypeptide in this study docked with TLR4 with a binding energy of $-317.92 \text{ kcal mol}^{-1}$. The low binding energy indicates a high binding affinity with human TLR4 (Fig. 6).

In conclusion, using an immunoinformatics approach, we identified epitopes that have potential as a multiepitope-based subunit vaccine against cancer. CTL, HTL, and B-cell lymphomas were screened and selected for vaccine design. Finally, the vaccine constructs consist of CTL, HTL, and B-cell epitopes, an adjuvant to elicit strong immune responses and fusion protein linkers. Conservancy analysis of the promiscuous epitopes was used to select epitopes with robust immune response against large variants or mutants of cancer. Molecular docking revealed stronger binding affinity with human TLR4. Our computational analysis suggests

that the constructed vaccine is highly immunogenic, safe, and stable in nature.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

FUNDING

None.

ACKNOWLEDGMENTS

The author wishes to Mr. Solomon Alile of the University of Benin for his technical support and advice.

ORCID

Louis Odinakaose Ezediuno

<https://orcid.org/0000-0002-9491-529X>

Michael Asebake Ockiya <https://orcid.org/0000-0002-6799-2679>

Luqman Oluwaseun Awoniyi

<https://orcid.org/0000-0001-7809-3268>

Adeola Oyepeju Sangodare

<https://orcid.org/0009-0004-3716-2238>

Kehinde Busuyi David

<https://orcid.org/0000-0001-8476-3408>

REFERENCES

- World Health Organization (WHO). Colorectal cancer [Internet]. WHO; c2023 [cited 2024 Dec 1]. Available from: <https://www.who.int/news-room/fact-sheets/detail/colorectal-cancer>
- Abbas Z, Rehman S. An overview of cancer treatment modalities. *Neoplasia* 2018;1:139-57.
- Marshall HT, Djamgoz MB. Immuno-oncology: emerging targets and combination therapies. *Front Oncol* 2018;8:315.
- Wolf P, Alzubi J, Gratzke C, Cathomen T. The potential of CAR T cell therapy for prostate cancer. *Nat Rev Urol* 2021;18:556-71.
- Aly HA. Cancer therapy and vaccination. *J Immunol Methods* 2012;382:1-23.
- Loisel-Meyer S, Felizardo T, Mariotti J, Mossoba ME, Foley JE, Kammerer R, et al. Potent induction of B- and T-cell immunity against human carcinoembryonic antigen-expressing tumors in human carcinoembryonic antigen transgenic mice mediated by direct lentivector injection. *Mol Cancer Ther* 2009;8:692-702.
- Bos R, van Duikeren S, Morreau H, Franken K, Schumacher TN, Haanen JB, et al. Balancing between antitumor efficacy and auto-immune pathology in T-cell-mediated targeting of carcinoembryonic antigen. *Cancer Res* 2008;68:8446-55.
- Larocca C, Schlom J. Viral vector-based therapeutic cancer vaccines. *Cancer J* 2011;17:359-71.
- Meraviglia-Crivelli D, Zheleva A, Barainka M, Moreno B, Villanueva H, Pastor F. Therapeutic strategies to enhance tumor antigenicity: making the tumor detectable by the immune system. *Bio-medicines* 2022;10:1842.
- Conry RM, Allen KO, Lee S, Moore SE, Shaw DR, LoBuglio AF. Human autoantibodies to carcinoembryonic antigen (CEA) induced by a vaccinia-CEA vaccine. *Clin Cancer Res* 2000;6:34-41.
- Yakar S, Leroith D, Brodt P. The role of the growth hormone/insulin-like growth factor axis in tumor growth and progression: lessons from animal models. *Cytokine Growth Factor Rev* 2005;16:407-20.
- Singh P, Alex JM, Bast F. Insulin receptor (IR) and insulin-like growth factor receptor 1 (IGF-1R) signaling systems: novel treatment strategies for cancer. *Med Oncol* 2014;31:805.
- Mahdavi M, Moreau V, Kheirollahi M. Identification of B and T cell epitope based peptide vaccine from IGF-1 receptor in breast cancer. *J Mol Graph Model* 2017;75:316-21.
- Hakam A, Fang Q, Karl R, Coppola D. Coexpression of IGF-1R and c-Src proteins in human pancreatic ductal adenocarcinoma. *Dig Dis Sci* 2003;48:1972-8.
- Xu JW, Wang TX, You L, Zheng LF, Shu H, Zhang TP, et al. Insulin-like growth factor 1 receptor (IGF-1R) as a target of miR-497 and plasma IGF-1R levels associated with TNM stage of pancreatic cancer. *PLoS One* 2014;9:e92847.
- Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. *Cell Mol Immunol* 2018;15:182-4.
- Aziz S, Almajhdi FN, Waqas M, Ullah I, Salim MA, Khan NA, et al. Contriving multi-epitope vaccine ensemble for monkeypox disease using an immunoinformatics approach. *Front Immunol* 2022;13:1004804.
- Ezediuno LO, Onile OS, Oladipo EK, Majolagbe ON, Jimah EM, Senbadejo TY. Designing multi-epitope subunit vaccine for ocular trachoma infection using *Chlamydia trachomatis* polymorphic membrane proteins G. *Inform Med Unlocked* 2021;26:100764.
- Shahsavani N, Sheikhha MH, Yousefi H, Sefid F. In silico homology modeling and epitope prediction of NadA as a potential vaccine candidate in *Neisseria meningitidis*. *Int J Mol Cell Med* 2018;7:53-68.
- Krishnamoorthy HR, Karuppasamy R. Design and in silico validation of a novel MZF-1-based multi-epitope vaccine to combat metastatic triple negative breast cancer. *Vaccines (Basel)* 2023;11:577.
- Samad A, Ahammad F, Nain Z, Alam R, Imon RR, Hasan M, et al. Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach. *J Biomol Struct Dyn* 2022;40:14-30.
- Malekshahi ZV, Rajabibazl M, Ebrahimizadeh W, Amani J, Negahdari B. Designing a chimeric vaccine against colorectal cancer. *Int J Cancer Manag* 2017;10:e7743.
- Shiratsuchi I, Akagi Y, Kawahara A, Kinugasa T, Romeo K, Yoshida T, et al. Expression of IGF-1 and IGF-1R and their relation to clinicopathological factors in colorectal cancer. *Anticancer Res* 2011;31:2541-5.
- Craig DB, Dombkowski AA. Disulfide by Design 2.0: a web-based tool for disulfide engineering in proteins. *BMC Bioinformatics* 2013;14:346.
- Zhu HQ, Wang DY, Xu LS, Chen JL, Chu EW, Zhou CJ. Diagnostic value of an enhanced MRI combined with serum CEA, CA19-9, CA125 and CA72-4 in the liver metastasis of colorectal cancer. *World J Surg Oncol* 2022;20:401.
- Shahab M, Hayat C, Sikandar R, Zheng G, Akter S. In silico designing of a multi-epitope vaccine against *Burkholderia pseudomallei*: reverse vaccinology and immunoinformatics. *J Genet Eng Biotechnol* 2022;20:100.
- Hu X, Fatima S, Chen M, Xu K, Huang C, Gong RH, et al. Toll-like receptor 4 is a master regulator for colorectal cancer growth under high-fat diet by programming cancer metabolism. *Cell Death Dis* 2021;12:791.
- Mohammadzadeh Hosseini Moghri SA, Ranjbar M, Hassannia H,

- Khakdan F. In silico analysis of the conserved surface-exposed epitopes to design novel multiepitope peptide vaccine for all variants of the SARS-CoV-2. *J Biomol Struct Dyn* 2023;41:7603-15.
29. Ul Rehman Z, Suleman M, Ashraf K, Ali S, Rahman S, Rashid MI. Immuno-bioinformatic approach for designing of multi-epitope merozoite surface antigen of *Babesia bigemina* and evaluation of its immunogenicity in inoculated calves. *Pol J Vet Sci* 2023;26:409-18.
30. Khalid K, Irum S, Ullah SR, Andleeb S. In-silico vaccine design based on a novel vaccine candidate against infections caused by *Acinetobacter baumannii*. *Int J Pept Res Ther* 2022;28:16.