Cell J
Cell Journal (Yakhteh)
Homepage: https://www.celljournal.org/
Review Article
Vol 26, No 7, July 2024, Pages: 403-426

A Comparative Analysis of Computational Strategies in Multi-Epitope Vaccine Design Against Human Papillomavirus and Cervical Cancer

Ali Najafi, Ph.D.[1], Mohammad Hossein Ataee, Ph.D.[1], Mahdieh Farzanehpour, Ph.D.[2], Hadi Esmaeili Gouvarchin Ghaleh, Ph.D.[2*]

1. Molecular Biology Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran
2. Applied Virology Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran

Abstract

Given the critical role of human papillomavirus (HPV) in the cause of cervical cancer and other malignancies, there is a need for innovative approaches to preventing this infection. It has been shown that immunoinformatics is an important strategy in computational vaccinology. It is used to design new multi-epitope vaccines against different types of HPV and subsequent cervical cancer. This paper reviews the scope of the entire computational pipeline of HPV vaccine design, starting from data analysis at the genomic and proteomic levels and continuing to epitope predictions of the innate and adaptive immune systems. The search strategy was based on investigating original articles published in “Google Scholar" and “PubMed" from 2015 to 2023-2024. The terms "Immunoinformatics", "Bioinformatics", "Human papillomavirus (HPV)", "Vaccine design", "In silico vaccine design", "Multi-epitope vaccine design", "Vaccinology" and "HPV vaccine" were used to for this purpose. We discussed various essential tools involved in the computational design of the vaccine process, e.g., sequence analysis, epitope prediction, conservancy analysis, tertiary structure modeling, refinement, molecular docking, molecular dynamics (MD) simulation, and in silico cloning. This review article describes immunoinformatics methods that facilitate the design of a multi-epitope vaccine against HPV. However, this pipeline can also be used to design novel chimeric vaccines for other pathogens.

Keywords: Human Papillomavirus, Immunoinformatics, Vaccinology

Citation: Najafi A, Ataee MH, Farzanehpour M, Esmaeili Gouvarchin Ghaleh H. A comparative analysis of computational strategies in multi-epitope vaccine design against human papillomavirus and cervical cancer. Cell J. 2024; 26(7): 403-426. doi: 10.22074/cellj.2024.2028622.1568

Introduction

Human papillomavirus (HPV) infection is still one of the most prominent public health problems, mainly due to its etiological role in many malignancies, primarily cervical cancer. The fact that HPV infections are invasive and have oncogenic potential speaks toward the high necessity for innovations and up to date strategies in prevention. Some effective vaccines against HPV, e.g., Gardasil, have been shown to play a significant protective role against some high-risk strains of HPV. Computational vaccinology can bring more coverage of different strains through finding the conserved identical epitopes between different strains. Also, recent advances in computational vaccinology bring great prospects toward overcoming the limitations that had traditionally been related to the formulation of vaccines. Immunoinformatics, which is defined as the application of bioinformatics techniques to understand immune system function and to design better vaccines and immunotherapies, have come to be a key technology for the design and, hence, evaluation and streamlining of multi-epitope HPV vaccines. Such a paradigm shift changed the way the immunogenic landscape of HPV was viewed and enabled us to come up with vaccines that would harness broad-spectrum efficacy across several genotypes of HPV, thus yielding expanded protection in the prophylaxis against HPV-related neoplasms (1, 2).

The development of Immunoinformatics has rather revolutionized the field of vaccine development; it gives scientists the ability to unravel complex immunological networks present in HPV infections. These computational tools and databases not only help in identifying, characterizing, and analyzing the conserved epitopes across strains of HPV but will also make an accurate prediction of the antigenic determinants to accurately model the vaccine constructs based on the in-silico evaluation of the immunogenicity and safety profiles. Together, these have provided strong evidence for the use of computational vaccinology in the second generation of HPV vaccine design, improved potency, cross-protective breadth, and rapid product development timelines (3, 4).

The current study aims to explain the complicated mathematical methods behind creating a multi-epitope

Received: 09/May/2024, Revised: 07/July/2024, Accepted: 27/July/2024
*Corresponding Address: P.O.Box: 1435915371, Applied Virology Research Center, Biomedicine Technologies Institute, Baqiyatallah University of Medical Sciences, Tehran, Iran
Email: esmaeili.h@bmsu.ac.ir

Royan Institute
Cell Journal (Yakhteh)

vaccine for HPV, list the most important advances and effects of recent research, and sketch out the future of HPV vaccine research. A detailed review of the whole computational pipeline, from epitope prediction to conservancy analysis and ending with molecular modeling and in-silico immunogenicity assessment, is beyond the scope of the article. This contribution would aim to put into perspective the current state of the question of computational vaccinology's contribution to the fight against HPV and its subsequent cancers.

Search strategy

We reviewed original articles published in "Google Scholar" and "PubMed" from 2015 to 2023-2024. For this purpose, the keywords "Immunoinformatics", "Bioinformatics", "Human papillomavirus (HPV)", "Vaccine design", "In silico vaccine design", "Multi-epitope vaccine design", "Vaccinology" and "HPV vaccine" were used.

Human papillomavirus background

HPV is a double-stranded DNA virus that belongs to the papillomavirus family. Approximately 200 HPVs have been identified that differ in their genomic sequence. Based on clinical manifestations, HPVs are classified into two main high-risk and low-risk groups. High-risk HPVs are responsible for malignancies, especially cervical cancer, and genital warts, while low-risk HPVs are responsible for benign lesions. HPV 16, 18, 31, 33, 35, and 45 have been isolated from about 90% of cervical cancer patients, so most studies have focused on the mentioned HPVs (2). The HPV genome contains early (E1, E2, E4, E5, E6, E7) and late (L1, L2) proteins (Fig.1). Early proteins participate in gene regulation, replication, and viral pathogenesis, while late proteins are involved in virus capsid assembly and transduction of HPV to host cells (Fig.2). Also, E6 and E7 oncoproteins have an undeniable role in the development of cervical cancer. For this reason, L1, E6 and E7 were the most preferred antigens for vaccine design. The function of HPV proteins is summarized in Figure S1 (See Supplementary Online Information at www.celljournal.org). The tumor protein p53 (p53) has various functions, including cell arrest, apoptosis induction, and antitumor activities. Also, E6-associated protein (E6AP) is an E3 ubiquitin ligase facilitating the degradation of intracellular proteins. Through the targeting of E6AP protein by E6, the ubiquitin from E6AP are transferred to p53 and cause its degradation by cytoplasmic proteasomes. Furthermore, through binding of E7 to retinoblastoma protein (Rb)-E2F tumor suppressor complexes, Rb protein and E2F are dissociated, and E2-F-independent activities lead to unabated cell growth. Therefore, due to the importance of E6 and E7 in HPV and cervical cancer pathogenesis, most studies focused on these two to design an efficient vaccine against HPV and subsequent cervical cancer (5).

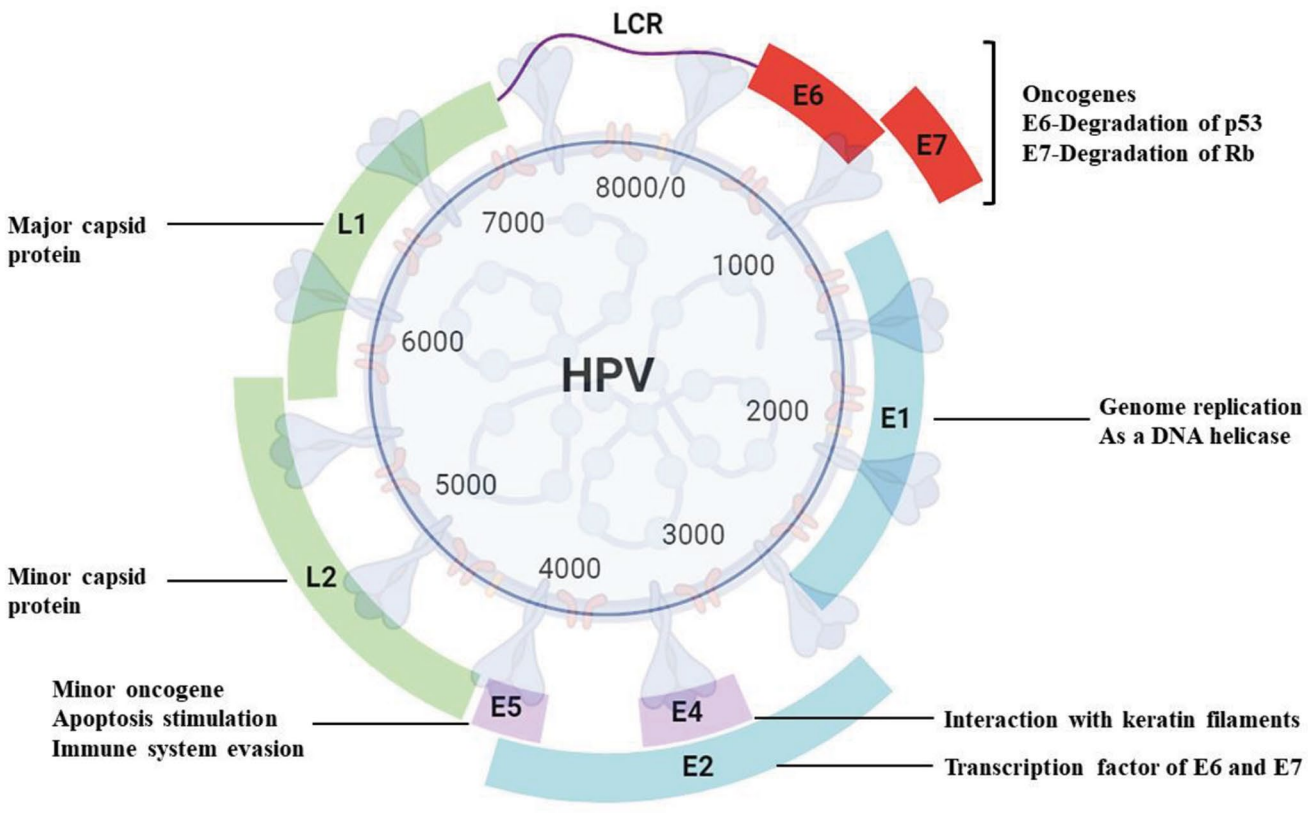

Fig.1: Early (E1, E2, E4, E5, E6, and E7) and late (L1 and L2) proteins of human papillomavirus (HPV). E6 and E7 are major oncoproteins. L1 and L2 are involved in capsid assembly.

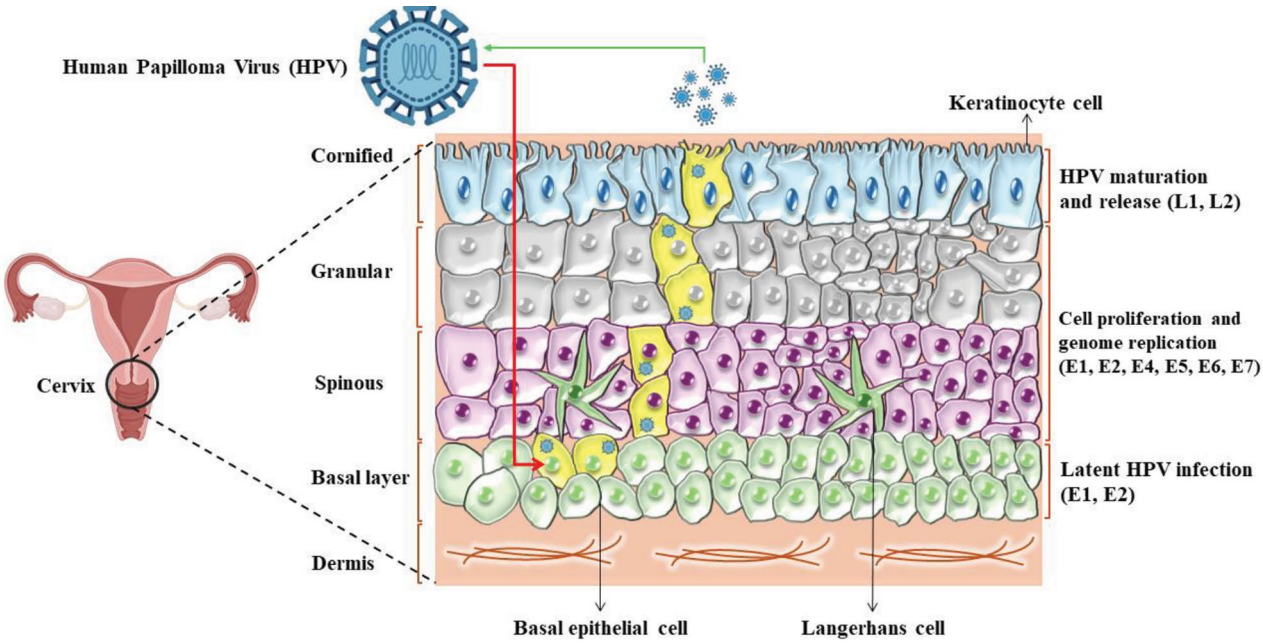

Fig.2: Pathogenesis of HPV infection. In the first step, the proliferation of HPV begins from basal epithelial cells. Up to the cornified layer, the virus increases the rate of proliferation. The keratinocytes cells are the last host of the virus which release assembled virions to initiate another infection cycle.

Computational pipeline for human papillomavirus vaccine design

The genomic and proteomic data of the relevant HPV strains form the basis of the computation pipeline in the development of HPV vaccines (Table 1). These computational tools help in locating the conserved part within the HPV genome that could be a promising target for a broad-spectrum vaccine. Next in the pipeline is in the center of predicting B-cell and T-cell epitopes using algorithms such as NetMHCpan and BepiPred. This utility analyzed the ability of peptide sequences in conserved epitopes of high-risk HPVs to induce immune responses, which in turn would give broad coverage by the vaccine. Immune epitope database (IEDB) can do this easily through epitope analysis tools that enable comparing the experimental epitopes against epitopes predicted by the researcher, hence helping in further narrowing the choices of vaccine candidates (6, 7).

Next, the conservancy and homology of the predicted peptide is checked. Some common tools, e.g., basic local alignment search tool for proteins (BLASTp) and Clustal W are used for the conservancy and homology analysis to find the highly conserved sequences in target HPVs. Meanwhile, we admit that low homology with human proteins minimizes the risk of strain-specific immunity, thereby avoiding autoimmunity (8).

Following the identification of the putative epitopes, the next step involves three-dimensional (3D) structure modeling and the complex formation of epitopes with immune receptors. The available commercial tools, such as I-TASSER for protein structure prediction and some free web-based molecular docking software (e.g., AutoDock Vina), aid in the physical interaction among epitopes and epitope- major histocompatibility complex (MHC) molecules, which are paramount to the evaluation of the immunogenic potential of vaccine constructs (9).

In-silico tools are used to predict the immunogenicity and safety profile of the designed vaccine before experimental validation. This is proved in this phase, where antigenicity, allergenicity, and toxicity were all predicted by the VaxiJen, AllerTOP, and ToxinPred tools, respectively, thus placing the candidate vaccine at a likely safe and immunogenic level (10).

Next, molecular dynamics (MD) simulations will provide insight into the behavior of the vaccine constructs and their stability within the body's given conditions. These tools, such as GROMACS, bestow flexibility, stability, and overall behavior information on the vaccine construct in the simulated biological environment, which is quite helpful during its refinement before experimental trials. Finally, the in-silico optimization of cloning and expression should be considered to achieve a high yield of expressed protein. In the next sections, we will describe all the sections mentioned above (Table 2, Fig.3).

Table 1: Human papillomavirus (HPV) antigens enrolled in vaccine development

HPVs	Antigen	Role in HPV	Importance for vaccine development	Degree of immunogenicity
HPV-6, 16, 18, 45 and 52	L1	Major capsid protein	Major protein of outer shell of the virus, required for virus entry into host cells	High
HPV-16, 18, 52 and 58	L2	Minor capsid protein	Essential for viral genome packaging and entry into host cells	Moderate
HPV-16	E1	Replication protein	Enrolled in viral DNA replication as a helicase	Low
HPV-18	E2	Regulatory protein	The transcription regulator of viral genes	Moderate
HPV-18	E4	Assembly and release	Entrapment of cytokeratin network required for virus assembly and release	Low
HPV-16, 18, 31 and 45	E5	Membrane protein and minor oncoprotein	Regulating the host cell's signaling pathways for enhancing cell transformation	Moderate
HPV-16, 18, 33 and 58	E6	Major oncoprotein	Enhancing the degradation of p53 leading to carcinogenesis	High
HPV-16, 18, 31, 33, 45 and 58	E7	Major oncoprotein	Enhancing the degradation of retinoblastoma protein (Rb) leading to carcinogenesis	High

Table 2: Servers and databases applicable in computational vaccine design for HPV

Application	Server name	Function	Advantages	Disadvantages	Ref
Sequence, conservancy, and homology analysis	National center for biotechnology information (NCBI)	Access to different data in terms of DNA, RNA, and protein sequences, through databases such as GenBank and PubMed.	Containing various comprehensive databases with user-friendly features.	Overwhelming for new users due to the high volume of data and tools available.	(8)
	Universal protein resource (UniProt)	A comprehensive resource for protein sequence and annotation data.	Containing functional information about proteins.	Overwhelming for new users due to the high volume of data.	(11)
	Papillomavirus episteme (PaVE)	A data resource for analyzing HPV sequences, offering tools for comparison, annotation, and phylogenetic studies.	Containing the detailed data, offering different tools for sequence analysis.	Limited the broader virological application due to the focus on HPV. The platform may lag in updates.	(12)
Epitope and cytokine prediction	NetMHC	Predicts the binding capacity of peptides to MHC, critical for T-cell recognition.	Supporting various MHC alleles.	Limited to the prediction of MHC-peptide binding.	(13)
	NetMHCpan	An extensive database for prediction of peptide-MHC binding.	Supporting various MHC alleles, vital for vaccine design for broad population coverage.	Due to the vastness of the data, it needs accurate interpretation for effective vaccine design.	(6)
	BepiPred	Uses a sequence-based prediction algorithm to identify linear B-cell epitopes which helps in the design of vaccines triggering humoral immunity.	Containing user-friendly features.	The accuracy of prediction of epitope-antibody can sometimes be faced with challenges due to the complexity.	(7)

Table 2: Continued

Application	Server name	Function	Advantages	Disadvantages	Ref
	ABCPred	Uses artificial neural networks in predicting B-cell epitopes.	Increasing the prediction accuracy through integration with machine learning.	Depending on validation datasets for optimal performance.	(14)
	DiscoTope	Assists the prediction of discontinuous B-cell epitopes based on structural information.	Eliminating the challenging aspects of conformational epitope prediction.	Requiring an accurate, well-characterized 3D structure of proteins.	(15)
	ElliPro	Predicts both linear and conformational B-cell epitopes.	Providing a more comparative overview on potential B-cell epitopes.	Dependent on the availability of the characterized 3D protein structures.	(16)
	IFNepitope	Predicts epitopes capable of eliciting IFN-gamma and supports the design of cellular immune response stimulator vaccines.	Focusing on the cellular immunity associated with therapeutic vaccines.	Limited to predicting epitopes associated with IFN-gamma, not other cytokines.	(17)
	Immune Epitope Database (IEDB)	Analysis of the epitopes for a detailed understanding of immune response.	Provides a rich dataset for epitope-based vaccine research.	Focused on curated experimental data, which may limit the real-time discovery of novel epitopes.	(18)
Conservancy and homology analysis	CLC sequence viewer	A free, useful software package for viewing and analysis of DNA, RNA, and protein sequences.	Simple user interface and quick analyses, most appropriate for educational purposes. Integration of more advanced analysis can be performed using other CLC bioinformatics tools.	Much less functional than the other very specialized tools, such as BioEdit or Phylogeny.fr.	(19)
	Clustal omega (Clustal O)	An advanced software for handling high-quality alignments of DNA, RNA, and protein sequences.	Better speed and improved accuracy compared to previous versions, e.g., Clustal W. Suitable for large-scale sequence datasets.	Requires some bioinformatics background for optimal use and interpretation of the results.	(20)
	Multiple sequence comparison by log-expectation (MUSCLE)	A suitable tool for multiple sequence alignments.	Appropriate for large-scale sequence datasets due to the speed and accuracy.	Focused predominantly on alignment, rather than on functional analysis.	(21)
	Molecular evolutionary genetics analysis (MEGA)	A comparative tool for analysis of homologous sequence alignment. However, they also contain the homology testing of evolutionary hypotheses, besides building the phylogenetic tree of the input sequence.	A wide range of tools for phylogenetic and evolutionary analysis is offered. The user interfaces are friendly, and good documentation is present with great tutorials.	Some of the analyzes may cost quite high forcomputation.	(22)
	ClustalW2	A popular tool used for various sequence alignments.	Proven and trustworthy, provided in different kinds of formats.	Newly available tools come with advanced features and improved algorithms.	(23)
	Basic local alignment search tool for proteins (BLASTp)	Allowing users to compare an amino acid query sequence against a protein sequence database.	Fast, widely used, and recognized, comprehensive databases.	Basic tools and more specialized tools may be needed for in-depth analysis.	(24)
Secondary structure and physiochemical Properties analysis	PSIPRED	Applying the profile-based prediction algorithm to predict the secondary structure of proteins.	User-friendly, and highly accurate.	Unable to perform a comprehensive analysis.	(25)

Table 2: Continued

Application	Server name	Function	Advantages	Disadvantages	Ref
	Expert protein analysis system (ExPASy)	Access to a variety of tools and databases for protein analysis	Comprehensive resource for bioinformatics tools.	Broad scope requires users to know which specific tools to use for their needs.	(26)
3D-Modeling, refinement, and molecular docking	Robetta	A full-chain protein structure prediction service.	Can model proteins without known homologs, useful for novel protein prediction.	Predictions may vary in accuracy, depending on the complexity of the protein.	(27)
	High ambiguity driven biomolecular docking (HADDOCK)	A suite of computation modeling for protein and biomolecule complexes, which take experimental results together with bioinformatics data in deriving a reliable predictive model of molecular interaction and complex structure.	Plays a significant role in the central prediction of a biomolecular complex with a high affinity from several experimental data and hence forms a basic tool in a study that characterizes the molecular interactions in structural biology.	Its requirements for input and parameters are too complex, hence challenging for new users.	(28)
	SCWRL	Used for predicting protein side-chain conformations.	It's widely recognized for its accuracy in modeling side-chain positions.	Its focus is limited to side-chain modeling and doesn't account for the backbone dynamics, which could be crucial for overall protein function.	(29)
	Swiss-Prot	A manually annotated and reviewed protein sequence database that is part of UniProt.	High-quality annotations and reliability due to manual curation.	Smaller database compared to automatically annotated ones; update frequency.	(30)
	Phyre2	Protein homology/analogy recognition engine that predicts and models the 3D structure of proteins.	User-friendly, provides detailed structural predictions, and is useful for understanding protein function and interactions.	Predictions can be less accurate for proteins with few known homologs.	(31)
	RaptorX	Predicts protein structures based on deep learning.	High accuracy in structure prediction, especially for proteins without known homologous structures.	Predictions may be less reliable for proteins with very few similar sequences.	(32)
	UCSF Chimera	A visualization tool for interactive molecular modeling.	Powerful visualization capabilities support a wide range of molecular formats.	It is primarily a visualization tool and does not perform computational modeling.	(33)
	I-TASSER	An integrated platform for automated protein structure and function prediction.	It combines various methods for improved prediction accuracy.	Different results achieved based on the quality and quantity of the input sequence data.	(9)
	PEP-FOLD	Predicts the structures of small peptides with high efficiency.	Fast and accurate for short peptide sequences.	Its application is limited to peptides and is not suitable for large protein complexes.	(34)
	SwarmDock and PatchDock	A molecular docking tool.	Effective in predicting protein-protein and protein-ligand interactions.	The accuracy of docking predictions can be affected by the chosen parameters and the complexity of the molecular system.	(35)
	AutoDock Vina	Excels in the automated docking of small molecules to biomolecular targets.	Known for its speed and accuracy in docking simulations.	While it is powerful, its accuracy can depend on the quality of the input structures and docking parameters.	(36)
	ClusPro	A tool specifically designed for protein-protein docking and simulation, allowing researchers to predict how proteins interact.	Makes the docking process available to non-experts, since this will be an automatic process.	The predictions may still stand in need of experimental validation for binding efficacy.	(37)

Table 2: Continued

Application	Server name	Function	Advantages	Disadvantages	Ref
	GalaxyRefine and ERRAT	Used for the refinement and verification of the protein structures, respectively. GalaxyRefine can enhance models of proteins by refining the structure. ERRAT checks the atomic interaction and non-bonded contacts in the protein structure.	Enhancing the accuracy of models that is critical for further analysis.	The refinement and validation processes may not capture all possible conformational states of a protein.	(38)
Molecular dynamics simulation	GROMACS	Widely known for its high performance and flexibility in simulating large biomolecular systems.	Introduces a number of broad functionalities for modeling protein dynamics that comprise support for complex workflows and high-throughput screening.	It requires a good knowledge of the principles of molecular dynamics combined with significant computational resources.	(39)
	AMBER	A cornerstone in molecular dynamics simulations, it is favored for its accurate force fields and comprehensive simulation capabilities.	It provides detailed modeling of biomolecular systems, especially for understanding protein-ligand interactions.	Similar to GROMACS, it demands substantial computational power and expertise in molecular dynamics.	(40)
	iMODS	Facilitates the analysis of internal motions in protein structures, which is essential for understanding their flexibility and function.	It provides insights into the dynamic properties of proteins that are critical for their biological function.	Its focus is more on internal motions and may need to be complemented with other types of simulation for a comprehensive analysis.	(41)
	Visual molecular dynamics (VMD)	A molecular visualization program for displaying, animating, and analyzing large biomolecular systems.	Powerful visualization capabilities support a wide range of molecular simulations.	Primarily a visualization tool, it does not perform sequence analysis.	(42)
Immunogenicity and safety	VaxiJen	Predicts antigens and epitopes based on the physicochemical properties of proteins, employing an alignment-independent approach.	It is useful across a spectrum of pathogens for antigen prediction.	As predictions are not based on immune processing, additional validation may be necessary.	(43)
	ProtParam	A tool on the ExPASy server that allows the computation of various physical and chemical parameters for a given protein stored in Swiss-Prot or TrEMBL or for a user-entered sequence.	Provides detailed information about the protein, such as molecular weight, theoretical pI, amino acid composition, atomic composition, etc. User-friendly features.	Limited to parameter calculation, it does not provide broader functional insights.	(44)
	AntigenPro	A server that predicts protein antigenicity through machine learning aids in identifying vaccine candidates and understanding immune recognition by evaluating the proteins' potential to trigger immune responses.	The server is for the critical assessment of protein antigenicity and serves to highlight promising candidates for vaccine target selection. This allows algorithmic approaches and provides for high-throughput screening.	The predictive accuracy will depend on the similarity of proteins to the training data, which might then limit the applicability of the method to new proteins.	(45)
	ToxinPred	It predicts and analyzes peptides with toxic properties, which could be helpful in designing novel peptides with specific bioactivities.	Offers complete predictions for peptide toxicity, which is a very important characteristic in drug design, while it provides the user with a friendly interface to facilitate even the most arduous of analyses.	Its predictions are limited to peptide sequences; therefore, it is not allowed in broad toxicological studies. Accuracy may vary depending on the peptide sequence and available data.	(46)

Table 2: Continued

Application	Server name	Function	Advantages	Disadvantages	Ref
	C-IMMSIM	It mimics human immune responses at the molecular and cellular levels; it provides insight on the immune process and vaccine designs.	provides a full simulation of immunity mechanisms towards the development and research in vaccine and immunotherapy.	Model predictions differ and require validation with experimental results.	(47)
	AllerTop	Dedicated server for the prediction of allergenic proteins and peptides needed in the design of hypoallergenic foods and therapeutics, research support, and allergen safety assessment.	Readily available and user-friendly; hence, very useful for universal applications.	Focuses on allergenicity and may fail to predict the whole spectrum of immunogenic responses. For critical applications, it should be confirmed with experimental evidence.	(10)
	AllergenFP	Prediction of allergenic potential proteins through fingerprinting.	An accurate and rapid way of making allergen predictions to boost product safety evaluation with user-friendly features.	The specificity of the method for allergenic properties may miss other immunogenic factors. Depending on the fingerprint patterns, it may lose out novel allergens with atypical profiles.	(48)
	admetSAR	A highly predictive method for pharmacokinetics, pharmacodynamics, and toxicity.	It is a huge scope of prediction that stretches from drug-likeness to safety profiles, all of which are very essential for drug development at its early stages. The interface is user-friendly, and users can easily navigate through the complex ADMET data.	Sometimes the predictions do not reconcile with the experimental result, but it requires careful validation. Overfocus on ADMET properties could overshadow other important areas within drug development.	(49)
In-silico cloning and codon optimization	Codon usage wrangler	Codon usage optimizer tool for expression in specific organisms	It covers various organisms as host, and has a user-friendly interface	It has limited advanced features compared to other tools; may not support complex optimization sets	(50)
	GeneScript Optimum Gene™	Expression optimization tools in terms of codon usage, GC content, and mRNA structure	It contains customizable features supporting a wide range of organisms	It needs a subscription for full features; not user-friendly	(51)
	EMBOSS	Containing different bioinformatics tools, including codon optimization.	It has an open-source access and is free, containing integrative features through other bioinformatic tools	It has a less user-friendly interface; may be confusing for new users	(52)
	JCat	Java codon adaptation tool (JCat) used for codon optimization	It is user-friendly and supports multiple host organisms. Also, it can provide codon adaptation index (CAI) scores	It contains just the basic optimization features and it is not conditional for in-depth customization options	(53)

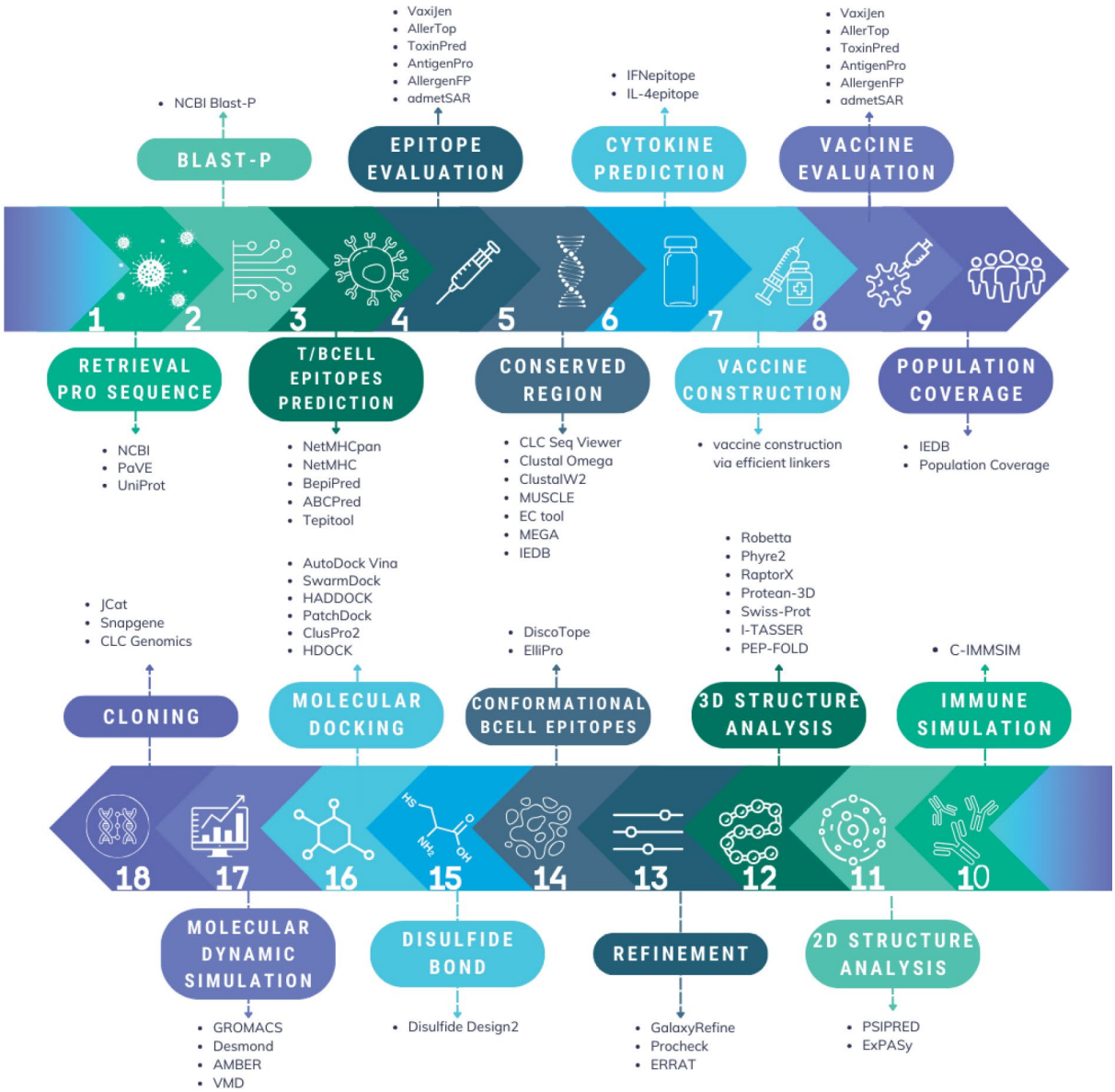

Fig.3: A diagram illustrating in-silico vaccine design steps.

Target protein sequence analysis for human papillomavirus vaccines

Databases for sequence retrieval

To get started on developing a vaccine, access to complete collections of the genetic sequences of HPV is needed. It should be noted that GenBank is a major repository representing information on nucleotide sequences of more than 300,000 organisms, and it has largely contributed to the availability of genetic data for HPV, which is imperative for the first steps of vaccine design. The UniProt offers a very rich database of protein sequences and functional information, which are quite important in understanding HPV proteomics and consequent knowledge on potential antigenic proteins. The papillomavirus episteme (PaVE) only caters to papillomavirus genomic sequences. This enables one to delve deeper into the different HPV strains that aid in the delimitation of conserved regions and potential areas of target for vaccine development (11, 12, 24).

Computational tools for sequence analysis

Since local similarity search was applied in the BLAST, the common conserved regions of the protein sequence among the different HPV strains are likely retrieved, and these serve as the potential regions applicable for vaccine design (24). Multiple sequence comparison by log-expectation (MUSCLE) is one of the tools used in the study for multiple alignments. MUSCLE has been the most preferable to use in the study for its very high speed and good accuracy, whereby it makes it easier for one to find the conservancy level among the sequences in HPV genotypes (21).

Conservancy and homology analysis of the designed peptide

Conservancy analysis

Analysis by conservancy has been able to identify epitopes that are highly conserved between different HPV genotypes, so they assure broad protection of the immunization against the virus with some particular relevance in the case of HPV, since it is one of the viruses with the most genetic variation. From the point of view of high conservancy, it is likely that there are high-conservancy epitopes in a number of HPV strains. These would be the best targets for making vaccines that would protect against the virus broadly. Bioinformatics tools, such as Clustal Omega for multiple sequence alignments and IEDB for conservancy analysis, are very essential. Such online resources assist researchers in identifying epitopes that have high conservation among many types of HPV and require high priority in vaccine constructs (20).

Minimizing autoimmunity through homology analysis

Homology analysis becomes fundamental for comparisons when working on the development of vaccines, more so in the evaluation of the similarity of human proteins with the proteins of the virus. This reduces the risks that could be involved, more so, from autoimmune responses that may result from the vaccine. The activation of the immune system by the vaccine may give rise to an autoimmune response in the way that it will identify and destroy the viral epitopes while mistakenly attacking the host cells, which bear similarity to those of the viral proteins. BLAST, Clustal Omega, and IPAS (Immunogenicity prediction and Analysis server) are used in this step (8). The IPAS predicts the risk of immune reactions, e.g., induction of an autoimmune response, through identifying similar sequences to host proteins. On a general overview, the formulated product should contain epitopes that do not show a very high degree of identity and might be cause for cross-reactivity. The maximum of 70% identity with human proteome is allowed for vaccine design (4).

Prediction of epitopes in human papillomavirus vaccine development

An epitope, a putative peptide sequence in an HPV protein where immune cells will interact, is identified through an epitope prediction. It is one of the most important steps in vaccine design, as it should assure specific, strong immune responses. Several computational tools and algorithms are used for epitope prediction in B cell and T cell manners.

B-cell epitope prediction

BepiPred is one of the bioinformatics tools that uses sequence-based information to improve linear B-cell epitope prediction. This method actually predicts the part or the regions in a given amino acid sequence of a protein that is likely to act as linear B-cell epitopes. These have short, linear, and relatively simple amino acid sequence motifs; they are exposed on the surface of a protein and are likely to bind a B-cell receptor directly. On the other hand, DiscoTope and ElliPro are servers used for epitope prediction and are 3D structure-based, with a different outlook from the above servers. They use the structural data in the prediction of the configuration of amino acids closely brought together and linking them in conformational epitopes of a B-cell later both features allow the prediction of conformational epitopes, which is an important criterion for predicting conformational epitopes in complex antigens like those found in HPV proteins (7).

T-cell epitope prediction

The NetMHC and NetMHCpan are some of the most recent computer-based algorithms that predict the binding of peptide sequences with MHC class I and II molecules. This is very critical since a linear correlation exists between binding and the potentiality of the peptide to trigger T-cell-mediated immunity. NetMHC is able to find the interactions between the peptide sequences and MHC class I molecule, which is important to identify the cytotoxic T-lymphocytes (CTL) epitopes. NetMHC figures out which parts of HPV antigens are most likely to bind strongly to MHC class I molecules. On the other hand, NetMHCpan includes a prediction of longer CTL epitopes, as well as an additional prediction for MHC class II molecules, which play a role in presenting antigens to helper T-lymphocytes (HTLs). HTLs secrete cytokines that help boost the activation and multiplication of CTLs, B-cells, and other immune cells. It does, therefore, have a lot of bearing in the identification of the peptides within HPV antigens that are capable of binding to MHC class II molecules, hence potentially acting as epitopes for HTL in developing immunotherapeutic strategies that aim at improving the capability of the host immune system against HPV (6).

In-silico evaluation of antigenicity, allergenicity, and safety

The epitopes of vaccine constructs for HPV will be an integral part and parcel of the edifice in terms of vaccine development, taking into account the development in computational design and predictive analytics in terms of being immunogenic. This stage takes it a notch higher than the theoretical plane and gets into empirical validation of the potential of the vaccine to elicit a desired immunological response. Research studies demonstrate the paradigm of using computational analysis to predict HPV vaccine efficacy This step should be performed for each epitope separately, as well as for a multi-epitope vaccine candidate (54).

C-IMMSIM, AntigenPro, and VaxiJen describe indispensable tools to look for insights into the antigenic potential profile of the constructs, meant to be part of a vaccine application for worldwide public health.

Additionally, it will illustrate MD simulation and its integration with HPV-related immunoinformatics. The C-IMMSIM server is a computational tool to model how the mammalian immune system responds to various immunological challenges; these could be vaccines, pathogens, or autoimmune diseases. The method to find its solutions, in this case, operates on sets of mathematical equations and algorithms that model complex interactions between a wide variety of immune cells (e.g., T-cells, B-cells, macrophages, dendritic cells), signaling molecules (cytokines, chemokines), and antigens. The C-IMMSIM model represents phenotypic molecular and cellular characteristics up to the organ level, providing a top-down view of immune response behavior. It mimics innate and adaptive immune responses. This would help in the prediction of the immunogenicity of such a vaccine and the likelihood of being able to raise a protective immune response by simulating an immune response toward a vaccine construct. The described approach could aid in the identification of the best antigen-adjuvant combination to elicit robust and long-lasting immunity (4).

AntigenPro applies the latest and most accurate machine learning models, trained from a massive dataset of known antigens and non-antigens. It will hence be making very accurate predictions of antigenicity on unidentified protein sequences. It is used in pathogen genome screening to find proteins with a higher propensity to be recognized by the immune system and hence have potential vaccine candidates. In this way, through comparative antigenicity analysis of different proteins, the risk of cross-reactivity may be evaluated. This evaluation is crucial to understanding the possibility of a vaccine inadvertently targeting host proteins, which could lead to autoimmunity (45).

VaxiJen is different in its function because it does not rely on sequence similarity and alignment, unlike most other bioinformatics tools. VaxiJen bases its prediction on the antigenicity of certain physicochemical properties of amino acids and their distribution within the protein. From there, it goes on to predict potential antigenicity in the absence of homologous sequences. The VaxiJen could then be applied for the comparative analysis of antigenic potentials between wild-type and mutant proteins, a very interesting application in the study of pathogen evolution and vaccine escape mutants (43).

Tools like ToxinPred predict high complexity for toxicity and immune response. On the other hand, AllerTop and AllergenFP are allergenicity predictors that help in scoring the allergenicity of the designed vaccine against HPV to come up with hypoallergenic HPV vaccine candidates (10, 48).

Multi-epitope vaccine construction and linkers

Hence, the multi-epitope vaccines are obtained using selected epitopes, and in turn, they are assembled in a strategic manner with the help of peptide linkers. Furthermore, these linkers play their part in such a way that the separation of epitopes becomes sufficient but not too much to lose recognition by the body's immune system. It is focusing mainly on five linkers with specificities: EAAAK, GPGPG, AAY, and KK, which have great importance in the assembly of epitopes into a coherent vaccine construct (55).

The EAAAK linker is an alpha-helical inflexible linker; hence, it allows good separation of functional domains, which have high alpha-helix content. This results in significant spatial separation between epitopes, thereby reducing steric hindrance and enabling each epitope to bind with high affinity to its cognate receptor. This linker is particularly beneficial for enhancing B-cell epitope presentation. In addition, the glycine-proline flexible linker (e.g., GPGPG) has junctional immunogenicity, hence being the best link between B-cell epitopes. This is because the rigid presence of proline residues in some sites creates structural rigidity that allows proper folding with adjacent epitopes, thereby raising their immunogenicity. The flexible AAY linker helps enhance processing and presentation of T-cell epitopes. The alanine residues are for flexibility, and it can also be that the tyrosine residue is used to increase immunogenicity within the epitope, given its potential to increase MHC class II binding affinity. Lysine residues in the KK linker serve to further separate epitopes, enhancing immune recognition. The lysine residues in the linker, being positively charged, will freely get coupled with some of the negatively charged receptor sites of the cell surface, thereby facilitating epitope uptake and presentation to form the most potent immune response by the antigen-presenting cells (55).

The choice of an appropriate linker is an important one in designing an efficient multi-epitope vaccine. The linker will not only affect the physical and chemical stability of the vaccine construction but also determine the processing and presentation of epitopes, leading to an effect on the immunogenicity of the vaccine. The above-mentioned linkers find suitable use in different designs of vaccines and hence manifest their utility in improving the immune response for the targeted pathogen by epitope optimization both in terms of recognition and processing by B-cells and T-cells.

Secondary structure analysis of vaccine and physiochemical assay

A stable and integrated vaccine construct with multi-epitopes provides immunogenic efficacy. This needs to secure the viability of the vaccine construct through a proper analysis of its secondary structure and physicochemical properties. We, in the present communication, outline the methods for a detailed analysis with advanced bioinformatics tools for the secondary structure and physicochemical properties.

PSIPRED is applicable for prediction of the secondary structure of proteins through a sequence. PSIPRED can detail some information about the helices, strands, and coils of the designed vaccine, which is crucial for utilizing the conformational structure. The analysis of B-cell

antigenic region prediction and epitope accessibility of the vaccine is very important since it helps guide a proper interaction with the immune system. Predicting secondary structures leads to better vaccine design because it ensures that epitopes are presented in the best way and stay stable (25).

The servers ProtParam and SOLpro were integrated to have a full comprehension of the physicochemical features of the vaccine. Using ProtParam, several parameters were computed, including the molecular weight, theoretical isoelectric point (pI), amino acid composition, aliphatic index, and stability index. These parameters indicate the solubility, stability, and suitability of vaccine antigens for large-scale production and administration (56).

SOLpro predicts the solubility of the vaccine based on its overexpression in E. coli. Thus, the prediction of solubility enables the identification of mutations that affect improved solubility, guiding the optimization process for efficient production (57).

3D modeling, refinement and molecular docking

In fact, such 3D modeling, refinement, and molecular docking methods are essential in the vaccine development process for HPV in the computational pipeline. These techniques provide molecular details down to the atomic level that allow the researcher to obtain epitope-receptor interactions, thereby setting the critical basis for the design of vaccines with optimized immunogenicity and specificity.

3D modeling

In short, 3D modeling plays an indispensable role in the prediction of viral protein structural arrangements and, in a way, sheds some light on epitope presentation and antigenicity. Iterative Threading ASSEmbly Refinement (I-TASSER) is a highly advanced computer simulation program that has become highly reputable in regard to predicting structure and function from amino acid sequences through iterative threading and assembly simulations (9). I-TASSER serves as a comprehensive server for predicting protein structure and function, achieving high accuracy in blind protein structure prediction contests (CASP). Primarily, it adopts the threading hierarchical approach to fill the gaps and is followed by structure modeling with the ab initio method.

Other servers carrying out similar functions are Protein Homology/Analogy Recognition Engine V 2.0 (Phyre2) and Swiss-Model, which provides a platform to accurately model 3D structures for the purpose of finding potential antigenic sites in the proteins of HPV. Phyre2 is a web tools for protein modeling that can model anything, from single sequences to complex hetero oligomers. In addition to predicting 3D structures, it uses state-of-the-art homology detection methods to offer detailed models of protein function and mutations. Swiss-Model relies on experimentally solved protein structures as templates for the model built with respect to the target sequences. It provides detailed quality assessment reports on computed models, as well as Global Model Quality Estimation (GMQE) and Qualitative Model Energy Analysis (QMEAN) scores (30, 31).

Robetta is the University of Washington Baker Lab's web server for automated protein structure prediction and analysis. It uses comparative modeling (homology modeling) to build the 3D structure of sequences with known homologous structures and offers a method of ab initio conformational sampling for sequences that do not have closely related homologs, thus applicable to different protein modeling problems. For proteins without an identified homolog, Robetta uses the Rosetta ab initio structure prediction method to generate 3D models based on principles of protein folding thermodynamics other than sequence identity. Homologous structures serve as a template, which is then used for model building to guide the structural conservation among related proteins when predicting the structure of an unknown protein (27, 28).

Refinement

The refinement of 3D models means the adjustments and improvements of the resolution in the initial 3D models. This step is crucially important for resembling the native structure for molecular docking and interaction studies. GalaxyWeb provides a suite of tools for protein structure prediction and refinement. Further, another part of the suite, GalaxyRefine, takes the model a step further in quality through side chain position, backbone structure refinement, and finally the model's overall geometry. The consequence of such improvements may be improved model precision and soundness based on the GalaxyRefine server's results. The result from GalaxyRefine is a set of refined structures showing high scores of improved geometry and minimal energy states, but more realistic physical interactions within the protein This gives the researcher an option on the protein sequence in terms of the level of detail, which can vary from fine to coarse in accordance with the researcher's needs. Improved model accuracy significantly influences epitope prediction and molecular docking, relying on a more dependable structure of the antigenic protein (58).

The *ERRAT* scores focus on scrutinizing the statistical quality of the model by computing the distribution of all types of atoms in reference to each other. High ERRAT scores express a well-resolved part of the model; hence, low scores express some regions that might need refinement in the model. A high ERRAT score would therefore indicate that the non-bonded contacts observed in the model are reasonable with respect to well-refined high-resolution structures and could allow for model acceptance. Low scores, on the other hand, do highlight problematic areas within the model that might require some touching up. This analysis is crucial to ensure the use of structural study models in designing the highest-quality vaccine. The minimum score applicable for refinement in *ERRAT*

is 50, while Ramachandran plot score should be at least 90% of residues (58).

PROCHECK analyzes the residue-by-residue geometry of structural conformation to assess the stereochemical quality. It is applicable for correction of the predicted dihedral angle outliers in the 3D model through plotting the Ramachandran. It allows the identification of areas that deviate from expected dihedral angle distributions, which indicates the further refinement. The outlier correction is crucial for developing an accurate model which predicts the interaction of the vaccine with the immune system or adjuvants (59).

Molecular docking

Molecular docking plays a very pivotal role throughout the process of vaccine design, especially toward the optimization of the formulation of the vaccine that may effectively be able to stimulate an immune response. Molecular docking is potent to predict the interaction of vaccine epitopes with immune system elements, e.g., adjuvants, MHC molecules, T cell receptor (TCR), and B cell receptor (BCR), which determine the immune recognition and reactions. This predictive capability is invaluable for designing vaccines that elicit a targeted and robust immune response. During this simulation, the exact mechanisms of how adjuvants modify the presentation of epitopes to the immune system may be established, indicating which adjuvant-epitope combination is most effective. Further docking of the adjuvant-epitope complex with MHC molecules or antibodies may provide some insight into the vaccine construct's likely immunogenicity. It could therefore also focus on designing a more stable vaccine formulation that is effective in delivering the antigen to the right immune cells. Also, MHC class I and II can be docked with presentable epitopes of L1 and L2 proteins in manner of the ligand-receptor complexes (60).

High Ambiguity Driven Docking (HADDOCK) is a docking server based on information, for the study of biomolecular structures. It models the relationship between proteins and other biomolecules and can be used in research on structural biology, design of drugs, and development of vaccines. HADDOCK is an integrative approach driven both by experimental data [e.g., nuclear magnetic resonance (NMR)] or other means (e.g., inter-residue contacts) obtained from other techniques with predefined secondary-structure elements. This approach leaves an open field for the use of a very wide range of biochemical and biophysical data that could well afford to allow them to make predictions of molecular interactions more accurately and realistically (28). On the other hand, AutoDock Vina is the most accurate and fast tool for simulating docking exercises, hence a favorite solution in the prediction of small molecule interaction with proteins and more relevant in docking studies of adjuvants (36).

However, some serious predictive accuracy, 3D modeling, refinement, and molecular docking computational barriers must be negotiated. The integrative powerful techniques based on artificial intelligence and machine learning hold promise in improving the predictive power and efficiency of the analyses.

MDs simulation in vaccine development

The described MD simulations would be very useful for studying in more detail how the vaccine constructs interact with their immune receptors at the atomic level and how they stay stable. In the fight against hard-to-tackle pathogens like HPV, MD simulations are of great help in providing a detailed picture of the interaction of vaccine candidates and their behavior with the immune system. In recent years, the discipline of immunoinformatics has used computational methods to develop and analyze the safety and efficacy of multi-epitope-based vaccines, as demonstrated by Sanami et al.'s (54) study of an HPV-16/18 vaccine candidate. The integration of computational predictions with empirical data allows them to anticipate immune responses and thus optimize vaccine constructs for maximal safety and efficacy.

The fusion of MD simulations with immunoinformatics marks a significant leap forward in vaccine research. This would certainly provide a platform for an extended and more comprehensive assessment of antigenic peptide vaccines, as portrayed by Jabbar et al. (61). These should maximize the good approach for choosing the most suitable vaccine candidate for treatment against HPV, owing to strong binding affinity evaluations This has been outlined as a framework through which simulation methods, together with computation techniques, would be applied in the process of refining the evaluation and optimization of vaccine constructs against a broad range of pathogens.

Among the computation tools for MD simulations, key execution strengths and features that are unique to each make a contribution to the detailed study of protein dynamics and their interactions: GROMACS and AMBER. GROMACS is celebrated for its adaptability, while AMBER is preferred for its precision in force fields. In complement, tools such as ClusPro and iMODS provide facilities to perform protein-protein dynamic simulations and motion analysis that are needed to assess the structural integrity and interaction dynamics of vaccine candidates (37, 40, 41).

The main tools that help in understanding the structural integrity and dynamics of proteins during vaccine development are root mean square deviation (RMSD) and root mean square fluctuation (RMSF). RMSD measures the average distance of atoms-mostly the backbone atoms-in a protein structure at different times within the simulation compared with a reference structure. That is one of the main indicators of maintaining the conformation of the protein over time. Further, the RMSD was helpful in demonstrating what effect the introduction of antigens, epitopes, or adjuvants may have on the overall construct of vaccine stability. A stable, low RMSD value over the simulation time would obviously assure that the structural integrity of the vaccine construct is maintained, which is

indispensable for the immunogenic efficacy of the vaccine. RMSF, on the other hand, gives a measure of the average fluctuation from the mean position of each atom, which will fluctuate and express the flexibility of the given region or regions of the molecule. The greater the value of RMSF, it reveals the regions of high flexibility or mobility. This will help to identify those parts that take part in epitope presentation or receptor interaction, evident through high flexibility by analyzing RMSF values. Knowledge about these dynamics can thus be helpful for the optimization of antigen presentation and later development of vaccines that would induce a particular immune response (54).

In fact, one of the things that makes current vaccine designs unique is that they use both experimental immunogenicity assays and computer predictions as part of the wider process of evaluating vaccines. This synergistic approach ensures that only the most promising candidates, carrying strong immunogenic potential with a very clean safety profile, are moved forward for clinical testing. These encompass the very threshold for precision medicine, through which the further advancement of next-generation HPV vaccines is laid down by the amalgamation of bioinformatics, immunoinformatics, and molecular biology (Table 3) (54).

Table 3: Review on computational studies for HPV vaccine design and its relevant steps

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Kumar et al. (62)	E5 protein of HPV-16	GenBank	dbMHC	IEDB	NM	NA	NA	ABCpred	JCat	Identification of potent T-cell and B-cell epitopes for E5 protein from HPV-16, which will be used in the course of the design of therapeutic vaccines.
Sarkar et al. (63)	E protein of DENV-1, L1 protein of HPV-16	GenBank, UniProt	NetMHCpan, BepiPred, MHCcluster	IEDB	NM	PEP-FOLD3, PatchDock, FireDock	iMODS	VaxiJen, AllerTOP, AllergenFP, ToxinPred	JCat	Epitope-based subunit vaccines designed against DENV-1 and HPV-16 through the complete genome of the target virus using reverse vaccinology and immunoinformatics approaches inactivated show good potential for further *in vitro* and *in vivo* studies.
Bahmani et al. (64)	HPV16-E7 protein	GenBank	NM	BioEdit, IEDB	NA	PEP-FOLD3, Swiss-model	NA	NA	NM	Effective peptides from the HPV-16-E7 protein were identified in the present study that could induce an immune response and are designed for high-coverage global putative vaccines. Further, molecular docking was carried out to evaluate peptide-MHC binding affinities.
Namvar et al. (65)	E5 and E7 proteins of HPV-16, HPV-18, HPV-31, HPV-45	NM	NM	BioEdit, IEDB	ProtParam, Predictprotein, SAVES	I-TASSER, GalaxyRefine	NM	IFNepitope, PA3P	NM	The study elaborately designed E5 and E7 peptide-based vaccines against HPV-16, HPV-18, HPV-31, and HPV-45, demonstrating their potential to stimulate immune responses and thus protect against tumorigenesis in the murine model.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Jabbar et al. (61)	E6 and E7 proteins of HPV-16 and HPV-18	GenBank	NetCTL, ElliPro, NetMHCpan	NM	ProtParam	Phyre2, PEP-FOLD3, FlexPepDock	AMBER	Prime-MMGBSA,	NM	Identification and epitope evaluation of a suitable vaccine candidate against HPV, with special emphasis on the computation approach in predicting an immunogenic peptide.
Sabah et al. (66)	E6 and E7 proteins of HPV-58	GenBank	NM	IEDB, MEGA Clustal W	NM	SOPMA, PFP	NM	AllerHunter, AllerTop	NM	The document introduces a methodological introduction to epitope-focused vaccine design for HPV-58, where a clear emphasis lies on the use of computational analysis for epitope prediction, conservancy, and an immunogenic evaluation that aims toward enhancing therapeutic vaccine development.
He et al. (67)	E6 and E7 proteins of HPV-33 and HPV-58	GenBank	ABCpred	ConSurf, IEDB	NM	Phyre2	NM	NM	NM	Identification of the variants and therapeutic epitopes in HPV-33/58 E6 and E7 proteins, focusing on the application of bioinformatics in the context of population-specific design of the vaccine.
Sanami et al. (54)	E5 and E7 proteins of HPV-16 and HPV-18	GenBank	NetCTL, NetMHCII, IFNepitope	NM	ProtParam, SOLpro, ANTIGENpro	PEP-FOLD, I-TASSER, GalaxyRefine	GROMACS, ClusPro	VaxiJen, AllerTOP, ToxinPred	JCat, GeneScript Optimum Gene™	The designed multi-epitope vaccine was observed to be structurally stable, non-allergenic, and non-toxic. Therefore, it would likely provide potential efficacy against cervical carcinoma upon further validation through *in vitro* and *in vivo* studies.
Soumia et al. (68)	E6 protein of HPV	NM	NA	NM	NA	AutoDock Vina, MGL	Desmond	admetSAR	NM	This study detects three compounds with potential inhibitory activity against HPV E6 protein, supported by virtual screening, molecular docking, and dynamics simulation, in order to pave the way for the development of targeted anti-HPV therapeutics.
Bahmani et al. (69)	E6 protein of HPV-16	GenBanK	NM	MEGA, Clustal W2, BioEdit, IEDB	NA	EP-FOLD3, Hex	NM	NM	NM	The designed vaccine candidate against HPV-16-E6 designed by in-silico analysis looks very promising, as the identified T-cell epitopes showed good binding affinities and good potential to cover large populations. The molecular docking results of computational analysis validated the interaction between selected epitopes and MHC molecules; further support was given to the probable effectiveness of the vaccine.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Rezaei et al. (70)	E7 protein of HPV-16, Homo sapiens/Mus musculus Hsp27 and Hsp70	GenBank	NetMHCpan, Syfpeithi, ProPred	IEDB	PA3P	Robetta, GalaxyPepDock,	NM	IL4pred, IL-10pred, IFNepitope VaxiJen, ToxinPred	NM	This study was able to design multi-epitope and whole sequence-based vaccine constructs using Hsp27 and Hsp70 linked to the HPV-16 E7 protein. It was alluded to in one study that the designed constructs may, therefore, display potential in eliciting strong immune responses highlighted by effective docking with TLRs and endocytic receptors, which might be an indication that these proteins may serve as an excellent vaccine candidate for HPV-related cancers.
Li et al. (71)	E1, E5, E7 proteins	UniProt	SCWRL	IEDB	NM	3D-QSAR, SCWRL	AMBER	NM	NM	The development of this unique pipeline makes it possible that computation and experimentation are used for the identification and validation of high-affinity T-cell epitopes in the HPV genome for the development of peptide vaccines for cervical cancer.
Mohamed et al. (72)	L1 protein of HPV-16	GenBank	NM	IEDB, BioEdit with ClustalW	NM	RaptorX, UCSF Chimera	NA	AllerTOP	NM	The conserved T and B cell epigraph for a prospective universal HPV-16 peptide vaccine.
Namvar et al. (73)	E5 and E7 proteins of HPV16,18,31,45	MUSCLE	BepiPred-, NetMHCpan, syfpeithi, ProPred	IEDB	ProtParam	CABS-dock, GalaxyPepDock	NM	VaxiJen, AllergenFP, Toxinpred	NM	Based on the conserved regions and designs here, constructs L1 and L2 showed good expression levels in HEK-293T cells and generated an effective immune response in mice, making them a potential universal vaccine candidate against high-risk HPVs.
Yazdani et al. (50)	L1 protein of HPV	GenBank	NetMHC, RANKPEP, CTLpred, BepiPred	CLC Sequence Viewer, IEDB	RAMPAGE, ERRAT	I-TASSER, GalaxyRefine	ClusPro	VaxiJen, ANTIGENpro, AllergenFP	Codon Usage Wrangler	An integrated study of immunoinformatics and structural vaccinology for the design of broad-spectrum HPV vaccines will, therefore, show one of the applications of bioinformatic tools in vaccine design.
Dehghani et al. (74)	L1 protein of HPV-16 and HPV-18	GenBank	Immuneepitope, BcePred,ABCpred, Bepipred	CLC Sequence Viewer	NA	I-TASSER, Phyre2, (PS)2	NM	AlgPred, VaxiJen	NM	The present study has reviewed the variants of HPV L1 protein, a possible target for vaccine, and their characteristic polymorphism that might be of importance in the development of an effective vaccine in Iran.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Kayyal et al. (75)	L1, L2, and E7 proteins of HPV-16 and HPV-18	GenBank, UniProt	NetMHCpan, NetMHCIIpan, NetCTL	IEDB	ProtParam, protein-sol web server, PSIPRED, RaptorX	I-TASSER, GalaxyRefine, GalaxyPepDock	ClusPro	AlgPred, ToxinPred web	NM	This study suggests two novel multiepitope DNA-based vaccine candidates for HPV-16 and HPV-18, which have shown substantial immunological responses and antitumor effects in mouse models, particularly after the addition of HSP70 epitopes.
Mahmoudvand et al. (76)	L1 protein of HPV-16	GenBank	netCTLpan, netMHCpan, BepiPred	IEDB	NA	ExPASy, PEP-FOLD, Molegro Virtual Docker	NA	AllergenFP, ToxinPred	NM	Among the T-cell epitopes predicted, some bind strongly to HLA-A0201 molecules and others to HLA-B3501 molecules, while one of the B-cell epitopes showed potential to elicit a neutralizing antibody response; hence, presenting them as pertinent targets for the development of an epitope-based vaccine for HPV-16.
Stolbikov et al. (77)	L1 protein of HPV-16 and HPV-6	GenBank	BepiPred, DiscoTope,	BioEdit	NM	PDB	NM	NM	NM	The research concluded that shared linear and 3D epitopes of HPV-16 L1 with the HPV6-L1 protein are identified, and thus cross-reactivity is raised as a possible foundation for the development of reactive antibodies. This cross-interaction was brought about due to the similarity in antigenic determinants of these HPVs, mainly towards T-cells, where they found a common antigenic determinant, which enabled moderate cross-reactivity of these HPV-16 L1 antigens' antibodies with antigens of HPV-6 L1.
Kumar et al. (4)	L1, E5, E6, and E7 proteins of HPV-16 and HPV-18	GenBank	IFNepitope	IEDB	ProtParam	I-TASSER, PatchDock, FireDock	GROMACS	AllerTOP, VaxiJen	JCat	This construct was introduced as a multi-domain, universal, chimeric recombinant vaccine for HPV-16 and HPV-18 with detailed immunoinformatic analyses, structure predictions, molecular docking, and dynamics simulations in silico. Some of the propitious physicochemical, antigenic, and immunogenic characteristics of the vaccine discussed here are its potentiality in further *in vitro* and *in vivo* investigations.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Firdaus et al. (78)	L1 and L2 proteins of HPV- 52	GenBank	ElliPro, BepiPred	IEDB, BioEdit	NM	PEP-FOLD 3, Swarmdock, BIOVIA Discovery Studio	NM	VaxiJen, AllerTOP, Toxinpred	NM	Several potent B and T cell epitopes were identified from HPV-52 L1 and L2 proteins, which displayed high conservancy and immunogenicity, suggesting their use in chimeric vaccine design for the effective prevention of HPV52 and the possibility of being cross-reactive with other HPVs.
Elshafei et al. (79)	L1 and L2 proteins of HPV-16	GenBank	ABCpred,	BioEdit, IEDB	ProtParam	NM	NM	AllerTOP, ToxinPred	NM	The design and evaluation of a multi-epitope vaccine based on L1 and L2 proteins of HPV-16 were carried out with the aim that such a construct would display properties showing promising physicochemical features and possible induction of immune response.
Karimah et al. (80)	L1 protein of HPV-45	GenBank, Uniprot	Ellipro	Clustal O	ProtParam	Swiss-Model	NM	VaxiJen, AllergenFP	NM	A hydrophilic vaccine candidate, which shows good immunogenicity against HPV-45 L1 protein, was further experimentally developed to have a stable conformation and also be non-allergenic.
Sulfianti et al. (81)	L1 protein of HPV-45	NM	NetMHCpan EL	IEDB	NM	Robetta	NM	VaxiJen	NM	The designed vaccine encoding the L1 protein of HPV-45 contains peptides with the capability to induce T-cell-mediated responses both within and across world populations, as predicted by in-silico HLA binding and antigenicity analysis.
Tîrziu et al. (82)	L1 protein of HPV-45	GenBank, Uniprot	Ellipro	Clustal O	ProtParam	Swiss-model	NM	VaxiJen, AllergenFP	NM	Using retained B-cell linear epitopes, a non-allergenic, stable, hydrophilic vaccine candidate design was achieved for the L1 protein of HPV-45, with potential good physicochemical characters for further experimental study.
Negahdaripour et al. (83)	L2 protein of HPV-16	UniProt	NetMHC, CTLPred, PAComplex, BepiPred	IEDB	ProtParam	I-TASSER, GalaxyRefine, 3Drefine, SwarmDock	NM	IFNepitope, VaxiJen, AllergenFP	Codon Usage Wrangler	The designed vaccine may provide humoral and cellular immunity against HPV, since the in-silico comprehensive analysis has indicated this.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Kaliamurthi et al. (52)	L2 protein of HPV-58	GenBank	NetMHC, CTLPred, Tepitool, IFNepitope, ABCPred	EC tool, IEDB	ProSA-web, RAMPAGE,	I-TASSER, GalaxyRefine, 3DRefine, SwarmDock, ERRAT	NA	VaxiJen, ANTIGENpro	GeneScript Optimum Gene™, EMBOSS	Design and evaluation of the SGD58 vaccine construct against cervical papilloma with detailed computational analysis for epitope prediction, vaccine evaluation, and molecular docking showed promising results.
Negahdaripour et al. (3)	L2 protein of HPV-16	Uniprot	DiscoTope	NM	ProtParam, PREAL	I-TASSER,	GROMACS, ClusPro	ANTIGENpro, VaxiJen, Allerdictor, PREAL, AllergenFP AllergenFP, AlgPred	Codon Usage Wrangler	The multi-epitope-based vaccine, which has been successfully designed, targets HPV with a greater emphasis on HPV16's L2 protein. The in-silico analysis by epitope prediction, structural modeling, and immunogenicity assay could recognize that the vaccine construct has all the desired properties to have very high immunogenic potential repeatedly and form important interactions with the immune receptors. This in-silico work paves the way for experimental validation and presents a way to develop a vaccine potentially effective against HPV.
Kaliamurthi et al. (84)	L2 protein of HPV-58	GenBank	NetMHC, CTLPred, Tepitool, IFNepitope, ABCPred	EC tool, IEDB	ProSA-web, RAMPAGE, ERRAT	I-TASSER, GalaxyRefine, 3DRefine, SwarmDock	NM	VaxiJen, ANTIGENpro	GeneScript Optimum Gene™, EMBOSS	Extensive computational analysis for epitope prediction, vaccine evaluation, and molecular docking exhibited good results in designing and evaluating the SGD58 vaccine construct against cervical papilloma. This paper describes a holistic insight into the methodologies applied to design a chimeric vaccine against HPV-58 and brings forward the potential of computational tools in vaccine development.
Mashhadi Abolghasem Shirazi et al. (85)	RG-1 epitope of L2 protein of HPV-16	GenBank, KEGG database	Rankpep	IEDB	DiANNA	NA	NM	C-IMMSIM, VaxiJen	NM	The designed construct of the vaccine, in which the RG-1 epitope was incorporated and equipped with the required adjuvants, can be a good construct of a preventive vaccine capable of eliciting good immune responses and holding out promise against diverse genotypes of HPV.

Table 3: Continued

Authors (Reference)	Main antigen/ strain	Sequence analysis	Epitope prediction	Homology analysis	Construct evaluation	3D-Modeling, Refinement, and Molecular Docking	Molecular dynamics simulation	Immunogenicity and Safety Evaluation	Virtual Cloning and Codon Optimization	Main findings
Shahab et al. (1)	L2 protein of HPV-16	UniProt, BLAST	NetCTL, IFNepitope	IEDB	ProtParam	trRosetta,	AMBER, ClusPro	C-IMMSIM, VaxiJen, AllergenFP, ToxinPred	NM	The vaccine construct designed thus seems like a very promising entity from all concepts of stability, immunogenic quality, and potential efficacy perspective to HPV-16, supported by the profound immunoinformatic analyses, molecular modeling, and in-silico evaluations.
Dorosti et al. (86)	L2 and E7 of HPV, HSV-2, and Chlamydia trachomatis	GenBank	MHCpred, RANKPEP, CTLpred, PAComplex, ABCpred, BcePREDS, DiscoTope	IEDB	ANTIGENpro, SOLpro, Protein-sol, ProtParam	I-TASSER, GalaxyTBM, Galaxyloop, GalaxyRefine	GROMACS	ANTIGENpro, VaxiJen, AlgPred, AllerTOP	NM	In the present study, a computationally designed multi-epitope vaccine targeting major causative agents of STDs has been proposed using a comprehensive immunoinformatics approach for its development.
Dharmawan et al. (87)	HPV-16 and HPV-18	GenBank	IEDB	NA	NM	Swiss-model, PatchDock, FireDock	NM	VaxiJen, AllerTOP, ToxinPred	NM	Identifies 16P1 and 18P4 as the most promising candidates for an HPV-16 and HPV-18-specific vaccine and recommends further experimental research to confirm their efficacy.
Gupta et al. (88)	E2, E4, E5, E6, E7, L1, and L2 proteins of HPV-18	GenBank	ABCpred	IEDB	NM	UCSF Chimera	NM	NM	NM	The results showed some potent T-cell and B-cell epitopes for HPV-18. They found 20, 87, and 94 potential epitopes for MHC class I, MHC class II, and B-cells, respectively.
Mashhadi Abolghasem Shirazi et al. (89)	RG1 L2 epitope of HPV-16	NM	NA	NA	ProtParam, SOLpro, ccSOL	I-TASSER, 3Drefine, SwarmDock	NM	VaxiJen, AllergenFP, Toxinpred	NM	Notably, the RP vaccine design induced more potent immune responses of the HPV RG1 epitope compared with adjuvanted TLR7 agonist or alum adjuvant, suggesting that this design may offer broader efficaciousness for HPV vaccines.

HPV; Human papillomavirus, dbMHC; Major Histocompatibility Complex database, IEDB; The Immune Epitope Database, NM; Not mentioned, NA; Not available, iMODS; Internal coordinates normal mode analysis server, SAVES; Structure Analysis and Verification Server, AMBER; Assisted Model Building and Energy Refinement, and TLRS; Toll-like receptors.

In-Silico approaches for epitope expression and cloning optimization

In-silico cloning of the nucleotide sequence into a plasmid vector is the first step for optimization of vaccines. This process leads to increasing the efficacy of expression of recombinant vaccines, immunogenicity, and stability. Choosing a suitable vector and host [based on the post-translational modifications (PTMs)], identifying the optimal cloning sites, and selecting the effective promoters are critical steps in in-silico optimization. Also,

copy number and recombination rates, which leads to optimizing both expression and stability, are two important properties of vectors which should be considered in in-silico optimization. This includes modifying promoters to induce expression at specific times or tissues, thereby increasing vaccine efficacy. PTMs such as glycosylation, phosphorylation, and cleavage play vital roles in the proper folding, stability, and function of the expressed epitopes. In-silico cloning can predict the effects of the chosen vector and host system on these modifications. Because certain PTMs significantly influence the vaccine's immunogenicity by affecting epitope presentation and immune system recognition, this predictive capability is critical for vaccine development (54).

Also, codon optimization emerges as a new strategic pillar for improving vaccine efficacy. One method of codon optimization involves altering the genetic code within vaccine constructs to increase protein expression to a level that contributes to enhancing immunogenicity. A study by Farzanehpour et al. (90) reveals how such optimization could significantly raise gene expression, hence making the vaccine have a higher potential to provoke a strong immune response. This paper is an example of codon optimization in vaccine design against *Plasmodium falciparum*; this abstract highlights a report of gene modification for mammalian expression, codon usage, Kozak sequence, and GC content optimization of two *Plasmodium falciparum* genes encoding the erythrocyte-binding antigen and the merozoite surface protein 1 (MSP1). This showed much higher levels of protein expression and, correspondingly, high immunogenicity. In this regard, 10 to 100 times less codon-optimized plasmid DNA is reportedly required for animal immunization to produce high antibody titers than plasmids containing non-optimized sequences.

Discussion

In this comparative review, we performed a literature review and evaluated the impact of different servers and databases, and their significant differences in computational design of HPV vaccines. In the first steps, databases like NCBI, UniProt, and PaVE are indispensable for sequence conservancy and homology analysis. Kumar et al. (4) used NCBI and UniProt for comprehensive sequence data retrieval. They also used BLASTp and Clustal Omega for selecting the conserved identical sequences among HPVs which are crucial for designing broad-spectrum vaccines. Similarly, Sarkar et al. (63) used Clustal Omega and IEDB for conservancy analysis aiming to find conserved epitopes between multiple HPVs. Clustal Omega also was used by Negahdaripour et al. (83) for the same purpose.

B-cell and T-cell epitopes should be predicted in the epitope prediction step. NetMHCpan, a T-cell epitope prediction tool that supports a wide variety of MHC alleles, covers a significant spectrum of world-wide population. Kumar et al. (4) used NetMHCpan for this purpose. Also, they used BepiPred for B-cell epitope prediction, which is critical in locating linear B-cell epitopes. Jabbar et al. (61) used ElliPro and DiscoTope to predict conformational and discontinuous B-cell epitopes, respectively. These servers perform a high-throughput analysis on putative epitopes. Sanami et al. (54) employed NetMHC and BepiPred for predicting epitopes with strong immunogenicity and binding affinity. In the same manner, Shahab et al.'s (1) study used NetMHCpan and BepiPred, for demonstrating the accuracy of the predicted epitopes to robust immune responses. High immunogenicity and binding capacity of the predicted epitopes in the mentioned servers demonstrate that reliable identification of epitopes inducing potent immune responses was made possible by the combination of at least two servers.

For structural modeling, AutoDock Vina and I-TASSER are used in various studies. In order to assess the epitopes-immunoreceptor interaction, Kumar et al. (4) used AutoDock Vina for molecular docking experiments. They also used I-TASSER for protein structure prediction. By using these strategies, the vaccine designs were improved and the immunogenic stable epitopes were provided. Also, Mahmoudvand et al. (39) used 3D model by I-TASSER. They demonstrated the precision of the predicted protein structures. Also in the next step, the accuracy and reliability of the structural models were significantly increased by the refinement and validation processes carried out by Sarkar et al. (63) and Shahab et al. (1) utilizing GalaxyRefine and ERRAT. These precautions guarantee that the vaccine constructions are able to retain their structural integrity and to interact with the immune system components in an efficient manner.

As mentioned previously, MD simulations is critical for evaluating the atomic-level stability and interactions of vaccine components. Kumar et al. (4) evaluated the stability of multi-epitope vaccine designs using MD simulations by GROMACS. They were able to monitor the flexibility and stability of epitopes by mimicking the vaccination structures in a biological setting, which confirms that the constructs maintained their structural integrity over time. Also, Mahmoudvand et al. (76) used MD simulations by GROMACS to evaluate the behavior and stability of their vaccine designs, confirming their eligibility for further experimental validation. In this regard, Jabbar et al. (61) used AMBER to study the dynamic behavior of HPV epitopes in association with MHC. Their evaluations estimated the stability and binding interactions between the predicted epitopes and MHC. Also, Sanami et al. (54) used MD simulations to evaluate the structural stability of HPV vaccine candidates.

Conclusion

Our scoping review study exhaustively analyzes the present status and future perspectives of computational strategies in the design of multi-epitope vaccines against HPV, which is the major etiological agent in cervical cancer and other malignancies. It emphasizes the changing power of immunoinformatics and computational vaccinology to lend a solution to the diversity and oncogenic potential

raised by HPV. The present paper attempts to give a detailed review of the complete computational pipeline, from sequence analysis and epitope prediction to conservancy analysis, molecular modeling, and in-silico evaluation of immunogenicity and safety, underlying the utmost importance of advanced computational tools in achieving the highest specificity, efficacy, and safety of the vaccine candidate. Using several bioinformatics databases and tools, the researchers have, therefore, succeeded in finding and mapping the conserved epitopes among the different genotypes of HPVs, predicting their possible immunogenicity potential, and evaluating the structural stability and safety profile of the vaccine constructs at the in-silico level before actual in-vivo or ex-vivo experimental validations are carried out. It leads to less time, higher success rate, and broader protection in development of multi-epitope HPV vaccines. The refinement and expansion in the repertoire of tools for vaccine research requires the advancement in computational vaccinology, which guarantees the accuracy, more speed, and cost-effectiveness. The pipelines covered in this scoping review article are critical in the development of a multi-epitope HPV vaccine and carry implications for the design and optimization of vaccines against a large variety of HPVs.

Acknowledgments

There is no financial support and conflict of interest in this study.

Authors' Contributions

A.N.; Investigation, Methodology,Validation, Resources, Data Curation, and Writing-original draft. M.H.E.; Conceptualization, Methodology, Project administration, and Supervision. M.H.A., M.F.; Visualization and Investigation. A.N., H.E.G.G.; Writing-reviewing and Editing. All authors read and approved the final manuscript.

References

1. Shahab M, Guo D, Zheng G, Zou Y. Design of a novel and potent multi-epitope chimeric vaccine against human papillomavirus (HPV): an immunoinformatics approach. Biomedicines. 2023; 11(5): 1493.
2. Kombe Kombe AJ, Li B, Zahid A, Mengist HM, Bounda GA, Zhou Y, et al. Epidemiology and burden of human papillomavirus and related diseases, molecular pathogenesis, and vaccine evaluation. Front Public Health. 2021; 8: 552028.
3. Negahdaripour M, Nezafat N, Eslami M, Ghoshoon MB, Shoolian E, Najafipour S, et al. Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. Infect Genet Evol. 2018; 58: 96-109.
4. Kumar A, Sahu U, Kumari P, Dixit A, Khare P. Designing of multi-epitope chimeric vaccine using immunoinformatic platform by targeting oncogenic strain HPV 16 and 18 against cervical cancer. Sci Rep. 2022; 12(1): 9521.
5. Mittal S, Banks L. Molecular mechanisms underlying human papillomavirus E6 and E7 oncoprotein-induced cell transformation. Mutat Res Rev Mutat Res. 2017; 772: 23-35.
6. NetMHCpan Server. Version 4.1.1. 2021. Available from: http://www.cbs.dtu.dk/services/NetMHCpan/ (21 Mar 2024).
7. BepiPred Server. Version 2.0. Available from: http://www.cbs.dtu.dk/services/BepiPred/ (21 Mar 2024).
8. National Center for Biotechnology Information (NCBI). 1988. Available from: https://www.ncbi.nlm.nih.gov/ (06 Apr 2017).
9. I-TASSER Server. Version 5.1. 2016. Available from: https://zhanglab.ccmb.med.umich.edu/I-TASSER/ (21 Mar 2024).
10. AllerTop Server. Version 2.0. 2014. Available from: http://www.ddg-pharmfac.net/AllerTOP/ (21 Mar 2024).
11. UniProt Consortium. UniProt: a worldwide hub of protein knowledge. Nucleic Acids Res. 2019; 47(D1): D506-D515.
12. Van Doorslaer K, Li Z, Xirasagar S, Maes P, Kaminsky D, Liou D, et al. The Papillomavirus Episteme: a major update to the papillomavirus sequence database. Nucleic Acids Res. 2017; 45(D1): D499-D506.
13. Andreatta M, Nielsen M. Gapped sequence alignment using artificial neural networks: application to the MHC class I system. Bioinformatics. 2016; 32(4): 511-517.
14. ABCPred Server . Version 2.0. Chandigarh, India: Institute of Microbial Technology. 2013. Available from: http://www.imtech.res.in/raghava/abcpred/ (21 March 2024).
15. Haste Andersen P, Nielsen M, Lund O. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. Protein Sci. 2006; 15(11): 2558-2567.
16. Ponomarenko J, Bui HH, Li W, Fusseder N, Bourne PE, Sette A, et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. BMC Bioinformatics. 2008; 9: 514.
17. IFNepitope Server. Version 1.0. 2019. Available from: http://crdd.osdd.net/raghava/ifnepitope/ (21 Mar 2024).
18. Immune Epitope Database . Version 3.0. 2008. Available from: https://www.iedb.org/ (21 March 2024).
19. CLC Sequence Viewer Software . Version 8.0. 2019. Available from: https://www.qiagenbioinformatics.com/products/clc-sequence-viewer/ (21 March 2024).
20. Clustal Omega Software. Version 1.2.4. 2011. Available from: https://www.ebi.ac.uk/Tools/msa/clustalo/ (21 Mar 2024).
21. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 2004; 32(5): 1792-1797.
22. Kumar S, Stecher G, Li M, Knyaz C, Tamura K. MEGA X: Molecular evolutionary genetics analysis across computing platforms. Mol Biol Evol. 2018; 35(6): 1547-1549.
23. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, et al. Clustal W and Clustal X version 2.0. Bioinformatics. 2007; 23(21): 2947-2948.
24. BLASTp: Basic Local Alignment Search Tool for proteins. 2023. Available from: https://blast.ncbi.nlm.nih.gov/Blast.cgi (21 Mar 2024).
25. PSIPRED Server. Version 4.0. 2018. Available from: http://bioinf.cs.ucl.ac.uk/psipred/ (21 Mar 2024).
26. Gasteiger E, Hoogland C, Gattiker A, Duvaud Se, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy server. In: Walker JM, editor. The proteomics protocols handbook. Totowa, NJ: Humana Press; 2005; 571-607.
27. Robetta Server. Version 3.0. 2015. Available from: http://robetta.bakerlab.org/ (21 Mar 2024).
28. Software. Version 2.4. 2010. Available from: http://www.bonvinlab.org/software/haddock2.4/ (21 Mar 2024).
29. Wang Q, Canutescu AA, Dunbrack RL Jr. SCWRL and MolIDE: computer programs for side-chain conformation prediction and homology modeling. Nat Protoc. 2008; 3(12): 1832-1847.
30. Boutet E, Lieberherr D, Tognolli M, Schneider M, Bairoch A. UniProtKB/Swiss-Prot. Methods Mol Biol. 2007; 406: 89-112.
31. Phyre2 Server. Version 2.0. 2015. Available from: http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index (21 Mar 2024).
32. RaptorX Server. Version 2.0. 2015. Available from: http://raptorx.uchicago.edu/ (21 Mar 2024).
33. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, et al. UCSF Chimera--a visualization system for exploratory research and analysis. J Comput Chem. 2004; 25(13): 1605-1612.
34. PEP-FOLD Server. Version 5.0. 2019. Available from: https://bioserv.rpbs.univ-paris-diderot.fr/services/PEP-FOLD3/ (21 Mar 2024).
35. Torchala M, Moal IH, Chaleil RA, Fernandez-Recio J, Bates PA. SwarmDock: a server for flexible protein-protein docking. Bioinformatics. 2013; 29(6): 807-809.
36. AutoDock Vina Software. Version 1.1.2. 2011. Available from: http://vina.scripps.edu/ (21 Mar 2024).
37. ClusPro Server. Version 2.0. 2015. Available from: https://cluspro.bu.edu/ (21 Mar 2024).
38. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993; 2(9): 1511-1519.
39. Mahmoudvand S, Esmaeili Gouvarchin Ghaleh H, Jalilian FA, Farzanehpour M, Dorostkar R. Design of a multi-epitope-based vaccine consisted of immunodominant epitopes of structural proteins

of SARS-CoV-2 using immunoinformatics approach. Biotechnol Appl Biochem. 2023; 70(3): 1189-1205.
40. Salomon-Ferrer R, Götz AW, Poole D, Le Grand S, Walker RC. Routine microsecond molecular dynamics simulations with AMBER on GPUs. 2. explicit solvent particle mesh ewald. J Chem Theory Comput. 2013; 9(9): 3878-3888.
41. iMODS Server. Version 1.0. 2012. Available from: http://imods.chaconlab.org/ (21 Mar 2024).
42. Visual Molecular Dynamics Software. Version 1.9.4. 2020. Available from: https://www.ks.uiuc.edu/Research/vmd/ (21 Mar 2024).
43. VaxiJen Server. Version 2.0. 2011. Available from: http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html (21 Mar 2024).
44. ProtParam Server. Version 2.0. 2020. Available from: https://web.expasy.org/protparam/ (21 Mar 2024).
45. Magnan CN, Zeller M, Kayala MA, Vigil A, Randall A, Felgner PL, et al. High-throughput prediction of protein antigenicity using protein microarray data. Bioinformatics. 2010; 26(23): 2936-2943.
46. Gupta S, Kapoor P, Chaudhary K, Gautam A, Kumar R, Raghava GP. Peptide toxicity prediction. Methods Mol Biol. 2015; 1268: 143-157.
47. Rapin N, Lund O, Bernaschi M, Castiglione F. Computational immunology meets bioinformatics: the use of prediction tools for molecular binding in the simulation of the immune system. PLoS One. 2010; 5(4): e9862.
48. AllergenFP Server. Version 1.0. 2017. Available from: https://ddg-pharmfac.net/AllergenFP/ (21 Mar 2024).
49. admetSAR Server. Version 1.0. Shanghai, China: Shanghai Institute of Organic Chemistry. 2018. Available from: http://lmmd.ecust.edu.cn/admetsar2/ (21 Mar 2024).
50. Yazdani Z, Rafiei A, Valadan R, Ashrafi H, Pasandi M, Kardan M. Designing a potent L1 protein-based HPV peptide vaccine: a bioinformatics approach. Comput Biol Chem. 2020; 85: 107209.
51. GeneScript Optimum GeneTM. 2024. Available from: https://www.genscript.com/codon-opt.html (6 Jul 2024).
52. Kaliamurthi S, Selvaraj G, Chinnasamy S, Wang Q, Nangraj AS, Cho WC, et al. Immunomics datasets and tools: to identify potential epitope segments for designing chimeric vaccine candidate to cervix papilloma. Data. 2019; 4(1): 31.
53. Grote A, Hiller K, Scheer M, Münch R, Nörtemann B, Hempel DC, et al. JCat: a novel tool to adapt codon usage of a target gene to its potential expression host. Nucleic Acids Res. 2005; 33(Web Server issue): W526-W531.
54. Sanami S, Rafieian-Kopaei M, Dehkordi KA, Pazoki-Toroudi H, Azadegan-Dehkordi F, Mobini GR, et al. In silico design of a multi-epitope vaccine against HPV16/18. BMC Bioinformatics. 2022; 23(1): 311.
55. Long Q, Wei M, Wang Y, Pang F. Design of a multi-epitope vaccine against goatpox virus using an immunoinformatics approach. Front Cell Infect Microbiol. 2024; 13: 1309096.
56. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A. ExPASy: The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003; 31(13): 3784-3788.
57. Magnan CN, Randall A, Baldi P. SOLpro: accurate sequence-based prediction of protein solubility. Bioinformatics. 2009; 25(17): 2200-2207.
58. GalaxyRefine and ERRAT Servers. Version 2.0. Seoul, South Korea: Seoul National University. 2019. Available from: http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=ERRAT (21 Mar 2024).
59. Laskowski RA, MacArthur MW, Moss DS, Thornton JM. PROCHECK: a program to check the stereochemical quality of protein structures. J Appl Cryst. 1993; 26(2): 283-291.
60. Saber AM, Aghamollaei H, Esmaeili Gouvarchin Ghaleh H, Mohammadi M, Sehri SY, Farnoosh G. Design and production of a chimeric enzyme with efficient antibacterial properties on staphylococcus aureus. Int J Pept Res Ther. 2024; 30(1): 7.
61. Jabbar B, Rafique S, Salo-Ahen OMH, Ali A, Munir M, Idrees M, et al. Antigenic peptide prediction from E6 and E7 oncoproteins of HPV types 16 and 18 for therapeutic vaccine design using immunoinformatics and MD simulation analysis. Front Immunol. 2018; 9: 3000.
62. Kumar A, Yadav IS, Hussain S, Das BC, Bharadwaj M. Identification of immunotherapeutic epitope of E5 protein of human papillomavirus-16: an in silico approach. Biologicals. 2015; 43(5): 344-348.
63. Sarkar B, Ullah MA, Araf Y. A systematic and reverse vaccinology approach to design novel subunit vaccines against Dengue virus type-1 (DENV-1) and human Papillomavirus-16 (HPV-16). Inform Med Unlocked. 2020; 19: 100343.
64. Bahmani B, Amini-Bayat Z, Ranjbar MM, Bakhtiari N, Zarnani AH. HPV16-E7 protein T cell epitope prediction and global therapeutic peptide vaccine design based on human leukocyte antigen frequency: an in-silico study. Int J Pept Res Ther. 2021; 27(1): 365-378.
65. Namvar A, Bolhassani A, Javadi G, Noormohammadi Z. In silico/In vivo analysis of high-risk papillomavirus L1 and L2 conserved sequences for development of cross-subtype prophylactic vaccine. Sci Rep. 2019; 9(1): 15225.
66. Sabah SN, Gazi MA, Sthity RA, Husain AB, Quyyum SA, Rahman M, et al. Designing of epitope-focused vaccine by targeting E6 and E7 conserved protein sequences: an immuno-informatics approach in human papillomavirus 58 isolates. Interdiscip Sci. 2018; 10(2): 251-260.
67. He J, Yang Y, Chen Z, Liu Y, Bao S, Zhao Y, et al. Identification of variants and therapeutic epitopes in HPV-33/HPV-58 E6 and E7 in Southwest China. Virol J. 2019; 16(1): 72.
68. Soumia M, Hajji H, El Mzibri M, Younes FZ, Mohammed B, Mohamed B, et al. In-Silico molecular modeling studies to identify novel potential inhibitors of HPV E6 protein. Vaccines (Basel). 2022; 10(9): 1452.
69. Bahmani B, Amini-bayat Z, Ranjbar MM, Masjedi MR, Zarnani AH. An immunoinformatic approach for a therapeutic epitope-based peptide vaccine design by targeting HPV16-E6: in silico study. Microbiology, Metabolites and Biotechnology. 2021; 4(1): 49-64.
70. Rezaei F, Namvar A, Akbari E, Heidarnejad F, Bolhassani A. Immunoinformatics studies of heat shock proteins 27 and 70: Development of potent therapeutic vaccine constructs against human papillomavirus-related cancers. Heliyon. 2023; 9(8): e19261.
71. Li B, Zheng X, Hu C, Cao Y. Human Papillomavirus genome-wide identification of T-Cell epitopes for peptide vaccine development against cervical cancer: an integration of computational analysis and experimental assay. J Comput Biol. 2015; 22(10): 962-974.
72. Mohamed SS, Osman MM, Sami SM, Elgailany HA, Bushara MK, Hassan HA, et al. Immunoinformatics approach for designing epitope-based peptides vaccine of L1 major capsid protein against HPV type 16. Int J Multidiscip Curr Res. 2016; 4: 879-896.
73. Namvar A, Panahi HA, Agi E, Bolhassani A. Development of $HPV_{16,18,31,45}$ E5 and E7 peptides-based vaccines predicted by immunoinformatics tools. Biotechnol Lett. 2020; 42(3): 403-418.
74. Dehghani B, Hasanshahi Z, Hashempour T, Motamedifar M. The possible regions to design Human Papilloma Viruses vaccine in Iranian L1 protein. Biologia (Bratisl). 2020; 75(5): 749-759.
75. Kayyal M, Bolhassani A, Noormohammadi Z, Sadeghizadeh M. In silico design and immunological studies of two novel multiepitope DNA-based vaccine candidates against high-risk human papillomaviruses. Mol Biotechnol. 2021; 63(12): 1192-1222.
76. Mahmoudvand S, Shokri S, Makvandi M, Taherkhani R, Rashno M, Jalilian FA, et al. In silico prediction of T-cell and B-cell epitopes of human papillomavirus type 16 L1 protein. Biotechnol Appl Biochem. 2022; 69(2): 514-525.
77. Stolbikov AS, Salyaev RK, Rekoslavskaya NI. A bioinformatics approach for identifying the probable cause of the cross-interaction of antibodies to the antigenic protein HPV16 L1 with the HPV6 L1 protein. Vavilovskii Zhurnal Genet Selektsii. 2021; 25(7): 787-792.
78. Firdaus MER, Mustopa AZ, Triratna L, Syahputra G, Nurfatwa M. Dissection of capsid protein HPV 52 to rationalize vaccine designs using computational approaches immunoinformatics and molecular docking. Asian Pac J Cancer Prev. 2022; 23(7): 2243-2253.
79. Elshafei AM, Mahmoud NA, Almofti YA. Development of multi-epitopes vaccine against human papilloma virus16 using the L1 and L2 proteins as immunogens. Biosci Biotech Res Asia. 2022; 19(3): 797-813.
80. Karimah N, Sulfianti A, Nurhasanah A. A bioinformatic approach towards designing a human papillomavirus vaccine based on L1 capsid protein sequence of HPV45. Indian J Biochem Biophys. 2022; 59: 927-935.
81. Sulfianti A, Karimah N, Nurhasanah A. In silico analysis of HLA-1 and HLA-2 recognition of a designed recombinant human papillomavirus vaccine based on L1 protein HPV subtype 45. J Genet Eng Biotechnol. 2023; 21(1): 167.
82. Tîrziu A, Avram S, Madă L, Cri-an-Vida M, Popovici C, Popovici D, et al. Design of a synthetic long peptide vaccine targeting HPV-16 and -18 using immunoinformatic methods. Pharmaceutics. 2023; 15(7): 1798.
83. Negahdaripour M, Eslami M, Nezafat N, Hajighahramani N, Ghoshoon MB, Shoolian E, et al. A novel HPV prophylactic peptide vaccine, designed by immunoinformatics and structural vaccinology approaches. Infect Genet Evol. 2017; 54: 402-416.
84. Kaliamurthi S, Selvaraj G, Chinnasamy S, Wang Q, Nangraj AS,

Cho WC, et al. Exploring the papillomaviral proteome to identify potential candidates for a chimeric vaccine against cervix papilloma using immunomics and computational structural vaccinology. Viruses. 2019; 11(1): 63.
85. Mashhadi Abolghasem Shirazi M, Arashkia A, Haghighat S, Roohvand F, Sadat SM. Immunoinformatics approach for designing an HPV epitope-based vaccine candidate harboring built-in adjuvants. Vac Res. 2021; 8(2): 28-38.
86. Dorosti H, Eskandari S, Zarei M, Nezafat N, Ghasemi Y. Design of a multi-epitope protein vaccine against herpes simplex virus, human papillomavirus and Chlamydia trachomatis as the main causes of sexually transmitted diseases. Infect Genet Evol. 2021; 96: 105136.
87. Dharmawan MA, Ansori ANM, Dian FA, Probojati RT, Tamam MB, Kharisma VD. Vaccine construction for human papilloma virus (HPV) type 16 and 18 Infection using in silico approach to combat cervical cancer. Genbinesia J Biol. 2021; 1(1): 13-22.
88. Gupta B, Kumar A, Sridevi P. A comprehensive in silico analysis for identification of immunotherapeutic epitopes of HPV-18. Int J Pept Res Ther. 2021; 27(4): 2717-2726.
89. Mashhadi Abolghasem Shirazi M, Sadat SM, Haghighat S, Roohvand F, Arashkia A. Alum and a TLR7 agonist combined with built-in TLR4 and 5 agonists synergistically enhance immune responses against HPV RG1 epitope. Sci Rep. 2023; 13(1): 16801.
90. Farzanehpour M, Shahriary A, Dorostkar R, Hamidinejad MA, Milani Fard AM, Esmaeili Gouvarchin Ghaleh H. New vaccine technologies for rapid response against emerging, reemerging infections and biological threats: lessons from COVID-19 for future. J Appl Biotechnol Rep. 2023; 10(1): 876-887.