Comprehensive Insights into HPV 16 Immunity: From T Cell Activation and Proliferation to Vaccine Design and Drug Discovery

Author's Name: Ehsan Soleymaninejadian

Mentored by: Professor Fausto Baldanti

Institution: University of Pavia

Department: Department of Biology and Biotechnology and Department of Clinical, Surgical, Diagnostics and Pediatric Sciences

Date: 31/12/2023

Dedications:

In the pursuit of knowledge, there are those whose guidance and support illuminate the path, making the journey not only possible but enriched with profound experiences. To these exceptional individuals, I dedicate this thesis.

Professor Baldanti Fausto

My sincere gratitude goes to Professor Baldanti Fausto, whose scholarly wisdom and unwavering support have been the foundation upon which this research stands. His guidance has been invaluable, shaping both my scientific understanding and approach to academia.

Dr. Daniele Lilleri and Dr. Irene Cassaniti

I extend my appreciation to Dr. Daniele Lilleri and Dr. Irene Cassaniti for their insightful discussions, encouragement, and the invaluable role they played in shaping my research journey.

Dr. Paola Zelini

To Paola Zelini, whose expertise and dedication in the laboratory were instrumental in honing my technical skills. I am grateful for her patience and generosity in sharing her knowledge.

Federica Bergami

I acknowledge Federica Bergami for her brief yet impactful contribution to my training. Her guidance added depth to my understanding of essential techniques.

Professor Mario Mondelli, Dr. Stefania Varchetta, and Dr. Dalila Mele

The time spent in Professor Mario Mondelli's lab was transformative, thanks to the guidance and support of Dr. Stefania Varchetta and Dr. Dalila Mele. Their encouragement instilled confidence in handling samples, antibodies, and assays.

Professor Davide Sassera

Special appreciation to Professor Davide Sassera, whose role as the coordinator of PhD students went beyond academic responsibilities. His generosity, including support for accommodation, reflects a commitment to the well-being of those under his guidance.

Professor Maurizio Zuccotti

I express gratitude to Professor Maurizio Zuccotti for his initial support and kindness, creating a welcoming environment during my early days in Pavia. His hospitality and friendship were cherished.

Nahid Raeisikia

To my beloved wife Nahid Raeisikia, whose unwavering support and understanding transcended the distance that separated us. Her patience and sacrifices allowed me to focus on reaching my academic goals.

Soraya Ghorbani

A heartfelt thank you to my mother, Soraya Ghorbani, who has been a constant source of support throughout my life. Her teachings of resilience and courage have shaped my journey, inspiring me to look forward and take each step, no matter how small.

Neda Soleymaninejadian

I am indebted to my sister, Neda Soleymaninejadian, for her unwavering support, both mentally and economically. Her generosity, including financial assistance during challenging times, has been a lifeline.

In this collective journey, I am grateful to these remarkable individuals for their influence, guidance, and unwavering support. This thesis stands as a testament to the collaborative spirit that defines the academic community.

Ehsan Soleymaninejadian

Table of Contents

Table of contents

Table of contents

Chapter 1: Abstract

Abstract:

This comprehensive thesis delves into the intricacies of the T cell immune response against human papillomavirus (HPV) 16 antigens (E6, E7, and L1) within both CD8 and CD4 T cell subsets. A particular emphasis is placed on exploring potential cross-reactivity, extending the investigation to include HPV18. The evaluation of traditional vaccines, Cervarix and Gardasil, is crucial to understanding their efficacy.

Utilizing advanced techniques such as AIM and LPA assays, the study meticulously analyzes T cell activation and proliferation. The research scope encompasses diverse participant groups, including HPV-vaccinated and non-vaccinated individuals, as well as those at different stages of HPV-induced cervical cancer.

A focal point of the investigation is on CD4 T cells expressing CXCR5, known for their pivotal role in B cell antibody production. The study aims to deepen this understanding through experimental analyses and immunoinformatics, culminating in the design of a sophisticated multi-epitope vaccine.

System immunology takes center stage, exploring gene expression changes in PBMCs before and after vaccination. Network analysis aims to identify key genes critical for an effective immune response. Additionally, drug design strategies, employing QSAR and pharmacophore modeling, seek to modulate the target gene, potentially enhancing vaccine efficacy.

The study's multifaceted objectives encompass T cell responses, follicular T cell dynamics, HPV clade differences, immunoinformatics-based vaccine design, gene expression alterations, target gene drugability, and computational drug design. Key findings underscore the importance of vaccination regimen optimization, provide rationale for L1 antigen selection, elucidate the representative roles of HPV 16 and 18, refine antigenic focus, showcase multi-epitope vaccine development, and offer insights from immune system simulations and gene expression dynamics.

Chapter 2: Abbreviations

Abbreviations

Absorption, Distribution, Metabolism, And Excretion	ADME
Absorption, Distribution, Metabolism, Excretion, And Toxicity	ADMET
Activation-Induced Marker	AIM
Adenine	A
Alanine	A Ala
Algorithm for the Prediction of Allergenic Proteins	AlgPred
Allophycocyanin	APC
An Integrated Method for The Prediction of T Cell Epitopes	NetTepi
Analysis Of Variance	ANOVA
Anogenital Warts	AGWs
Antigen Database	Antigen DB
Antigen-Induced Proliferation	AIM
Antigen-Presenting Cells	APCs
Arginine	R Arg
Asparagine	N Asn
Aspartic Acid	D Asp
Aspartic Acid or Asparagine	B Asx

Basic Local Alignment Search Tool	BLAST
Bayes Factors	BF10
B-Cell Epitope Prediction	Bcepred
Beta-Mercaptoethanol	β-mercaptoethanol
Binding Database	Binding DB
Blood-Brain Barrier	BBB
Burkitt Lymphocyte Receptor 1	BLR1
Cama-Separated Values	CSV
Canadian Organization for The Advancement of Computers In Health	COACH
Carbon Dioxide	CO2
CC Chemokine Receptor	CCR6
Cervical Intraepithelial Neoplasia	CIN
Chemistry At Harvard Macromolecular Mechanics	CHARMM
Chemokine	C-C motif ligand 20
Chemokine (C-C Motif) Ligand 20	CCL20
C-ImmSim Server	C-ImmSim
Cluster of Differentiation	CD
Codon Adaptation Index	CAI
Codon Adaptation Index	CAI

Computed Atlas of Surface Topography of Proteins	CASTp
Confidence Interval	CI
Consensus Method for Prediction of Catalytic Residues	CSmetaPred
Coronavirus	COVID-19
Cross-Terms Map	CMAP
C-X-C Chemokine Receptor 5	CXCR5
Cysteine	C Cys
Cytosine-Phosphate-Guanine	CpG
Cytotoxic T Cells	CTCs
Cytotoxic T-Cell Lymphocytes	CTLs
Database Integrated Servers Using Constraints	DiscoTop
Database Of Genotypes and Phenotypes	dbGaP
Deoxyribonucleic Acid	DNA
Disease Gene Network	DisGeNet
Disease-Modifying Anti-Rheumatic Drugs	DMARDs
Early Proteins	E
Ethylenediaminetetraacetic Acid	EDTA
European Molecular Biology Open Software Suite	EMBOSS
Fast-All	FASTA

Fetal Calf Serum	FCS
G Protein-Coupled Receptor 9	GPR9
Gamma Delta T Cells	γδ T cells
Gastrointestinal	GI
Gene Expression Omnibus	GEO
Glutamic Acid	E Glu
Glutamic Acid or Glutamine	Z Glx
Glutamine	Q Gln
Glycine	G Gly
Grand Average of Hydropathy	GRAVY
Groningen Machine for Chemical Simulations	GROMACS
Guanine	G
Half-Maximal Inhibitory Concentration	IC50
Head And Neck Cancers	HNCs
Head And Neck Squamous Cell Carcinoma	HNSCC
Helper T Cell	Th or TH
Helper T Lymphocyte	HTL
High-Risk	HR
Histidine	H His

Honestly Significant Difference	HSD
Hugo Gene Nomenclature Committee	HGNC
Human Immunodeficiency Virus	HIV
Human Leucocyte Antigen	HLA
Human Papillomavirus	HPV
Hydrogen Bond Acceptors	HBA
Hydrogen Bond Donors	HBD
interferon	IFN
Immunoglobulin G	IgG
Indoleamine 2,3-Dioxygenase	IDO-1
Inducible T-Cell CO-Stimulator	ICOS
Interactive Tree of Life	iTOL
Interferon-Gamma	IFN-γ
Interleukin	IL
Interleukin-1 Receptor-Associated Kinase	IRAK
Isoleucine	I Ile
Janus Kinase	JAK
Java Codon Adaptation Tool	JCat
Kilodaltons	KD

Kyoto Encyclopedia of Genes and Genomes	KEGG
Late Protein	L
Leucine	L Leu
Linear Constraint Solver	LINCS
Log2(FC)	LogFC
Low-Risk	LR
Lymphocyte Proliferation Assay	LPA
Lysine	K Lys
Major Histocompatibility Complex	MHC
Methionine	M Met
MHC Class I and II Binding Peptide Prediction Server	ProPred I and II
Minimal Detectable Change	MDC
Minimum Essential Medium	MEM
Molecular Dynamics	MD
Molecular Graphics Laboratory	MGL
National Center for Biotechnology Information	NCBI
Natural Killer	NK
Net Major Histocompatibility Complex Class II Pan-Specific Binding Prediction	NetMHCIIpan
Non-vacs	Non vaccinated

Nuclear Factor Kappa-Light-Chain-Enhancer of Activated B Cells	NF-κB
Nuclear Magnetic Resonance Spectrometer	NMR
Number Of Particles, System Pressure, Temperature	NPT
Number Of Particles, System Volume, Temperature	NVT
Odds Ratios	OR
Oligodeoxynucleotides	ODN
Online Mendelian Inheritance in Man	OMIM™
Open Reading Frames	ORFs
Optimized Potentials for Liquid Simulations	OPLS
Oral Squamous Cell Carcinoma	OSCC
Papillomavirus Episteme	PaVE
Paraformaldehyde	PF
Partial Least Square Regression	PLS
Particle Mesh Ewald	PME
Pathogen-Associated Molecular Patterns	PAMPs
PDZ-Binding Motifs	PBMs
Penicillin-Streptomycin-Glutamine	PSG
Peptide Structure Prediction Server	PEP-FOLD
Peripheral Blood Mononuclear Cells	PBMC

Phenylalanine	F Phe
Phosphate-Buffered Saline	PBS
Phycoerythrin	PE
Phycoerythrin-Cyanine 5	PECY5
Pocket Druggability Prediction	PockDrug
Polar Surface Area	PSA
Postsynaptic Density Protein 95	PSD95
Prediction Of Protective Antigens and Subunit Vaccines	VaxiJen
Programmed Cell Death Protein-1	PD-1
Proline	P Pro
Protein Data Bank	PDB
PSD95-Disc Large-Zonula Occludens	PDZ
P-Type Adenosine Triphosphatase	ATPases
Pubchem Compound Identification	CID
Public Chemical Database	PubChem
Quantitative Structure-Activity Relationships	QSAR
Recurrent Respiratory Papillomatosis	RRP
Reference Number	REF
Regulatory T Cells	Treg

Ribonucleic Acid	RNA
Root Mean Square Deviation	RMSD
Root Mean Square Fluctuation	RMSF
Roswell Park Memorial Institute	RPMI
Search Tool for Interactions of Chemicals	STICH
Search Tool for Retrieval of Interacting Genes/Proteins	STRING
Selenocysteine	U* Sec
Serine	S Ser
Simplified Molecular Input Line Entry System	SMILES
Sodium Pyruvate	NaPyruvate
Solvent Accessible Surface Area	SASA
Standard Error of Measurement	SEM
Standardized Mean Differences	SMD
Stepwise Multiple Linear Regression	SMLR
Structure-Data File	SDF
T Cell Receptors	TCRs
T Follicular Helper	TFH or Tfh
Tab-Separated Values	TSV
The Food and Drug Administration	FDA

The Immune Epitope Database	IEDB
The Long Control Region	LCR
The Molecular Evolutionary Genetics Analysis	MEGA
The Self-Optimized Prediction Method with Alignment	SOPMA
Theoretical Isoelectric Point	pI
Three-Site Model	TIP 3
Threonine	T Thr
Thymine	T
Toll-Like Receptors	TLRs
Transforming Growth Factor-Beta	TGF-β
Tryptophan	W Trp
Tumor Necrosis Factor Alpha	TNF-α
Tumor Necrosis Factor Ligand Superfamily Member 9	4-1BB or CD137
Tumor-Necrosis Factor Receptor	TNFR
Tyramide Signal Amplification	TSA
Tyrosine	Y Tyr
Unknown Or 'Other' Amino Acid	X** Xaa
Upstream Regulatory Region	URR
Uracil	U

Vacs	Vaccinated
Valine	V Val
Viral Protein Interaction Database	VirusMINT
Virus-Like Particles	VLPs
Web Server Designed to Predict Protein Allergenicity Potential	AllerCatPro
ZINC Is Not Commercial	ZINC

Chapter 3: Introduction

Introduction

Human papillomavirus (HPV) stands as the prevailing sexually transmitted infection (STI) on a global scale, and it largely operates in a subclinical manner, often without symptomatic manifestation among infected individuals [1]. Nevertheless, when diagnosed, either through clinical tests or the presentation of symptoms, patients encounter multifaceted challenges, including not only the physical aspects of treatment but also the intricate psychosocial dimensions that accompany the condition. These psychosocial ramifications can profoundly impact the patient's well-being, often impeding both the monitoring of the disease's progression and the implementation of appropriate therapeutic interventions.

To gauge the psychosocial effects of HPV comprehensively, researchers have employed tools such as the Psycho-Estampa Scale. This scale serves as a valuable instrument for quantifying the psychosocial burden associated with HPV infection. Notably, an escalation in Psycho-Estampa Scale scores correlates significantly with an increased likelihood of abnormal cell proliferation and the presence of cancer-associated HPV strains. Furthermore, it has been observed that individuals with elevated scores on the Psycho-Estampa Scale exhibit a higher tendency to discontinue follow-up examinations and treatments [2].

HPV demonstrates a remarkable capacity to infect both cutaneous and mucosal epithelial tissues, including the cervical and anogenital mucosae. The virus's replication cycle is intricately intertwined with the process of epithelial differentiation. On a global scale, the lifetime risk of HPV infection for both men and women stands at a staggering 50% [3]. Importantly, HPV infections tend to persist over extended periods, potentially giving rise to chronic lesions that can culminate in malignancy [4].

Among the spectrum of HPV-associated cancers, cervical cancer predominates as the most prevalent, constituting a substantial 84% of HPV-related cancer cases worldwide [5]. However, an intriguing trend has emerged in high-income countries, notably the United States, where the incidence of HPV-related head and neck squamous cell carcinoma (HNSCC) is on a marked upward trajectory relative to HPV-related cervical cancer [6]. Additionally, HPV has been linked to the development of anogenital cancers, including vulval, vaginal, and penile cancers [7]. Notably, a connection between HPV infection and breast cancer has also been suggested, with HPV prevalence observed in breast cancer patients ranging from 4% to 86%, with specific emphasis on the prevalence of HPV-16 and 18 exceeding 50% among Iranian women with breast cancer [8–10].

The intricate interplay between the immune system and HPV infection is a pivotal determinant of disease outcome. An appropriately timed and targeted immune response can effectively mitigate post-infection sequelae, while impaired or overly aggressive immune responses can precipitate severe pathological consequences [11]. Furthermore,

HPV-related lesions and cancers exhibit heightened prevalence among immunocompromised individuals. Approximately 30% of women receiving immunosuppressive medications or grappling with autoimmune conditions harbor high-risk HPV strains, particularly HPV-16 [12]. For instance, in the context of Crohn's disease management, the administration of gut-selective immunosuppressive agents, such as Vedolizumab, over an extended duration has been shown to foster HPV replication. Notably, a case study illuminated the progression from HPV-16 replication in a woman with Crohn's disease to severe vulvar intraepithelial neoplasia, with subsequent observations of intraepithelial neoplasia occurring in novel locations, including the cervix [13].

Moreover, individuals diagnosed with the human immunodeficiency virus (HIV) are at heightened susceptibility to HPV-related cervical cancers. The risk ratio for cervical cancer in HIV-infected patients exceeds six, with approximately 6% of cervical cancer cases in 2018 being HIV-positive [14].

In light of these complex interactions between HPV and the immune system, this comprehensive review endeavors to provide an overview of recent publications related to the immunology of HPV. Furthermore, it explores potential immunotherapeutic strategies that hold promise for enhancing the overall management of HPV-associated cancers, aiming to shed light on innovative approaches to combat this prevalent and impactful group of diseases.

Structure of Human Papillomavirus (HPV)

Human Papillomavirus (HPV) is a member of the Papillomaviridae family, characterized by a double-stranded, closed, circular DNA genome of approximately 8 kb and a nonenveloped icosahedral capsid [15]. The HPV genome comprises 8 protein-coding genes organized into three functional regions: the upstream regulatory region (URR), early (E), and late (L) regions.

The URR, also known as the long control region (LCR), is situated between the L1 and E6 open reading frames (ORFs). It houses the early promoter and regulatory elements crucial for viral DNA replication and transcription. The early region encodes proteins E1, E2, E4, E5, E6, and E7, which are central to viral gene expression, replication, and survival. In contrast, the late region contains genes responsible for the viral envelope and the structural proteins of the capsid (L1 and L2).

Methylation of the LCR plays a pivotal role in the expression of HPV genes. Analysis of cervical samples from HPV-16-infected patients, including those with cervical cancer, intraepithelial neoplasia, and asymptomatic individuals, underscores the significance of LCR methylation in regulating HPV-16 gene expression. Methylation events occurring on promoter and enhancer regions of HPV-16 genes hold clinical and pathological relevance. LCR methylation is prevalent in cervical carcinoma (over 80%) and asymptomatic patients (around 70%), with a prevalence of approximately 40% in intraepithelial neoplasia patients. Moreover, HPV-16 LCR CpG islands predominantly

exhibit methylation on gene promoters [16]. This methylation pattern can serve as a biomarker for monitoring the progression of HPV-16-associated cancer.

The major viral oncoproteins, namely E5, E6, and E7, play pivotal roles in the initiation and progression of cancer by modulating cell cycle regulation [15]. The expression of E6 and E7 proteins is closely associated with the integration of viral DNA into the host genome, driving malignant transformation and ultimately cancer development [17]. These oncoproteins, E6 and E7, are the primary agents responsible for viral transformation and possess the ability to evade immune responses by manipulating cytokine expression, thereby influencing cell proliferation and interferon responses [18].

HPVs are classified into five genera (α, β, γ, μ, and ν), with over 200 types identified [19–21]. Among these, the α group is the largest, comprising 64 HPVs, primarily infecting mucosal epithelia. A subset of 17 mucosal HPVs (16, 18, 23, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 66, 68, 73, and 82) is categorized as high-risk (HR) HPV types. Conversely, 10 other HPV types (40, 42, 43, 44, 53, 54, 61, 72, 73, and 81) are classified as low-risk (LR) or non-oncogenic, typically causing benign lesions [22]. HR HPV types are implicated in various cancers, including those affecting the cervix, vagina, vulva, anus, penis, and a subset of head and neck cancers (HNCs), particularly oropharyngeal cancers [23, 24]. In contrast, LR HPV types are associated with anogenital warts (AGWs), certain cutaneous warts [25], and recurrent respiratory papillomatosis (RRP) [26].

The β group of HPVs mainly infects cutaneous epithelia and encompasses 50 characterized types. HPVs from the remaining three groups (γ, μ, and ν) typically induce benign diseases [21]. This diverse classification underscores the multifaceted nature of HPV infections and their significant impact on human health.

HPV Life Cycle

HPV exhibits a versatile life cycle, capable of infecting cutaneous epithelial cells or mucosal tissues, categorizing them into either cutaneous or mucosal types based on their tropism [27]. The journey of the viral particle begins as it reaches epithelial cells through micro-wounds in the tissue or by interacting with cell surface receptors, such as integrin α6, commonly found in basal cells and epithelial stem cells [28]. Subsequently, the viral protein L1 binds to cellular receptors, undergoing structural modifications that are essential for the endocytosis of the virion [29, 30]. During this transition along the endosomal pathway, L1 dissociates from the viral genome, and the protein L2 mediates the viral egress from the endosomes. This process guides HPV vesicles along microtubules into the nucleus. Once inside the dividing cells of the basal layer, viral early transcription is initiated, marked by the expression of the early proteins E1 and E2 [31]. The productive life cycle of HPV can be categorized into three distinct phases: establishment, maintenance replication, and vegetative or productive amplification [32].

The establishment phase revolves around viral transcription and genome amplification, following nuclear entry. Viral genomes remain episomal within host cells for extended periods. During this phase, early viral proteins E1, E2, and E4 play essential roles in

increasing viral genome replication. Simultaneously, E6 and E7 promote host cell proliferation while preventing apoptosis, as E2 limits the expression of E6 and E7. The loss of E2 repression function results in the deregulation of viral E6 and E7 oncogenes [33].

Following the establishment phase, the maintenance phase commences. In this phase, the virus works to maintain a constant number of viral genomes, establishing a persistent infection. Finally, the last step involves vegetative or productive viral replication, resulting in the production of progeny virions [34, 35]. During this stage, oncoproteins E6 and E7, expressed at relatively low levels in differentiated cells, play a crucial role. They inactivate tumor suppressor proteins like p53 and retinoblastoma protein (pRb) while activating signal transduction pathways, ensuring that infected cells remain active and progress to the S phase of the cell cycle. Moreover, vegetative amplification, beyond increasing HPV genome copy numbers, is accompanied by the expression of structural proteins L1 and L2, as depicted in Figure 1 . This multifaceted life cycle underscores the intricacies of HPV infection and its potential consequences:

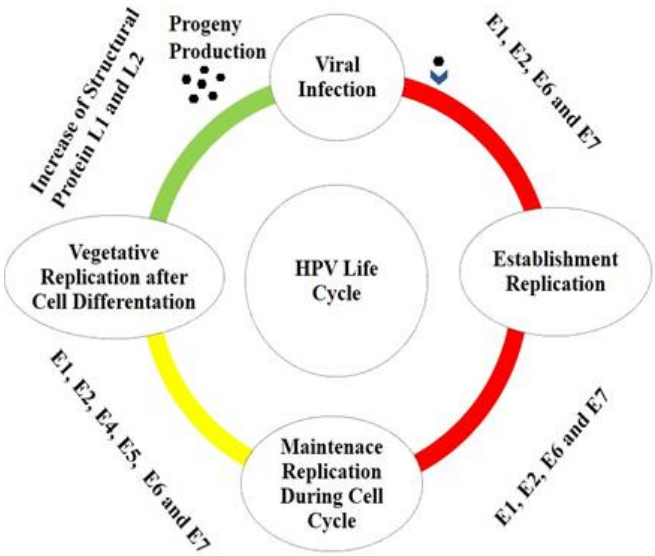


Figure 1 Illustrates the intricate HPV life cycle, showcasing the key stages and processes involved.

Establishment Replication: This initial step of the HPV replication cycle focuses on the maintenance of a consistent number of episomal copies of the viral genome within host cells. To achieve this, a set of essential viral proteins including E1, E2, E6, and E7 come into play. They collaborate to ensure the stable episomal maintenance of HPV.

Maintenance Phase: Following the establishment phase, the maintenance phase takes over. Here, the primary objective is to create conditions that sustain a constant number of

viral genomes within the nuclei of undifferentiated basal cells. This persistence contributes to the establishment of a long-lasting infection, a hallmark of HPV.

Vegetative or Productive Viral Replication: The final phase marks the initiation of vegetative or productive viral replication, culminating in the generation of progeny virions. During this phase, the oncoproteins E6 and E7, expressed at relatively low levels in differentiated cells, assume a pivotal role. They are responsible for inactivating tumor suppressor proteins like p53 and retinoblastoma protein (pRb) while activating signal transduction pathways. This ensures that infected cells remain active and progress to the S phase of the cell cycle [36].

This comprehensive depiction of the HPV life cycle underscores the multifaceted nature of the virus's replication strategy, with each phase contributing to its persistence and potential impact on host cells.

HPV Association with Cancers

The role of Human Papillomavirus (HPV) in cancer development is a topic of critical importance, and this section provides an insightful overview.

HPV Tropism: HPV exhibits the ability to infect both cutaneous epithelial cells and mucosal tissues, as highlighted in Table 1. The diversity in its infection sites yields varying clinical implications.

Cutaneous β-HPV Infections: Often contracted through skin-to-skin contact in young children, cutaneous β-HPV types typically result in asymptomatic infections at cutaneous sites. However, they can occasionally lead to debilitating papillomatosis with an associated risk of cancer [37]. The precise role of cutaneous β-HPV types in cancer development remains somewhat unclear [38].

Mucosal α-HPV Infections: Mucosal infections, primarily associated with α-HPV types, generally occur during the initial sexual exposures in early adulthood, although non-sexual transmission is also possible [39, 40]. α-HPV types are implicated in several forms of anogenital cancers, apart from cervical cancer. These include vulvar, penile, and anal cancers. However, it is noteworthy that the population-level impact of these cancers is relatively small when compared to cervical cancer [41].

Cervical Intraepithelial Neoplasia (CIN): High-risk HPVs (HR-HPVs) play a crucial role in the initiation and progression of cervical intraepithelial neoplasia (CIN) [44,45]. Interestingly, the majority of cervical HPV infections (>90%) are effectively resolved by the host immune system within 1–2 years, without resulting in chronic infection [42]. Nonetheless, a minority of HPV infections persist, significantly elevating the risk of developing epithelial cell abnormalities or cancer. Notably, the persistent and chronic infection of HR-HPV, particularly type 16, stands out as the principal risk factor for the initiation and development of squamous cell carcinoma [43].

This section provides a comprehensive understanding of how HPV infections can lead to various cancers, shedding light on the intricate relationship between HPV types, transmission routes, and cancer development.

Table 1 HPV Types and Their Associated Lesions

HPV Risk Type	HPV Types
High-Risk (HR)	16, 18, 23, 31, 33, 35, 39, 45, 51, 52, 53, 56, 58, 66, 68, 73, and 82
Low-Risk (LR)	40, 42, 43, 44, 53, 54, 61, 72, 73, and 81
Lesions	Intraepithelial neoplasia and cervical cancer (HR)
	Intraepithelial neoplasia, genital or cutaneous warts (LR)

HPV and T Cells: A Complex Interaction

The interplay between Human Papillomavirus (HPV) and T cells is a pivotal aspect of HPV-related cancers, and this section delves into its intricate dynamics.

T Cells in HPV-Related Cancers: T cells emerge as critical players in HPV-related cancers. Approximately 80% of samples obtained from HPV-16-related cervical cancers exhibit T cell infiltration [44]. Additionally, an immunogenomic study involving 119 HPV-positive head and neck squamous cell carcinoma (HNSCC) patients underscores the significance of T cell infiltration within the tumor microenvironment [45]. A case-control study involving Egyptian women emphasizes the predictive value of CD4 and CD8 T cells in HPV-related cancers such as breast and head-neck carcinoma [46]. Interestingly, the absence of co-stimulatory receptors like CD28 on T cells can render HPV more aggressive, a phenomenon witnessed in patients with CD28-deficient T cells, leading to conditions like "tree-man syndrome" due to HPV-2 and HPV-4 infections. In individuals with normal CD28+ T cell frequencies, HPV-2 and HPV-4-related warts struggle to thrive, highlighting the importance of these receptors [47].

T Cell Subpopulations in Different HPV Infections: The composition of T cell subpopulations infiltrating cutaneous and mucosal tissues varies depending on the HPV-infected organs. In HPV-16-infected cervical cancer patients, T CD8 cells predominate, with a concomitant decrease in CD4 T cell frequency [44]. Three subsets of T CD8 cells are active during HPV-related head-neck cancers, all expressing programmed cell death protein-1 (PD-1). These subsets exhibit distinct characteristics, including stem cell-like properties and varying responses to HPV antigens [48]. CD8 T cells expressing CD103

are found within the tumor microenvironment of HPV-related oropharyngeal squamous cell carcinoma (OSCC) and are vital for cancer cell eradication and patient surveillance [49]. Furthermore, stem-like memory CD8 T cells (Tscm) have demonstrated the ability to eliminate HPV-16-induced tumors in vivo and in vitro, displaying heightened responsiveness in the presence of CD40L activation [50].

T Cell Infiltration and HPV: In samples obtained from the epithelial layer, CD8 T cells predominate, while the stromal layer sees a higher frequency of CD4 T cells [51]. Cross-talk between HPV-infected tissues, T cells, and other immune cells within the tumor microenvironment often favors tumor survival. Systematic reviews and meta-analyses have revealed a reduction in the number of T cell subtypes as HPV-infected tissues progress toward cancer, with T regulatory (Treg) cells becoming the predominant population. Treg cells exert their influence by suppressing CD8 T cells, with higher Treg populations observed in common HPV families such as HPV-16 and HPV-18 [46]. Cross-talk between Tim+Treg cells and Galectin-9+ monocytes in HPV-related cervical cancer promotes the secretion of immunosuppressive factors while inhibiting key immune responses, ultimately enhancing tumor aggressiveness and mortality [52]. Strategies to counteract Treg effects include T-win technology, Treg cell receptor inhibitors, immune checkpoint blockers, and photodynamic therapy [53]. Moreover, a substantial portion of CD8 T cells in the infected area exhibit exhaustion, a challenge that can be mitigated by manipulating T cells with PD-1 blockers and Indoleamine 2,3-dioxygenase (IDO-1) inhibitors, thereby enhancing CD8 T cell cytotoxicity [44].

Peripheral T Cell Response: Peripheral T cell responses against HPV-related antigens require further exploration. A cohort study involving oral squamous cell carcinoma (OSCC) patients revealed a robust T central memory response against HPV16 L1 and E6 proteins. However, the response against E7 was deficient due to overexpression of PD-L1 induced by HPV16 E7. This deficiency in peripheral T cell activity against E7 contributes to tumor escape mechanisms, highlighting the potential for new immunotherapeutic approaches The Role of Regulatory T Cells in Pathogenesis and Therapy of Human Papillomavirus-Related Diseases, Especially in Cancer [53].

γδ T Cells in HPV-Related Breast Cancer: In addition to conventional T cells, gamma delta T cells (γδ T cells), particularly a rare subset that produces IL-17, play a pivotal role in the progression of HPV-related breast cancer. This specific population acts as a metastasis promoter by producing high levels of IL-17 [54].

This section unravels the intricate relationship between T cells and HPV-related cancers, shedding light on the diverse T cell subpopulations, their functions, and their impact on tumor progression, Figure 2.

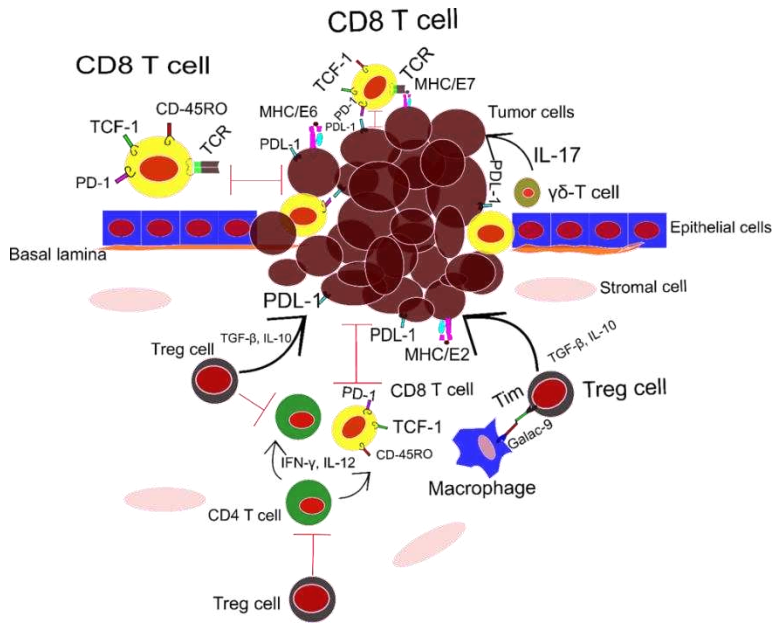


Figure 2 This figure illustrates the key T cell subpopulations involved in HPV-related cancers, highlighting their distinct roles within the tumor microenvironment. CD8 T Cells: Predominantly located in the epithelial layer, CD8 T cells play a pivotal role in recognizing viral antigens presented by MHC-I on tumor cell surfaces. This recognition facilitates the eradication of tumor cells. However, tumor cells expressing PDL-1 receptors interact with PD-1 on CD8 T cells, compromising their function. CD4 T Cells: Found in the stromal layer, the majority of CD4 T cells are regulatory T cells (Treg cells). Treg cells promote tumor development by producing immunosuppressive cytokines, such as IL-10 and TGF-β, which induce tumor tolerance in other T cell types. γδ T Cells: Unconventional γδ T cells are prominent in the epithelial layer and are associated with tumor progression. These cells contribute to tumor advancement by producing IL-17. For a comprehensive understanding of these T cell subsets and their roles in HPV-related cancers, refer to the main text.

Immunoinformatics

Immunoinformatics, a multidisciplinary field at the intersection of immunology and bioinformatics, plays a pivotal role in modern vaccine design and development. By utilizing computational approaches, immunoinformatics accelerates the identification, evaluation, and design of epitopes, which are the critical components of vaccines. It harnesses various tools and databases to discover epitopes, predict their compatibility with the human immune system, and evaluate their safety. Immunoinformatics significantly enhances the speed and accuracy of vaccine development, enabling researchers to create innovative, targeted, and effective vaccines against a wide range of pathogens, ultimately contributing to the advancement of public health by combating infectious diseases more efficiently [55].

Using traditional vaccine development methods, vaccines are typically designed with large proteins. However, using inappropriate antigens in vaccine formulations can increase the risk of allergic reactions. In contrast, a multi-epitope vaccine based on peptides containing short antigenic fragments, known as epitopes, has the potential to overcome these limitations. Epitopes represent the antigenic components of pathogens

that are recognized by the host immune system. Conventional approaches to vaccine development have historically relied on the use of large proteins as the primary antigenic components. These proteins, often derived from the targeted pathogen, serve as the cornerstone of vaccine formulations. While this traditional method has yielded numerous successful vaccines that have protected populations from a wide range of infectious diseases, it is not without its limitations and potential drawbacks. For example, Cervarix and Gardasil utilize the L1 antigen of HPV, which presents certain limitations. L1-based vaccines primarily target specific HPV strains (e.g., types 16 and 18). This limits their efficacy against a broader range of HPV strains, leaving individuals susceptible to non-covered types [56]. Cross-protection against non-vaccine HPV types is limited, contributing to potential gaps in immunity [57].

In recent years, there has been a growing interest in exploring alternative approaches to vaccine design. One such approach involves the use of peptide-based multi-epitope vaccines. These vaccines are distinct from traditional vaccines in that they contain short antigenic peptide fragments, known as epitopes. Epitopes are the minimal units of an antigen that can be recognized by the immune system. By utilizing epitopes, peptide-based vaccines offer a unique set of advantages. Peptide-based vaccines have the potential to overcome some of the limitations associated with traditional vaccines. Their use of short, specific epitopes minimizes the risk of allergic reactions, as the immune response is directed towards well-defined molecular targets [58]. This precision in targeting reduces the likelihood of unwanted immune responses, making peptide-based vaccines a promising avenue for enhancing vaccine safety. Moreover, peptide-based vaccines are highly customizable. Vaccine developers can select epitopes that are unique to the pathogen while excluding those that may be associated with allergic responses. This level of control allows for the creation of safer and more tailored vaccines, reducing the potential for adverse reactions. Designing multi-epitope vaccines using immunoinformatics allows the inclusion of epitopes from various HPV proteins, expanding protection beyond L1. This strategy enhances cross-protection and addresses strain-specific limitations. Immunoinformatics enables in silico analysis to predict antigenic epitopes, assess binding affinities, and optimize vaccine formulations. This accelerates the identification of potential vaccine candidates with broader efficacy [59].

To increase the efficacy of multiepitope vaccines, an adjuvant part is added to main vaccine. For example adding a CpG DNA sequence for binding to TLR9. They play a crucial role in eliciting innate immunity and cell-mediated immune responses. The cell-mediated immune response relies heavily on pattern recognition receptors that detect pathogen-associated molecular patterns (PAMPs) on the pathogen. Within the family of pattern recognition receptors, Toll-like receptors (TLRs) stand out as key players. There are eleven different TLR proteins, each with a unique ability to interact with various PAMPs. These receptors are expressed on the surface of cells and contribute to the detection of pathogens. For instance, Toll-like receptor 9 (TLR9) serves as a critical pattern recognition receptor primarily localized intracellularly within immune cells. Toll-like receptor 9 (TLR9) plays a pivotal role in the realm of pattern recognition receptors, primarily occupying intracellular compartments within a diverse array of immune cells.

These immune sentinels include dendritic cells, macrophages, natural killer cells, and an assortment of antigen-presenting cells (APCs). The principal actors on TLR9's stage are the unmethylated cytidine phosphate guanosine (CpG) oligodeoxynucleotides (ODN). Upon encountering its agonists, TLR9 orchestrates an intricate symphony of molecular events, setting in motion a cascade of inflammatory responses. These responses serve as a powerful call to arms for the immune system, inciting a series of reactions with profound consequences. Among these consequences are heightened phagocytosis, a process by which cells engulf and consume invading microorganisms and malignant cancer cells. This increased phagocytic activity effectively bolsters the body's innate defenses, enhancing its ability to combat microbial intruders and suppress the growth and spread of cancerous cells. Yet, the influence of TLR9 extends beyond the innate immune system's immediate response. TLR9 activation serves as a critical trigger for the development of adaptive immune responses. This means that TLR9's role is not confined to the initial containment and destruction of threats but extends to the orchestration of long-term immunity [60].

The adaptive immune system, a powerful arm of the body's defense mechanisms, relies on memory and specificity to recognize and neutralize pathogens. TLR9, by virtue of its activation, contributes to the generation of adaptive immunity [60]. It assists in the formation of immunological memory, ensuring that the body can recognize and respond more effectively to the same threat upon subsequent encounters. This process is central to the development of vaccines and the body's ability to mount a swift and targeted defense against infections. Following infection, cytotoxic T-cell lymphocytes (CTLs) become activated and target and eliminate infected cells. This immune recognition is made possible by antigens bound to major histocompatibility complex (MHC) molecules, which are presented on the surface of infected cells. The MHC class I molecules display cytosolic peptide antigens from the infected cells, while MHC class II molecules present antigens from phagocytosed materials.CD4 + T cells are responsible for recognizing antigenic peptides displayed by class II MHC molecules, whereas cytotoxic T-cell lymphocytes (CTLs) interact with class I MHC-peptide complexes. Once activated, CD4 + T cells secrete cytokines that further activate B cells, leading to the production of specific antibodies. This orchestration of immune responses is critical for an effective defense against pathogens. Multiepitope vaccines play a crucial role in stimulating a robust immune response involving CD8+, CD4+ T cells, and B cells. These vaccines are designed to contain multiple epitopes, enhancing their effectiveness. Multiepitope vaccines include epitopes recognized by cytotoxic T-cell lymphocytes (CTLs) and engage CD4+ T cells, promoting their activation. Also, Multiepitope vaccines often include B-cell epitopes. In addition, CD4+ T cell activation, facilitated by the vaccine, supports B cell activation [61].

One of the key challenges associated with traditional vaccine design lies in the selection of appropriate antigens. The choice of antigens is a critical decision in vaccine development, as it directly influences the vaccine's effectiveness and safety. Inappropriate selection of antigens can lead to various issues, and one of the most concerning is the potential for allergic reactions. Allergies are immune responses that are hypersensitive

and exaggerated in nature. When a vaccine contains antigens that can trigger allergic responses in certain individuals, it poses a significant health risk [62]. These allergic reactions can range from mild discomfort to severe, life-threatening anaphylactic responses, making the safety of the vaccine a paramount concern. To mitigate the risk of allergic reactions, vaccine developers must carefully consider the antigens they choose for inclusion in their formulations. The goal is to identify antigens that are specific to the pathogen of interest, thus ensuring that the immune response is focused on the intended target. This specificity is crucial for an effective vaccine, as it enables the immune system to recognize and remember the pathogen, providing protection upon subsequent encounters. On the other hand, Immunoinformatics utilizes computational methods to analyze pathogen data and predict antigenic epitopes[63]. Also, advanced algorithms help identify non-allergenic and non-toxic epitopes, minimizing the risk of allergic reactions in vaccine recipients. Additionally, multiepitope vaccines incorporate multiple epitopes, providing broader coverage of pathogen variability. By including specific antigens relevant to the pathogen of interest, multiepitope vaccines promote a targeted immune response [64].

Thus, while traditional vaccine development has achieved remarkable successes in protecting against infectious diseases, the risk of allergic reactions due to inappropriate antigen selection remains a concern. Exploring alternative approaches, such as peptide-based multi-epitope vaccines, holds the promise of mitigating this risk and enhancing vaccine safety. As the field of vaccinology continues to evolve, the quest for safer and more effective vaccines remains a critical priority in safeguarding public health.

The integration of immunoinformatics into vaccine design not only expedites the discovery of novel epitopes but also facilitates the development of innovative, safe, and effective vaccines. By focusing on epitopes with the greatest potential for inducing protective immunity while minimizing the risk of adverse reactions, immunoinformatics is poised to revolutionize viral vaccination strategies. As the field continues to advance, the synergy between computational tools and experimental validation holds the promise of a new era in vaccine development, enhancing our ability to combat viral infections and safeguard global health.

System immunology

System immunology, the holistic study of the immune system and its complex interactions, holds profound importance in the quest to identify and develop targets for immune-based vaccines. The immune system is a dynamic network of cells, molecules, and pathways that collaboratively defend the body against pathogens, including viruses, bacteria, and cancer cells. By comprehensively understanding the system-level behavior of the immune system, researchers can pinpoint key components and mechanisms crucial for an effective immune response. System immunology allows us to explore the intricate interplay between innate and adaptive immunity, uncover patterns of immune activation and regulation, and elucidate the dynamics of immune cell populations. This deep insight enables the identification of specific antigens, epitopes, and immune-related molecules

that are essential for triggering protective immune responses [65]. Furthermore, system immunology helps in characterizing the diverse immune responses across individuals, taking into account factors like genetics, age, and health status. This personalized approach aids in tailoring immune-based vaccines to different population groups, enhancing their efficacy and safety. The field also plays a pivotal role in understanding the concept of immunological memory, which is fundamental in vaccine development. By dissecting the memory responses of the immune system, researchers can identify targets that induce long-lasting immunity, ensuring the durability of vaccine protection [66].

In the context of infectious diseases and emerging pathogens, system immunology provides a comprehensive framework to study host-pathogen interactions. It enables the identification of conserved and vulnerable points in the pathogen life cycle that can be targeted for vaccine development. System immunology sheds light on the complex interplay between the host and HPV, providing insights into the molecular mechanisms of infection and immune responses [67]. Also, systemic insights into how immunodeficiency conditions influence HPV infections contribute to understanding the broader context of pathogen-host interactions [68].

Moreover, the field allows researchers to assess the immunodominance of specific antigens and prioritize them for vaccine candidates. This approach streamlines the selection of the most potent and relevant vaccine targets, expediting the vaccine development process. System immunology also facilitates the exploration of novel vaccine platforms and adjuvants. Understanding how the immune system responds to different delivery systems and adjuvants is critical for enhancing vaccine efficacy. By characterizing the immune responses triggered by these elements, researchers can fine-tune their formulations to achieve the desired outcomes [69].

Additionally, the integration of 'omics' technologies, such as genomics, transcriptomics, proteomics, and metabolomics, into system immunology provides an unprecedented depth of information. These technologies allow for the profiling of immune responses at various molecular levels, offering insights into the underlying mechanisms of vaccine-induced immunity. This high-dimensional data can be harnessed to identify biomarkers, predict vaccine responses, and optimize vaccine designs. System immunology plays a crucial role in the identification of key cytokine genes with pathological effects. Cytokines are signaling molecules that regulate immune responses, and their dysregulation can lead to various diseases. System immunology leverages comprehensive approaches to unravel the intricate interplay between immune cells, cytokines, and their downstream effects [70, 71]. Gene network analysis plays a crucial role in system immunology, offering insights into the complex interactions within the immune system. A study utilizes co-expression gene network analysis to uncover novel insights into system biology and transcriptomic analysis related to the immune system. It contributes to understanding the intricate regulatory networks governing immune responses [72]. Hernández-Gea and team employ co-expression gene network analysis to unravel novel insights into system biology and transcriptomic analysis in the realm of the immune

system. This study significantly contributes to unraveling the intricate regulatory networks governing immune responses, providing a deeper understanding of the underlying mechanisms at play [72]. Another work explores the intricate interplay between the immunome and disease-gene networks, shedding light on the dysregulation involved in immune-mediated inflammatory diseases. The study significantly contributes to unraveling the complex connections between the immune system and genetic factors, providing valuable insights for furthering our understanding of immune-related disorders [73].

Pharmacophore modeling, Quantitative Structure-Activity Relationship (QSAR) models, docking simulations, and molecular dynamics (MD) simulations are integral tools for targeting cytokines in drug discovery and development. They assist in understanding cytokine interactions, designing potential inhibitors or modulators, and predicting their behavior [74].

Pharmacophore Modeling

Pharmacophore modeling plays a vital role in various aspects of medical science. Pharmacophore modeling is extensively used in drug discovery. It enhances virtual screening performance, elucidates structure-property relationships, and derives pharmacophores for apo-protein sites, contributing to the development of new therapeutic agents [75]. Also, it is escribed as an innovative technology for exploring and extracting potential interactions between drugs or ligands and target proteins. This application is fundamental in understanding drug-target interactions in medical research [76]. The most common application of pharmacophore-based virtual screening is demonstrated in lead identification. This cherry-picking approach is a valuable strategy in drug discovery, aiding in the selection of potential drug candidates [77].

Pharmacophore modeling is instrumental in discovering ligands for cytokines and immune proteins. A study employed pharmacophore modeling based on hot-spot residues of the interaction site between interleukin-6 (IL-6) and IL-6 receptor alpha (IL-6Rα). This approach aids in identifying ligands targeting cytokine-receptor interactions [78]. Also, in a virtual screening strategy, including pharmacophore modeling, was employed to identify small molecules binding to human Inducible T-cell CO-Stimulator (ICOS). This approach aids in finding ligands relevant to immune response modulation [79]. Moreover, pharmacophore models were developed to describe molecular features and spatial arrangements of ligand-protein interactions, contributing to the discovery of novel Janus kinase (JAK) JAK1 inhibitors. This highlights the importance of pharmacophore modeling in immune-related drug development [80].

In the case of viral infection especially HPV vaccination, pharmacophore modeling is crucial in the development of effective vaccines for HPV. As HPV is a double-stranded DNA virus, understanding its genetic structure through pharmacophore modeling is necessary for designing effective vaccines [81]. Furthermore, Pharmacophore modeling and molecular docking studies are employed to identify potential inhibitors to the E6 PBM-PDZ, PDZ(PSD-95/Dlg/ZO-1)-binding motif (PBM), interaction in HPV. This

approach aids in discovering compounds with inhibitory activity, contributing to vaccine development [82].

Quantitative Structure-Activity Relationship (QSAR) Models

Quantitative Structure-Activity Relationship (QSAR) models play a crucial role in immunology by aiding in the identification of the most effective compounds, molecules, or small proteins for various components of the immune system, including T cells, B cells, cytokines, complements, and other immune structures [83]. Also, in drug discovery for immunotherapies, QSAR models guide the lead optimization process. They help in refining the chemical structures of compounds to enhance their affinity and specificity for immune system targets, maximizing therapeutic efficacy [83]. Additionally, QSAR models can quantitatively predict binding parameters of small molecules to immune system components. This is valuable in understanding the thermodynamics and kinetics of interactions, providing insights into the mechanism of action [84]. With QSAR models, researchers can efficiently screen large databases of compounds to identify those with the desired immunomodulatory properties. This accelerates the process of discovering potential candidates for further experimental validation [85]. QSAR helps address knowledge gaps in immune response and immunotherapy by providing a quantitative understanding of the relationship between compound structure and immunomodulatory properties [86].

For instance, Fukunishi et al. 2017 utilized QSAR models by employing protein-drug docking simulations. These models were applied to public affinity data to predict binding free energy and docking scores [87]. Also, another study conducted a QSAR investigation on a dataset of 224 compounds, including clinically installed molecules. The study aimed to understand the quantitative relationships between compound structures and their activities, facilitating the identification of compounds with immunomodulatory potential [88]. In the case of cancer immunotherapy, a group of researchers emphasized the use of molecular docking, dynamics, QSAR, and similarity searching in smart nanoparticles for cancer therapy. These computational approaches aid in identifying compounds suitable for cancer immunotherapy [89]. Furthermore, a multidisciplinary approach involving QSAR was employed in a study related to coronavirus disease. This approach utilized various computational methods to identify potential compounds for intervention [90]. QSAR was not only employed to cure Coronavirus. QSAR was employed to study the structure-activity relations of HPV6-E1 helicase ATPase inhibitors. This research investigated the inhibitory effects of compounds and their structural features, aiding in the design of potential HPV therapeutics [91].

These examples underscore the diverse applications of QSAR models in identifying compounds for immune system manipulation, ranging from cancer therapy to addressing infectious diseases like COVID-19. QSAR's ability to predict and understand the relationships between compound structures and immunomodulatory activities contributes significantly to the discovery of novel therapeutic agents.

Molecular docking

The field of drug discovery has witnessed a transformative shift with the advent of computational techniques, particularly molecular docking, offering unprecedented insights into ligand-protein interactions. This methodology has proven instrumental in accelerating drug development processes, with profound implications for both traditional drug discovery and emerging immunotherapies [92]. In drug discovery, molecular docking plays a pivotal role in predicting the binding affinity of small molecules to specific proteins or receptors, enabling the identification of potential drug candidates. This predictive capability significantly expedites the initial stages of drug development by narrowing down the pool of compounds for further experimental validation. As drug discovery increasingly leans towards precision medicine and targeted therapies, molecular docking emerges as a crucial tool in the arsenal, facilitating the design of drugs tailored to interact with specific biomolecular targets [93].

Immunotherapy, a revolutionary approach harnessing the body's immune system to combat diseases, has gained prominence in recent years. Molecular docking proves indispensable in this realm, facilitating the identification of potential immune targets. It allows for the exploration of ligand-protein interactions involved in modulating immune responses. Albumin/vaccine nanocomplexes designed through molecular docking have demonstrated improved assembly in vivo, showcasing the potential for enhanced vaccine delivery [94]. Notably, the potential targets for immunomodulation are as diverse as the components of the immune system itself, offering a myriad of possibilities for therapeutic intervention. The application of molecular docking in finding targets for cytokines, chemokines, and other immune system proteins constitutes a critical facet of immunotherapy research. Understanding the intricate interactions between these proteins and potential ligands enables the design of interventions that can finely tune immune responses. This targeted modulation holds the promise of developing therapies with enhanced specificity and reduced off-target effects [95]. For example, Docking analysis has been employed to study the binding affinity of vaccines with Toll-like Receptor 4 (TLR-4) and Toll-like Receptor 8 (TLR-8), providing insights into the immunogenicity of vaccines [96]. To prevent the HPV-E6 antigen, molecular docking studies identified cidofovir and jaceosidin as potent inhibitors of the E6 protein of HPV. These compounds were selected based on high docking scores, showcasing the utility of molecular docking in screening for potential vaccine candidates [97]. Moreover, Autodock Vina and the Molecular Graphics Laboratory (MGL) tools were employed in docking studies to identify 28 candidate molecules for HPV16 E6 treatment. This approach resulted in the selection of three promising compounds, demonstrating the efficacy of molecular docking in the early stages of vaccine discovery [98]. Screening of small molecules for HPV-E6 by molecular docking, using the X-ray crystal structure of HPV-16 E6, aided in screening small molecular compounds. This approach contributes to the identification of potential vaccine candidates by understanding the interaction between compounds and viral proteins [99].

Looking forward, the future of immunotherapy lies in the continued refinement and advancement of computational techniques like molecular docking. With the increasing understanding of immune system complexities, the potential for discovering novel targets and designing tailored interventions expands. The integration of artificial intelligence, machine learning, and big data analytics into docking simulations opens avenues for more accurate predictions and the identification of subtle nuances in ligand-protein interactions [100]. In this case, the development of multi-epitope vaccines against diseases like cancer has benefited from molecular docking strategies, optimizing the interaction between vaccine components and immune receptors [101].

As the field evolves, exploring the synergies between experimental and computational approaches will be paramount. Collaboration between computational biologists, immunologists, and experimental researchers will enhance the translational potential of findings, ultimately contributing to the development of more effective and personalized immunotherapies. The journey of molecular docking in drug discovery and immunotherapy is poised for exciting developments, holding the promise of revolutionizing therapeutic interventions in the years to come.

Molecular Dynamics Simulations:

Molecular dynamics plays a crucial role in elucidating the interactions between interleukins (ILs) and their target medications or small compounds. Understanding the dynamic motions of atoms in protein targets and their complexes is essential for drug discovery and design. The significance of molecular dynamics lies in its ability to provide insights into the stability, flexibility, and behavior of ILs during interactions with ligands. This computational technique allows researchers to simulate the movement of atoms over time, offering a dynamic perspective on the IL-target interactions [102].

Research studies have provided valuable insights into the interactions between Interleukin-6 (IL-6) and various ligands. Molecular docking analyses indicate comparable affinities for positively charged positions on the surface of IL-6 for two ligands [103].

A computational-experimental investigation delved into the mode of binding of the IL6-piperine complex, utilizing experimental and computational molecular biophysical methods [104]. Another study aimed to discover IL-6 inhibitors based on protein-protein interactions with a novel camelid Fab fragment (68F2) within a crystal protein complex [105]. Furthermore, Structure-based 3D-pharmacophore modeling was employed to discover novel IL-6 inhibitors. Pharmacophore models were constructed based on key residues of IL-6, contributing to lead screening efforts [78].

Molecular dynamics analysis of the binding of human IL-6 with the IL-6 α-receptor revealed an interplay of electrostatic, hydrophobic, hydrogen bonding, and aromatic stacking interactions, facilitating the formation of the hIL-6/IL-6Rα complex. A related study further explored the molecular dynamics analysis of this binding [106].

By integrating system immunology with computational approaches, researchers can better understand the immune response in various diseases, identify critical cytokine targets, and design drug candidates that modulate the immune system for therapeutic purposes. This integrative approach accelerates the drug discovery process, increases the likelihood of success, and contributes to the development of novel immunotherapies and treatments for a wide range of conditions, including autoimmune diseases, cancer, and infectious diseases. Molecular Dynamics (MD) simulations play a pivotal role in unraveling the dynamic behavior of cytokine-receptor complexes over time, offering valuable insights into structural changes and the stability of binding. These simulations also delve into allosteric modulation, uncovering crucial allosteric sites and mechanisms that influence cytokine activity indirectly. Additionally, MD simulations consider solvent effects, providing a comprehensive understanding of molecule behavior in biologically relevant environments [107]. To have better vaccines for HPV, Molecular dynamics simulations and modeling were involved in predicting antigenic peptides from HPV E6 and E7 proteins. This information aids in the design of vaccines targeting specific antigenic regions for effective immune response [108]. To understand the binding affinity of HPV vaccines, MD contributes to understanding the interactions between potential vaccine candidates and viral proteins, guiding the development of effective HPV vaccines [109]. To find better HPV vaccines, in silico screening and molecular dynamics simulations were employed to identify potential anti-HPV hits. The PubChem database screening, followed by ADMET predictions and molecular docking, enhances the selection of candidates for further vaccine development [97]

These techniques are often intertwined for a more comprehensive approach. For instance, pharmacophore models guide compound selection for virtual screening, while docking simulations predict binding modes of selected compounds. Molecular dynamics simulations further refine and analyze these binding interactions. Quantitative Structure-Activity Relationship (QSAR) models contribute insights into the structure-activity relationships of selected compounds [110].

Integrating system immunology with these methodologies significantly enhances drug discovery, particularly in developing immunotherapies and cytokine-based treatments. Through systematic data gathering and analysis in system immunology, researchers collect extensive immunological data, identifying key cytokines, receptors, and signaling pathways involved in immune responses. Utilizing insights from system immunology, pharmacophore models specific to identified cytokine targets are created, determining essential structural features and chemical properties for potential drug candidates. QSAR models, incorporating data from system immunology, predict the biological activity of potential drug candidates against target cytokines. Docking simulations, guided by system immunology insights, aid in predicting binding modes of drug candidates to cytokines, contributing to a more informed drug discovery process [111].

In the pursuit of developing novel vaccines with heightened effectiveness, our research followed a multifaceted approach. We initiated our study by investigating the activation and proliferation of specific T cell subpopulations, including CD8, CD4, and follicular T

cells, in response to HPV virus antigens, such as E6, E7, and L1. Subsequently, we harnessed the power of immunoinformatics to design new epitopes, a crucial step in the creation of potential vaccines. To further our understanding and expand our focus, we conducted systematic investigations using publicly available gene expression data from the Gene Expression Omnibus (GEO) database, with an emphasis on post-vaccination alterations in gene expression within peripheral blood mononuclear cells (PBMCs). Upon identifying target genes, we employed pharmacophore and QSAR modeling techniques to discover suitable ligands. These ligands were then sourced from comprehensive databases, including ZINC and PubChem. In the final stages of our research, we conducted molecular docking and dynamics simulations to evaluate the interactions between our target genes and their respective ligands. This comprehensive approach integrates immunological, bioinformatics, and computational strategies to potentially pave the way for the development of innovative vaccine.

Chapter 4: Aim of study

Aim of study

Introduction

Human papillomavirus (HPV) infections are a pressing global health concern, linked to the development of various cancers, including cervical cancer. Understanding the intricate interplay between the immune system and HPV is crucial for devising effective strategies to combat HPV-related diseases. This chapter delineates the primary aims and objectives of our study, which employed a multifaceted approach, incorporating Lymphocyte Proliferation Assay (LPA), Activated Induced Marker Assay (AIM). The study aimed to investigate T-cell activation, proliferation, and interferon-gamma (IFN-γ) production in response to HPV antigens, specifically HPV early 6 (E6), early 7 (E7), and L1 proteins, across three distinct groups of women: those without vaccination, vaccinated individuals, and women infected with HPV.

Rationale for the Study

Human Papillomavirus (HPV) is an extremely common sexually transmitted infection that affects a large portion of the global population. It is responsible for a variety of health issues, including genital warts, but its most concerning association is with certain high-risk HPV types that are strongly linked to the development of cervical cancer and other malignancies, such as anal, penile, and oreopharyngeal cancers.

In the battle against HPV and its associated cancers, the immune system plays a crucial role, with T cells taking center stage. Among these T cells, CD4 and CD8 T cells are particularly vital components. They are instrumental in recognizing and eliminating cells in the body that have been infected by HPV.

Understanding the mechanisms that govern T-cell activation and proliferation in response to HPV antigens is of paramount importance. Here's why:

Host Immunity

T cells are key players in the immune response. When they encounter HPV-infected cells, they are responsible for recognizing these infected cells as foreign invaders and mounting an immune response to eliminate them. Understanding how T cells respond to HPV is essential in deciphering the complex interplay between the virus and the immune system.

Therapeutic Interventions

Research into T-cell responses to HPV can lead to the development of novel therapeutic interventions. This includes immunotherapies that enhance the body's natural ability to target and destroy HPV-infected cells. By understanding the precise mechanisms involved, researchers can design more effective treatments that boost T-cell responses.

Vaccine Development:

Vaccines have been highly effective in preventing certain types of HPV infections. A deeper understanding of how T cells interact with HPV antigens can inform the development of more potent and targeted vaccines. This is critical for preventing HPV infections and, consequently, reducing the incidence of HPV-related cancers.

In summary, unraveling the intricacies of T-cell activation and proliferation in response to HPV is essential for advancing our comprehension of how the immune system responds to this virus. This knowledge not only contributes to our understanding of host immunity to HPV but also guides the development of therapeutic strategies and vaccines that can better combat HPV infections and related malignancies, ultimately improving public health outcomes.

Study Objectives

Objective 1: Investigate and Compare T-Cell Activation and Proliferation in Response to HPV Antigens

One of the primary goals of our study was to delve into the intricate world of T-cell responses to HPV antigens, specifically targeting the E6, E7, and L1 proteins. We set out to achieve the following:

Comparison Across Three Distinct Groups:

To comprehensively understand the immune response to HPV, we compared and contrasted the activation and proliferation of CD4 and CD8 T cells in three distinct groups of individuals. These groups represented different scenarios of HPV exposure and immune interaction: unvaccinated women, vaccinated women, and women who were actively infected with HPV.

Participant Criteria for this Study

To be eligible for participation in the HPV study, participants must meet specific criteria based on vaccination status and HPV infection. The criteria are as follows:

Non-Vaccinated Participants:

- Must not be infected with high-risk HPVs.
- Should not have any type of HPV infection for the last ten years.
- Must not have received any doses of HPV vaccines.

Vaccinated Participants:

- Must have received at least one dose of the following vaccines:

- Gardasil 9-valent or 4-valent.
- Cervarix.
- Other commercially available vaccines containing L1.

- Should not be currently infected with HPV.

Early Stage Patients:

- Must show abnormal cells in the cervix.
- Abnormal cells should be localized to the cervix and not spread to other tissues [112].

Understanding T-Cell Behavior:

By focusing on these critical subsets of T cells, we aimed to decipher how they behaved in response to the presence of HPV antigens. This knowledge was crucial in unraveling the nuances of the immune response, potentially shedding light on differences in immune defense mechanisms among the three groups.

Specific Antigens:

The choice of specific HPV antigens (E6, E7, and L1 proteins) was deliberate, as these are known to be central players in the immune recognition of HPV infection. Analyzing how T cells responded to these antigens provided insight into the immune system's ability to recognize and mount a defense against HPV.

Objective 2: Immunoinformatics for Epitope Discovery:

In this phase, our primary objective was to systematically explore and assess the immunogenicity of antigenic proteins for the identification of potential epitopes recognized by T cells and B cells. This also involved the evaluation of epitope binding to MHC class I and II molecules. Furthermore, we aimed to determine the antigenicity of protein sequences and assess their suitability as vaccine candidates. Specifically, we sought to:

- Identify epitope regions within protein sequences and assess their immunogenic potential.
- Predict and evaluate the binding of epitopes to MHC class I molecules, with a special focus on epitopes of varying lengths, including those with 9mer peptides and above.

- Utilize the resources to comprehensively explore epitope prediction for MHC class I and II molecules and B cell antigens.

- Harness the capabilities of the immunomedicine group's antigen peptide prediction server for the identification and analysis of antigenic epitopes.

- Extend our investigation to predict MHC class II epitopes and T cytotoxic cell epitopes, thus enabling a comprehensive epitope discovery for various immune responses.

- Predict B cell antigen epitopes and analyze the tertiary structure of antigens targeted by B cells.

Epitope Evaluation:

In this stage of the study, we assessed the antigenicity and allergenicity of predicted epitopes to ensure the selection of safe and immunogenic epitopes. Additionally, our objectives included:

- Determining the population coverage of identified epitopes and evaluating the percentage of the population exhibiting stimulated immune responses.

- Generating 3D structures of epitopes to facilitate docking studies and assessing the solubility of proteins and epitopes. This step was essential to ensure the feasibility of these components in vaccine development.

In Silico Cloning:

A significant component of our study involved the in silico cloning of proteins. We achieved this by concatenating protein sequences with epitopes and adapting DNA sequences for expression in specific organisms. This process enabled the virtual production of proteins.

Plasmid Creation:

To support vaccine development, we created plasmids and explored a range of plasmid options and expression machinery in different organisms through.

Overall, our study established a comprehensive immunoinformatics framework for epitope discovery, evaluation, and in silico cloning. Our objectives were realized, and this framework has the potential to facilitate the design of multi-epitope-based vaccines. By doing so, our research contributes to the advancement of immunological knowledge and the development of innovative vaccines.

Objective 3: Finding a key target gene through system immunology

The primary objectives of this study were focused on employing a systems biology and systems immunology approach to investigate immune responses in women who had previously received HPV vaccines containing the L1 antigen. These investigations were carried out in comparison to a control group of women who had received a placebo (media) instead of the vaccine. The study placed specific emphasis on the following key objectives:

Data Analysis:

This objective entailed the retrieval and comprehensive analysis of gene expression data derived from peripheral blood mononuclear cells (PBMCs) of women who had undergone HPV vaccination, as well as control subjects. During this phase, our focus was on the following key actions:

- Retrieving and scrutinizing gene expression data.
- Filtering genes based on specific criteria, including adjusted p-values and fold change.

Gene Network Construction:

The study sought to construct gene networks using the STRING online tool, which facilitated the exploration of potential interactions and functional relationships among differentially expressed genes.

Network Analysis:

The analysis of these networks was conducted through the use of Cytoscape. Key aspects of this objective included:

- Employing Cytoscape for network analysis.
- Focusing on metrics such as centrality, betweenness, and degree to identify and emphasize genes within the network that played pivotal roles in the immune response.

Modularity and Eigenvector Analysis:

Gephi software was employed to delve deeper into the structural aspects of the network, particularly concentrating on modularity and Eigenvector centrality. This analysis was aimed at identifying gene clusters and influential nodes within the network structure.

These objectives collectively formed the one core of our study, designed to provide insights into immune responses following HPV vaccination. Through our efforts, we aimed to advance our understanding of the intricate mechanisms involved in these

immune responses and contribute to the broader body of knowledge in the fields of systems biology and systems immunology.

Objective 4: The objective of this part of study was to ascertain a unique ligand for the target protein.

obtained from the system immunology section, by employing a combination of pharmacophore modeling and Quantitative Structure-Activity Relationship (QSAR) techniques. Furthermore, the research aimed to thoroughly evaluate the ligand's compatibility with the target protein through the application of docking procedures and molecular dynamics simulations.

Data Retrieval and Analysis: The study began by collecting and analyzing gene expression data of the target protein, which was obtained from the system immunology section. The analysis involved filtering genes based on adjusted p-values ($p < 0.05$) and fold change (≥ 2.5). This initial step provided the foundation for further ligand discovery.

Database Exploration:

To explore potential ligands, the study delved into a range of databases, including but not limited to Clinical Trials.gov, STITCH database, KEGG pathways, KEGG disease, KEGG animal, BioCarta, Panther pathways, HGNC, Gene Expression Atlas, The Human Protein Atlas, OMIM, DisGeNet, dbGaP, DRUGBANK, Therapeutic Target Database, KEGG Drug Database, BRENDA-Enzyme database, Vector Alignment Search Tool, SwissADME, PubChem, ADMETlab, DGIdb, PockDrug, CASTp, CSmetaPred, COACH, and PHAROS. These resources were instrumental in exploring potential ligands and understanding their interactions, targets, and suitability.

Ligand Screening and Analysis:

Once potential ligands were identified, the study employed software tools like pyrex, Binding DB, OpenBabel, PaDLe, SMLR, Chemoface, and Schrodinger for ligand screening, analysis, and modeling. This phase aimed to assess the ligand's characteristics, descriptors, and suitability for binding to the target protein. The study also employed screening and modeling techniques to predict ligand-protein interactions and structural compatibility.

ADME Property Evaluation:

In addition to ligand screening, the study delved into ADME (Absorption, Distribution, Metabolism, and Excretion) properties of the identified ligands. Utilizing tools like QikPro, the study analyzed the ligand's ADME characteristics to understand their potential for therapeutic applications.

Structure Building and Optimization:

The study also involved the use of software tools like Minerva Sketch and Chimera to create molecular structures based on the SMILE format. These tools allowed for the construction of ligand structures and subsequent minimization to refine their conformations.

Overall, the study's aim was to comprehensively investigate potential ligands for the target protein by leveraging a wide array of bioinformatics tools, databases, and modeling techniques. This approach sought to identify a ligand with suitable characteristics for binding to the target protein, potentially paving the way for further research in drug development and therapeutic applications.

Study Methods

Lymphocyte Proliferation Assay (LPA):

In our research, we utilized a comprehensive set of methodologies to investigate the immune response to HPV antigens among individuals from different groups. The Lymphocyte Proliferation Assay (LPA) played a pivotal role in this endeavor. This assay involved several key steps:

Isolation and Characterization of CD4 and CD8 T cells:

We first isolated and characterized CD4 and CD8 T cells from individuals representing three distinct groups. This step allowed us to obtain a clear picture of the T-cell populations within each group.

Stimulation of T Cells with HPV Antigens:

To evaluate the specific response of T cells to HPV, we exposed these isolated T cells to HPV antigens, specifically targeting the E6, E7, and L1 proteins. These proteins are known to be crucial in the immune response to HPV infection.

Assessment of T-Cell Proliferation Responses:

The primary focus of the LPA was to assess how T cells from different groups proliferated in response to HPV antigens. We paid particular attention to follicular and non-follicular CD4 T cells and CD8 T cells, as these subsets may exhibit distinct responses.

Activated Induced Marker Assay (AIM):

In addition to the LPA, we employed the Activated Induced Marker Assay (AIM) as another key research tool. This assay followed a similar approach:

Isolation of T Cells:

T cells were isolated from participants in each of the three groups under investigation.

Exposure to HPV Antigens:

These isolated T cells were then exposed to HPV antigens to mimic the conditions of an HPV encounter in the body.

Evaluation of T-Cell Activation Markers:

The AIM assay allowed us to assess the activation status of T cells by examining specific markers. This provided valuable insights into how T cells responded to HPV antigens, complementing the data obtained through the LPA.

Immunoinfromatics

One of the primary objective of this research is to expedite the development of effective vaccines against a diverse range of pathogens by harnessing the power of immunoinformatics. This multidisciplinary field at the intersection of immunology and bioinformatics serves as the cornerstone for epitope discovery, design, and evaluation.

Epitope Discovery and Evaluation:

In pursuit of this aim, a comprehensive suite of online tools and databases was utilized to facilitate the identification and evaluation of epitopes. The selection of epitope databases and resources was guided by their potential to expedite vaccine development.

- AntigenDB: Assisting in the analysis of antigenicity and epitope identification within protein sequences [113].
- VaxiJen: Determining the antigenic nature of given sequences [114].
- EMBOSS Antigenic: Identifying epitope regions within protein sequences along with associated scores [115].
- ProPred-I-MHC1: Predicting epitope interactions with MHC class I molecules [116].
- NetTepi 1.0 Server: Facilitating the identification of epitopes, particularly those of at least 9mer peptides [117].
- The Immune Epitope Database (IEDB): A valuable resource for locating epitopes for MHC class I and II and B cells [118].
- Net MHCIIpan 4.0: Dedicated to MHC class II epitope prediction [119].
- CTLPred: Predicting epitopes for T cytotoxic cells [120].
- NetCTL 1.2 Server: Predicting epitopes for various supertypes [121].

- Bcepred: Focused on predicting B cell antigen epitopes [122].
- DiscoTop: Providing tertiary structure information for B cell antigens [123].

Epitope Evaluation:

The next pivotal step in this endeavor involves the evaluation of epitopes for their compatibility and suitability. A careful selection of evaluation tools was conducted to ensure that only the most promising epitopes are considered.

- **VaxiJen**: Employed for the assessment of epitope antigenicity.
- **AlgPred**: Checking for the allergenicity of epitopes [124].
- **AllerCatPro**: Assessing the allergic potential of epitopes [125].
- **The Immune Epitope Database (IEDB)**: Evaluating population coverage to estimate the percentage of the population that the epitopes may stimulate [118].
- **Pepfold 4**: Instrumental in generating 3D structures of epitopes for subsequent docking studies [126].
- **ClusPro**: Utilized for epitope-HLA docking [127].
- **Scratch Protein Predictor**: Aiding in the assessment of protein solubility [128].

Plasmid and Vaccine Design:

Beyond epitope discovery and evaluation, the study also delves into plasmid design for vaccine development. This critical aspect is supported by resources such as **Addgene** for the exploration of various plasmids and their expression machinery in different organisms, as well as Benchling for the design of plasmids.

This multifaceted approach, encompassing epitope discovery, evaluation, and plasmid design, is aimed at the accelerated development of effective vaccines targeting a diverse array of pathogens.

Vaccine effects on immune system:

In the simulations of the vaccine construct's impact on the immune system, it was observed that administering the vaccine through three injections had the potential to stimulate various immunoglobulins. The initial response manifested in an elevated level of IgM, while the subsequent response included increased levels of IgM + IgG, IgG1 + IgG2, IgG1, IgG2, and B-cell populations. Following three vaccine injections, there was a decline in antigen levels. Both T cell populations (CTL and HTL) exhibited an enhanced response, indicating the immunogenicity of T cell epitopes within the

vaccine construct. Each exposure led to heightened macrophage activity, with consistent NK cell activity throughout the period. Subsequent exposures resulted in a significant increase in IFN-gamma, IL-10, IL-23, and IL-12 levels. After 12 repeated vaccine injections at regular intervals, antigen levels peaked similarly, accompanied by a notable increase in IgM + IgG, IgG1 + IgG2 levels. Persistent elevation in memory cells for B and T cells was observed throughout the exposures, while IFN-gamma levels remained consistently high from the first to the last exposure. This indicates that the vaccine elicited a robust immune response in both short and subsequent repeated exposures.

System Immunology:

The study aimed to investigate immune responses in women who had received HPV vaccines containing the L1 antigen, in comparison to a control group of women who had received a placebo. The study followed a systematic method for achieving its key objectives:

Data Analysis:

Data Retrieval: The primary step involved the retrieval of gene expression data from peripheral blood mononuclear cells (PBMCs) of women who had undergone HPV vaccination and control subjects. Data was collected from repositories such as Gene Expression Omnibus (GEO) on the National Center for Biotechnology Information (NCBI) platform, which provided access to datasets with cel dataframes.

Data Scrutiny: The collected data was meticulously scrutinized to ensure its quality and relevance. Any discrepancies or issues with the data were addressed to ensure the reliability of the subsequent analyses.

Data Filtering: To identify differentially expressed genes, genes were filtered based on specific criteria, including adjusted p-values and fold change. This filtering process aimed to isolate genes that exhibited significant changes in expression following HPV vaccination.

Gene Network Construction:

Network Building: The study utilized the STRING online tool, a bioinformatics resource that facilitates the construction of protein-protein interaction networks. Gene symbols from the differentially expressed genes were input into STRING to create a gene network.

Interaction Data: STRING provided information on potential interactions and functional relationships among the genes within the network. This data included confidence scores, which helped identify the strength of interactions.

Network Analysis:

Cytoscape Implementation: Cytoscape, a widely used software platform for visualizing and analyzing complex networks, was employed for in-depth network analysis. The constructed gene network was imported into Cytoscape for further investigation.

Metric Assessment: The analysis focused on key network metrics such as centrality, betweenness, and degree. These metrics were utilized to identify genes within the network that played pivotal roles in the immune response. Genes with high centrality, betweenness, and degree values were considered influential in the network.

Modularity and Eigenvector Analysis:

Gephi Software: Gephi, an open-source network analysis and visualization software, was used for modularity and Eigenvector centrality analysis.

Modularity Assessment: The study concentrated on the modularity of the network, aiming to identify gene clusters with distinct functionalities or roles within the network structure. Modularity analysis provided insights into the organization of genes based on their interactions.

Eigenvector Centrality: Eigenvector centrality analysis was conducted to identify influential nodes or genes within the network. Genes with high Eigenvector centrality values were considered central to the network structure.

The study method encompassed data retrieval and analysis, network construction, network analysis, and modularity and Eigenvector analysis. These steps collectively formed a comprehensive approach to gain insights into immune responses following HPV vaccination. The study method was designed to advance the understanding of complex immune mechanisms within the context of systems biology and systems immunology, contributing to the broader body of knowledge in these fields.

Ligand Discovery:

The study employed a systematic approach to discover a unique ligand for the target protein obtained from the system immunology section. The method encompassed several key steps and utilized various bioinformatics tools and databases. The study's primary focus was on identifying a ligand with compatibility for the target protein. The method was divided into the following phases:

Data Retrieval and Analysis:

Data Collection: The study initiated by collecting gene expression data related to the target protein. This data was obtained from the system immunology section.

Data Analysis: Gene expression data was subjected to comprehensive analysis. Genes were filtered based on specific criteria, including adjusted p-values ($p < 0.05$) and fold change (≥ 2.5). This analysis provided a foundation for further ligand discovery.

Database Exploration:

Exploration of Databases: A wide array of bioinformatics databases and resources were explored. This included databases such as Clinical Trials.gov, STITCH database, PhID database, KEGG pathways, KEGG disease, KEGG animal, BioCarta, Panther pathways,

HGNC, Gene Expression Atlas, The Human Protein Atlas, OMIM, DisGeNet, dbGaP, DRUGBANK, Therapeutic Target Database, KEGG Drug Database, BRENDA-Enzyme database, Vector Alignment Search Tool, SwissADME, PubChem, ADMETlab, DGIdb, PockDrug, CASTp, CSmetaPred, COACH, and PHAROS.

Ligand Screening and Analysis, QSAR and pharmacophore modeling:

Ligand Identification: Potential ligands were identified using various bioinformatics software tools, including pyrex, Binding DB, OpenBabel, PaDLe, SMLR, Chemoface, and Schrodinger. These tools facilitated the screening and analysis of potential ligands.

Characteristics Assessment: Identified ligands were subjected to a detailed assessment of their characteristics and descriptors. This included evaluating their suitability for binding to the target protein.

Prediction of Ligand-Protein Interactions: Screening and modeling techniques were employed to predict interactions between the identified ligands and the target protein. This phase aimed to assess the structural compatibility of the ligands.

Unveiling Protein-Protein Interaction Dynamics

My research employs a multifaceted strategy to unravel the complexities of protein-protein interactions. Utilizing molecular docking, I visualize ligand binding affinities, identifying potential interaction hotspots. Complemented by 2D-QSAR modeling, the study predicts structure-activity relationships, enhancing the design of targeted interventions. This synergistic approach provides a comprehensive understanding of the molecular underpinnings of protein-protein interactions.

The investigation extends to molecular dynamics simulations, offering a real-time dynamic perspective. By integrating docking, free energy calculations, and simulations, the study aims to unveil intricate interaction details, contributing valuable insights to high-impact journals and advancing our understanding of these vital biological processes.

ADME Property Evaluation:

ADME Analysis: ADME (Absorption, Distribution, Metabolism, and Excretion) properties of the identified ligands were examined. Tools such as QikPro were utilized to assess these characteristics, providing insights into the potential for therapeutic applications.

Structure Building and Optimization:

Molecular Structure Construction: Molecular structures of the ligands were constructed based on the SMILE format. Software tools such as Minerva Sketch and Chimera were used for this purpose.

Structure Refinement: The constructed molecular structures were subjected to a minimization process to refine their conformations, ensuring structural optimization.

The study method encompassed a systematic approach that combined data analysis, database exploration, ligand screening, ADME property evaluation, and structural optimization to identify a ligand with compatibility for the target protein. This method served as a comprehensive framework for ligand discovery, which could have significant implications for drug development and therapeutic applications.

Expected Contributions

Our study holds the promise of making substantial contributions to the field of HPV immunology and, by extension, the broader realm of infectious disease research. The anticipated contributions are as follows:

Nuanced Understanding of T-Cell Dynamics:

By meticulously investigating T-cell activation, proliferation, and cytokine production in response to HPV antigens across different clinical contexts, our study aims to provide a nuanced understanding. This nuanced insight is critical, as it delves beyond the surface-level understanding of immune responses. By deciphering the intricate dance of immune cells in the presence of HPV, we anticipate revealing subtleties that could be pivotal in understanding how the immune system recognizes, responds to, and combats HPV infections.

Insights into Host Immune Responses:

Our research endeavors to offer deep insights into the dynamics of host immune responses concerning HPV. Understanding these responses at a granular level is invaluable, not only for HPV but also for similar infectious agents. By elucidating the complexities of T-cell behavior, our study aims to contribute foundational knowledge that could be applicable to the study of other viral infections, potentially paving the way for broader advancements in immunological research.

Guiding Targeted Therapeutic Interventions:

One of the primary goals of our study is to identify key factors influencing T-cell-mediated immunity in individuals with varying HPV exposures and vaccination histories. This understanding is instrumental in the development of targeted therapeutic interventions. Armed with this knowledge, researchers and healthcare professionals can design therapies that specifically enhance the immune response in individuals who might be at higher risk due to specific exposure scenarios. This targeted approach holds the potential to significantly improve the effectiveness of treatments, leading to better outcomes for patients.

Influencing HPV Vaccine Development:

The insights gleaned from our research could have a transformative impact on the development of HPV vaccines. By identifying the factors that enhance or hinder T-cell-mediated immunity, our study may provide crucial data for refining existing vaccines or designing new, more effective ones. Understanding how different populations respond to

HPV antigens can inform vaccine formulations, dosages, and administration strategies, ultimately bolstering the global effort to prevent HPV-related diseases.

Improved Vaccine Development:

Insights into immune responses can lead to more effective HPV vaccines, offering better protection against HPV infection and related diseases, such as cervical cancer.

Enhanced Vaccine Safety:

Identification of potential adverse reactions or safety concerns associated with HPV vaccines can ensure the safety of patients who receive the vaccination.

Identification of Biomarkers:

Specific genes, such as IL6, can serve as biomarkers to assess vaccine effectiveness, enabling tailored vaccination strategies and improved outcomes for patients.

Personalized Medicine:

Understanding regulatory mechanisms in immune responses can pave the way for personalized vaccination approaches, considering the unique needs of patients.

Treatment Strategies:

Findings may influence the development of more effective therapies for HPV-related diseases, benefiting patients who are already infected with HPV.

Public Health Impact:

Improving vaccine efficacy and safety can lead to increased vaccination rates, reducing the prevalence of HPV-related diseases and their impact on public health.

Patient Education:

Study findings can contribute to patient education and informed decision-making about HPV vaccination, empowering patients to make well-informed choices about their health.

In summary, our study's expected contributions extend beyond the laboratory, potentially shaping the future landscape of HPV research, therapeutic interventions, and preventive strategies. By unraveling the complexities of the immune response to HPV, we aim to empower the scientific community with knowledge that can directly influence clinical practices, improve patient outcomes, and contribute to the global fight against HPV-related illnesses.

Chapter 4: Materials and methods

Materials and Methods:

Activation-Induced Marker (AIM) Assay:

Thawing and Preparation of PBMCs:

In this critical initial phase of the experiment, we undertook a series of meticulous steps to ensure the viability and functionality of Peripheral Blood Mononuclear Cells (PBMCs):

Thawing Process:

- The process commenced with the careful thawing of the PBMCs, which were retrieved from deep freeze storage. These cells are exceptionally sensitive and preserving their integrity is paramount for the success of the experiment.

- To initiate the revival of the PBMCs, they were thawed with precision and then carefully placed within a controlled Carbon Dioxide (CO2) incubator, Thermo Scientific™ Heracell™ 150i. This controlled environment is vital for maintaining the optimal conditions required for cellular activity.

Overnight Incubation:

- Throughout the night, the PBMCs were allowed to gradually acclimate to the incubator's controlled atmosphere. This extended incubation period played a crucial role in reactivating the cells and ensuring they were fully reanimated from their frozen state.

Quantification and Concentration Determination:

- The following morning marked a pivotal stage in the process. We meticulously quantified the PBMCs after centrifugation at 1800 for 5 minutes (Thermo Scientific™ Megafuge™ Series 16, TX-400 4 x 400ml Swinging Bucket Rotor-75003629), employing a hemocytometer, a precise instrument designed for the enumeration of cells.

- The hemocytometer enabled us to calculate the precise concentration of the PBMCs. This step was instrumental in ensuring that we had an accurate understanding of the number of cells at our disposal, enabling us to proceed with confidence.

Resuspension in RPMI Medium:

- Once the concentration was determined, we proceeded to resuspend the PBMCs in a specially formulated RPMI 1640 medium, Euroclone, Milano, Italy. This medium was meticulously prepared with the inclusion of 10% Fetal Calf Serum

(FCS), Sigma, St. Louis, MO, US), and 1% Penicillin-Streptomycin-Glutamine (PSG), Euroclone.

- The RPMI medium served as the ideal environment to nurture and support the PBMCs during the ensuing experiments. It provided the necessary nutrients and growth factors required to maintain the cells' vitality and functionality.

Achieving 100 µL per Well:

- Ensuring consistency and precision in our experimental setup, we carefully measured and adjusted the volume of resuspended PBMCs to achieve a uniform 100 µL for each well in the experiment. This standardization was crucial for accurate and reliable results across all aspects of the study.

In summary, the thawing and preparation of PBMCs constituted a meticulous and critically important phase of the experiment. The careful steps taken during this process were geared towards reviving and optimizing the PBMCs' functionality, thereby laying a robust foundation for the subsequent stages of our investigation into T-cell responses to HPV peptides.

Preparation of Stimuli:

In this phase, the stimuli, which play a central role in stimulating the immune response, were prepared with precision and attention to detail:

HPV Peptides (E6, E7, L1):

- The specific HPV peptides used in this study, namely E6 (PepMix™ HPV16-E6, PM-HPV16-E6), E7 (PepMix™ HPV16-E7, PM-HPV16-E7), and L1 (PepMix™ HPV16-L1, PM-HPV16-L1) bought from innovative peptide solution company , were pivotal components. They were employed at a meticulously determined concentration of 1 µg/mL, and each well was endowed with a final volume of 200 µL.

- The careful choice of this concentration was based on prior research and knowledge of the peptides' potency in inducing a T-cell response. It was essential to strike a balance between providing a sufficient stimulus and avoiding overstimulation, ensuring that the immune responses observed were physiologically relevant.

- To achieve this, stimuli were prepared at a concentration of 2 µg/mL. This concentration was meticulously calculated to deliver precisely 100 µL of the stimulus per well, achieving consistency and uniformity in our experimental setup.

Actin as the Negative Control:

- As a crucial element of the experiment, Actin (15 mers, overlapping by 10 amino acids, Pepscan, Lelystad, The Netherlands), serving as the negative control, was handled with great care. Actin was initially stored at -20°C to maintain its stability and integrity.

- Prior to use, Actin was prepared at a concentration of 2 µg/mL, aligning with the same concentration used for the HPV peptides. This careful standardization allowed Actin to serve as a baseline control, ensuring that any observed responses were specific to the HPV peptides and not due to other factors.

Dilution in RPMI Medium with FCS and PSG:

- To facilitate the precise delivery of these stimuli and maintain the cells in an optimal environment, the peptides, including the HPV peptides and Actin, were meticulously diluted. This dilution process was carried out in RPMI 1640 medium enriched with 10% Fetal Calf Serum (FCS) and 1% Penicillin-Streptomycin-Glutamine (PSG).

- The choice of RPMI medium with FCS and PSG was strategic. It provided an ideal milieu, ensuring that the cells were nourished and supported during the experimental process. This medium not only promoted the survival of cells but also allowed for the accurate delivery of stimuli to each well.

In summary, the preparation of stimuli involved the precise calculation and dilution of HPV peptides and the negative control Actin. These carefully prepared stimuli were essential components of the study, enabling us to probe the immune responses of T cells to HPV peptides with accuracy and consistency.

Peptide Storage:

The storage of HPV peptides, including E6, E7, and L1, is a critical aspect of maintaining the integrity and potency of these essential components throughout the course of the experiment. Here's a detailed account of how this storage was managed:

Temperature Control:

- To safeguard the stability and biological activity of the HPV peptides, a controlled temperature environment was maintained. The peptides were diligently stored at -20°C, a temperature that is well-suited for preserving the long-term viability of biomolecules.

Use of Aliquots:

- To minimize the risk of degradation and maintain the peptides in their pristine condition, a prudent approach was adopted. The peptides were divided into small,

manageable portions known as aliquots. Each aliquot contained precisely 10 µL of the peptide solution.

- The use of aliquots is a standard practice in laboratory settings for several compelling reasons. Firstly, it prevents repeated freeze-thaw cycles, which can compromise the integrity of the peptides. Secondly, it minimizes the potential for contamination, ensuring that each aliquot remains uncontaminated and ready for use.

- By dividing the peptides into aliquots, we ensured that only the required amount of the peptide solution was thawed and utilized for each experiment, preserving the quality and potency of the remaining peptide stock.

Precision and Consistency:

- The careful management of peptide storage, with the use of aliquots and adherence to a consistent storage temperature, guaranteed that the peptides retained their full biological activity. This precision and consistency were paramount, as even minor variations in storage conditions could impact the reliability and reproducibility of experimental results.

In summary, the storage of HPV peptides (E6, E7, L1) was meticulously executed, with a focus on maintaining a controlled temperature environment and using aliquots to preserve the integrity and potency of these vital components. This ensured that the peptides remained in optimal condition for use in our experiments, enabling us to conduct our research with precision and confidence.

Stimulation and Incubation:

The stimulation and incubation phase is a crucial step in our experimental process. During this stage, we carefully initiated the interaction between Peripheral Blood Mononuclear Cells (PBMCs) and the stimuli, ensuring optimal conditions for cellular activation and observation:

Stimuli and Co-Stimulation Molecules:

- With precision, we added 100 µL of PBMCs and 100 µL of the prepared stimulus to each well of the experimental plate. This step initiated the immune response within the wells, as PBMCs began to interact with the HPV peptides (E6, E7, L1) or the negative control, Actin.

- To enhance the cellular response, we introduced co-stimulation molecules, CD28 (0.8 µL) and CD49d (0.8 µL). These molecules facilitated the activation of T cells, optimizing their responsiveness to the stimuli.

Verification of Cell Distribution:

- To ensure uniformity across the wells, we carefully inspected the distribution of cells within each well under a microscope. This step confirmed that PBMCs were evenly dispersed throughout the well, minimizing experimental variability.

Incubation Duration:

- The experimental plate was placed in an incubator, providing a stable temperature and atmosphere conducive to cellular activity.

- The incubation period lasted 20-24 hours, chosen precisely to allow for the activation of CD137, a critical marker of T-cell activation.

- During this incubation period, PBMCs underwent complex immune responses driven by their interactions with HPV peptides or the negative control. These interactions were closely monitored, providing valuable insights into immune dynamics in response to HPV antigens.

In summary, the stimulation and incubation phase were conducted meticulously, involving the addition of PBMCs, stimuli, and co-stimulation molecules, along with verification of uniform cell distribution. The incubation period enabled the activation of immune responses, setting the stage for subsequent stages of our investigation into T-cell behavior in response to HPV antigens.

Staining:

The staining phase is a pivotal step in our experimental protocol, where we prepare the cells for analysis and further investigation:

Supernatant Collection for Cytokine Detection:

- If required for specific analysis, we meticulously collected 100 µL of supernatant from each well after centrifugation. This supernatant contained valuable information about cytokine production by the cells and was preserved at a frigid temperature of -80°C to maintain its integrity for future assessment.

Centrifugation at 1500 rpm for 2 Minutes:

- To separate the cellular components from the surrounding liquid, we conducted centrifugation (Thermo Scientific™ Megafuge™ Series 16, M-20 Microplate Swinging Bucket Rotor-75003624) at 1500 rotations per minute (rpm) for a duration of 2 minutes. This step ensured that the cells were isolated and ready for further processing.

Washing of Cells with PBS Containing 2 mM EDTA:

- Post-centrifugation, the cells were washed meticulously with 200 µL of a phosphate-buffered saline (PBS) solution containing 2 millimolars (mM) of ethylenediaminetetraacetic acid (EDTA), Euroclone. This step served a dual purpose - it helped remove any residual substances and prepared the cells for subsequent staining.

Application of Violet Die Staining:

- To visualize and label the cells, we introduced Violet Die ((Invitrogen, Waltham, Massachusetts, USA), which had been diluted to a specific concentration of 1:1000 in PBS. This meticulously prepared staining solution was added to each well, with a precise volume of 100 µL.

- The plate, now containing the stained cells, was maintained at a temperature of 4°C for a carefully controlled duration of 30 minutes. During this incubation, the Violet Die interacted with the cells, allowing for their visualization and subsequent analysis.

Final Centrifugation and PBS Wash:

- Following the staining incubation, the plate was subjected to another round of centrifugation at 1500 rpm for 2 minutes. This step ensured that any excess staining solution was removed, leaving us with cells that had taken up the Violet Die and were ready for downstream analysis.

- To conclude this phase, we washed the cells once more with PBS, ensuring that they were free of any residual staining solution and prepared for the final stages of our experimental analysis.

In summary, the staining process was carried out meticulously to prepare the cells for analysis, with a focus on preserving and visualizing critical cellular characteristics for our investigation into T-cell behavior in response to HPV antigens.

Antibody Mixing and Cell Preparation:

This phase is pivotal in our experimental process, as it involves the precise preparation of the cells for flow cytometry analysis, enabling the assessment of T-cell activation and marker expression:

Antibody Mix Preparation:

- To initiate this step, we meticulously prepared an antibody mix within each well. This mix included PBS with 5% Fetal Calf Serum (FCS), creating a medium conducive to antibody interactions with the cells. Each well received 50 µL of this prepared medium, facilitating the subsequent steps.

CD4 Antibody Introduction:

- For the identification and characterization of CD4-positive T cells, we introduced 1.3 µL of CD4 APC Cy7 antibody into each well.
- CD4 Antibody Information:
    - Supplier: BD Pharmingen
    - Type: Mouse anti-Human
    - Clone: RPA-T4
    - Catalog Number: 557871
    - Quantity: 100 tests
    - Volume: 0.5 ml
    - Concentration: 5 µL per test
    - Storage: 2-8°C

CD8 Antibody Introduction:

- To distinguish and assess CD8-positive T cells, we added 1.5 µL of CD8 V500 antibody to each well.
- CD8 Antibody Information:
    - Supplier: BD Horizon
    - Type: Mouse anti-Human
    - Clone: RPA-T8
    - Catalog Number: 560774
    - Quantity: 100 tests
    - Volume: 0.5 ml
    - Concentration: 5 µL per test
    - Concentration: 200 µg/ml

- Storage: 2-8°C

CD137 Antibody Introduction:

- To evaluate T-cell activation, we included 2.5 µL of CD137 PECY5 antibody in each well.
- CD137 Antibody Information:
    - Supplier: BD Pharmingen
    - Type: Mouse anti-Human
    - Catalog Number: 551137
    - Quantity: 100 tests
    - Volume: 2 ml
    - Storage: 4°C

Note: CD137, 4-1BB, is a member of the TNF superfamily and serves as an indicator of T-cell activation [129].

CD25 Antibody Introduction:

- To monitor T-cell activation and immune responses, we introduced 1.3 µL of CD25 PECY7 antibody into each well.
- CD25 Antibody Information:
    - Supplier: BD Pharmingen
    - Type: Mouse anti-Human
    - Catalog Number: 557741
    - Quantity: 100 tests
    - Volume: 0.5 ml
    - Storage: 4°C

Note: CD25 expression serves as an indicator of T-cell activation [130].

Uniform Mixing:

- The antibody mix, now containing specific antibodies, was added to each well, ensuring a uniform distribution. A precise volume of 56.6 µL was added to each well.

Incubation at 4°C:

- Subsequently, the plate was maintained at a controlled temperature of 4°C for 30 minutes. During this incubation, the antibodies bound to their respective markers on the T cells, facilitating their subsequent identification and analysis.

In summary, the antibody mixing and cell preparation process was conducted meticulously, ensuring that the cells were labeled with specific antibodies and prepared for flow cytometry analysis. This phase played a pivotal role in enabling us to assess T-cell activation and marker expression in response to HPV antigens, providing valuable insights into the immune dynamics under investigation.

Final Steps:

In the concluding phase of our experiment, we executed critical steps to prepare the cells for analysis and initiate the assessment of T-cell activation and marker expression:

Cell Pelleting and Washing:

- Initially, the cells were subjected to a process of pelleting, followed by a thorough wash with phosphate-buffered saline (PBS) enriched with 5% Fetal Calf Serum (FCS). This meticulous washing step ensured that any residual substances or contaminants were removed, leaving us with clean and well-prepared cells.

Cell Resuspension in Paraformaldehyde (PF 1%):

- Following the wash, the cells were resuspended in a solution of 1% paraformaldehyde (PF), Euroclone-31628.01 . This crucial step served to fix the cells in their current state, preserving their characteristics and preventing any further changes before analysis.

Data Acquisition with Flow Cytometry:

- The culmination of our experimental journey involved the data acquisition phase, where we utilized a flow cytometer (BD FACSLyric™ and BD FACSuite™ software) protocol designed specifically for the AIM-HPV assay. This advanced technology allowed us to perform precise and comprehensive analyses of the cells.

- The ultimate goal of this data acquisition was to assess T-cell activation and marker expression. We meticulously recorded and analyzed the data generated

by the flow cytometer to gain insights into how T cells responded to HPV peptides and how various markers indicative of activation were expressed.

This rigorous and systematic experimental procedure was adhered to with the utmost care and precision. It was meticulously followed to investigate T-cell responses to HPV peptides and to assess T-cell activation through AIM markers. By doing so, our research contributed significantly to the broader understanding of immune dynamics in the context of HPV exposure [131].

Lymphocyte proliferation Assay (LPR)

Preparation a Day Before the Experiment:

1. Ensured an adequate supply of LPR medium a day before the experiment. In case of insufficient quantity, LPR medium with 5% human serum was prepared in advance.

2. Calculated the required amounts of antigens and reagent concentrations a day ahead and made a record for easy reference on the day of the experiment.

3. Reviewed the protocol and visualized the experiment steps a day in advance to mentally prepare for the procedure.

4. Considered potential challenges or issues that might arise during the experiment.

On the Day of the Experiment:

1. First, removed the LPR medium from storage and allowed the antigens to reach room temperature.

Preparing medium used for plating the cells (Pmedium):

To create the plating medium, also known as Pmedium, the following components were assembled:

1. RPMI (without FCS): This was a basal medium providing essential nutrients and buffering capacity for cell growth.

2. 1% PSG (Penicillin-Streptomycin-Glutamine): PSG was added to prevent bacterial and fungal contamination while also providing L-glutamine, which was vital for cell survival and growth.

3. 10% Human serum: Human serum, Sigma, was rich in growth factors, hormones, and proteins, making it an essential component for supporting cell growth and proliferation. It provided the necessary nutrients and signaling molecules to sustain cells during the experiment.

4. 1% NaPyruvate: Sodium pyruvate (Gibco, Grand Island, NY, USA) was included to enhance the energy metabolism of the cells, as it could be readily converted into ATP, the cellular energy currency.

5. 1% MEM (Minimum Essential Medium): MEM (Gibco, Grand Island, NY, USA) contained additional amino acids, vitamins, and minerals, which contributed to the overall nutritional support of the cells.

6. 0.1% β-mercaptoethanol: β-mercaptoethanol (Gibco, Grand Island, NY, USA) acted as a reducing agent and helped maintain the reducing environment within the cell culture, which was important for the stability of cellular components.

Preparing medium used for resuspending stimulating peptides or viral antigens (Smedium):

For creating the resuspending medium, referred to as Smedium, the following components were gathered:

1. RPMI (without FCS): This served as the base medium for resuspending and diluting the stimulating peptides or viral antigens.

2. 1% PSG (Penicillin-Streptomycin-Glutamine): PSG was included to maintain a sterile environment and support the overall health of the cells during the experiment.

3. 1% NaPyruvate: Sodium pyruvate was included to enhance cellular energy production and metabolism, which was crucial for maintaining cell viability and functionality.

4. 1% MEM (Minimum Essential Medium): MEM provided additional nutrients and essential factors that helped sustain cell health and function.

5. 0.1% β-mercaptoethanol: β-mercaptoethanol contributed to the stability of cellular components by maintaining a reducing environment within the cell culture.

The distinguishing factor between Pmedium and Smedium was the presence of 10% human serum in Pmedium. To align both mediums, a single 5% human serum solution was prepared and used for both applications. This ensured consistency in the growth and stimulation conditions for the cells, allowing for meaningful and reproducible experimental results [132].

If the Cells were Stored in Liquid Nitrogen:

When cells were stored in liquid nitrogen and needed preparation for an experiment, the following steps were meticulously followed:

1. Thawed the Cells Rapidly: The vial containing the frozen cells was swiftly thawed by placing it in a 37°C water bath. The warm water bath facilitated the quick thawing process, preventing potential cell damage that could occur during slow thawing.

2. Incubated Cells Overnight: After thawing, the vial was transferred to the controlled environment of a CO2 incubator set at 37°C. The cells were left to incubate overnight. This extended incubation period helped the cells recover and adjust to their new environment, allowing them to regain their optimal physiological state.

3. Cell Counting: Once the cells had ample time to recover, their concentration was determined through cell counting using a suitable method, aiming for a target of 200,000 cells per well. Precise cell counting ensured consistent seeding and contributed to the reliability of experimental results.

4. Resuspended Cells in Pmedium: To prepare the cells for plating, they were resuspended in the appropriate volume of Pmedium. Each well typically required 100 µl of cell suspension. For instance, when seeding cells in 21 wells, the cells were calculated and resuspended in a total volume of 2100 µl of Pmedium. Ensuring the cells were properly suspended in the medium was crucial for their even distribution and successful culture in the experimental wells.

By following these steps meticulously, the recovery and vitality of cells from liquid nitrogen storage were optimized, setting the stage for a successful lymphoproliferation assay (LPR) [131].

Preparing the HPV Antigens (E6, E7, and L1) or Stimuli:

For accurate and effective preparation of HPV antigens or other stimuli, the following guidelines were followed:

Stimuli Concentration:

A consistent concentration of 0.1 µg in a final volume of 200 µl was maintained for the experiments. To achieve this precise concentration, the stimuli were initially prepared at a concentration of 0.2 µg/ml. This concentration was suitable for plating 100 µl per well, achieved by adding 1 µl of the peptide to every 100 µl of Smedium.

Actin Concentration:

Actin, with an initial concentration of 810 µg/ml, was diluted to a final concentration of 0.2 µg/ml. This concentration was optimal for the experimental requirements.

By adhering to these concentration specifications, the stimuli and Actin were appropriately prepared for the lymphoproliferation assay (LPR). This precision in preparation ensured reliable and meaningful results in the experiments conducted.

Seeding the cells was conducted as follows:

1. Each well was seeded with 100 µl of cells and 100 µl of peptide stimulus.

2. To prevent evaporation, PBS was added around the empty wells.

To mitigate the occurrence of the Edge-effect, wherein results from cells at the edges differ from those in the middle of the plate, phosphate buffer was added to the plate wells.

1. The cells were meticulously examined under a microscope to ensure their even distribution within the wells.

2. Subsequently, the plate was placed in an incubator set at 37°C for a duration of 7 days [133, 134].

Staining:

Preparation the Day Before Staining:

1. A day before the experiment, the required quantities of antibodies were meticulously calculated. This involved determining the necessary amounts of both the medium and antibodies, and these calculations were meticulously documented for future reference.

2. To ensure preparedness for the experiment on the following morning, the antibodies were systematically organized and placed in a separate rack.

3. The availability of essential reagents, including PBS, PBS-FBS, PBS-EDTA, and others, was verified. In cases where quantities were found to be insufficient, the required preparations were made to ensure an adequate supply.

4. In order to optimize the execution of the staining procedure, mental visualization was employed to anticipate potential issues that might arise during the experiment. This proactive approach helped to identify and address challenges in advance.

On the Staining Day:

Cell Examination:

- Cells were examined under a microscope to detect cell aggregates, indicating potential cell proliferation.

Centrifugation and Washing:

- A centrifugation step at 1500 rpm for 2 minutes was performed.
- Cells were washed with PBS containing 0.5 mM EDTA.

Staining Initiation:

- Cells were resuspended in 100 µl of PBS per well.

- Violet Dye (1:1000) was added, and cells were incubated at 4°C for 30 minutes.
- Following incubation, another washing step with PBS was performed.

Antibody Master Mix Preparation:

- Antibody master mix (1:160 µl of CXCR5) was prepared in 50 µl of PBS with 5% FCS per well.
- Cells were incubated for an additional 20 minutes at room temperature.
- Post-incubation, another thorough wash was carried out.

Note: CXCR5, also known as CD185 or Burkitt lymphocyte receptor 1 (BLR1), is expressed on both CD4 and CD8 T cells, with particular prominence as a defining marker for T Follicular Helper (Tfh) cells [135]

Second Antibody Master Mix:

- Antibody master mix (Anti-IgG2b Biotin at 1:1000) was prepared in 100 µl of PBS with 5% FCS per well.
- Cells underwent a 20-minute incubation at room temperature.
- Subsequently, they were washed again.

Note: Anti-IgGb serves as a signal amplifier for fluorescent signals in various immunostaining applications. The technique involving Anti-IgGb, often associated with Tyramide Signal Amplification (TSA), enhances the sensitivity of detection, making it effective for identifying low-abundance target [136].

Antibody Characteristics:

- Manufacturer: SouthernBiotech
- Size: 2ml
- Concentration: 0.5 mg/ml
- Catalog Number (Cat): 1090-08
- Type: Goat Anti-Mouse
- Storage Conditions: Store at 2-8 °C.

Final Antibody Mix:

- Final antibody mix (Streptavidin Brilliant Violet 510 at 1:800) was prepared in 50 µl of PBS with 5% FCS per well.

Streptavidin Characteristics:

- Supplier: Biolegend
- Catalog Number: 405233
- Size: 100 µl
- Concentration: 0.1 mg/ml
- Storage: Store at 2-8°C.

Note: Streptavidin serves as a signal amplifier for fluorescent signals , particularly beneficial for visualizing medium- and low-abundance targets [137].

- Antibodies Used:

  CD3 PerCP.Cy 5.5, CD25 PeCy7, CD4 APC Cy-7, CD8 FITC, CCR6 APCR 700, ICOS APC, CXCR3 PE.

CD3 PerCP.Cy 5.5 Characteristics:

- Brand: BD Pharmingen
- Type: Mouse Anti-Human
- Clone: UCHT1
- Catalog Number: 560835
- Size: 50 Tests (5ul/test), 0.25 ml
- Storage: 4 °C

CD25 PeCy7 Characteristics:

- Brand: BD Pharmingen
- Type: Mouse Anti-Human
- Clone: M-A251
- Catalog Number: 557741
- Size: 100 Tests, 0.5 ml
- Storage: 4 °C

CD4 APC Cy-7 Characteristics:

- Brand: BD Pharmingen
- Type: Mouse Anti-Human
- Clone: RPA-T4
- Catalog Number: 557871
- Size: 100 Tests, 0.5 ml
- Volume per Test: 0.5 µl/test
- Storage: 2-8 °C

CD8 FITC Characteristics:

- Brand: BD
- Clone: SK1
- Reference Number (REF): 345772
- Size: 100 Tests, 20 µl
- Storage: 2-8 °C

CCR6 APCR 700 Characteristics:

- Brand: BD Horizon
- Clone: 11A9
- Catalog Number (Cat): 565173
- Size: 50 Tests, 0.25 ml
- Volume per Test: 5 µl
- Concentration: 200 µg/ml
- Storage: 2-8 °C

ICOS APC Characteristics:

- Brand: Invitrogen, eBioscience
- Catalog Number (REF): 17-9948-42

- Clone: ISA-3
- Size: 100 Tests
- Volume per Test: 5 µl or 0.125 µg
- Storage: 4 °C

CXCR3 PE Characteristics:

- Brand: Biolegend
- Clone: G025H7
- Catalog Number (Cat): 353706
- Size: 100 Tests
- Volume per Test: 5 µl
- Concentration: 100 µg/ml
- Storage: 2-8 °C

Note: Th17 cells predominantly express the CC chemokine receptor (CCR6) and produce its ligand, CCL20. This chemokine receptor plays a role in the migration and recruitment of Th17 cells to inflammatory and inflamed sites, indicating its significance in the immune response [138].

Note: ICOS, also known as 4IBB or CD278, is a member of the CD28 and CTLA-4 cell-surface receptor family. ICOS is involved in immune responses and is required for the effective production of various cytokines, including IL-2, IL-4, IL-5, and IFNγ, from activated T cells [139].

Note: CXCR3, also known as GPR9/CD183, is a chemokine receptor expressed on Th1 cells. Th1 cells are a subset of T helper cells that play a crucial role in cell-mediated immunity. The expression of CXCR3 on Th1 cells is associated with Th1-induced inflammatory conditions, and its role extends to skin inflammation and immune responses in various contexts [140].

Incubation and Final Steps:

- Cells were incubated with the final antibody mix for 30 minutes.
- Afterward, cells were washed with PBS containing 5% FCS.
- Finally, cells were resuspended in 250 µl of PF%.

- This comprehensive flow cytometry staining protocol ensures accurate analysis of T cell subsets, including the expression of CXCR5 [131].

Statistical Analysis

All statistical analyses were conducted using Python 3.13, employing a diverse set of libraries for various analytical aspects:

Data Preparation and Import:

- Utilized the pandas library for efficient data import and structuring, ensuring the dataset's readiness for subsequent analyses [141, 142].

Mathematical and Statistical Analysis:

- Applied numpy, scipy.stats, and penguin libraries for mathematical and statistical analyses. These tools offered functions for hypothesis testing, effect size estimation, and Bayesian analysis, providing a comprehensive exploration of the dataset [143, 144].

Data Visualization:

- Employed matplotlib [145] and seaborn [146] libraries for graphical representation, enabling the creation of informative visualizations such as graphs and plots. The seaborn library further enhanced visualizations through annotation capabilities [146].

Power Analysis and Effect Size Estimation:

- Utilized the statsmodels.stats.power [147] and penguin libraries [148] for assessing statistical power and effect size. This approach facilitated a thorough evaluation of statistical significance and the magnitude of observed effects, with the interpretation of Bayes Factor (Bf10) guided by established guidelines Table 1,2.

Table 1 Bayes Factor Interpretation:

Bf10	**Interpretation**
> 100	Extreme evidence for H1
30-100	Very strong evidence for H1
10-30	Strong evidence for H1
3-10	Moderate evidence for H1
1-3	Anecdotal evidence for H1
1	Equal evidence for H1 and H0
1/3 - 1	Anecdotal evidence for H0
1/10 - 1/3	Moderate evidence for H0
1/30 - 1/10	Strong evidence for H0
1/100 - 1/30	Very strong evidence for H0
<1/100	Extreme evidence for H0

Table 2 Effect Size Estimation (Cohen's d):

Relative Size	**Effect Size**	**% of the control group below the mean of the experimental group**
0.0	small	50
0.2	Medium	58

0.5	Large	69
0.8	Very Large	79
1.4	Extremely large	92

For parametric data in the AIM assay, the student-t-test was employed. Normality of data was assessed using the Shapiro-Wilk test and pp-plot. Non-parametric data, on the other hand, was analyzed using the Mann-Whitney test, ensuring robust statistical comparisons between the two groups.

This comprehensive statistical approach allowed for a rigorous examination of the dataset, ensuring the validity and reliability of the study's findings.

Calculation of Minimal Detectable Change (MDC) using Distribution-Based Method

In the assessment of the Minimal Detectable Change (MDC), a distribution-based method was employed to quantify the smallest change in measurements that can be considered beyond the scope of random measurement error. This method involves the utilization of the Standard Error of Measurement (SEM), which is calculated as the standard deviation of the differences between repeated measurements divided by the square root of 2. The formula for MDC using the 95% confidence interval is expressed as:

$$MDC = SEM \times 1.96$$

Where:

- *SEM* is the Standard Error of Measurement.

This computation provides a threshold for change that surpasses the inherent variability in the measurements, offering a reliable indicator of meaningful alterations in the assessed parameter [149].

Immunoinformatics

Retrieval and Comparative Analysis of Amino Acid Sequences

HPV Genotype Classification:

1. High-Risk Genotypes: HPV genotypes associated with a high risk of infection and potential carcinogenicity, including 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58, 59, and 68.

2. Probable High-Risk Genotypes: Genotypes with a likelihood of high risk, pending further research to confirm their role, comprise 26, 53, 66, 70, 73, and 82.

3. Low-Risk Genotypes: Genotypes considered low risk for infection and carcinogenicity encompass 6, 11, 40, 42, 43, 44, 54, 61, and 81.

4. Unknown Risk Genotypes: Genotypes with uncertain risk profiles necessitating additional investigation include 62, 67, 83, and 89.

Amino Acid Sequence Retrieval and Analysis:

The research embarked on a comprehensive analysis of the genetic composition and evolutionary affinities of HPV proteins, specifically E6, E7, and L1. The following sources were instrumental in this endeavor:

1. National Center for Biotechnology Information (NCBI): A prime repository of biological data, NCBI facilitated access to amino acid sequences of target proteins in FASTA format, ensuring compatibility with computational analyses [150].

2. UniProt: The Universal Protein Knowledgebase, known for its comprehensive and high-quality protein sequence information, played a crucial role. UniProt contributed essential protein sequences, enhancing the robustness of the analysis [151].

3. Papillomavirus Episteme (PaVE): Acknowledging the significance of specialized resources, PaVE, dedicated to papillomaviruses, provided vital sequences, aligning with the specific focus of the study [21].

Incorporating Evolutionary Perspective:

To comprehensively explore evolutionary relationships, our study expanded the dataset to include amino acid sequences from diverse members of the papillomavirus family.

Sequence Alignment and Conservancy Analysis:

1. Tool Utilized: The Clustal Omega multiple sequence alignment tool played a pivotal role in our analysis. Renowned for its accuracy and efficiency, it facilitated a meticulous alignment of amino acid sequences [152].

2. Analysis Process:

    - Alignment Precision: Clustal Omega's capabilities enabled precise sequence alignment, uncovering both similarities and differences among the retrieved amino acid sequences.

    - Conserved Regions: The tool effectively highlighted conserved regions within the target proteins, offering insights into areas with retained structural and functional significance.

    - Variation Identification: Rigorous analysis unveiled variations within the target proteins, providing valuable information about regions subjected to evolutionary changes.

3. Insights Gained:

    - Sequence Conservancy: The alignment process yielded a nuanced understanding of the sequence conservancy among the target proteins.

    - Functional Importance: Identification of conserved regions shed light on areas crucial for the functional aspects of the proteins.

    - Evolutionary Constraints: Variations within the sequences revealed regions subjected to evolutionary constraints, contributing to a comprehensive understanding of the proteins' evolution.

Phylogenetic Analysis:

To delve into the evolutionary relationships among the sequences, a phylogenetic analysis was executed. This analysis provided a phylogenetic tree that illustrated the evolutionary divergence and clustering of the target proteins.

To perform this task, Tools Employed:

- MEGA 11

- Interactive Tree of Life (iTOL) Online Server

Execution

Process: The execution involved the construction of a phylogenetic tree, providing a visual representation of the evolutionary dynamics within the papillomavirus family.

Visualization: iTOL, a state-of-the-art online tool, enabled the visualization of the intricate evolutionary relationships among the target proteins [153].

Framework for Investigation

Comprehensive Approach:

- Sequence Retrieval: Diverse data sources were integrated to gather a comprehensive set of sequences for analysis.
- Analysis: MEGA 11 [154] and iTOL were employed synergistically to analyze the sequences, ensuring a multifaceted exploration of evolutionary patterns.

Research Foundation: This approach laid a robust foundation for our study, facilitating an in-depth investigation into the genetic underpinnings and evolutionary context of the target proteins.

Contribution to Understanding: The integration of diverse data sources and cutting-edge bioinformatics tools elevated the depth and rigor of our research. This, in turn, fostered a profound understanding of papillomavirus biology and its broader implications.

Physicochemical and Secondary Structural Analysis of Target HPV antigens

In-depth characterization of the target HPV E6, E7, and L1 and newl0ey designed vaccine required a comprehensive assessment of their physicochemical attributes and secondary structural elements. This analytical endeavor was facilitated by utilizing specialized bioinformatics tools and adhering to stringent parameters for the derivation of meaningful results.

Physicochemical Analysis:

Tool Utilization: The Expasy Protparam tool (Expasy Protparam) was employed, renowned for its reliability. It calculated crucial parameters including molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and GRAVY, collectively providing a comprehensive portrait of the antigens' physicochemical characteristics [155].

Antigenicity Assessment

Importance: A pivotal aspect of the analysis was the determination of antigenicity for each protein, addressing immunological relevance and vaccine candidacy.

Tool Utilization: The VaxiJen 2.0 server, known for proficiency in antigen prediction, was employed with a judiciously set antigenicity threshold of 0.4.

This ensured a stringent assessment of antigenic potential, crucial for discerning immunological relevance [114].

Secondary Structural Prediction

Task Complexity: Prediction of secondary structural elements (helices, sheets, turns, and coils) required precision.

Tool Utilization**:** The SOPMA secondary structure analysis tool, distinguished for reliability and accuracy, was utilized. Parameters such as the number of conformational states, similarity threshold, and window width remained unaltered, ensuring standardized and rigorous analysis of secondary structural components [156].

Overall Impact

Insights: Meticulous analysis of physicochemical attributes and secondary structural features provided a comprehensive understanding of intrinsic properties. This analysis played a pivotal role in evaluating the potential utility of viral antigens in immunogenicity and therapeutic development.

Enrichment: The integration of these insights significantly enriched the investigation, offering a deeper comprehension of the structural and functional attributes of the target HPV antigens.

Homology Modeling and Structural Validation

A pivotal phase in this study involved the construction and validation of three-dimensional (3D) structural models for the target E6, E7, and L1 antigens. This crucial task was executed with the aid of established homology modeling tools, characterized by their utility and reliability. The resultant models underwent rigorous scrutiny to ensure their quality, integrity, and suitability for subsequent investigations.

Homology Modeling:

1. Tool Utilization: Phyre 2, a widely acknowledged homology modeling tool, was employed to craft accurate 3D structures. It infers protein structures through sequence alignment with homologous proteins of known structures [157].

Structural Refinement

2. Addressing Distortions: Resultant structural models underwent meticulous refinement using the 3Drefine tool. This tool, known for enhancing structural accuracy, was applied as needed based on the Ramachandran plot [158]. The refinement process aimed to ameliorate any distortions that may have arisen during the homology modeling process. It focused on optimizing models, aligning them with the inherent structural parameters and constraints of the target viral antigens.

Structural Validation

3. Quality Assessment: Thorough validation of refined structural models was conducted using the PDBsum server, accessible at PDBsum [159].

4. Principle Utilized: The PROCHECK principle was integrated into the validation process, with a specific focus on the Ramachandran plot analysis. This plot assesses the stereochemical quality of protein structures, segregating Glycine and Proline residues [160].

Model Selection

Optimal Results: Structural models demonstrating optimal results through rigorous analysis were thoughtfully selected for subsequent utilization.

Impact of Integration: The meticulous integration of homology modeling and validation procedures significantly fortified the integrity of the structural models. This enhancement increased their potential utility in advancing investigative pursuits.

Prediction of T Cell Epitopes, Including CTL Epitopes

The identification and characterization of T cell epitopes, particularly cytotoxic T lymphocyte (CTL) epitopes, are fundamental components of this investigation. The prediction of these epitopes, which play a pivotal role in immune responses, was executed with precision and rigor. A stringent methodology, guided by recognized bioinformatics tools, ensured the comprehensive prediction of epitopes.

CTL Epitope Prediction

1. Objective: The investigation focuses on the identification and characterization of cytotoxic T lymphocyte (CTL) epitopes, crucial components of immune responses.

2. Methodology: The NetCTL 1.2 server was employed for the precision prediction of 9-mer T cell epitopes. This prediction specifically targeted commonly occurring HLA Class I supertypes: HLA-A, HLA-B, HLA-C.

3. Crucial Parameters:

    - Transporter Associated with Antigen Processing (TAP) transport efficiency threshold: 0.05

    - Proteasomal C-terminal cleavage threshold: 0.15

    - Epitope identification threshold: 0.75 These thresholds ensured a robust evaluation of epitope candidacy within the NetCTL 1.2 framework [121].

Expanding Epitope Coverage

Inclusive Approach: In addition to the mentioned supertypes, the study extended epitope prediction to HLA Class I alleles: HLA-B*07:02, HLA-C*06:02, HLA-C*12:02, HLA-C*12:03, HLA-A*02:01, HLA-A*11:01, HLA-A*24:02, HLA-A*68:01, HLA-B*27:01, HLA-B*57:01, HLA-B*52:01, HLA-B*58:01, HLA-C*07:01, and HLA-C*04:01. The Immune Epitope Database-Consensus (IEDB) method was strategically employed for this purpose [118].

Coverage Significance**:** Encompassing a diverse array of HLA Class I epitopes addressed immunological diversity globally, covering more than 90% of the population.

Selection of Strong Binders

Robust Criteria: Only peptides with a consensus score of ≤2 were considered strong binders for CTL epitopes, ensuring the selection of epitopes with the highest binding affinities.

Significance: This stringent criterion played a crucial role in identifying epitopes with high binding affinities, emphasizing their potential significance in immune responses and vaccine development [161].

Comprehensive Understanding

Meticulous Approach: The comprehensive epitope prediction process, driven by a judicious selection of parameters and a wide-ranging array of HLA Class I alleles, fortifies the study's capacity to unravel immunologically relevant epitopes.

Contribution: This meticulous approach contributes to the identification of epitopes with high binding affinities, advancing our understanding of their potential roles in immune responses and immunotherapeutic applications.

Prediction of Helper T Lymphocyte (HTL) Epitopes

The meticulous prediction of HTL epitopes, specifically those of 15-mer length, is an indispensable component of this study, aimed at comprehensively understanding immunological responses. The prediction of these epitopes, pivotal for their role in the immune system, was conducted with a high degree of precision and adherence to established bioinformatics tools.

HTL Epitope Prediction:

Objective: The study prioritizes the meticulous prediction of 15-mer HTL epitopes, crucial for a comprehensive understanding of immunological responses and their role in the immune system.

Tools Used:

- Net MHC II pan 4.1 Server [119].
- Immune Epitope Database (IEDB) [118]

HLA Class II DRB1 Alleles Considered:

- 01:01, 03:01, 04:01, 07:01, 08:03, 10:01, 11:01, 12:01, 13:02, 14:01, 15:01, HLA-DQB1*06:02

Global Coverage:

Alleles were carefully selected to ensure coverage of more than 95% of the global population, reflecting their prevalence across diverse human populations.

Categorization of Epitopes

Assessment Criteria:

- Strong binders: Percentile rank of 2% or less
- Intermediate binders: Percentile rank of 10%
- Non-binders: Percentile rank greater than 10%

Rigorous Evaluation: The stringent categorization allowed for a comprehensive evaluation of epitope candidacy within the study's scope, reflecting binding affinities and immunological relevance.

Study Fortification: The comprehensive epitope prediction process, guided by the selection of a diverse set of HLA Class II DRB1 alleles and adherence to strict categorization criteria, fortifies the study's ability to identify epitopes with strong binding affinities.

Significance in Vaccine Development

Instrumental Epitopes**:** The resulting epitopes are instrumental in advancing our understanding of immune responses and hold potential applications in the context of immunology and vaccine development [162].

Identification of Promiscuous and Overlapping T Cell Epitopes

Promiscuous Epitope Analysis

Significance in Vaccine Design: Promiscuous T cell epitopes, capable of binding with high affinity to multiple Human Leukocyte Antigens (HLAs), are crucial in vaccine design. They play a pivotal role in generating robust and comprehensive immune responses, owing to their ability to interact with various allelic forms of HLAs.

Identification Criteria: In this study, T cell epitopes from both HLA Class I and Class II were meticulously selected based on their high binding affinity to multiple HLAs. These epitopes, recognized for their promiscuity, were earmarked as pivotal candidates for further investigation in the context of immunogenicity and vaccine development [163].

Overlapping Epitope Assessment

Special Interest in Overlapping Epitopes: Overlapping epitopes, activating both cytotoxic T cells (CTLs) and helper T cells (HTLs), were a focus due to their integral sequences comprising both CTL and HTL epitopes. HTL epitopes with high binding affinities were scrutinized for overlaps with strong binding affinity CTL epitopes.

Unified Peptide Fragment: Identified overlaps were thoughtfully listed and amalgamated into a unified peptide fragment. This strategic merging of epitopes facilitated a more integrated and comprehensive understanding of their functional roles in immune responses and vaccine design [161].

Significance and Potential Impact

Enriching Insights: The identification and categorization of promiscuous and overlapping T cell epitopes significantly enrich our insights into the nuanced dynamics of immunogenicity, paving the way for innovative vaccine strategies.

Potential for Revolutionizing Vaccine Development: These epitopes, distinguished by their unique attributes, hold the potential to revolutionize vaccine development and therapeutic interventions[164].

Prediction of Peptide Immunogenicity

VaxiJen v2.0 Tool: In this study, the antigenicity of predicted promiscuous epitopes was assessed using the VaxiJen v2.0 tool. This tool is renowned for its proficiency in antigenicity prediction, offering a comprehensive analysis of epitopes' potential to induce an immune response.

Stringent Antigenicity Prediction Threshold: A crucial aspect of the analysis involved the meticulous setting of a stringent antigenicity prediction threshold at

0.4. This threshold served as a critical criterion for distinguishing epitopes with significant antigenic potential. The rigorous assessment ensured a comprehensive understanding of the immunological relevance of the predicted promiscuous epitopes.

Immunological Insights and Vaccine Design

Enriching Comprehension: The assessment of epitope antigenicity, facilitated by the VaxiJen v2.0 tool, significantly enriched the comprehension of the immunological attributes of the predicted promiscuous epitopes. This insight is instrumental in evaluating their suitability for vaccine development and understanding their potential roles in therapeutic interventions.

The strategic use of the VaxiJen v2.0 tool, with a stringent threshold, contributes valuable insights for selecting epitopes with significant antigenic potential, enhancing their potential utility in vaccine design and therapeutic applications [114].

Identification of B Cell Epitopes: Linear and Conformational

In this study, the comprehensive identification of B cell epitopes, including both linear and conformational epitopes, was an essential undertaking. These epitopes play a pivotal role in stimulating B cell lymphocytes, triggering their differentiation into memory B cells and plasma cells, thereby contributing to the overall immune response.

Linear B Cell Epitope Identification:

1. BCpred 2.0 and IEDB Servers: The study utilized BCpred 2.0 [165] and IEDB servers [118] to predict linear B cell epitopes within the antigen sequence. This involved a comprehensive analysis considering beta turn prediction, surface accessibility, flexibility, antigenicity, and hydrophilicity. These characteristics provided critical insights into the attributes and potential immunogenicity of the identified linear B cell epitopes .

Conformational B Cell Epitope Prediction:

2. IEDB-Ellipro Server**:** For predicting conformational B cell epitopes, the IEDB-Ellipro server was employed. This server, chosen for its reliability, defined prediction parameters with a minimum score and maximum distance (measured in Angstroms) set at 0.7 and 6 Å, respectively. These parameters ensured precise identification of conformational epitopes and insights into their spatial distribution within the antigen structure [166].

Peptides Conservation Analysis and Population Coverage:

IEDB Conservancy Analysis Tool: To assess epitope conservation, the Immune Epitope Database (IEDB) Conservancy Analysis tool was used. This tool is known for determining the extent of epitope conservation within a range of

protein sequences, identifying epitopes with 100% conservancy for further analysis [167].

IEDB Population Coverage Analysis Tool: The study evaluated population coverage using the IEDB Population Coverage Analysis tool. This tool estimated the fraction of individuals predicted to respond to the screened epitopes based on the frequencies of their Human Leukocyte Antigen (HLA) genotypes. This analysis provided critical insights into epitopes' capacity to elicit immune responses and their coverage within diverse human populations [168].

The dual analysis of epitope conservation and population coverage enriches our understanding of the immunological relevance and potential utility of these epitopes. It aids in identifying highly conserved epitopes with the potential to elicit immune responses across diverse human populations, contributing to the advancement of immunology and vaccine design.

Characterization of Predicted Epitopes in Immunology and Vaccine Development

Physicochemical Properties Analysis:

Expasy Protparam Tool: The physicochemical properties of target HPV antigens were analyzed using the Expasy Protparam tool. This tool facilitated the calculation of crucial parameters, including molecular weight, theoretical isoelectric point (pI), instability index, aliphatic index, and grand average of hydropathicity (GRAVY). These parameters collectively provided a comprehensive perspective on the physicochemical characteristics of the target HPV antigens, enhancing our understanding of their structural and functional properties [155].

Autoimmunity Mitigation:

BLASTP Search Against Human Proteome: To mitigate the risk of inducing autoimmunity, identified epitopes underwent a stringent BLASTP-NCBI search against the human proteome. This search aimed to identify any similarity between the epitopes and human proteins. Epitopes showing similarity to any human protein were meticulously eliminated from further consideration. This step ensures the safety of potential immunotherapeutic interventions by reducing the risk of triggering unwanted immune responses against host tissues [169].

The thorough characterization of predicted epitopes, encompassing physicochemical attributes and safety measures against autoimmunity, ensures a comprehensive assessment of their potential utility in immunological applications. This approach not only enriches our understanding of epitope properties but also contributes to the safe and effective design of immunotherapeutic interventions.

Peptides and HLA Interaction Pattern Analysis: Modeling of Peptides and HLA Alleles

Understanding the interaction patterns between peptides and Human Leukocyte Antigens (HLAs) is a pivotal aspect of this study, as it sheds light on the structural and functional aspects of epitope recognition. This analysis encompasses the modeling of peptide and HLA structures, enabling a comprehensive evaluation of their three-dimensional configurations and interaction dynamics.

Peptide 3D Structure Generation:

PEPFOLD 4 Server: The online accessible PEPFOLD 4 server was employed for generating 3D structures and PDB files of epitopes. This server is recognized for its efficacy in creating spatial representations, enhancing our understanding of epitope structural conformations and properties [126].

HLA Allele 3D Structure Retrieval:

HLA Alleles: X-ray crystallographic structures of three common HLA alleles—HLA-DQ, HLA-DRB115 (HLA Class II alleles), and HLA-A*02:01 (HLA Class I allele)—were retrieved from the Protein Data Bank (PDB) with PDB IDs: 1S9V, 5V4M, and 7RTD, respectively. These structures served as foundational templates for subsequent analyses.

Energy Minimization and Ligand Removal:

Chimera 1.17.1 Software: An energy minimization process using Chimera 1.17.1 software optimized the energy profiles of HLA alleles, removing any previously bound ligands. This ensured stable and energetically favorable configurations for the receptors, free from interfering ligands.

Molecular Docking Analysis:

HDOCK Server:

For protein-protein interaction patterns, the Hdock server was utilized. It offers an integrated platform for efficient protein–protein docking, employing a hybrid algorithm for interaction prediction. Noteworthy features include support for amino acid sequences, a unique hybrid docking strategy, and compatibility with experimental information like protein–protein binding sites [170, 171].

Optimal Model and Visualization:

Hydrophobic Environment: The docking analysis was conducted in a hydrophobic environment, ensuring a realistic representation of interaction conditions.

PyMOL Molecular Graphics System: The optimal model, characterized by epitopes binding within the peptide binding groove of HLA alleles, was selected.

The resulting docked complex was visualized using the PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC.

This comprehensive analysis reveals binding affinities and interaction patterns between epitopes and HLA alleles, crucial for understanding the molecular basis of immune responses and informing the design of immunotherapeutic strategies.

Toll-Like Receptor 9 (TLR-9) - Homology Modeling and Structure Validation

Homology modeling and structure validation of Toll-Like Receptor 9 (TLR-9) play a pivotal role in understanding its three-dimensional conformation and functional attributes. The following methodology was employed for homology modeling using the Phyre2 [157] and SWISS-MODEL servers [172]:

Amino Acid Sequence Retrieval:

UniProt Database: The amino acid sequence of TLR-9 was retrieved from the UniProt database [151] using its specific identifiers - UniProt ID Q9NR96 and PDB ID 5ZLN. This sequence serves as the foundational blueprint for subsequent modeling [1].

Homology Modeling with SWISS-MODEL:

User Input: SWISS-MODEL requires minimal user input, often only the amino acid sequence of the target protein .

SWISS-MODEL Workspace:

Each user is provided with a personal web-based workspace where protein homology models can be built, and results are stored .

Homology Modeling:

The server utilizes automated comparative modeling techniques to generate three-dimensional protein structures.

Continuous Updates:

SWISS-MODEL implements a continuous release mechanism, ensuring models are regularly updated based on the latest template information.

Workspace Features:

A personal web-based environment allows users to conduct multiple modeling projects simultaneously .

In essence, the SWISS-MODEL server streamlines the homology modeling process, providing a user-friendly platform for inputting protein sequences and obtaining 3D structural models through automated comparative modeling [172].

HDOCK Server - Docking

The HDOCK server is a computational platform specializing in protein-protein and protein-ligand docking. The following outlines the general methodology likely employed by HDOCK for docking simulations:

Ligand and Receptor Preparation:

- Ligand: Newly designed multi-epitope vaccine peptides are prepared in PDB or PDBQT format, considering atomic coordinates and necessary parameters.
- Receptor: Target proteins, including antigenic proteins or specific receptors, undergo similar preparation, ensuring correct conformation and relevant details.

Search and Sampling:

- Global Search: HDOCK likely employs global search algorithms to explore the vast conformational space of ligand-receptor complexes, aiming to find the global energy minimum among potential binding poses.
- Sampling Techniques: Techniques like Monte Carlo-based methods or genetic algorithms may be used for sampling various conformations and orientations of ligand and receptor molecules.

Scoring Function: Energy Evaluation:

- HDOCK evaluates binding energies using scoring functions, considering van der Waals interactions, electrostatic interactions, hydrogen bonding, and desolvation energies. The scoring function quantifies the fitness of each ligand-receptor complex.

Clustering and Analysis:

- Clustering: Generated docking solutions are clustered based on structural similarities, identifying distinct binding modes and ranking them by prevalence.
- Analysis: Post-docking analysis tools are employed to analyze interactions, such as hydrogen bonds, hydrophobic contacts, and electrostatic interactions. Visualization tools aid in understanding binding poses and critical residues.

Ranking and Selection:

- Ranking: Complexes are ranked based on binding energies or scores, where lower energies or more negative scores indicate more favorable binding interactions.

- Selection: Best-docked complexes are chosen based on energy scores and structural feasibility, representing the most likely binding configurations between the designed multi-epitope vaccine peptides and target proteins.

By employing these methodologies, the HDOCK server provides valuable insights into potential binding modes and interactions of newly designed multi-epitope vaccine peptides with their target proteins [170, 171].

Construction of the Multi-Epitope Vaccine Sequence

In developing the final multi-epitope vaccine sequence, a meticulous process ensured efficacy and safety, focusing on critical immunogenic properties of Cytotoxic T Lymphocyte (CTL) and Helper T Lymphocyte (HTL).

Epitope Selection:

- Epitopes were selected based on key criteria: high promiscuity, ability to overlap, demonstrated immunogenicity, and proven lack of allergenicity.

Linker Utilization:

- AAY and GPGPG linkers were employed to fuse CTL and HTL epitopes, respectively.

    - *AAY Linker:* Acts as a proteasomal cleavage site, separating epitopes during antigen processing, influencing stability, and enhancing immunogenicity [161] .

    - *GPGPG Linker:* A flexible, glycine-rich spacer preventing steric hindrance, allowing independent epitope folding, optimizing exposure to the immune system [162, 173].

Adjuvant Integration:

- CpG-containing oligodeoxynucleotides, recognized by Toll-Like Receptor 9 (TLR-9), were integrated to enhance immunogenicity.

    - K-type CpG-ODNs, prominent in human clinical trials, served as effective adjuvants [162].

Connection with Linker:

- An EAAAK linker facilitated the connection between the multi-epitope sequence and CpG-containing oligodeoxynucleotides, ensuring a cohesive and functional vaccine design.

Immunological Approach:

- The construction adhered to stringent criteria for a robust and safe immunization strategy.

In summary, AAY and GPGPG linkers, along with strategically integrated CpG-containing oligodeoxynucleotides, contribute to the overall effectiveness of multi-epitope vaccines. They influence protein stability, reduce immunogenicity issues, and ensure optimal spacing between epitopes for an enhanced immune response.

Prediction of Antigenicity, Allergenicity, and Physicochemical Properties

To ensure the efficacy and safety of the final vaccine construct, a comprehensive evaluation was conducted, encompassing various critical aspects. This analysis involved the prediction of antigenicity and allergenicity, as well as the assessment of several key physicochemical properties.

Antigenicity Prediction:

The antigenicity of the final vaccine construct was assessed using the VaxiJen v2.0 tool [114]. This tool, acknowledged for its accuracy, quantifies the likelihood of a protein or peptide to elicit an immune response. The threshold for antigenicity prediction was set at 0.4, ensuring that the vaccine construct possessed the necessary attributes to trigger an immune response effectively.

Allergenicity Evaluation:

To gauge the allergenic potential of the vaccine, the AlerCatPro server was employed [125]. This server specializes in predicting allergenic properties, which is a crucial consideration to mitigate potential allergic reactions among vaccine recipients.

Physicochemical Parameter Analysis:

The ProtParam server played a pivotal role in assessing a spectrum of physicochemical properties that contribute to the overall characteristics of the vaccine [155]. The parameters under scrutiny included the theoretical isoelectric point (pI), in vitro and in vivo half-life, amino acid composition, molecular weight, instability, aliphatic index, and the grand average of hydropathicity (GRAVY). These parameters collectively inform our understanding of the vaccine's stability, composition, and potential interactions within the biological system.

By subjecting the vaccine construct to this meticulous evaluation, we ensure that it aligns with essential criteria for immunogenicity, safety, and physicochemical suitability, strengthening its potential as an effective and reliable vaccination strategy.

Vaccine Structure Modeling, Refinement, and Validation

In the quest for a resilient and dependable vaccine, a systematic approach was adopted, encompassing intricate processes of structural modeling, refinement, and stringent validation.

Prediction of Secondary Structural Properties:

The SOPMA server (SOPMA) was instrumental in evaluating the secondary structural properties of the final vaccine construct [156]. This critical analysis provided valuable insights into the organization of secondary structure elements within the vaccine, unveiling its conformational characteristics.

3D Model Generation:

The SwissModel, a specialized online server for homology modeling, skillfully crafted the three-dimensional model of the vaccine [172]. This process involved leveraging homologous structures to ensure compatibility with known protein structures, reinforcing the model's reliability.

Model Refinement:

To heighten the structural integrity and quality of the model, the SwissModel server employed a meticulous procedure. Identifying suitable template structures, aligning target protein sequences, and subsequently constructing a 3D model based on the template, the server then executed energy minimization and refinement. SwissModel's user-friendly interface and robust algorithms make it invaluable for researchers exploring protein structures and functions.

Validation of Tertiary Structure:

The tertiary structure of the vaccine underwent validation through ProSA-web, a prominent server evaluating overall 3D protein model quality. The assessment, expressed as a z-score, meticulously considered deviations from the characteristic range of native proteins, addressing any structural anomalies [174].

Ramachandran Plot Analysis:

Integral to structural validation, the Ramachandran plot analysis, conducted via the PDBsum server, comprehensively evaluated the vaccine model's quality. Assessing the distribution of amino acid dihedral angles, it served as a critical indicator of structural soundness [159].

These meticulous steps ensure the final vaccine construct adheres to the highest standards of structural integrity and quality. The outcome is a theoretically robust vaccine design, fortified by rigorous computational validation.

Immunoinformatics Analysis for Vaccine Immunogenicity

In our pursuit of understanding the vaccine's immunogenicity, I employed cutting-edge tools to predict epitopes within the vaccine construct:

Linear B Cell Epitopes:

Utilized BCpred 2.0 [165] and IEDB servers [118] for predicting linear or continuous B cell epitopes.

These tools provided insights into specific antigenic regions within the vaccine.

Conformational B Cell Epitopes:

Leveraged IEDB-ElliPro server to identify conformational or discontinuous B cell epitopes.

Crucial for understanding how the vaccine interacts with the immune system in a three-dimensional context.

Holistic Immunological Perspective:

The analyses collectively offered a holistic view of potential immunological interactions facilitated by the vaccine.

This comprehensive approach contributes to a well-rounded assessment of the vaccine's immunogenic properties.

This integration of advanced immunoinformatics tools provides valuable insights into the antigenic characteristics and potential efficacy of the vaccine.

In Silico Cloning and Optimization

In our meticulous journey to optimize the expression of the multi-epitope vaccine, we leveraged advanced computational tools.

Java Codon Adaptation Tool (JCat):

JCat played a crucial role by employing reverse translation and codon optimization.

The strategic use of JCat ensured seamless integration of the vaccine into a suitable expression vector.

Codon Adaptation Index (CAI) Optimization:

JCat provided a comprehensive report, including the vital Codon Adaptation Index (CAI).

Achieving an ideal CAI score above 0.8, potentially reaching 1.0, was imperative for optimizing protein expression.

GC Content Assessment:

JCat meticulously assessed the GC content of the genetic insert.

Maintaining GC content within the recommended range of 30% to 70% was crucial for high-level protein expression [175].

Benchling Integration:

The final optimized multi-epitope vaccine sequence was seamlessly assembled using Benchling server [176].

Benchling expertly integrated the optimized sequence into the pET-28a(+) vector.

pET-28a(+) Vector from Addgene:

The pET-28a(+) vector for expression was sourced from Addgene server, Deposit lab: Andrew Millar (Plasmid #141289) [176].

This comprehensive approach, from codon optimization to vector integration, sets the stage for a successful expression process of the multi-epitope vaccine in Escherichia coli, specifically the K12 strain [177].

Computational Immunology with C-ImmSim for the HPV Multi-Epitope Vaccine Design

To unravel the intricate immunogenic profiles and immune responses induced by a multi-epitope vaccine, we employed the C-ImmSim server (C-ImmSim). This powerful tool utilizes Position-specific scoring matrix (PSSM) and machine learning methodologies to predict immune epitopes and their interactions.

Anatomical Replication:

C-ImmSim replicates three vital anatomical domains of mammals: Bone marrow, Thymus, and tertiary lymphatic organs. This comprehensive simulation ensures a holistic understanding of the immune response.

Simulation Protocol:

The immune simulation protocol involved three injections at four-week intervals.

Each injection comprised 1000 vaccine molecules, providing a robust representation of the vaccine's impact.

Essential Parameters:

Standardized parameters included a random seed of 12,345, a simulation volume of 10 µl, and a simulation step of 1050.

These parameters play a crucial role in determining the accuracy and reliability of the simulated immune responses.

Temporal Sequence:

Adhering to established research practices, a minimum four-week interval between injections was maintained.

The temporal sequence for the three injections followed time steps of 1, 84, and 168, where each time step represented an 8-hour interval in real-life scenarios.

By leveraging C-ImmSim with meticulous attention to these parameters, we gained valuable insights into the anticipated immunogenic profiles and temporal dynamics of the multi-epitope vaccine, contributing to informed vaccine design and development [178–180].

System Immunology

Unveiling Gene Expression Dynamics in PBMCs During HPV Vaccination:

Our endeavor to comprehend the intricacies of immune responses in HPV-related cervical cancer patients led us through a systematic exploration of the Gene Expression Omnibus (GEO) database [181, 182]. Focused on human datasets, our aim was to uncover insights into gene expression changes in peripheral blood mononuclear cells (PBMCs) pre and post-human papillomavirus (HPV) vaccination.

Database Exploration:

Meticulous scrutiny of the GEO database for datasets relevant to our objectives.

A quest for papers providing data on gene expression differences in PBMCs from individuals undergoing HPV vaccination.

Format Preference:

Specific interest in datasets available in CEL format for analytical purposes.

Paper Selection:

Careful screening of potential candidates based on alignment with research goals.

Rigorous assessment leading to the selection of papers providing access to desired CEL format databases.

Data Refinement:

Extraction of genes with a minimum 2.5-fold change in expression, guided by the Log Fold Change (LogFC) metric.

Focus on genes with statistical significance ($p < 0.05$), ensuring a robust selection process.

Insights and Analyses:

Identification of genes undergoing substantial expression alterations during cervical cancer vaccination.

Meticulously curated genes forming the foundation for in-depth analyses, contributing invaluable insights into the complex interplay between HPV vaccination and immune responses within PBMCs.

This methodical journey paved the way for a comprehensive understanding of gene expression dynamics during HPV vaccination, providing a solid foundation for subsequent analyses and furthering our insights into the intricate relationship between vaccination and immune responses [183].

Deciphering Functional Interactions in Differentially Expressed Genes through Network Analysis

Our investigation into differentially expressed genes, marked by a substantial 2.5-fold increase and stringent statistical criteria, led us to unveil the intricate web of functional interactions among these genes. Employing the STRING website (STRING), we delved into a comprehensive exploration of protein-protein interactions, gaining insights into the complex relationships within our gene set.

STRING Platform:

Operates by amalgamating diverse biological knowledge sources, including experimental data, computational predictions, and text mining.

Allows users to input a list of genes, initiating a thorough examination of known and predicted interactions.

Network Visualization:

STRING orchestrates the presentation of an intricate network diagram.

Edges between nodes represent various interaction types, with thickness and color indicating the strength and reliability of interactions[184].

Functional Enrichments:

Provides valuable insights into Gene Ontology terms and KEGG pathways (KEGG) [185].

Offers a broader understanding of the biological contexts in which the genes operate.

Export to Cytoscape:

Data generated in STRING was exported in a format compatible with Cytoscape software version 3.10.1.

Cytoscape excels in visualizing and analyzing biological networks, allowing a deeper exploration of functional relationships and potential pathways among the genes.

This meticulous approach sheds light on the intricate molecular mechanisms underlying the identified differential gene expressions in our study, providing a holistic understanding of the functional landscape [186].

Unraveling Key Genes through Topological Centrality Measures in Cytoscape

In our quest to decipher the intricacies of gene interaction networks, Cytoscape emerged as an invaluable tool for comprehensive biological network analysis. Leveraging data from the STRING website (STRING), our focus was on identifying central genes within the network using a diverse set of topological centrality measures.

Betweenness Centrality:

Identifies genes crucial for network connectivity, serving as pivotal bridges facilitating efficient communication between components.

Closeness Centrality:

Assesses the proximity of genes to other network nodes, spotlighting those closely connected to a majority of components, indicating potential for rapid information transmission.

Degree Centrality:

Examines the number of direct connections a gene possesses, identifying hubs that interact directly with several other genes.

Eigenvector Centrality:

Considers both the number and quality of connections a gene has. Genes with high eigenvector centrality values are well-connected and linked to other well-connected genes.

Implementation in Cytoscape involved adhering to the platform's manual and guidelines for network analysis. This meticulous approach allowed us to pinpoint central genes with pivotal roles in our study. These genes, strategically positioned within the network, not only exhibited strong connectivity but also revealed potential significance in the broader biological context of our investigation [187].

Advanced Network Analysis using Gephi 0.10 Software

To delve deeper into our network of genes and explore their significance, we harnessed the power of Gephi 0.10, a powerful network analysis software. Gephi facilitated an extensive examination of key network metrics, including eigenvector centrality, modularity, and additional centrality measures, ultimately enabling the identification of pivotal genes in our study.

Gephi operates through a structured workflow:

Network Import:

The software allows for the import of network data, where nodes represent genes and edges denote their interactions or relationships. In our case, the network was established based on the criteria of degree, closeness centrality, and betweenness centrality.

Eigenvector Centrality Analysis:

Gephi excels in the computation of eigenvector centrality, a crucial metric that determines the influence of a node in a network. This measure identifies genes with the highest centrality scores, signifying their significance within the network.

Modularity Analysis:

Modularity analysis is a fundamental step in understanding network organization. Gephi provides tools for assessing modularity, which aids in identifying distinct groups or modules within the network. This feature is particularly valuable for segregating genes with shared functionality or properties.

Visualization:

Gephi offers interactive visualization, allowing researchers to explore the network graphically, facilitating a comprehensive understanding of gene interactions and centrality measures.

The software's modularity analyzer, in particular, plays a vital role in dissecting the network into coherent subgroups, revealing intricate relationships among genes. By applying this approach, we not only uncovered genes with high eigenvector centrality, degree, closeness centrality, and betweenness centrality but also discerned their participation in distinct functional modules.

The utilization of Gephi in our analysis allowed for a more in-depth investigation of the gene network's structure and dynamics. This method provided valuable insights into the interconnectedness of genes and identified key players within the network. Moreover, the software's user-friendly interface and detailed documentation make it accessible for reproducibility and further investigations in the realm of network biology [188].

Functional Enrichment Analysis in Enricher: Unraveling Biological Significance

To shed light on the potential biological roles and pathways associated with the selected genes derived from our Gephi analysis, we embarked on a journey into the realm of functional enrichment analysis. This essential step was conducted utilizing the Enricher platform, an invaluable resource developed by Ma'ayan Laboratory for Systems Biology(Enricher) [189], renowned for its capability to elucidate the roles genes play in specific biological processes and pathways.

The Enricher methodology hinges on leveraging an expansive knowledge base and an amalgamation of biological data resources. Users provide a list of genes of interest, such as the central genes identified in our study, with the aim of understanding their involvement in particular biological contexts. The platform operates through a series of steps:

Gene Set Collection:

Enricher harnesses a multitude of gene set databases, curating an extensive collection of gene sets related to biological pathways, functions, diseases, and more.

Ranking Genes:

The provided gene list is meticulously examined, with genes scored or ranked based on their relevance to specific biological processes.

Statistical Analysis:

Enricher employs statistical methods, such as hypergeometric testing, to discern if the genes in the input list are significantly overrepresented in any of the gene sets from the collection.

Pathway Enrichment:

This analysis provides insights into the biological pathways in which the selected genes are actively participating.

Function Enrichment:

The platform also uncovers the functional roles of the genes, linking them to specific biological processes, molecular functions, and cellular components.

Visualization:

Enricher offers an intuitive visualization of the results, typically presented in the form of enriched pathways, gene ontologies, and relevant statistical scores.

By conducting functional enrichment analysis through the Enricher platform, we were able to unravel the biological significance of the selected genes. This method allowed us to explore their involvement in specific biological processes and pathways, contributing to a comprehensive understanding of the broader molecular context of our research. The reproducibility of this method enables other researchers to undertake similar investigations, providing valuable insights into the functional roles of genes within their own studies.

Protein Structure Preparation Using UCSF Chimera 1.17.3

Following the identification of our target protein via comprehensive system immunology analysis, we embarked on the crucial step of acquiring its three-dimensional structure from the Protein Data Bank (PDB) [190]. Our methodology for protein selection adhered to stringent criteria, focusing on models with a resolution of 2 angstroms or less. Additionally, we leveraged the R-work value, calculated as the resolution divided by 10, as a pivotal indicator for model selection, prioritizing lower R-work values indicative of higher model quality.

Subsequently, we turned to UCSF Chimera software version 1.17.3 to meticulously prepare the obtained protein structure from the PDB. The following key steps were undertaken in this preparatory process:

Removal of Ligands and Solvent Molecules:

To ensure the fidelity of subsequent analyses, we meticulously removed ligands and solvent molecules from the protein structure. This step aimed to retain only the core protein structure for in-depth investigations.

Refinement and Saving:

The refined protein structure, purged of extraneous elements, was then saved. This preparatory step served as a pivotal foundation for further in-depth analysis and investigations into the target protein.

This meticulous approach to protein structure preparation using UCSF Chimera 1.17.3 was instrumental in ensuring the integrity and quality of the subsequent stages of our research [191].

Identification of Lead-like Compounds from Databases

In the pursuit of identifying potential lead compounds, we conducted extensive searches in renowned databases, specifically ZINC [192] and PubChem [193]. The goal was to select substances that adhered to Lipinski's criteria for lead-like drugs. Lipinski's rules, a fundamental guideline for drug discovery, stipulate that selected compounds should have:

1. A molecular weight of less than 350 g/mol,
2. A calculated logP (partition coefficient) less than 3.5,
3. A net charge ranging from -5 to +5,
4. Fewer than five hydrogen bond donors and acceptors,
5. A polar desolvation energy not exceeding 1 kcal/mol,
6. An apolar desolvation energy of no more than 40 kcal/mol, and
7. A polar surface area less than 200 Å².

Following rigorous screening, all compounds meeting these criteria were meticulously saved in the Structure-Data File (SDF) format for further analysis. This meticulous process is pivotal in the initial stages of our research and lays the foundation for subsequent investigations [192].

Retrieval of Ligands Using BindingDB

In the pursuit of identifying ligands for our target protein, we employed Binding Data Bank (BindingDB), a valuable resource for accessing binding data. The specific procedure we followed can be summarized as follows:

Selection of IC50 Icon: Within the BindingDB platform, the IC50 icon was selected as the criterion for filtering ligands. This choice allowed us to prioritize

compounds with the lowest IC50 values, indicative of their strong binding affinity.

Search for the Target Protein: We initiated the ligand discovery process by searching for the specific protein of interest within the BindingDB database.

Choosing an IC50 Threshold: To refine our search, we set an IC50 threshold of not more than 50 to narrow down the list of potential ligands, ensuring that the selected compounds met our desired criteria.

Obtaining 3D SDF Files: We opted to retrieve 3D structure data files (SDF) for the selected ligands. These files are essential for further computational analysis using specialized software.

Providing Authentication: For access to the requested data and files, we provided the necessary authentication by entering our email and password.

SDF and TSV File Downloads: Subsequently, the SDF files were downloaded, enabling us to proceed with in-depth computational analysis. Additionally, tab-separated values (TSV) files, which contain structured data, were also obtained for comprehensive record-keeping and analysis purposes.

This meticulous process of data retrieval within the BindingDB platform served as a critical step in our research, facilitating the subsequent stages of ligand evaluation and analysis [194].

Generating SDF Files Using Marvin 23.14 Software

In cases where we encountered difficulties in finding SDF files, we utilized Marvin software to draw and export structures as SDF files. Here's a detailed outline of the process:

1. Begin by drawing the molecule using Marvin and save it in Simplified Molecular Input Line Entry System (SMILE) format.

2. Open the saved SMILE format using word processing software, such as Word, and copy the SMILE format of the molecule.

3. Proceed to Chimera: Tools - Build Structure - SMILES String.

4. Paste the copied SMILE format using 'Ctrl-V.'

5. Click 'Apply' (only once).

6. To minimize the structure, access 'Tools - Minimize Structure - Minimize - OK - Amber - OK.'

7. Save the structure, ensuring the appropriate nomenclature is used to distinguish the generated structure.

This method allowed us to create SDF files for molecules when the original files were not readily available, facilitating their integration into our computational analysis and drug discovery research [195].

Descriptor Analysis Using PaDEL-Descriptor 2.21

In order to gain a comprehensive understanding of the ligands obtained from both the ZINC database and BindingDB, we employed PaDEL-Descriptor 2.21, a valuable software tool for calculating molecular descriptors. The method we followed, along with a brief explanation of the software's functionality, is elucidated below:

Input SDF Files:

The SDF files obtained from both ZINC and BindingDB served as the input for this analysis. These files contain the 3D structures of ligands, and it is vital to extract molecular descriptors from them.

PaDEL-Descriptor 2.21 Software:

PaDEL-Descriptor 2.21 is a versatile software designed for the calculation of a wide range of molecular descriptors. It accepts SDF files as input and generates output in the form of CSV and Excel files containing descriptor values.

Descriptor Calculation:

Once the SDF files were loaded into the software, PaDEL-Descriptor 2.21 performed a thorough analysis of the chemical structures. The software calculated a multitude of descriptors, including but not limited to constitutional descriptors, topological descriptors, and 3D descriptors. These descriptors offer valuable insights into various molecular properties, such as size, shape, polarity, and electronic characteristics.

CSV and Excel Outputs:

After the descriptor calculations were completed, the software provided the results in the form of CSV and Excel files. These output files contain detailed information on each descriptor's value for every ligand in the dataset.

Software Functionality:

PaDEL-Descriptor 2.21 is a user-friendly and powerful tool for descriptor calculation. It leverages established algorithms and methodologies to compute a wide array of molecular descriptors. The software's extensive descriptor set aids in characterizing chemical compounds comprehensively, making it a valuable asset for ligand analysis in drug discovery and computational chemistry.

1. Launch the PaDEL-Descriptor 2.21 software.

2. Load the input SDF files containing ligand structures.

3. Initiate the descriptor calculation process.

4. Once the analysis is complete, export the results in the form of CSV and Excel files for further examination and analysis.

This method of descriptor analysis using PaDEL-Descriptor 2.21 significantly enhances our ability to evaluate the ligands retrieved from ZINC and BindingDB, providing critical insights into their molecular properties [196, 197].

Data Preprocessing for Descriptor Analysis with SMLR 1.1.0

In our quest to identify the most relevant descriptors for further analysis, we employed the SMLR 1.1.0 software, which harnesses machine learning techniques. However, before we could utilize SMLR for this purpose, it was essential to prepare the descriptor outcomes obtained from PaDEL-Descriptor 2.2. Below, we outline the data preprocessing steps and the subsequent actions taken with SMLR 1.1.0:

Column Name Adjustment:

To distinguish between data obtained from BindingDB and ZINC12, the name of the first column was replaced with a numerical identifier. The respective identifiers for each dataset were saved in separate text files, allowing clear differentiation.

Elimination of Zero Values:

The second column was carefully examined, and if it contained only zero values, it was removed from the dataset.

Handling the Last Column:

Similar scrutiny was applied to the last column. If any non-numeric entries were present, they were removed, and zeros were added in their place.

Blank Cell Treatment:

1. Blank cells within the dataset were identified using the 'Find and Select' function and subsequently replaced with zeros.

Conversion to Excel Format:

The processed data was saved in Excel format, ensuring its compatibility with SMLR 1.1.0.

Data Pretreatment:

The prepared data was then imported into the SMLR 1.1.0 software, facilitating data pretreatment.

Alpha Value Application:

The software allowed us to apply alpha values as part of the analysis process, which was essential for our study.

Process Validation:

We thoroughly validated the data and analysis parameters to ensure accuracy and reliability.

Submission:

Upon verification, the dataset was submitted for analysis.

Result Assessment:

After processing, we assessed the results. In cases where specific descriptors had values of zero in tests, we applied a minor adjustment by changing them to a very small number, typically 0.00001.

This data preprocessing and analysis strategy, incorporating SMLR 1.1.0, was a crucial step in identifying the most informative descriptors for our ligand analysis. The machine learning capabilities of SMLR, combined with our data preparation, allowed us to extract valuable insights for our research in drug discovery and computational chemistry [198].

QSAR Analysis Using Chemoface 1.65

In our pursuit of conducting Quantitative Structure-Activity Relationship (QSAR) analysis, we employed the Chemoface 1.65 software. The following is a detailed account of the procedure undertaken during our QSAR analysis:

Selecting Descriptors from SMLR Results:

We initiated the analysis by choosing descriptor rows derived from BindingDB data obtained through SMLR. These rows were copied without their headers.

Data Import:

In Chemoface, we navigated to 'File' and selected 'Paste X.' This step involved pasting the descriptor data obtained from SMLR.

Data Integrity Check:

1. We thoroughly examined the pasted data to ensure that no 'Non' values were present. If any were found, we replaced one of the numbers in the zero column with 0.09.

Part 2: Data Preparation: 4. IC50- Data Selection:

We returned to the TSV files previously saved from BindingDB. Specifically, we focused on the IC50- values for the descriptors selected from the SMLR results. These IC50- values were copied and pasted into a new Excel sheet.

Log Transformation: To prepare the data, we calculated the logarithm (log) of the IC50- values for each descriptor using the formula = -log(a1:a21). In cases where a log was not computed, we assigned a value of 0.9.

Data Copy and Paste: The log-transformed IC50- values were then copied and pasted into Chemoface under 'File' and 'Paste Y.'

Part 3: Data Preprocessing: 7. Autoscaling:

We proceeded to preprocess the data by selecting 'Pretreatment' and applying autoscaling.

Test Set Selection:

A test set consisting of 20% of the data was chosen, typically 4 from 20, for validation. This selection could be done manually or automatically using the software.

Part 4: Partial Least Square Regression (PLS): 9. PLS Model Run:

We conducted Partial Least Square Regression (PLS) analysis by running cross-validation on the data.

Visualization:

We generated plots to visualize the analysis results, ensuring they provided meaningful insights into the data.

Model Saving:

To preserve the model generated, we selected 'File' and 'Save Model.'

Part 5: Model Validation and Prediction:

Model Validation:

Having saved the model, we proceeded to test it by introducing molecules without IC50- values. This allowed us to assess whether these molecules aligned with the regression line based on the model.

Data Import:

In this case, we returned to the SMLR results and copied and pasted the descriptors obtained from the ZINC dataset into Chemoface as 'Prediction.'

Predictions: Using the same criteria, we conducted predictions by clicking on the 'Predict' option in the last column.

Our QSAR analysis workflow with Chemoface 1.65 enabled us to derive insights into the relationships between chemical structures and biological activities. The preparation and analysis of our data provided a foundation for our research in computational chemistry and drug discovery, ultimately contributing to the identification of promising candidate molecules [199, 200].

3D-QSAR Modeling Using Schrodinger-Maestro 2023.4

In our endeavor to perform 3D-QSAR modeling, we employed Schrodinger-Maestro 2023.4 software. Here is a detailed description of the steps undertaken during the 3D-QSAR modeling process:

Part 1: Ligand Preparation:

1. Launch Schrodinger software.
2. Task: Browse to 'Ligprep' and select 'List Structure Form.'
3. Select 'File' and open the file containing Binding DB ligands.
4. Set 'OPLS2005' as the force field.
5. Initiate the process by selecting 'Neutralize.'
6. Ensure 'Desalt' remains unchecked.
7. Keep 'Generate Tautomers' unchecked.
8. Set 'Generate at most' to 10.
9. Ensure 'Maestro' is selected.
10. Specify a job name, such as "Ligprep - Name of Protein Binding DB."
11. Under 'Setting,' ensure the 'Job Setting' is set to yellow.
12. Execute the job by clicking 'Run.'
13. Wait for the process to finish and close the small window.

Part 2: Pharmacophore Hypothesis Generation:

1. On the left menu, close the scroll-down menu.
2. Click to turn it green.
3. In the 'Ligand' section on the left menu, examine the ligand features, such as rings or ions.
4. Return to the 'Task' and navigate to 'Browse-Phase-Develop Pharmacophore Hypothesis.'
5. Select 'Multiple Ligands' and choose the relevant ligands.
6. Define 'Show Family' (all).
7. For 'Select IC50-,' designate 'Active' if the activity is below (1) and 'Inactive' if the activity is above (1).
8. Click 'Apply' and then 'Ok.'
9. Set 'Hypothesis Settings' to 'Hypothesis should match best (80%).'
10. Specify the 'Number of Features in the Hypothesis' (4-5).
11. Define 'Preferred Minimum Number of Features' as 4.
12. Save the hypothesis.
13. Examine the feature symbols in the 'Hypothesis Settings,' and modify them if necessary.
14. Ensure 'Generate Conformers' remains unchecked.
15. Set a job name like "Chase-Pharm-Name of Protein-Pharmacophore Modeling."
16. Execute the job by clicking 'Run.'
17. In the main window on the far left, the results will appear. Click the drop-down menu (blue circle) to view the pharmacophore model in the main black window.
18. Right-click on each feature to view its dimensions in X, Y, and Z coordinates, and make note of them.
19. If multiple models are available, click the 'Table' icon in the upper right to access a table displaying the survival rates, with the most significant model being the one with the highest rate.

20. Save your project at the end of this part: 'File-Save Project.'

Part 3: Molecule Screening:

1. Task: Navigate to 'Browse-Phase-Ligand and Database Screening.'
2. Select 'File' and open the molecules you wish to screen (not from Binding DB).
3. Add the pharmacophore hypothesis previously created.
4. Under 'Screening Settings,' select 'Conformers,' and use existing conformers. Save these settings.
5. Specify a job name, e.g., "Phase-Screen-Name of Dataset."
6. Execute the job by clicking 'Run.'
7. In the main window on the far left, click the molecules that match our model (checkmark icon).

Part 4: Saving Fitting Molecules:

1. If you reopen the 'Table' icon, you can view the molecules that fit the model in the table, and their names are listed in the last column.

Part 5: ADME Property Analysis:

1. Browse to 'ADME and Molecule Properties' and select 'Ligand-Based ADME.'
2. Use 'QikPro' and import the structure file by selecting 'File-Browse the Molecules.'
3. Specify a job name, such as "QikPro File Name."
4. Execute the job by clicking 'Run.'
5. Open the 'Table' icon to view the results in the table.

This 3D-QSAR modeling process with Schrodinger-Maestro 2023.4 software equipped us to analyze ligands, generate pharmacophore hypotheses, screen molecules, and evaluate ADME properties, contributing significantly to our research in computational drug discovery and molecular design [201, 202].

2D-QSAR Modeling of Protein-Ligand Interactions Using BIOVIA Studio Visualizer

2D-QSAR Modeling:

The protein structure was imported into BIOVIA Studio Visualizer to initiate the 2D-QSAR modeling process.

Docking Model Analysis:

Docked models obtained from PyRx software were imported into BIOVIA Studio Visualizer for in-depth analysis of interactions between the protein and ligands.

Receptor-Ligand Interaction Exploration:

Receptor-ligand interactions were established within BIOVIA Studio Visualizer, providing insights into the binding dynamics of the molecular complex.

Selection of the Best Docking Model:

The best docking model, characterized by the lowest affinity energy and RMSD for ligands, was chosen as the basis for further analysis.

Receptor-Ligand Interaction Feature Utilization:

Utilizing the Receptor-Ligand Interaction feature, the chosen protein was designated as the receptor, facilitating a detailed examination of ligand interactions.

Incorporation of Hydrogen Bonds:

Hydrogen bonds were incorporated into the analysis to understand the specific molecular interactions influencing binding.

Amino Acid Identification and Labeling:

Amino acids involved in the interactions were labeled for comprehensive identification and interpretation [203].

Comprehensive Molecular Dynamics Simulation Workflow for Ligand-Protein Interaction Studies

Topology Generation:

The topology of ligand was generated using SwissParam, ensuring compatibility with the CHARMM all atoms force field [204, 205].

Protein Topology Setup:

The protein's topology was established using GROMACS utilities with the CHARMM27 all-atom force field, incorporating CHARMM22 plus CMAP for proteins. The water model was set to TIP 3-point [206].

System Configuration:

The Ligand-protein complex structure was defined within a unit cell box under periodic boundary conditions. The box, with a triclinic shape, was filled with water, followed by neutralization using Cl- or Na+ counter ions [207].

Minimization and Equilibration:

Steepest descent energy minimization was performed, and the system underwent equilibration under NVT conditions for 50 ns at 300 K. Subsequently, NPT simulation and MD run were conducted for 50 ns and 10 ns respectively [208].

Constraint Handling and Interaction Treatment:

Covalent bonds were constrained using the LINCS algorithm, and electrostatic interactions were treated via the Particle Mesh Ewald (PME) method. Cut-off radii for Coulomb and van der Waals interactions were set to 10.0 and 14.0 Å, respectively [209].

Trajectory Analysis:

Trajectories were analyzed for RMSD, RMSF, Rg, the number of H-bonds, and SASA of the protein-ligand interaction, utilizing GROMACS utilities such as “gmx rms”, “gmx rmsf”, “gmx gyrate”, “gmx hbond”, and “gmx sasa” cammands [210].

Ligand-Protein Stability Assessment:

Ligand-protein stability was assessed by monitoring hydrogen bond dynamics over time. XMgrace was employed to prepare graphical representations.

Screening of Ligand Potency:

To screen the ligand's potency, compounds forming three or more hydrogen bonds with the protein were validated through a short 10 ns simulation run with consistent parameters [210].

Chapter 6: Results

Results:

AIM

AIM Assay Results for CD4/CD8 T Cells and HPV-16 Antigens

The AIM assay results for CD4/CD8 T cells targeting HPV-16 antigens (E6, E7, and L1) are presented in this section, providing insights into the immune response following vaccination.

CD4 and CD8 T Cells Activation for HPV-E6

Comparison between cases (vaccinated) and controls (non-vaccinated), demographic of the people participate in the study are shown in Table 2, reveals no significant difference in CD4 and CD8 T cell activation for HPV-E6, Figure a, b.

CD4 and CD8 T Cells Activation for HPV-E7

Only CD4 T cells show a significant activation for HPV-E7, as indicated by the asterisk (*) denoting significance. This observation is supported by statistical analysis: Th-E7 p-value: 0.039058, CI95%: [-0.26, -0.01], Cohen-d: 2.248314, BF10: 2.96.

CD4 and CD8 T Cells Activation for HPV-L1:

The results for HPV-L1 (Figure 1 d, f) exhibit no significant difference in activation between CD4 and CD8 T cells.

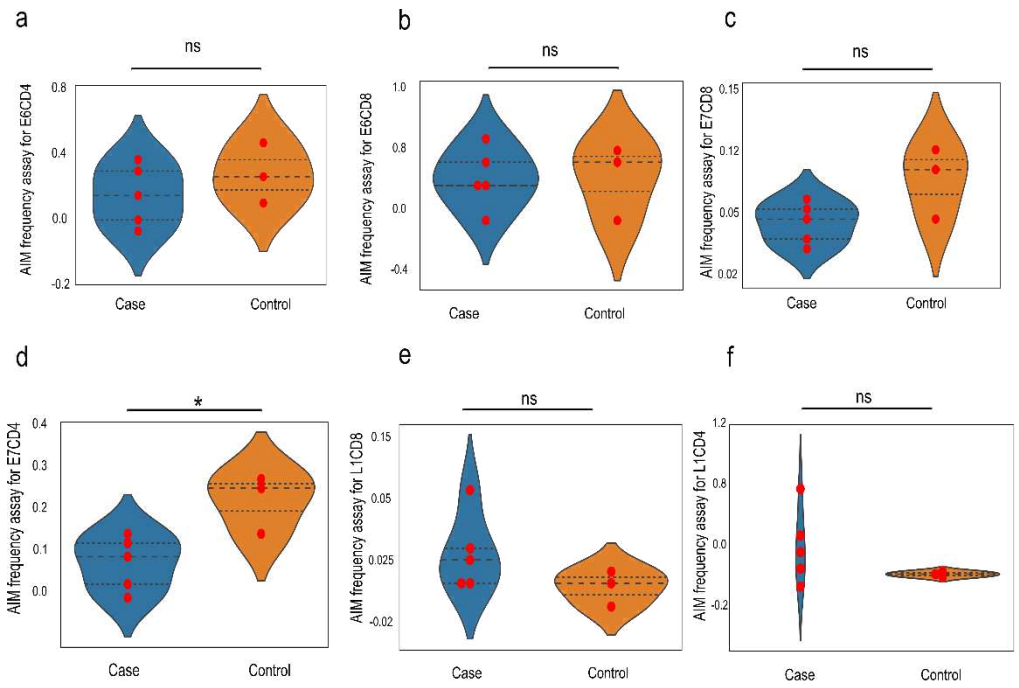


Figure 1 The graph illustrates the frequency of AIM assays for CD4/CD8 T cells responding to HPV antigens (E6, E7, and L1). Subfigures (a, b) compare vaccinated cases with non-vaccinated controls, showcasing CD4

and CD8 T cell activity for HPV-E6. Similarly, subfigures (c, d) demonstrate CD4 and CD8 T cell activation for HPV-EV, while (d, f) present results for HPV-L1.

Table 1 summarizes the T cells' activation-inducing molecule assay for HPV-16 proteins (E6, E7, and L1), providing p-values, confidence intervals (CI95%), Cohen's d effect size, and Bayes Factors (BF10). T cells activity and corresponding statistical parameters including p-value, confidence interval (CI95%), Cohen's d, and Bayes Factor (BF10) for different HPV proteins (Th-E6, CTC-E6, Th-E7, CTC-E7, Th-L1, CTC-L1) are included in the table.

Table 1 T Cells Activity and Statistical Parameters, confidence interval (CI95%), Cohen's d, and Bayes Factor (BF10)

T cells activity	p-value	CI95%	Cohen-d	BF10
Th-E6	0.38	[-0.2, 0.43]	0.699583	0.689
CTC-E6	0.97	[-0.06, 0.06]	0.023603	0.532
Th-E7	0.039058	[-0.26, -0.01]	2.248314	2.96
CTC-E7	0.154991	[-0.12, 0.03]	1.639783	1.285
Th-L1	0.171323	[-0.13, 0.54]	0.91118	1.056
CTC-L1	0.1389	[-0.01, 0.07]	1.038533	1.096

However, these findings highlight the selective activation of CD4 T cells in response to HPV-E7, underscoring the nuanced T cell responses to specific antigens. Notably, the relative T cell activity towards HPV-16 antigens (E6, E7, and L1) did not exhibit a notable increase when compared to T cell activity targeting actin.

Actin served as a reference protein for assessing T cell activation, providing valuable insights into the nuanced and selective nature of T cell responses to specific HPV antigens. The use of Human Actin peptide as an irrelevant antigen control serves as a baseline or reference in the study. By including an irrelevant antigen like actin, researchers can distinguish specific immune responses to peptides derived from HPV antigens (E6 and E7) from non-specific reactions [211].

LPA Results for CD4/CD8 T Cells and HPV-16 Antigens

In response to the limitations encountered in the initial Aim, the study strategically shifted focus to Lymphocyte Proliferation assay (LPA) analysis. To enhance the scope and depth of investigation, the participant pool was expanded to include vaccinated and non-vaccinated individuals, along with patients in the early stages of cervical cancer. This comprehensive approach aimed to discern potential disparities in the proliferation of CD4

T helper cells and CD8 T cells among the diverse participant categories. The demographic table of the study participants includes three groups: non-vaccinated individuals (non-vacs), patients in the first stage of HPV-related cervical cancer (patients), and vaccinated participants who received Gardasil and Cervarix (vacs). The study conducted descriptive analyses of the study population, Table 2.

Table 2 The demographic table of the participants in this study. n: number, SD: standard deviation

		Grouped by Samples				
		Missing	**Overall**	**non-vacs**	**patient**	**vacs**
n			48	14	20	14
Age, mean (SD)		0	34.9 (10.5)	39.6 (2.5)	43.9 (11.2)	28.6 (1.8)
Samples, n (%)	**non-vacs**	0	14 (29.2)	14 (100.0)		
	patient		20 (41.7)		20 (100.0)	
	vacs		14 (29.2)			14 (100.0)
Doses, n (%)	**Zero**	0	34 (70.8)	14 (100.0)	20 (100.0)	
	Two		4 (8.3)			4 (28.6)
	Three		10 (20.8)			10 (71.4)
Last dose since sampling, n (%)	**No-doses**	0	35 (72.9)	14 (100.0)	20 (100.0)	1 (7.1)
	six		8 (16.7)			8 (57.1)
	year		5 (10.4)			5 (35.7)

The Tucky statistical method played a pivotal role in facilitating Pairwise comparisons between each pair of groups, offering a robust statistical framework for in-depth analysis [212]. Concurrently, the assessment of effect size was conducted through the application of the standard mean difference, providing valuable insights into the practical significance of observed differences.

CD4 T Cell Activity against E6 Antigen: Comparative Analysis

The analysis of CD4 T cell activity against the E6 antigen involved an analysis of variance (ANOVA) to explore potential differences among different participant groups, non-vacs: non vaccinated, vacs: vaccinated participants, and patient means the group of patients with first stage of HPV related cervical cancer. The ANOVA table revealed the following statistical information:

- F-Statistic (F): 1.846144
- p-value (p-unc): 0.182561

- Partial Eta Squared (np2): 0.149532

While the ANOVA did not indicate a statistically significant difference among the groups (p-value = 0.182561), further analyses were conducted to explore specific pairwise comparisons Figure 2.

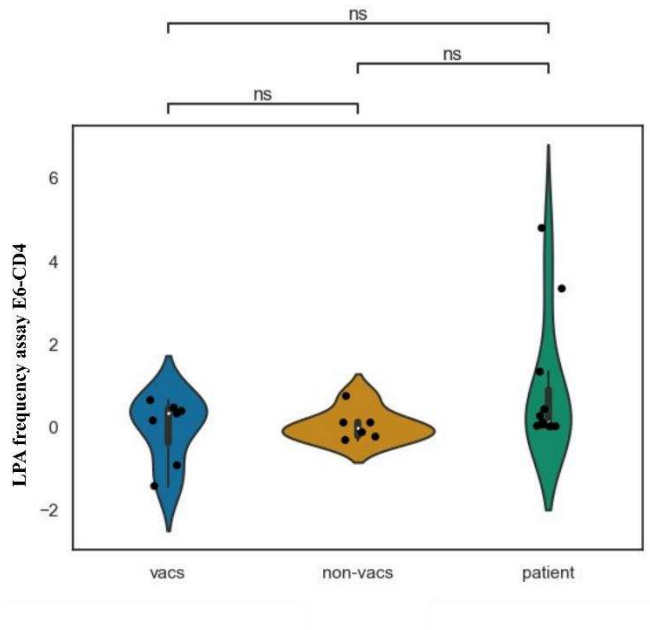


Figure 2 ANOVA did not indicate a statistically significant difference among the groups for CD4 T cells for HPV16-E6. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients

Tukey HSD Post Hoc Test

The Tukey Honestly Significant Difference (HSD) test was employed for pairwise comparisons. The results are summarized below:

- non-vacs vs. patient: Mean Difference = 0.8974, p-adj = 0.3304, 95% CI [-0.6533, 2.4481]
- non-vacs vs. vacs: Mean Difference = -0.1002, p-adj = 0.9879, 95% CI [-1.8001, 1.5997]
- patient vs. vacs: Mean Difference = -0.9977, p-adj = 0.2278, 95% CI [-2.475, 0.4796]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To further understand the practical significance of the observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD4 activity-E6-vacs-nonvacs: SMD = 0.1565, 95% CI [-0.5878, 0.9009]
- Minimal Detectable Change (MDC) for CD4 activity-E6-vacs-nonvacs: 13.944122342898568
- CD4 activity-E6-vacs-patients: SMD = 0.7271, 95% CI [-0.4897, 1.9438]
- Overall MDC for CD4 activity-E6-vacs-patients: 9.21917776238874
- CD4 activity-E6-nonvacs-patients: SMD = 0.6679, 95% CI [-0.4273, 1.7630]
- Overall MDC for CD4 activity-E6-nonvacs-patients: 11.942026896533873

These effect size measures provide insights into the magnitude of differences, suggesting modest to moderate effects without reaching statistical significance.

Study Power and Sample Size

The study's statistical power was calculated at 0.108, indicating a relatively low power to detect differences. The required sample size for adequate power was estimated to be 353 participants.

CD4 T Cell Activity against E7 Antigen: Comparative Analysis

The analysis of CD4 T cell activity against the E7 antigen involved an analysis of variance (ANOVA) to explore potential differences among different participant groups, non-vacs: non vaccinated, vacs: vaccinated participants, and patient means the group of patients with first stage of HPV related cervical cancer. The ANOVA table revealed the following statistical information:

- F-Statistic (F): 1.08313
- p-value (p-unc): 0.35671
- Partial Eta Squared (np2): 0.093509

Although the p-value is greater than the conventional significance level (0.05), suggesting no statistically significant difference among groups, further analyses were conducted to explore specific pairwise comparisons Figure3.

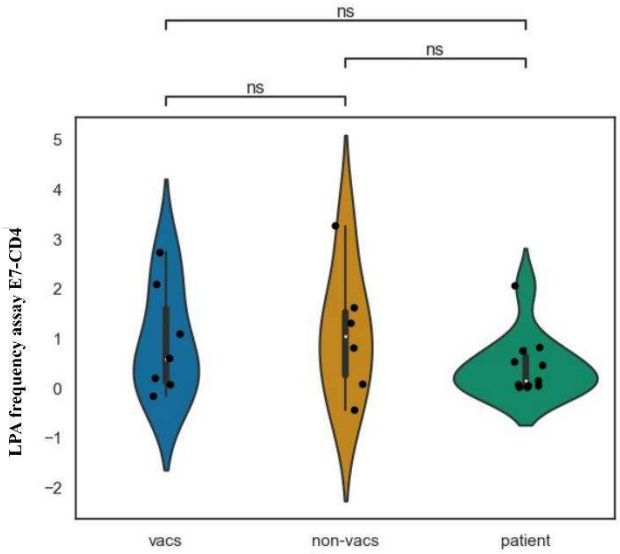


Figure 3 ANOVA did not indicate a statistically significant difference among the groups for CD4 T cells for HPV16-E7. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients

Tukey HSD Post Hoc Test

The Tukey Honestly Significant Difference (HSD) test was employed for pairwise comparisons. The results are summarized below:

- non-vacs vs. patient: Mean Difference = -0.6554, p-adj = 0.3887, 95% CI [-1.8867, 0.5758]
- non-vacs vs. vacs: Mean Difference = -0.1626, p-adj = 0.9506, 95% CI [-1.5123, 1.1871]
- patient vs. vacs: Mean Difference = 0.4928, p-adj = 0.5492, 95% CI [-0.6802, 1.6658]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To further understand the practical significance of the observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD4 activity-E7-vacs-nonvacs: SMD = 0.1361, 95% CI [-1.3484, 1.6206]
- CD4 activity-E7-vacs-nonvacs, overall MDC: 13.006690536508211
- CD4 activity-E7-vacs-patients: SMD = -0.5961, 95% CI [-1.5561, 0.3639]
- CD4 activity-E7-vacs-patients, Overall MDC: 14.95889165338539

- CD4 activity-E7-nonvacs-patients: SMD = -0.7245, 95% CI [-1.9267, 0.4776]
- CD4 activity-E7-nonvacs-patients, overall MDC: 18.720007849781386

These effect size measures provide insights into the magnitude of differences, suggesting minimal to moderate effects without reaching statistical significance.

Study Power and Sample Size

The study's statistical power was calculated at 0.072, indicating a relatively low power to detect differences. The required sample size for adequate power was estimated to be 900 participants.

CD4 T Cell Activity against L1 Antigen: Comparative Analysis

The investigation into CD4 T cell activity against the L1 antigen involved an analysis of variance (ANOVA) to explore potential differences among participant groups, non-vacs: non vaccinated, vacs: vaccinated participants, and patient means the group of patients with first stage of HPV related cervical cancer. The ANOVA table revealed the following statistical information:

- F-Statistic (F): 2.577069
- p-value (p-unc): 0.0998
- Partial Eta Squared (np2): 0.197068

Although the p-value is greater than the conventional significance level (0.05), suggesting no statistically significant difference among groups, further analyses were conducted to explore specific pairwise comparisons Figure 4.

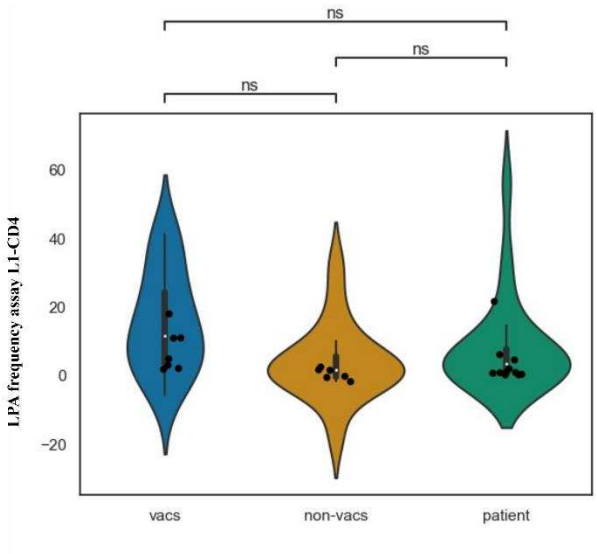


Figure 4 ANOVA did not indicate a statistically significant difference among the groups for CD4 T cells for HPV16-L1. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients.

Tukey HSD Post Hoc Test

The Tukey Honestly Significant Difference (HSD) test was employed for pairwise comparisons. The results are summarized below:

- non-vacs vs. patient: Mean Difference = 2.9646, p-adj = 0.5462, 95% CI [-4.0592, 9.9885]
- non-vacs vs. vacs: Mean Difference = 6.8636, p-adj = 0.0863, 95% CI [-0.836, 14.5632]
- patient vs. vacs: Mean Difference = 3.8989, p-adj = 0.3256, 95% CI [-2.7924, 10.5903]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To further understand the practical significance of the observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD4 activity-L1-vacs-nonvacs: SMD = -1.4829590145261706, 95% CI [-6.7541, 3.7882]
- CD4 activity-L1-vacs-nonvacs, overall MDC: 28.26603503928943
- CD4 activity-L1-vacs-patients: SMD = -0.6263735332142737, 95% CI [-6.9543, 5.7016]
- CD4 activity-L1-vacs-patients, overall MDC: 30.513590807077826
- CD4 activity-L1-nonvacs-patients: SMD = 0.5666329586944453, 95% CI [-3.7237, 4.857]
- CD4 activity-L1-nonvacs-patients, overall MDC: 24.62997918527744

These effect size measures provide insights into the magnitude of differences, suggesting moderate to minimal effects without reaching statistical significance.

Study Power and Sample Size

The study's statistical power was calculated at 0.152, indicating a relatively low power to detect differences. The required sample size for adequate power was estimated to be 204 participants.

CD8 T Cell Activity against HPV E6 Antigen: Comparative Analysis

The examination of CD8 T cell activity against the HPV E6 antigen involved an analysis of variance (ANOVA) to assess potential group differences, non-vacs: non vaccinated, vacs: vaccinated participants, and patient means the group of patients with first stage of HPV related cervical cancer. The ANOVA results are summarized as follows:

- F-Statistic (F): 1.131613
- p-value (p-unc): 0.341404
- Partial Eta Squared (np2): 0.097288

The p-value suggests no statistically significant differences among participant groups in CD8 T cell activity against the E6 antigen Figure 5.

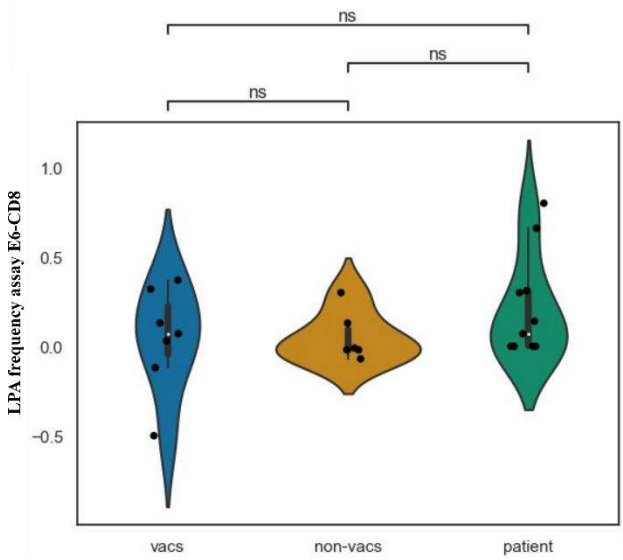


Figure 5 ANOVA did not indicate a statistically significant difference among the groups for CD8 T cells for HPV16-E6. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients.

Tukey HSD Post Hoc Test

To further explore pairwise group differences, the Tukey Honestly Significant Difference (HSD) test was conducted. The results are presented below:

- non-vacs vs. patient: Mean Difference = 2.9646, p-adj = 0.5462, 95% CI [-4.0592, 9.9885]
- non-vacs vs. vacs: Mean Difference = 6.8636, p-adj = 0.0863, 95% CI [-0.836, 14.5632]
- patient vs. vacs: Mean Difference = 3.8989, p-adj = 0.3256, 95% CI [-2.7924, 10.5903]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To understand the practical significance of observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD8 activity-E6-vacs-nonvacs: SMD = 0.0374, 95% CI [-0.2362, 0.3110]
- CD8 activity-E6-vacs-nonvacs, overall MDC: 4.708838267585426
- CD8 activity-E6-vacs-patients: SMD = 0.5709, 95% CI [0.2739, 0.8678]
- CD8 activity-E6-vacs-patients, Overall MDC: 1.9553189701938658
- CD8 activity-E6-nonvacs-patients: SMD = 0.6317, 95% CI [0.4120, 0.8513]
- CD8 activity-E6-nonvacs-patients, overall MDC: 4.380901865352174

These effect size measures suggest small to moderate differences in CD8 T cell activity against the E6 antigen, with the most notable difference observed between the vaccinated (vacs) and non-vaccinated (non-vacs) participants.

Study Power and Sample Size

The study's statistical power was calculated at 0.074, indicating a low power to detect differences. The required sample size for adequate power was estimated to be 831 participants.

CD8 T Cell Activity against HPV E7 Antigen: Comparative Analysis

The investigation into CD8 T cell activity against the HPV E7 antigen involved an analysis of variance (ANOVA) to examine potential group differences, non-vacs: non vaccinated, vacs: vaccinated participants, and patient means the group of patients with first stage of HPV related cervical cancer. The ANOVA results are summarized as follows:

- F-Statistic (F): 0.882503
- p-value (p-unc): 0.42854
- Partial Eta Squared (np2): 0.077532

The p-value suggests no statistically significant differences among participant groups in CD8 T cell activity against the E7 antigen Figure 6.

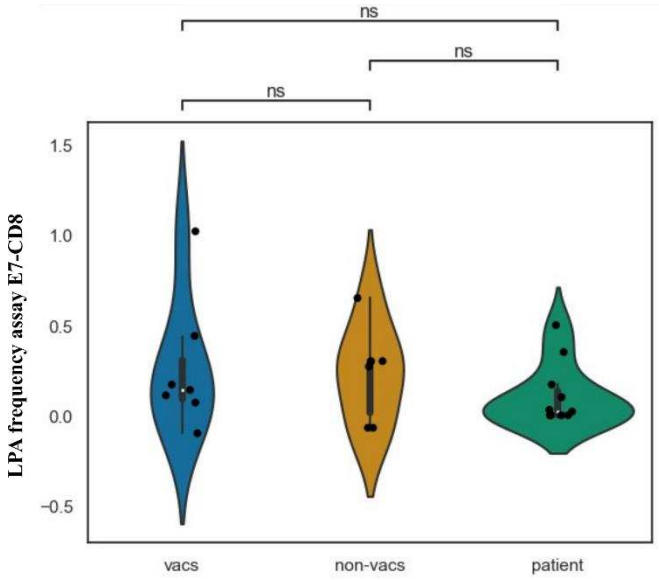


Figure 6 ANOVA did not indicate a statistically significant difference among the groups for CD8 T cells for HPV16-E7. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients.

Tukey HSD Post Hoc Test

Further exploration of pairwise group differences was conducted using the Tukey Honestly Significant Difference (HSD) test. The results are presented below:

- non-vacs vs. patient: Mean Difference = -0.1236, p-adj = 0.6348, 95% CI [-0.4629, 0.2156]
- non-vacs vs. vacs: Mean Difference = 0.0343, p-adj = 0.9707, 95% CI [-0.3376, 0.4061]
- patient vs. vacs: Mean Difference = 0.1579, p-adj = 0.4483, 95% CI [-0.1652, 0.4811]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To understand the practical significance of observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD8 activity-E7-vacs-nonvacs: SMD = -0.1043, 95% CI [-0.4968, 0.2882]
- CD8 activity-E7-vacs-nonvacs, overall MDC: 3.761121592716725
- CD8 activity-E7-vacs-patients: SMD = -0.5999, 95% CI [-0.9154, -0.2843]

- CD8 activity-E7-vacs-patients, overall MDC: 2.6937371743417673
- CD8 activity-E7-nonvacs-patients: SMD = -0.5912, 95% CI [-0.8512, -0.3312]
- CD8 activity-E7-nonvacs-patients, overall MDC: 2.9969096857622155

These effect size measures suggest small to moderate differences in CD8 T cell activity against the E7 antigen, with the most notable difference observed between vaccinated (vacs) and non-vaccinated (non-vacs) participants.

Study Power and Sample Size

The study's statistical power was calculated at 0.065, indicating a low power to detect differences. The required sample size for adequate power was estimated to be 1308 participants.

CD8 T Cell Activity against HPV L1 Antigen: Comparative Analysis

The investigation into CD8 T cell activity against the HPV L1 antigen involved an analysis of variance (ANOVA) to explore potential group differences. The ANOVA results are summarized as follows:

- F-Statistic (F): 3.242844
- p-value (p-unc): 0.059249
- Partial Eta Squared (np2): 0.235966

While the p-value is above the conventional significance level (0.05), the effect size (np2) indicates a moderate influence of the group variable on CD8 T cell activity against the L1 antigen Figure 7.

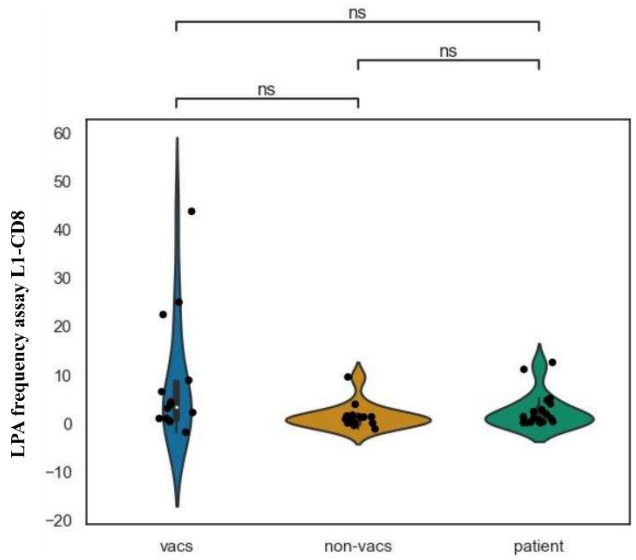


Figure 7 ANOVA did not indicate a statistically significant difference among the groups for CD8 T cells for HPV16-L1. non-vacs: non vaccinated, vacs: vaccinated participants, and patient: first stage HPV related cervical cancer patients.

Tukey HSD Post Hoc Test

To further explore pairwise group differences, the Tukey Honestly Significant Difference (HSD) test was conducted. The results are presented below:

- non-vacs vs. patient: Mean Difference = -0.1236, p-adj = 0.6348, 95% CI [-0.4629, 0.2156]
- non-vacs vs. vacs: Mean Difference = 0.0343, p-adj = 0.9707, 95% CI [-0.3376, 0.4061]
- patient vs. vacs: Mean Difference = 0.1579, p-adj = 0.4483, 95% CI [-0.1652, 0.4811]

The p-adj values indicate that none of the pairwise comparisons reached statistical significance.

Effect Size Measures

To understand the practical significance of observed differences, standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI):

- CD8 activity-L1-vacs-nonvacs: SMD = -1.2840, 95% CI [-3.9429, 1.3748]
- CD8 activity-L1-vacs-nonvacs, overall MDC: 19.51745331059357
- CD8 activity-L1-vacs-patients: SMD = -0.8037, 95% CI [-3.6201, 2.0127]
- CD8 activity-L1-vacs-patients, overall MDC: 18.126661343984658

- CD8 activity-L1-nonvacs-patients: SMD = 0.5847, 95% CI [-0.8495, 2.0190]
- CD8 activity-L1-nonvacs-patients, overall MDC: 6.425104320015117

Study Power and Sample Size

The study's statistical power was calculated at 0.198, suggesting a low-to-moderate power to detect differences. The required sample size for a higher power (0.80) was estimated to be 143 participants.

The SMD confidence intervals for CD4 and CD8 activation, comparing non-vaccinated vs. vaccinated, patients vs. vaccinated, and non-vaccinated vs. patients, demonstrated considerable width. This suggests substantial uncertainty in the effect estimates. Notably, none of the intervals conformed to conventional Cohen's d effect size thresholds, including 0.2, 0.5, 0.8, or exceeding 1.4. These findings are visually depicted in Figure 8.

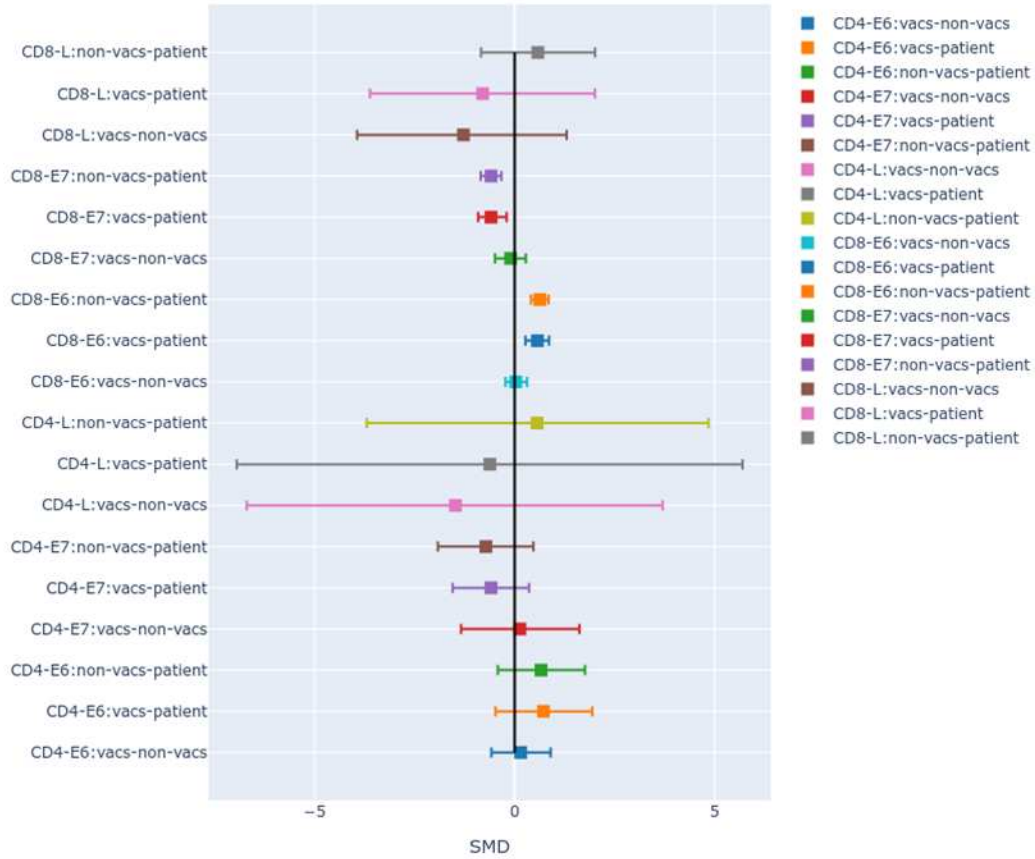


Figure 8 The forest plot visually presents the Standardized Mean Differences (SMDs) and their corresponding confidence intervals for CD4 and CD8 proliferation. The comparisons include non-vaccinated vs. vaccinated, patients vs. vaccinated, and non-vaccinated vs. patients. Notably, the forest plot illustrates the uncertainty reflected in the wide confidence intervals, indicating variability in the effect size estimates based on Cohen's d. This uncertainty underscores the challenges in precisely determining the impact of vaccination and patient status on CD4 and CD8 proliferation.

Comparative Analysis of T Cell Activity to Actin Across Groups

In this study, an examination was conducted to assess the relative fold increase in T cell activity to actin among different groups. Notably, the investigation revealed a distinctive pattern in the fold increase in T cell activity, specifically observed for HPV16-E7 and HPV16-L1 for vaccinated participants.

The results, presented graphically in the accompanying figure, distinctly illustrate the enhanced T cell activity concerning these specific antigens. Noteworthy is the observation that the fold increase in T cell activity to actin was significant only for HPV16-E7 and HPV16-L1, suggesting a targeted and selective response of T cells to these antigens.

This finding holds significant implications for our understanding of the immune response dynamics within the studied groups. The depicted results in the figure serve as a visual representation of the observed patterns, providing a clear and concise overview of the comparative T cell responses across the specified antigens Figure 9. In this context, the hierarchy of fold increases is noteworthy:

- TH1-L1 exhibits the highest fold increase with 40-fold enhancement.
- TFH1-L1 follows with a substantial 28-fold increase.
- CXCR5-TH1-L1 demonstrates a noteworthy 20-fold enhancement.
- CD4-T-L1 displays a 15-fold increase in activity.
- TFH exhibits a 13-fold rise in T cell activity.
- TFH1-E7 shows a significant 9-fold increase.
- CD8-T-L1 displays an 8-fold enhancement.
- CXCR5-L1 reveals a 5-fold increase.
- CXCR5-E7 exhibits a 4-fold enhancement.

These findings provide a detailed perspective on the specific T cell responses to different HPV16 antigens, shedding light on the variations in their respective fold increases.

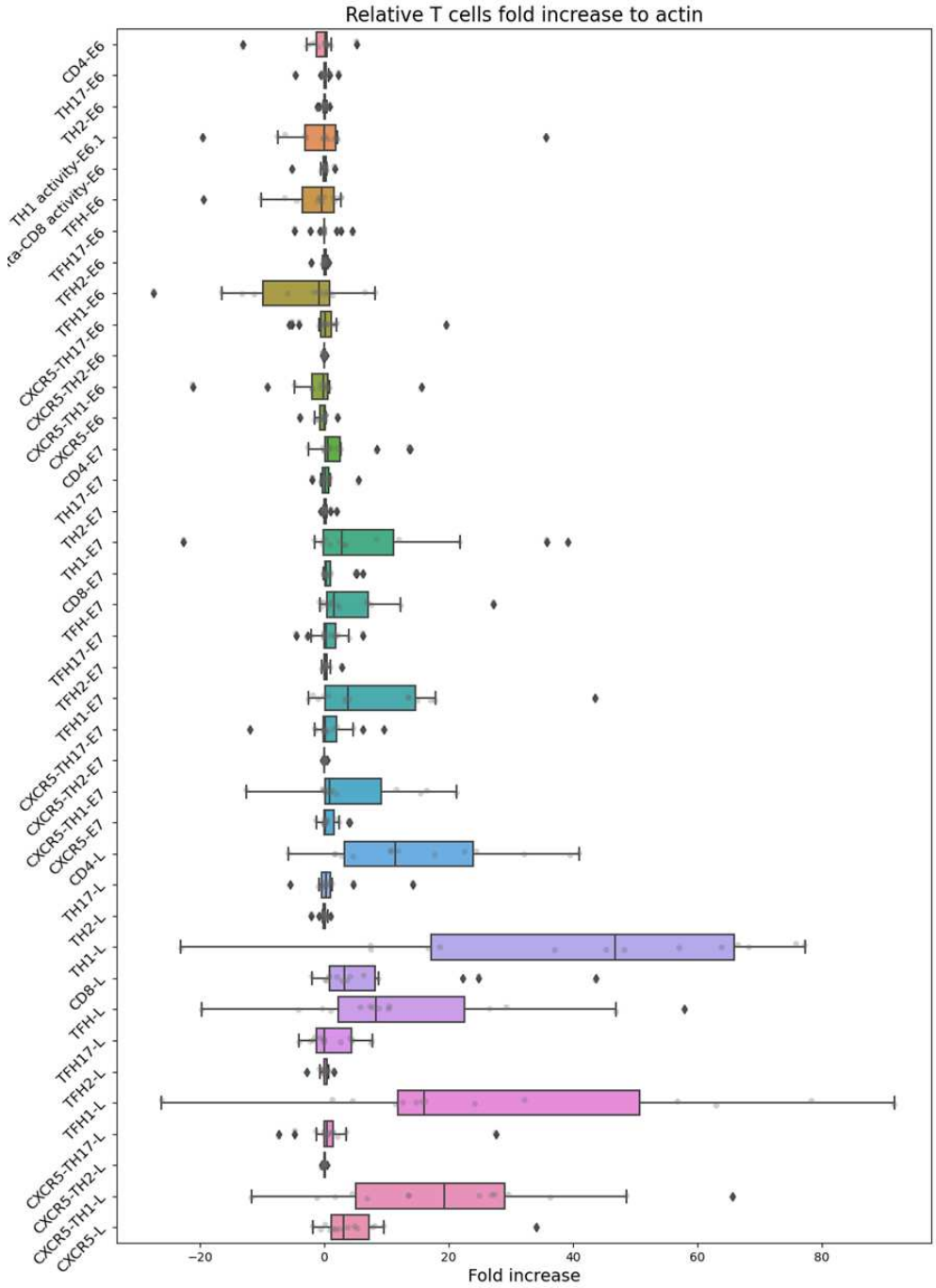


Figure 9 This graph illustrates the fold increase in T cell proliferation specific to HPV16 antigens (E6, E7, and L1) in comparison to T cell proliferation for actin. Notably, a discernible fold proliferation increase is observed exclusively for T cells targeting E7 and L1 antigens. In this case, TH1-L1 with 40-fold increase was the first, TFH1-L1 28-fold, CXCR5-TH1-L1 20-fold, CD4-T-l1 15-fold, TFH 13-fold, TFH1-E7 9-fold, CD8-T-L1 8-fold, CXCR5-L1 5-fold, and CXCR5-E7 4-fold.

Impact of Vaccination Parameters on T Cell Activity Fold Increase

The logistic regression analysis comparing vaccinated and non-vaccinated individuals reveals that the brand name of the vaccine, whether 9-valent, 4-valent, or 2-valent, does not significantly impact the fold increase in T cell activity. The focus shifts to key factors influencing T cell activation, as indicated by odds ratios (OR) and associated statistics.

Logistic Regression Results:

- Intercept:
    - OR: 9.175791e+29
    - z-value: 0.999801
    - 2.5%: 0.000000
    - 97.5%: inf
- Doses:
    - OR: 1.310651e+10
    - z-value: 0.999614
    - 2.5%: 0.000000
    - 97.5%: inf
- Vaccine Name:
    - OR: 9.154724e-11
    - z-value: 0.999867
    - 2.5%: 0.000000
    - 97.5%: inf
- Age:
    - OR: 4.582071e-01
    - z-value: 0.228859
    - 2.5%: 0.128527

- 97.5%: 1.633544

These results underscore the significance of the intercept, number of vaccine doses, vaccine brand name, and age in influencing the activation of T cells. Notably, vaccine brand specificity does not contribute to variations in T cell activity fold increase, highlighting the robustness of T cell response across different vaccine formulations.

Immunoinformatics:

Sequence Alignment and Conserved Areas Across HPV Groups

The sequence alignment of HPV species across various risk groups—high risk, low risk, probably high risk, and unknown risk—reveals limited conserved areas in the E6, E7, and L1 regions, Table 3. This table delineates conserved regions within the E6, E7, and L1 protein sequences for distinct HPV groups. Notably, 'x' designates sequences categorized as unknown risk, uppercase letters highlight conserved areas and amino acids, and lowercase letters signify inconsistencies observed among aligned protein sequences.

Table 3 Conserved Regions and Variabilities in E6, E7, and L1 Protein Sequences Across HPV Groups. 'x' denotes sequences categorized as unknown risk, uppercase letters signify conserved areas and amino acids, while lowercase letters indicate inconsistencies observed among aligned protein sequences.

Antigens	**Preserved sequences**
E6	**MxqxxxxxxRxxxexxxxxxxxxxxxxLxxxxxxxkxxxxxxxxxxfxxxxIVxxxxNxxxxxxxxxx xxxxxxxxxYxYxxxxxxxxxxxxxxxxxxxLxxxxxxxxxxxxxKEKxxxxxKRrxxxxxxxxxxx xxxxxxxxxxxqxxxxxxxxxxxxxxxxxxxxxxxxxx**
E7	**MHGxxpTLKDIVLDLQPExxxxxxxxQxxDSSEEEDEDdxxxxxxxxxQxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxExLxxxxxILxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxx**
L1	**MALxxxSxxxVYLPPxxVSxVVSTDEYVxRTxIYYxAxSSRLLxVxHPYxxVxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxD xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx xxxxxxxxxxxxxxxxxxxxVxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxYxxxxxxxxx xxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxTxS xxxxxxxSxxSxxxxxxxSxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxNxxxxxxxxxxxxx xxxxxxxxxxxxxxSxxxxxQxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx xxxxKxxASSSxxxxxxxxKxKxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxxx**

Phylogenetic Analysis of HPV Species without Antigen-Based Categorization

Protein sequences for distinct groups of HPV species (high risk, low risk, probably high risk, and unknown risk) were sourced from NCBI, Uniprot, and PaVe databases. After analysis, a consensus-based approach identified unique sequences for E6, E7, and L1. Phylogenetic trees were then constructed for each protein across all species collectively, as illustrated in the accompanying figure 10.

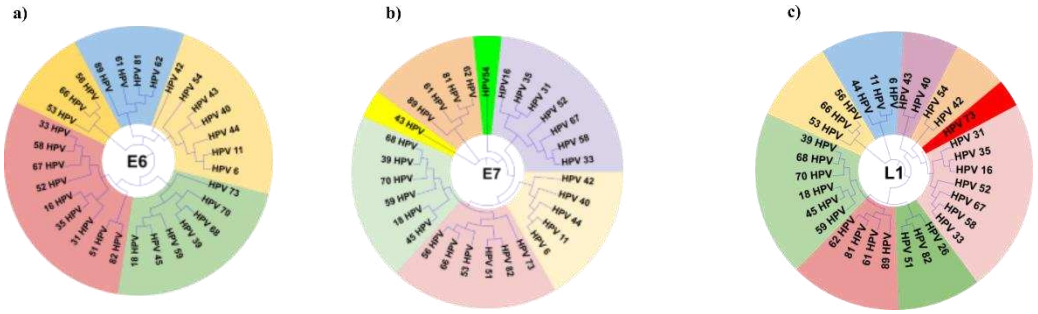


Figure 10 Phylogenetic analysis encompassing all HPV species without subgroup classification. a) Phylogenetic tree based on E6 antigen, b) Phylogenetic tree based on E7 antigen, c) Phylogenetic tree based on L1 antigen.

Phylogenetic Analysis of High-Risk HPV Clades Based on E6, E7, and L1 Antigens

Conducting a phylogenetic analysis on high-risk HPV clades utilizing E6, E7, and L1 antigens elucidates the evolutionary relationships among these high-risk species. The resulting figure illustrates the distinct clades, providing valuable insights into the grouping of HPV species with shared genetic characteristics, Figure 11.

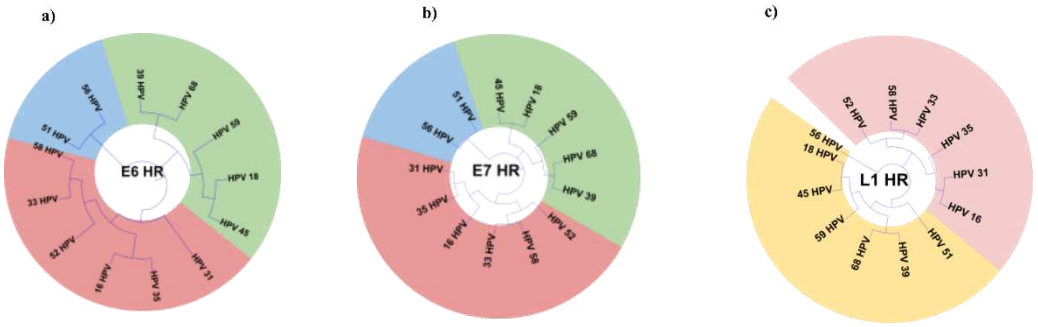


Figure 11 Phylogenetic analysis depicting high-risk HPV species using E6, E7, and L1 antigens. a) Phylogenetic tree based on E6, b) Phylogenetic tree based on E7, c) Phylogenetic tree based on L1. HR: High Risk.

Phylogenetic Analysis of Low-Risk HPV Clades Based on E6, E7, and L1 Antigens

Conducting a phylogenetic analysis on low-risk HPV clades using E6, E7, and L1 antigens unveils the relationships among these species. The resulting figure illustrates the distinct clades within the low-risk group, shedding light on which HPV species share a common evolutionary lineage, Figure 12.

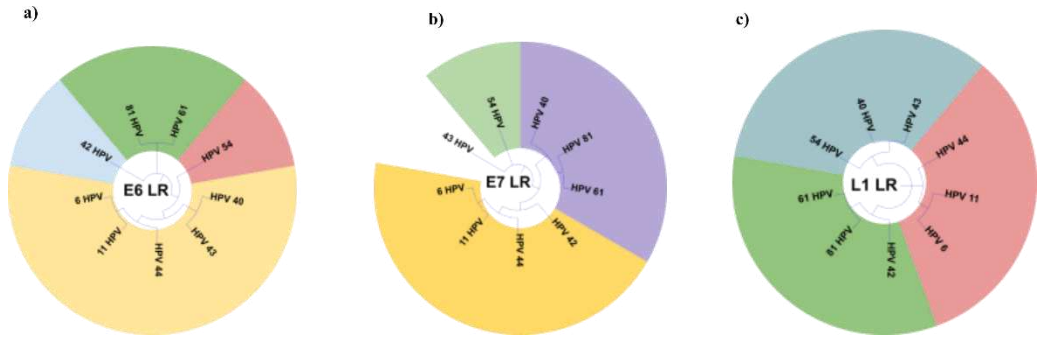


Figure 12 Phylogenetic tree analysis depicting the low-risk HPV clades using E6, E7, and L1 antigens. a) Phylogenetic tree based on E6, b) Phylogenetic tree based on E7, c) Phylogenetic tree based on L1. LR: low risk

Phylogenetic Analysis of Likely High-Risk HPV Species Based on E7 and L1 Antigens

Classifying HPV species into the high-risk group elucidates the shared characteristics among them, as depicted in the accompanying figure. Notably, the phylogenetic tree generated for E6 within this high-risk group demonstrates distinct differences between each subgroup, indicating diverse evolutionary paths, Figure 13.

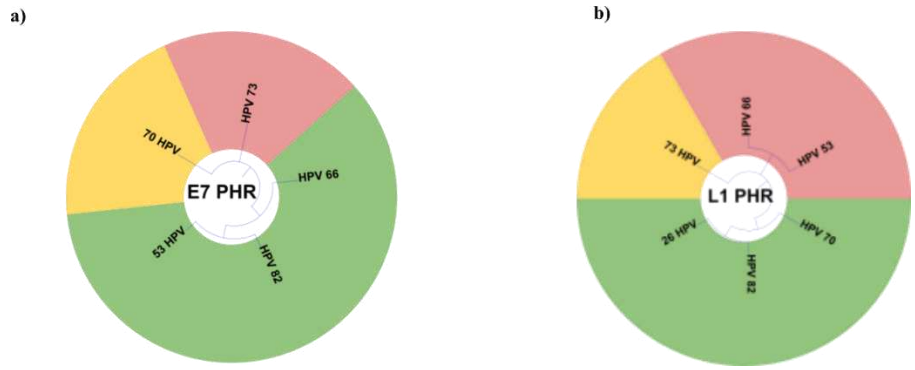


Figure 13 the phylogenetic tree analysis explores potentially high-risk HPV clades through examination of E7 and L1 antigens. PHR: potentially high risk

Phylogenetic Analysis of Unknown-Risk HPV Species Using E7 and L1 Antigens

Unfortunately, conducting a phylogenetic tree analysis within this group proved challenging due to the substantial diversity present. The intricate genetic variations among the unknown-risk HPV species, particularly in the E7 and L1 antigens, hindered the establishment of a meaningful phylogenetic tree. The extensive diversity observed in this set of HPV species posed a significant obstacle, preventing a conclusive phylogenetic analysis.

Exploration of High-Risk HPV Antigens (HPV-16 and HPV-18)

Given the significance of the high-risk HPV group and the observed phylogenetic inconsistencies in other groups, our focus shifted towards investigating selected antigens in HPV-16 and HPV-18. These chosen antigens are representative of two distinct classes within the high-risk HPV group, a critical consideration, particularly for classical vaccines primarily composed of the HPV-L1 antigen.

To initiate our analysis, we systematically extracted the sequences corresponding to E6, E7, and L1 for both HPV-16 and HPV-18. This extraction process was carried out meticulously using reliable databases, as detailed in Table 4.

Table 4 The examination of protein sequences for the HPV:16,18 E6, E7, and L1 antigens through comprehensive searches on NCBI, UniProt, and PaVe databases..

Species	**Antigen**	**Database**	**ID**
HPV16	E6	NCBI	AYV61474.1
HPV16	E6	Uniprot	P03126
HPV16	E6	PaVe	HPV16-E6
HPV16	E7	NCBI	2002324A
HPV16	E7	Uniprot	P03129
HPV16	E7	PaVe	HPV16-E7
HPV16	L1	NCBI	NP_041332.2
HPV16	L1	Uniprot	P03101
HPV16	L1	PaVe	HPV16-L1
HPV18	E6	NCBI	Non
HPV18	E6	Uniprot	A0A7D5IFD8
HPV18	E6	PaVe	HPV18-E6
HPV18	E7	NCBI	Non
HPV18	E7	Uniprot	A0A291PQC6
HPV18	E7	PaVe	HPV18-E7
HPV18	L1	NCBI	AMJ21969.1
HPV18	L1	Uniprot	Q5G240
HPV18	L1	PaVe	HPV18-L1

In the sequence alignment of Human Papillomavirus (HPV) types 16 and 18, distinct conserved regions are observed within the genetic sequences of the E6, E7, and L1 antigens, as depicted in the accompanying Figure 14. This alignment highlights the shared elements between the two HPV types, offering valuable insights into the preserved areas of these critical antigens.

The conserved regions identified in the sequence alignment are indicative of potential functional significance, suggesting common structural or functional motifs crucial for the biological activities of E6, E7, and L1, Figure 14.

Figure 14 a comprehensive multiple alignment is presented, depicting the genetic sequences of Human Papillomavirus types 16 (HPV16) and 18 (HPV18) for three crucial genes: a) E6, b) E7, and c) L1. The alignment provides insights into the similarities and differences in the nucleotide sequences of these genes between the two HPV types.

Physicochemical and Secondary Structural Analysis of Target HPV antigens

After obtaining sequences from databases, we conducted a thorough analysis of the physicochemical and secondary structural properties of HPV proteins, using Expacy ProtParam [155], as outlined in Table 3.

Theoretical pI (Isoelectric Point) is the pH at which the protein carries no net electrical charge, representing its neutral state [213]. For HPV16-L1 and HPV18-L1, the theoretical pI is 8.55 and 8.52, respectively. Instability Index is a predictor of protein stability. Values above 40 generally indicate an unstable protein [214]. HPV16-L1 has an instability index of 36.57, while HPV18-L1 has an index of 49. Estimated Half-life (Mammalian Reticulocytes) is the time it takes for half of the protein to be degraded in mammalian reticulocyte cells, expressed in hours [215]. Both HPV16-L1 and HPV18-L1 have an estimated half-life of 30 hours. Aliphatic Index represents the relative volume occupied by aliphatic side chains, indicating thermostability [216]. HPV16-L1 and HPV18-L1 exhibit aliphatic indices of 74.32 and 70.16, respectively. GRAVY (Grand Average of Hydropathicity) is A measure of the protein's hydrophobicity [217]. Negative values suggest hydrophilicity, while positive values indicate hydrophobicity. HPV16-L1 and HPV18-L1 have GRAVY values of -0.35 and -0.43, respectively.

These physicochemical properties provide insights into the stability, degradation, and hydrophobicity of the major capsid proteins or late proteins (L1) of HPV16 and HPV18, crucial information for understanding their functional characteristics, Table 5.

Table 5 Physicochemical Properties of Structural Proteins L1 and E for HPV16 and HPV18 (Analyzed with ProtPram tool)

Proteins	**Molecular weight**	**Theoretical pI**	**Instability index**	**Estimated half-life**	**Aliphatic index**	**GRAVY**
HPV16-L1	56308.21	8.55	36.57	30 hours	74.32	-0.35
HPV18-L1	56506.94	8.52	49	30 hours	70.16	-0.43
HPV16-E7	11022.32	4.2	63	30 hours	78.57	-0.41
HPV16-E6	18334.27	9.01	73.25	30 hours	70.99	-0.73
HPV18-E7	11995.57	4.7	72.93	30 hours	86.38	-0.38

Antigenicity Assessment

Antigenicity assessment was conducted utilizing the Vaxijen 2.0 server [114] for HPV18 and HPV 16, specifically targeting the E7, E6, and L1 proteins. The outcomes of this analysis are presented in the accompanying Table 6.

Table 6 Analysis of HPV 18 and 16 Antigenicity with VaxiJen 2.0

Antigens	Protein ID (PaVE)	Size	Antigenicity
HPV-16-E6	HPV-16-E6	151	0.6733
HPV-16-E7	HPV-16-E7	56	0.5765
HPV-16-L1	HPV-16-L1	505	0.52
HPV-18-E6	HPV-18-E6	158	0.54
HPV-18-E7	HPV-18-E7	105	0.49
HPV-18-L1	HPV-18-L1	507	0.52

Secondary Structure Prediction:

The protein's secondary structure analysis provides a lucid depiction of potential secondary structural elements, enhancing our understanding of protein structure (see Figure 15). This analysis aids in unraveling the intricacies of the protein's structural composition.

Figure 15 the [SOPMA tool](#) [156] illustrates the distribution of structural elements, including the percentage of Alpha helix, Beta sheets or Extended strands, Beta turn, and Random coil, in the target proteins. Specifically, panels (a) to (f) represent the results for HPV16-E6, HPV16-E7, HPV18-E6, HPV18-E7, HPV16-L1, and HPV18-L1, respectively.

Homology Modeling and Structural Validation

Phyre2 generated five models for each protein, from which the best model with 100% confidence and over 90% coverage was selected, Figure 16. The structural validation involved testing all selected models using the Ramachandran plot. Detailed results are presented in the accompanying Table 7.

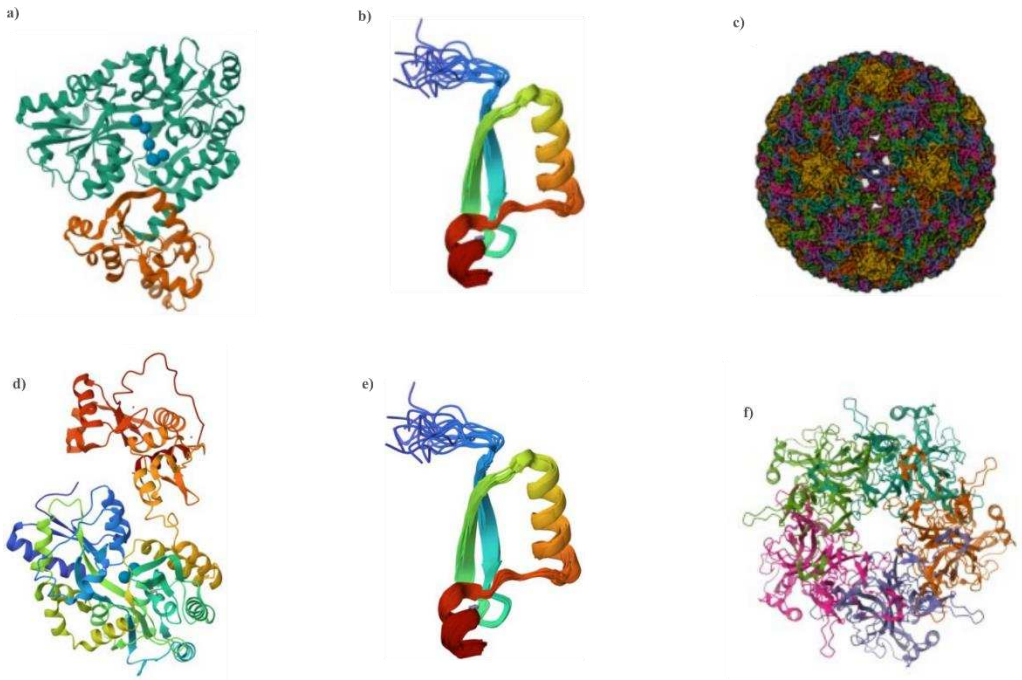


Figure 16 Illustrates the homology modeling process conducted with the Phyre2 tool. The models for (a) HPV16-E6, (b) HPV16-E7, and (c) HPV16-L1 were generated using the PDB ID numbers 4GIZ, 2EWL, and 3J6R, respectively. Similarly, models for (d) HPV18-E6, (e) HPV18-E7, and (f) HPV18-L1 were created based on the PDB ID numbers 6SLM, 2EWL, and 2R5I.

Table 7 presents models generated with Phyre2, which underwent refinement and analysis through the Ramachandran plot.

Protein	**Template**	**Ramachandran Plot**		
		Favored region	Allowed region	Disallowed region
HPV-16-E6	4GIZ	92.80%	7.20%	0.00%
HPV-16-E7	2EWL	72.5%	25.50%	2.00%
HPV-16-L1	3J6R	81.10%	17.90%	1.1%
HPV-18-E6	6SLM	90.80%	9.20%	0.00%
HPV-18-E7	2EWL	72.5%	25.50%	2.00%
HPV-18-L1	2R5I	80.2%	19.6%	0.00%

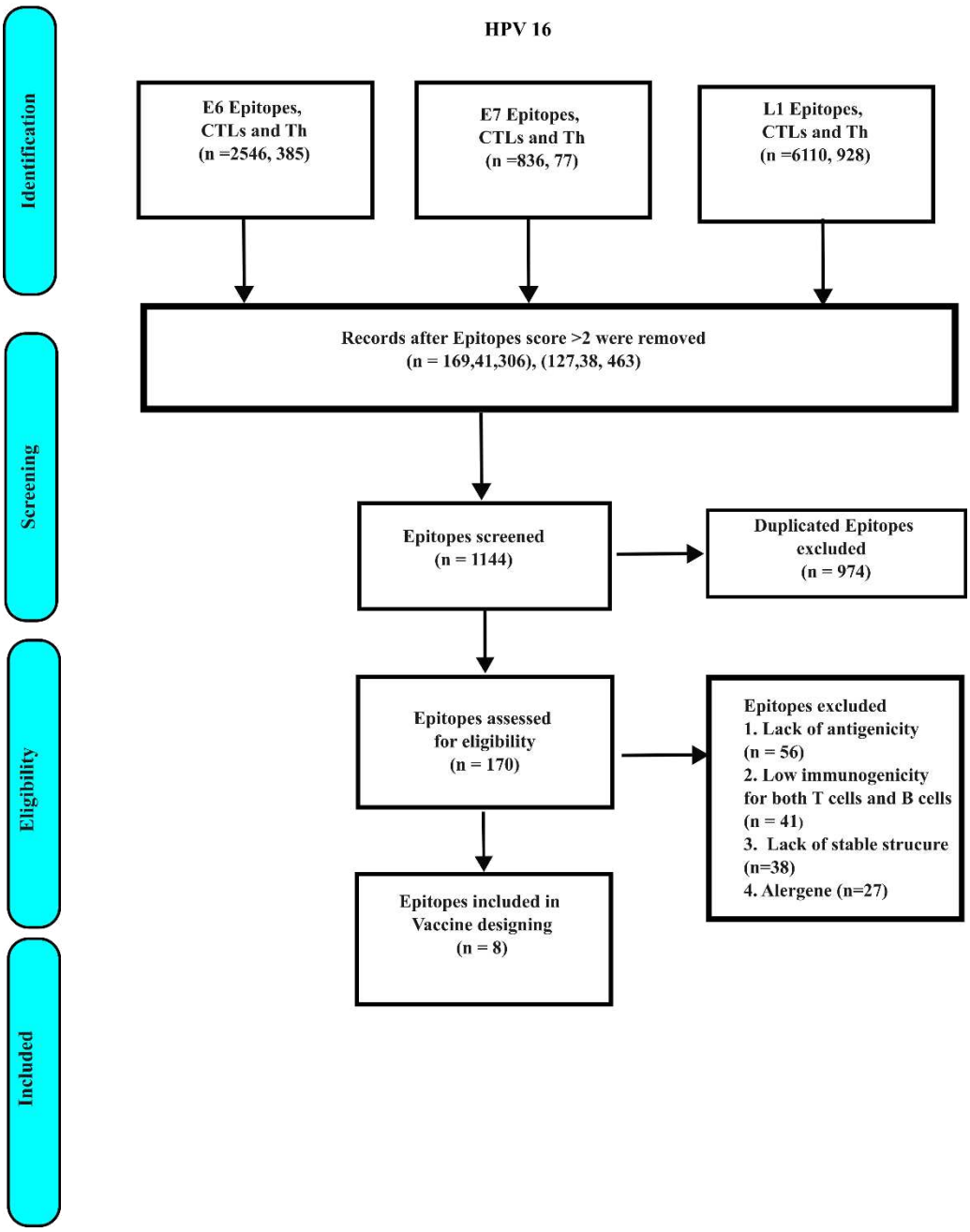
Prediction of T Cell Epitopes for HPV16

Utilizing the IEDB [118] and NetCell 1.2 [218] servers with a modeled version of HPV16-E6, E7, and L1, epitopes were identified for both T helper cells and T cytotoxic cells. A systematic screening process was implemented, evaluating each epitope for antigenicity, duplications, and immunogenicity for both types of T cells. The comprehensive results of this analysis are depicted in Figure.

Prediction of T Cell Epitopes for HPV18

A parallel approach was undertaken for HPV18-E6, E7, and L1 refined models. Epitopes for both T helper cells and cytotoxic T cells (CTCs) were meticulously extracted and subjected to a stepwise filtration process. This method aimed to attain the optimal number of epitopes suitable for inclusion in the vaccine, as illustrated in Figure 17.

a)


Identification
Screening
Eligibility
Included
HPV 16
E6 Epitopes, CTLs and Th (n =2546, 385)
E7 Epitopes, CTLs and Th (n =836, 77)
L1 Epitopes, CTLs and Th (n =6110, 928)
Records after Epitopes score >2 were removed (n = 169,41,306), (127,38, 463)
Epitopes screened (n = 1144)
Duplicated Epitopes excluded (n = 974)
Epitopes assessed for eligibility (n = 170)
Epitopes excluded
1. Lack of antigenicity (n = 56)
2. Low immunogenicity for both T cells and B cells (n = 41)
3. Lack of stable strucure (n=38)
4. Alergene (n=27)
Epitopes included in Vaccine designing (n = 8)

b)

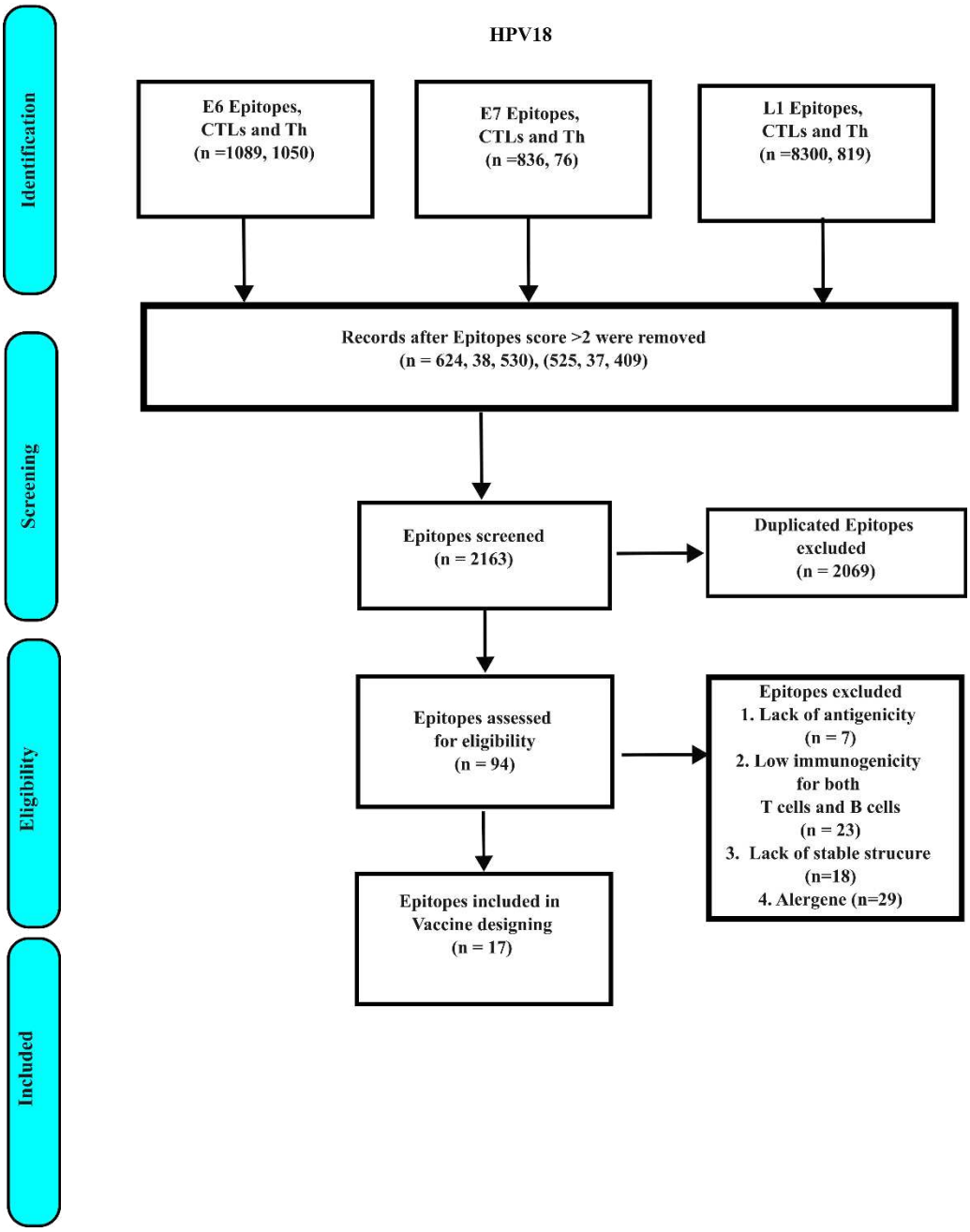


Figure 17 illustrates the epitope selection process for potential inclusion in our vaccine. (a) presents the flowchart detailing epitope selection for HPV16, and (b) demonstrates a similar process for HPV18. Initially, candidate epitopes for Th cells and CTCs underwent filtration based on scores provided by IEDB and NetCell. Subsequently, the process involved checking for epitope duplication and removing any duplicates. Finally, the selected epitopes were assessed for stability of structure, antigenicity, and allergenicity.

Epitopes meeting criteria for duplication, antigenicity, allergenicity, and structural stability were incorporated into the prospective vaccine. Subsequently, the population coverage, based on IEDB-Population Coverage , assessment revealed that all selected epitopes exhibited coverage exceeding 90%. Detailed information on these epitopes is presented in Table 8.

Table 8 The epitopes that successfully met the filtering criteria are presented for both HPV16 and HPV18, specifically targeting Th cells and cytotoxic T cells (Tc).

HPV types	**T cytotoxic epitopes**	**T-helper epitopes**
HPV-16	SYSLYGTTL(E6) SAYAANAGV(L1) TTLEQQYNK(E6) YGTTLEQQY(E6)	KKQRFHNIRGRWTGR(E6) LDDTENASAYAANAG(L1) SYSLYGTTLEQQYNK(E6) DTENASAYAANAGVD(L1)
HPV-18	KPLGAVALK(E6) AVNYGVTVL(E6) LTNKGISDL(E6) SAVNYGVTV(E6) DTDYSIAEA(E6) DYSIAEAAF(E6) AGAKAGLTF(E6) GAKAGLTFL(E6) QTVDAALAA(E6) ALAAAQTNA(E6) GVTVLPTFK(E6) NADTDYSIA(E6) GRQTVDAAL(E6)	KDKPLGAVALKSYEE(E6) LGAVALKSYEEELAK(E6) SLYIKGTGMRASPGS(L1) PQSLYIKGTGMRASP(L1)

Subsequently, the identified epitopes underwent mapping onto the E6 and L1 proteins of both HPV16 and HPV18. Notably, during this process, it was observed that none of the E7 epitopes met the stringent criteria to qualify for inclusion in the vaccine formulation.

To visualize the mapped antigens, Chimera, a molecular visualization tool, was employed. The results of this mapping process are graphically depicted in Figure 18. This visual representation aids in comprehending the spatial distribution and interaction of the identified epitopes on the respective HPV proteins.

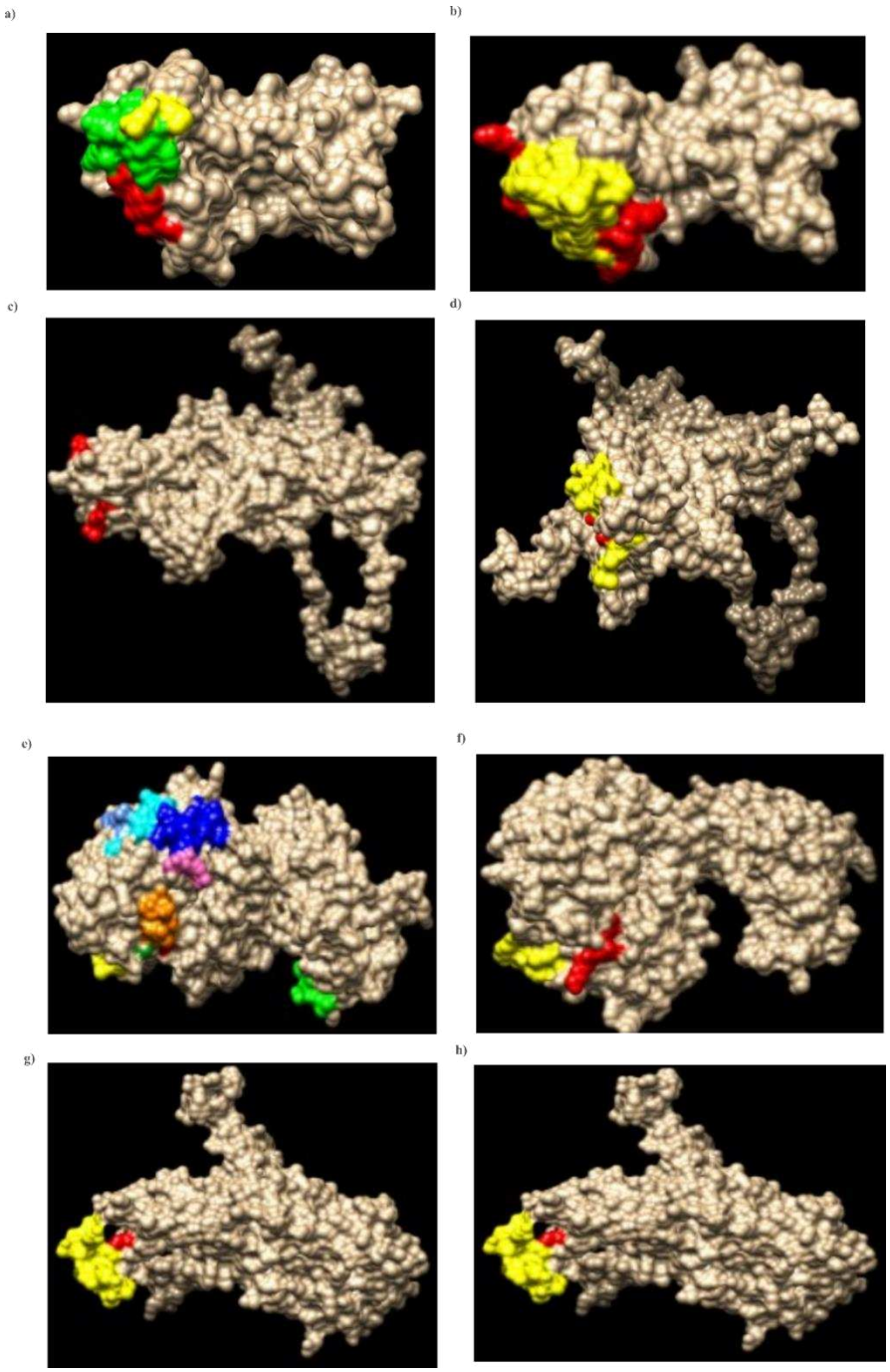


*Figure 18 Illustrates the mapped epitopes for C*ytotoxic T *Cells (CTCs) and T helper cells (Th cells) on HPV16 and HPV18, focusing on E6 and L1 proteins. The specific mappings are as follows: a) CTCs epitopes mapped on HPV16-E6, b) Th cells epitopes mapped on HPV16-E6, c) CTCs epitopes mapped on HPV16-L1, d) Th cells epitopes mapped on HPV16-L, e) CTCs epitopes mapped on HPV18-E6, f) Th cells epitopes mapped on HPV18-E6, g) CTCs epitopes mapped on HPV18-L1, h) Th cells epitopes mapped on HPV18-L1. This comprehensive depiction provides insights into the antigenic regions recognized by CTCs and Th cells for each specific HPV protein, aiding in the understanding of immune responses.*

B cell-Stimulating Epitopes

Linear and conformational epitopes were extracted from HPV16 and HPV18, specifically from E6, E7, and L1 proteins. The process involved identifying two types of epitopes capable of stimulating B cells: linear and conformational. These epitopes play a crucial role in triggering antibody production within B cells. Detailed information on the extracted linear and conformational epitopes can be found in Tables 9, 9.1, 10 and 10.1, respectively. Additionally, the spatial distribution of conformational epitopes on HPV antigens is visually represented in Figure 20.

Table 9 Displays Linear B cell epitopes along with their respective source antigens

No	**Peptides**	**Origin**
1	LSPEEKQRHLD	HPV16-E6
2	GLKAKPKFTLG	HPV16-L1
3	LEDGDM	HPV18-L1
4	PQVAAT	HPV18-E6
5	YAFKYENGKYDIKDVGVDNA	HPV18-E6
6	TLQDTKCE	HPV18-L1
7	TPAAAFQDK	HPV18-E6
8	RAGTMGDTVPQSLYIKGTGMRASP	HPV18-L1
9	LQMSADPY	HPV18-L1
10	LTDEGLEAVNKDKPLGAVALKSYEEELAKDP	HPV18-E6
11	YRHYCYSLYGTTLEQQYNKPLCDL	HPV16-E6
12	FGFPDTSFYNPDT	HPV16-L1
13	STILEDWNFGLQPPPGGTLEDTYRFVTS	HPV16-L1

Table 9.1 Continue with displays Linear B cell epitopes along with their respective source antigens

1	TSETTYKNTNFKEYLRHGEE	HPV16-L1
2	YIKMVSEPY	HPV16-L1
3	QKGEIMPNIPQMSA	HPV18-E6
4	DLKEKFSLDLDQYPL	HPV18-L1
5	FYNKLDDTESSHAATSNVSEDVR	HPV18-L1
6	PVPVSKVVSTDEYV	HPV16-L1
7	WRRPRTETQVGSSGAAESSELTLQEL	HPV18-E6
8	LEDWNNKDPYDKL	HPV18-L1
9	HMNADTDYSI	HPV18-E6

Table 10 Displays conformational B cell epitopes along with their respective source antigens.

No	Peptides	Origin
1	PTFKGQP	HPV18-E6
2	KPLSPEEKQRHLD	HPV16-E6
3	KYENGKYDIKDV	HPV18-E6
4	AEITPAAAFQDKLYP	HPV18-E6
5	PLSPEEKQRHLDKKKRFHNIGGRWTGRCIACWRRPRTETQVGSSGAAESSE	HPV18-E6
6	KGTACKSRPLSQGDCPPLE	HPV18-L1
7	LEDWNFGLQPPPGGTLEDTYRFVTSQAIACQKHTPPAPKE	HPV16-L1
8	VCTKCLRFYSKVSEFRWYRYSVYGTTLEKLTNKGIS	HPV18-E6
9	GSPCTQVAVQPGDCPP	HPV16-L1
10	FSADLDQFPLGRKFLLQLGLKAKPKFTLGKRK	HPV16-L1
11	PNKFGFPDTSFYNPDTQ	HPV16-L1
12	MSLCAAISTSETTYKNTNFKEYLR	HPV16-L1
13	LEIPYDELRLNCVYCKGQLTETEVLD	HPV18-E6
14	HPLLNKLDDTENASAYAANAGVDNRECI	HPV16-L1
15	IDTSAVNY	HPV18-E6
16	SGRQTVDA	HPV18-E6
17	NKLDDTESSHAATSNVSEDVRDN	HPV18-L1
18	AAFNKGETA	HPV18-E6
19	RVPAGGGNKQDIP	HPV18-L1

Table 10.1 Continue with displays conformational B cell epitopes along with their respective source antigens.

20	SMVTSD	HPV16-L1
21	GSIVTSD	HPV18-L1
22	ATVYLPPVPVSKVVSTDEYVARTNIYYH	HPV16-L1
23	YNKDLLPNPPKTWEEIPALDKELKAKGKSA	HPV18-E6

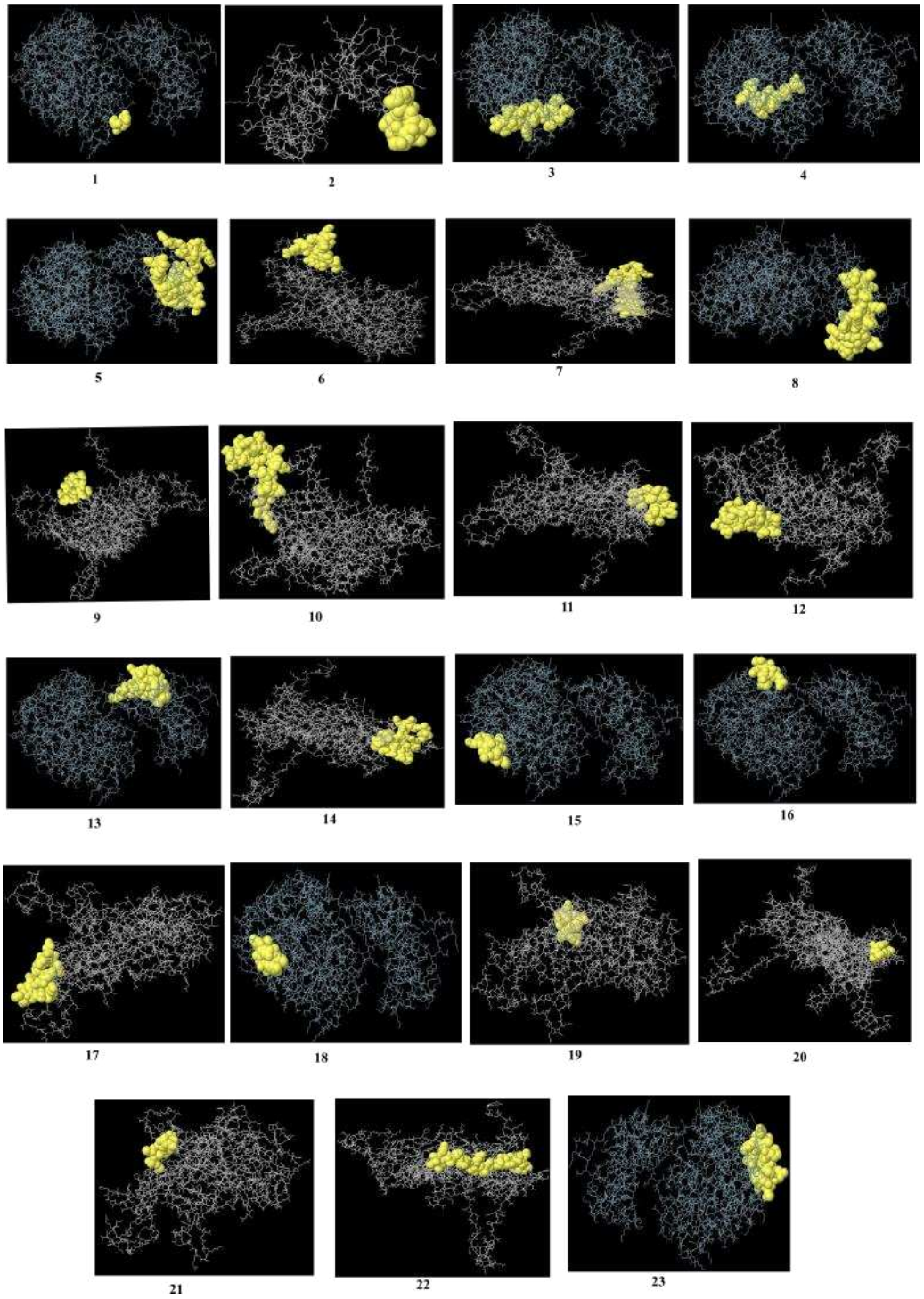
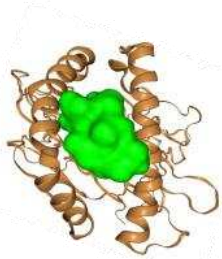
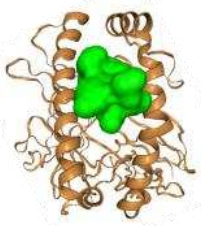


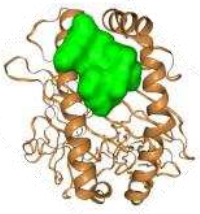
Figure 19 This figure illustrates the organization of conformational B cells epitopes for HPV16 and HPV18 based on the epitope counts provided in Table X.

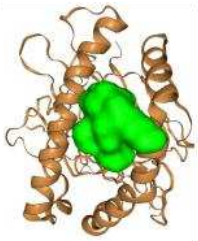
Molecular docking of epitopes with MHC-I and MHC-II

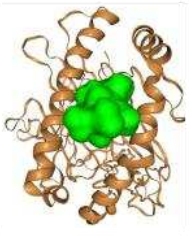
After the epitopes successfully cleared all filters and demonstrated overlaps with B cell epitopes, the 3D structures of the peptides obtained from the PEPFOLD 4 server [126] underwent docking with HLA-A, HLA-DBR, and HLA-DQ, guided by their recognition by T cytotoxic or T helper cells. The outcomes of these docking interactions are visually represented in the accompanying Figure 20.

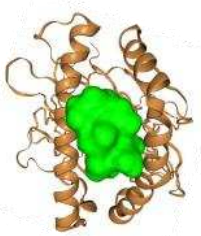

AVNYGVTVL-HPV18-E6

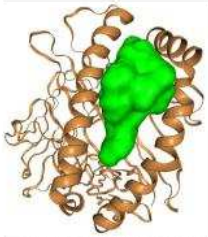

DTDYSIAEA-HPV18-E6

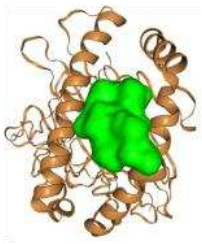

TTLEQQYNK-HPV16-E6

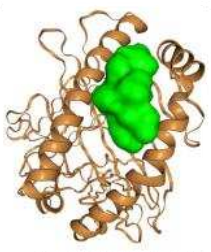

AGAKAGLTF-HPV18-E6

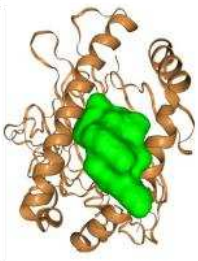

ALAAAQTNA-HPV18-E6

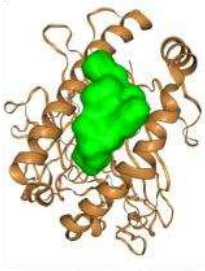

SYSLYGTTL-HPV16-E6

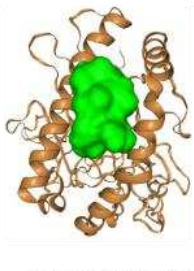

YGTTLEQQY-HPV16-E6

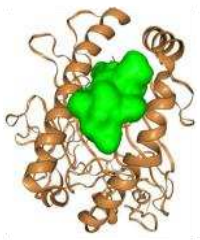

GVTVLPTFK-HPV18-E6

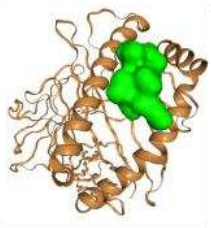

DYSIAEAAF-HPV18-E6

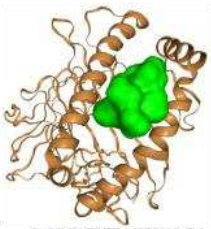

GAKAGLTFL-HPV18-E6

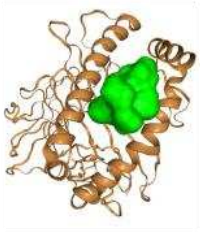

LTNKGISDL-HPV18-E6

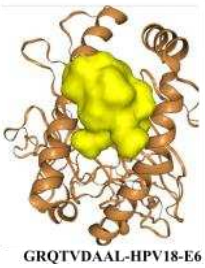

NADTDYSIA-HPV18-E6

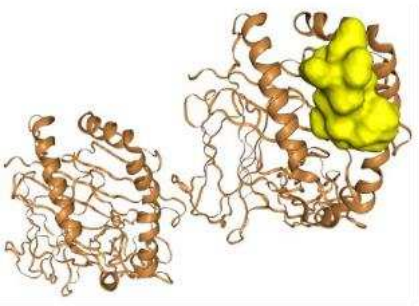

KPLGAVALK-HPV18-E6


QTVDAALAA-HPV18-E6

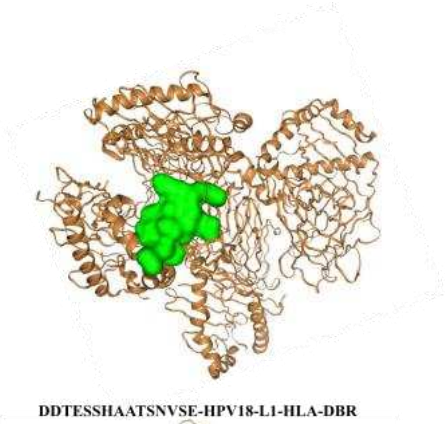

SAVNYGVTV-HPV18-E6


SAYAANAGV-HPV16-L1

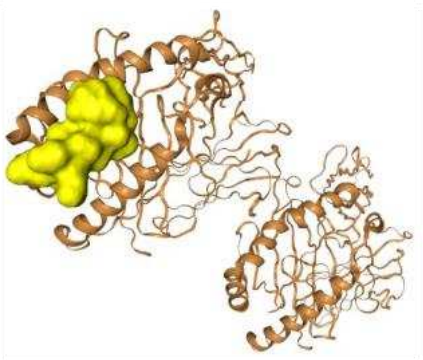

GRQTVDAAL-HPV18-E6



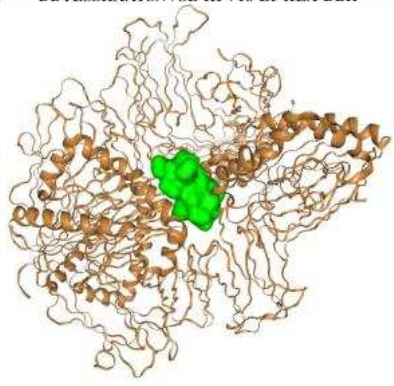
DDTESSHAATSNVSE-HPV18-L1-HLA-DQ



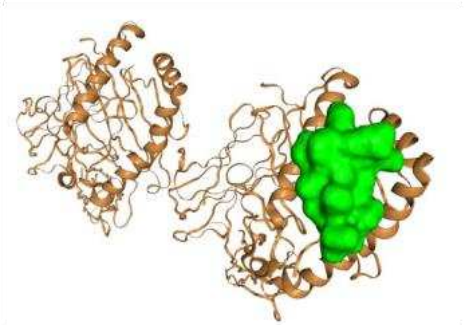
DDTESSHAATSNVSE-HPV18-L1-HLA-DBR



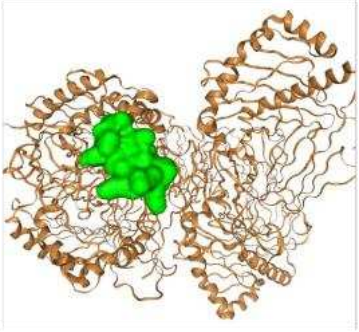
DTENASAYAANAGVD-HPV16-L1-HLA-DQ



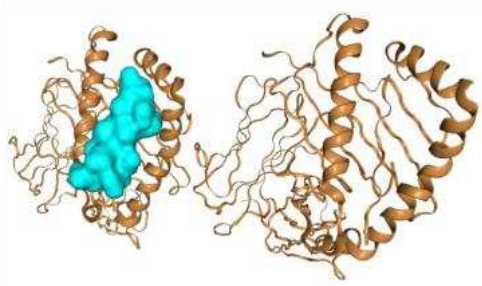
DTENASAYAANAGVD-HPV16-L1-HLA-DBR



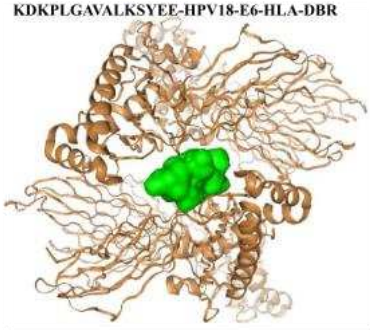
KDKPLGAVALKSYEE-HPV18-E6-HLA-DQ



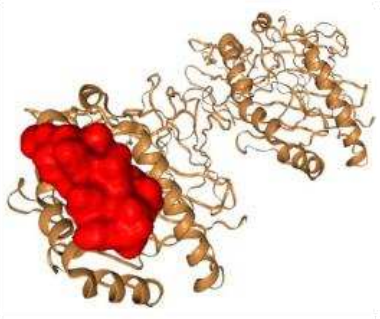
KDKPLGAVALKSYEE-HPV18-E6-HLA-DBR



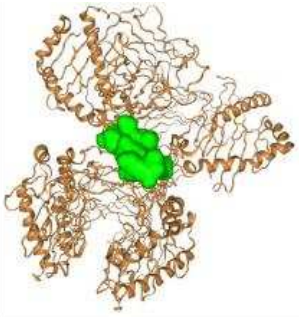
LDDTENASAYAANAG-HPV16-L1 -HLA-DQ



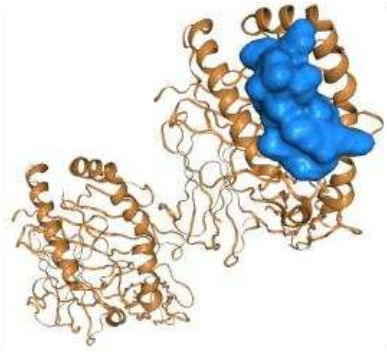
LDDTENASAYAANAG-HPV16-L1-HLA-DBR



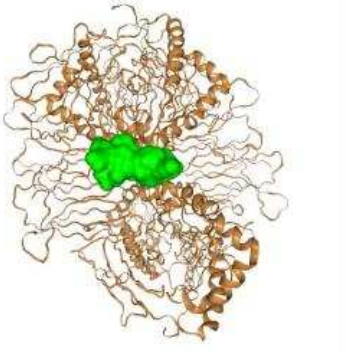
LGAVALKSYEEELAK-HPV18-E6-HLA-DQ



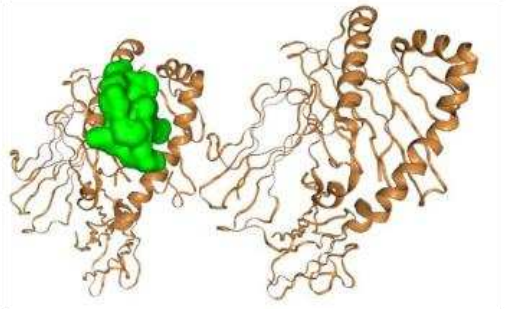
LGAVALKSYEEELAK-HPV18-E6-HLA-DBR



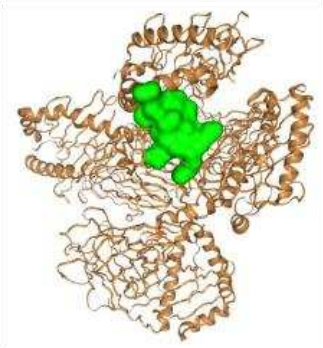
PQSLYIKGTGMRASP-HPV18-L1-HLA-DQ



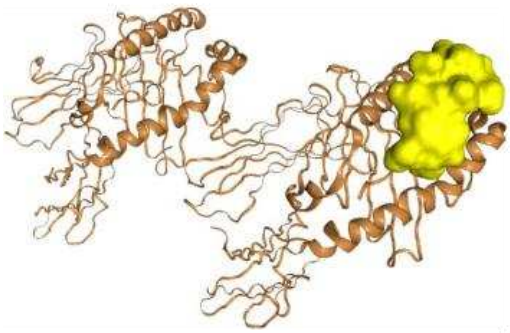
PQSLYIKGTGMRASP-HPV18-L1-HLA-DBR



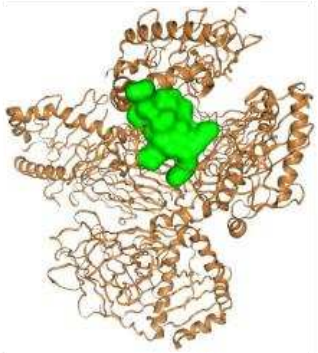
SLYIKGTGMRASPGS-HPV18-L1-HLA-DQ



SLYIKGTGMRASPGS-HPV18-L1-HLA-DBR



SYSLYGTTLEQQYNK-HPV16-E6-HLA-DQ



SYSLYGTTLEQQYNK-HPV16-E6-HLA-DBR

Figure 20 Peptides, having cleared filters and demonstrated overlaps with B cell epitopes, underwent docking with HLA-A, HLA-DRB, and HLA-DQ

Vaccine construction

The vaccine was developed by incorporating the CpG adjuvant and employing EAAAK, AAY, and GPGPG linkers. Detailed physicochemical characteristics of the vaccine are presented in the accompanying Table 11.

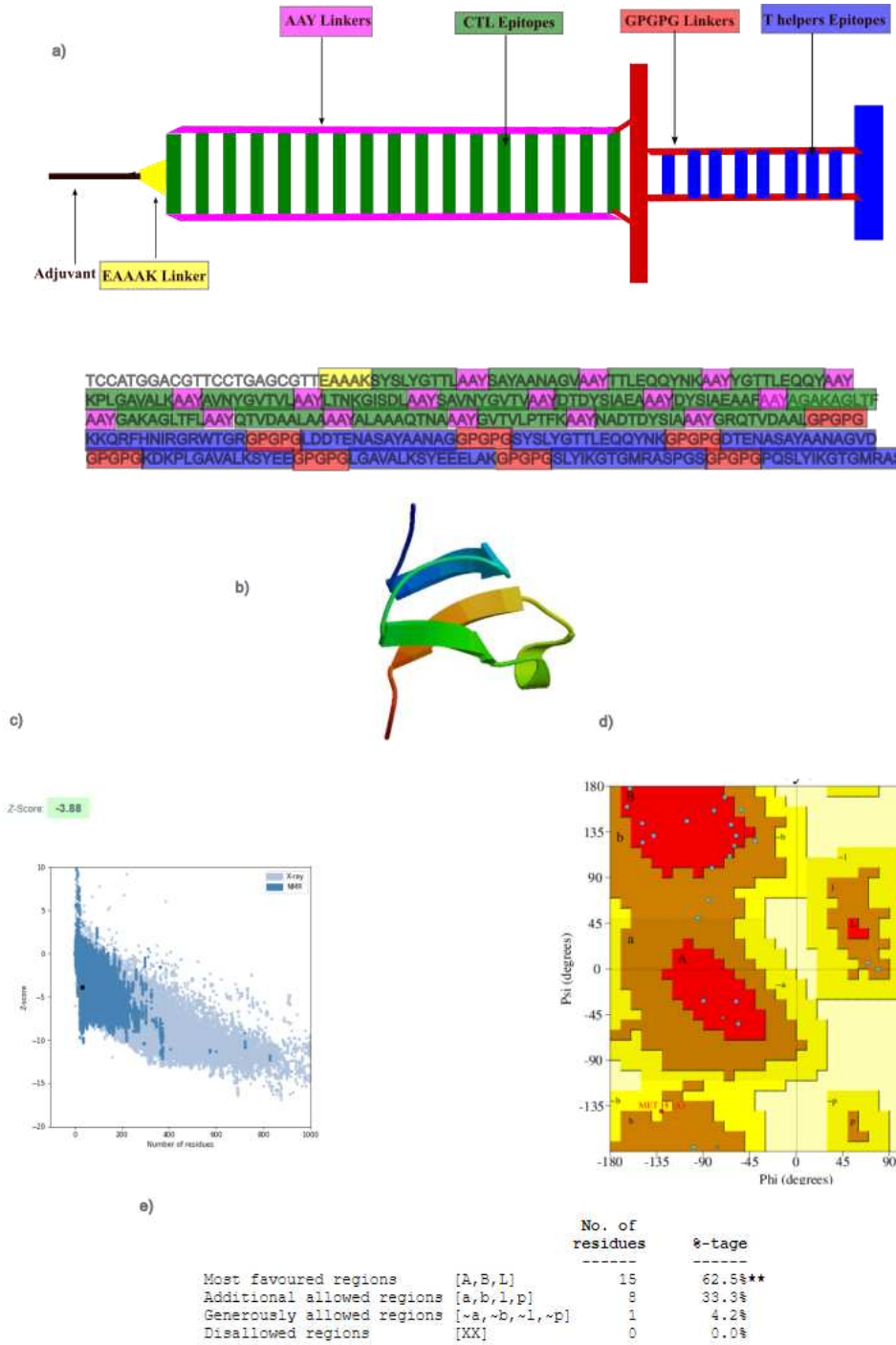
Table 11 Physicochemical structure of designed multiepitope vaccine.

Proteins	**Molecular weight**	**Theoretical pI**	**Instability index**	**Estimated half-life** (mammalian reticulocytes)	**Aliphatic index**	**GRAVY**
Vaccine	38798.99	5.93	20.37	7.2 hours	66.55	-0.107

According to AllerCatPro [125] analysis, the structure of the vaccine is non-allergenic. Experimental assessments were conducted, including tests for gluten, IgE production, similarity to common allergens, and autoimmune-stimulating allergens. VaxiJen 2.0 yielded a predicted score of 0.64, indicating the vaccine's antigenic nature. The vaccine's sequence is as follows:

“TCC**ATGGA**CG**TTCCT**GAGCGTTEAAAKSYSLYGTTLAAYSAYAANAGVAAYTTLEQQYNKAAYYGTTLEQ
QYAAYKPLGAVALKAAYAVNYGVTVLAAYLTNKGISDLAAYSAVNYGVTVAAYDTDYSIAEAAAYDYSIAE
AAFAAYAGAKAGLTFAAYGAKAGLTFLAAYQTVDAALAAAAYALAAAQTNAAAYGVTVLPTFKAAYNADT
DYSIAAAYGRQTVDAALGPGPGKKQRFHNIRGRWTGRGPGPGLDDTENASAYAANAGGPGPGSYSLYGTTLE
QQYNKGPGPGDTENASAYAANAGVDGPGPGKDKPLGAVALKSYEEGPGPGLGAVALKSYEEELAKGPGPGSL
YIKGTGMRASPGSGPGPGPQSLYIKGTGMRASP”

The schematic representation of the designed vaccine, along with its physicochemical properties, is illustrated in Figure 21.



		No. of residues	%-tage
Most favoured regions	[A,B,L]	15	62.5%**
Additional allowed regions	[a,b,l,p]	8	33.3%
Generously allowed regions	[~a,~b,~l,~p]	1	4.2%
Disallowed regions	[XX]	0	0.0%

Figure 21 The vaccine was constructed utilizing EAAAK, AAY, and GPGPG linkers between CpG, CTCs, and Th cell epitopes. a) Schematic representation of the protein structure. b) Three-dimensional structure of the protein obtained from SwissModel. c) The Z score of the protein, positioning it within the range observed for NMR proteins. d) The Ramachandran plot depicting the protein's conformational quality.

Linear and conformational B cells epitopes of the vaccine

Linear and conformational epitopes of the vaccine were identified through the IEDB server. A total of 13 linear epitopes have been delineated and are detailed in the accompanying Table 12. Additionally, the presence of two discontinuous or conformational B-cell epitopes is illustrated in the corresponding Figure 22.

Table 12 Predicted Linear epitopes by IEDB for vaccine

Linear B cells epitopes	**Score**
GCGTTEAAAK (27)	0.53
QYNKAAYY (64)	0.513
TLEQQYAAY (75)	0.519
IS (106)	0.515
YSI (129)	0.504
A (192)	0.505
AAYNADT (211)	0.527
GRQTVDAALGPGPGKKQRFHNIRGRWTGRGPGPGLDDTENASAYAANAGGPGPGS (274)	0.590
EQQYNKGPGPGDTENASAYAANAGVDGPGPGKDKPL (318)	0.615
KSYEEGPGP (332)	0.517
EEELAKGPGP (352)	0.604
GTGMRASPGSGPGPGP (374)	0.620
KGTGMR (385)	0.520

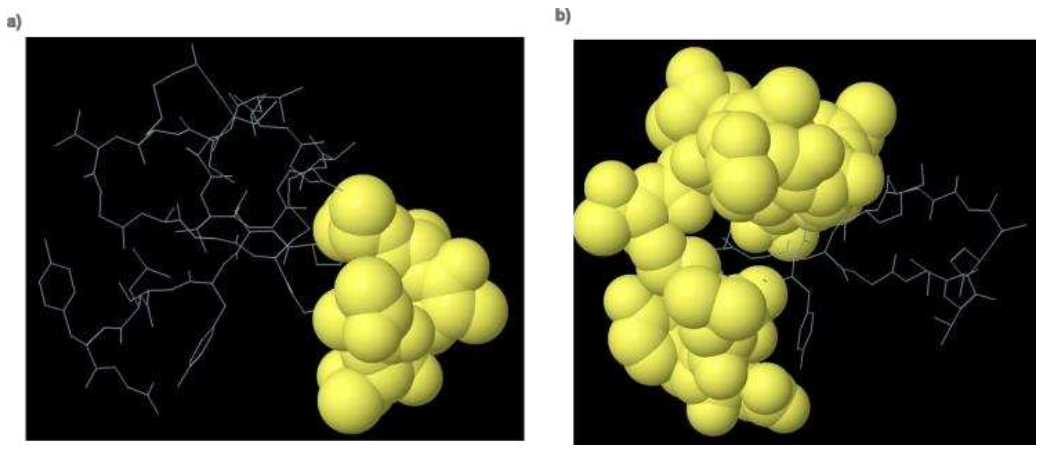


Figure 22 Two conformational or discontinuous B cells epitopes have been predicted using the IEDB-Ellipro tool.

Interaction of TLR9 with designed vaccine

The interaction between TLR9 and the designed vaccine was investigated by first constructing the TLR9 using the Swiss-Model server (SwissModel) [172]. Subsequently, the HDOCK server (HDOCK) [170, 171] was employed to perform docking simulations between the vaccine and TLR9. Ten models were generated, and, consistent with HDOCK server protocols, the model displaying the lowest energy, recorded at -232.47 kcal/mol, was chosen among them. The outcomes of this interaction are visually presented in the accompanying Figure 23.

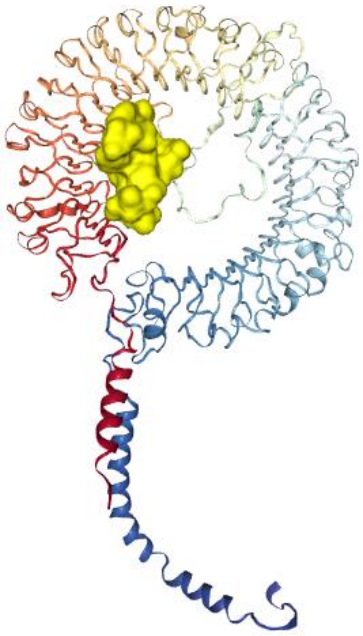


Figure 23 The docking model depicting the interaction between the vaccine and TLR9 was generated using the HDOCK server. In the visual representation, TLR9 is color-coded in a rainbow spectrum, while the vaccine is highlighted in yellow.

In silico coloning

Following the reverse translation of the vaccine into a DNA sequence using EMBOSS Backtranseq (EMBOSS Backtranseq) [219]:

“ACCTGCTGCGCCACCGGCGGCGCCTGCGGCACCACCTGCTGCACCGGCGCCGGCTGCGGC

ACCACCGAGGCCGCCGCCAAGAGCTACAGCCTGTACGGCACCACCCTGGCCGCCTACAGC

GCCTACGCCGCCAACGCCGGCGTGGCCGCCTACACCACCCTGGAGCAGCAGTACAACAAG

GCCGCCTACTACGGCACCACCCTGGAGCAGCAGTACGCCGCCTACAAGCCCCTGGGCGCC

GTGGCCCTGAAGGCCGCCTACGCCGTGAACTACGGCGTGACCGTGCTGGCCGCCTACCTG

ACCAACAAGGGCATCAGCGACCTGGCCGCCTACAGCGCCGTGAACTACGGCGTGACCGTG

GCCGCCTACGACACCGACTACAGCATCGCCGAGGCCGCCGCCTACGACTACAGCATCGCC

GAGGCCGCCTTCGCCGCCTACGCCGGCGCCAAGGCCGGCCTGACCTTCGCCGCCTACGGC

GCCAAGGCCGGCCTGACCTTCCTGGCCGCCTACCAGACCGTGGACGCCGCCCTGGCCGCC

GCCGCCTACGCCCTGGCCGCCGCCCAGACCAACGCCGCCGCCTACGGCGTGACCGTGCTG

CCCACCTTCAAGGCCGCCTACAACGCCGACACCGACTACAGCATCGCCGCCGCCTACGGC

AGGCAGACCGTGGACGCCGCCCTGGGCCCCGGCCCCGGCAAGAAGCAGAGGTTCCACAAC

ATCAGGGGCAGGTGGACCGGCAGGGGCCCCGGCCCCGGCCTGGACGACACCGAGAACGCC

AGCGCCTACGCCGCCAACGCCGGCGGCCCCGGCCCCGGCAGCTACAGCCTGTACGGCACC

ACCCTGGAGCAGCAGTACAACAAGGGCCCCGGCCCCGGCGACACCGAGAACGCCAGCGCC

TACGCCGCCAACGCCGGCGTGGACGGCCCCGGCCCCGGCAAGGACAAGCCCCTGGGCGCC

GTGGCCCTGAAGAGCTACGAGGAGGGCCCCGGCCCCGGCCTGGGCGCCGTGGCCCTGAAG

AGCTACGAGGAGGAGCTGGCCAAGGGCCCCGGCCCCGGCAGCCTGTACATCAAGGGCACC

GGCATGAGGGCCAGCCCCGGCAGCGGCCCCGGCCCCGGCCCCCAGAGCCTGTACATCAAG

GGCACCGGCATGAGGGCCAGCCCC”

The Java Codon Adaptation tool (JCat) [175] was utilized to tailor the vaccine code for production in the bacterial system, specifically *E. coli* (K12). Post-optimization, the vaccine DNA structure comprised 1164 nucleotides, exhibiting a GC content of 56.01%. The Codon Adaptation Index (CAI) for the improved sequence reached 1.0. Subsequently, the pET28a(+) vector was chosen from the Addgene server (addgene) [220], imported into the Benchling server (Benchling), and the plasmid design was finalized, Figure 24.

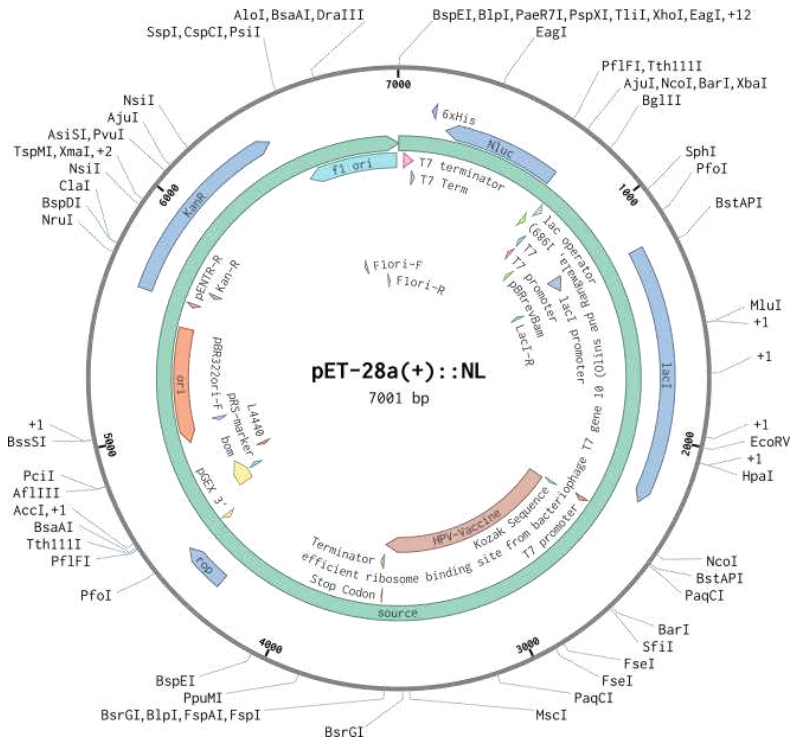
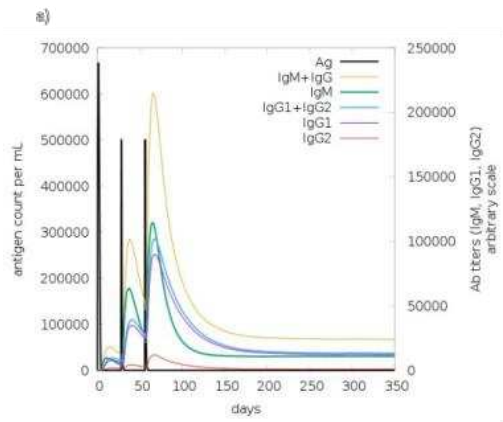


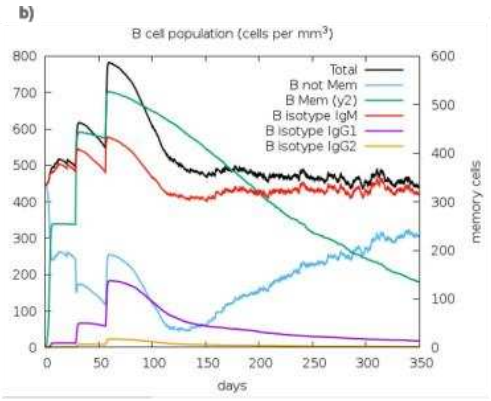
Figure 24 The multiepitope HPV vaccine was integrated into the pET28a(+) plasmid, denoted as "HPV-Vaccine." The synthesis of the vaccine utilized the T7 promoter and terminator sequences. The termination codon TGA marked the conclusion of the coding sequence. Additionally, the Kozak sequence, represented by GCCGCCACCAUGG, played a crucial role in optimizing translational initiation.

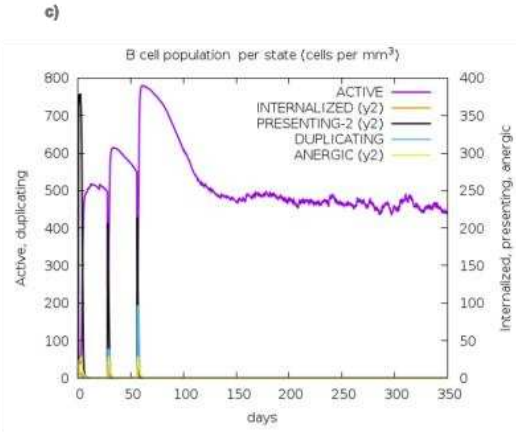
Immune Simulations of Constructed Vaccine

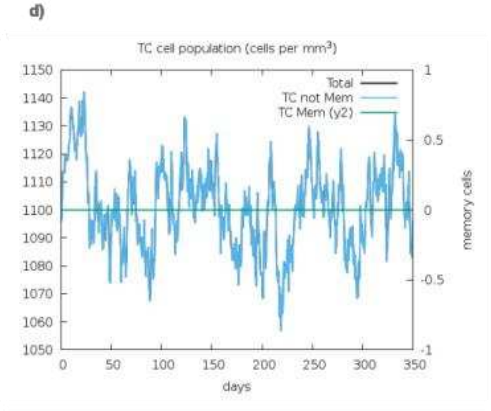
The immune simulations of the vaccine construct revealed that a three-injection regimen effectively elicited diverse immunoglobulins. The primary response manifested as an elevated IgM level, while the secondary response exhibited increased IgM + IgG, IgG1 + IgG2, IgG1, IgG2, and B-cell populations. Subsequent exposures led to a decline in antigen levels. Both CTL and HTL T cell populations exhibited heightened responses, indicating the immunogenicity of T cell epitopes within the vaccine. Macrophage activity increased with each exposure, and NK cell activity remained consistent. Notably, subsequent exposures resulted in significant increases in IFN-γ, IL-10, IL-23, and IL-12 levels.

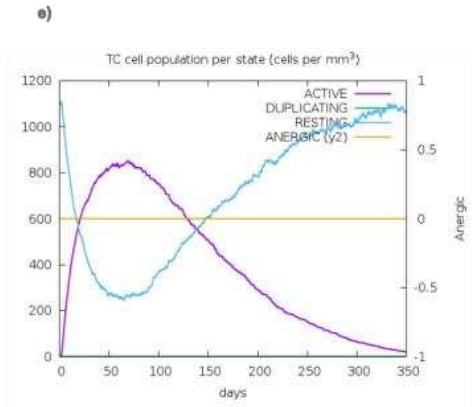
Upon 12 repeated exposures at regular intervals, the antigen levels reached a similar peak, accompanied by a remarkable increase in IgM + IgG, IgG1 + IgG2 levels. Persistent elevation in B and T cell memory was observed throughout the exposure, while IFN-γ levels remained consistently high from the first to the last exposure. These findings underscore the vaccine's ability to generate a robust immune response in short exposures, with continued enhancement upon subsequent repeated exposures, Figure 25.

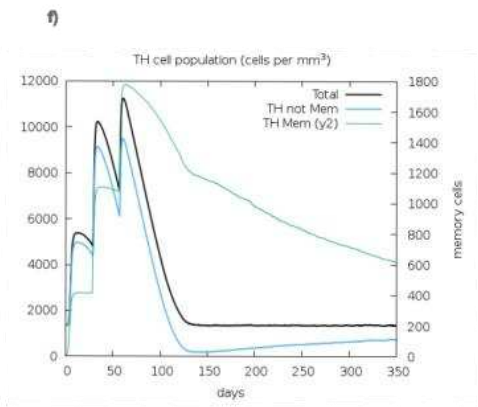

a)
Ag
IgM+IgG
IgM
IgG1+IgG2
IgG1
IgG2
antigen count per mL
Ab titers (IgM, IgG1, IgG2) arbitrary scale
days


b)
B cell population (cells per mm³)
Total
B not Mem
B Mem (y2)
B isotype IgM
B isotype IgG1
B isotype IgG2
memory cells
days


c)
B cell population per state (cells per mm³)
ACTIVE
INTERNALIZED (y2)
PRESENTING-2 (y2)
DUPLICATING
ANERGIC (y2)
Active, duplicating
Internalized, presenting, anergic
days


d)
TC cell population (cells per mm³)
Total
TC not Mem
TC Mem (y2)
memory cells
days


e)
TC cell population per state (cells per mm³)
ACTIVE
DUPLICATING
RESTING
ANERGIC (y2)
Anergic
days


f)
TH cell population (cells per mm³)
Total
TH not Mem
TH Mem (y2)
memory cells
days

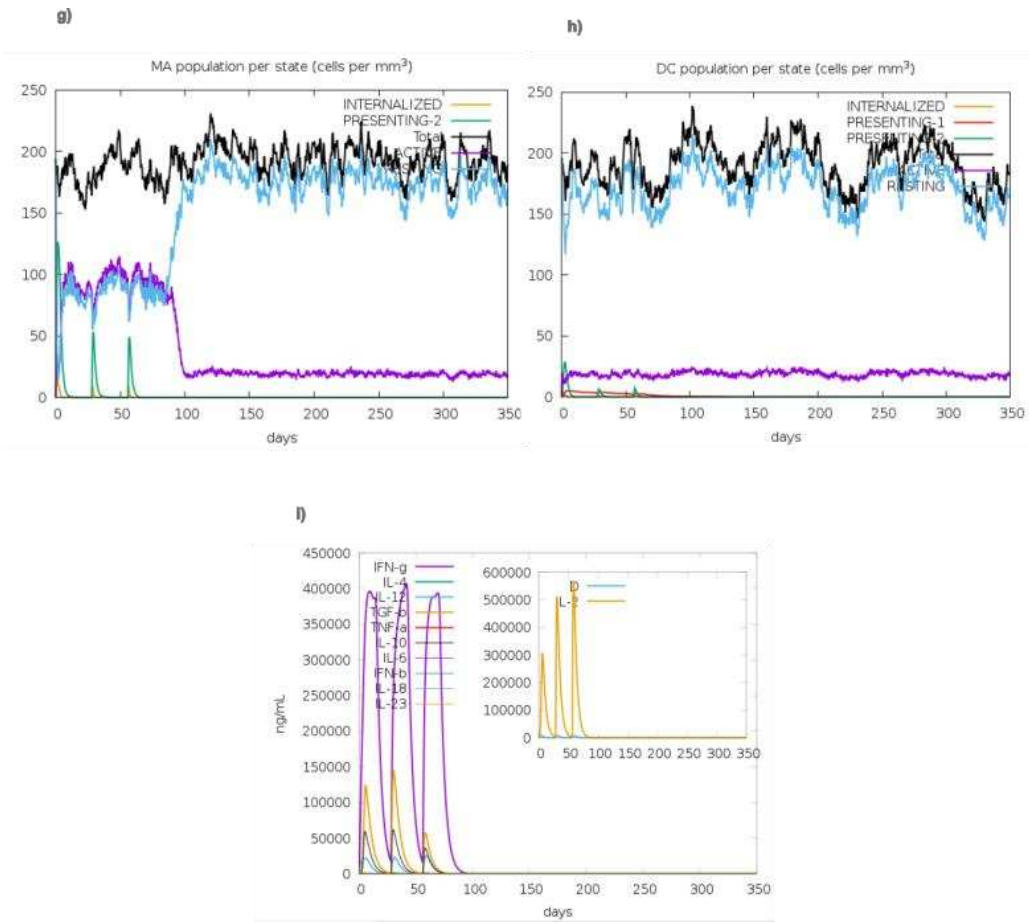


Figure 25 In Silico Simulation of Immune Response Using Vaccine as an Antigen. The in silico simulation of immune response following three subsequent injections of the vaccine, used as an antigen, revealed multifaceted dynamics: Antigen and Immunoglobins a) the simulation depicted variations in antigen levels and immunoglobulin responses, providing insights into the vaccine's impact on humoral immunity. B-cell Population. b) The B-cell population exhibited dynamic changes, highlighting the vaccine's ability to modulate B-cell responses crucial for adaptive immunity. B-cell Population per State. c) Detailed analysis of B-cell population states provided nuanced information about the vaccine-induced states of B cells. Cytotoxic T-cell Population. d) The simulation showcased alterations in the cytotoxic T-cell population, indicating the vaccine's influence on cellular immunity. Cytotoxic T-cell Population per State. e) Specific states of cytotoxic T cells were explored, offering insights into the diverse effects of the vaccine on this crucial immune cell subset. Helper T-cell Population. f) The helper T-cell population demonstrated fluctuations, indicating the vaccine's impact on orchestrating immune responses. Macrophages Population per State. g) The simulation highlighted changes in macrophage states, suggesting the vaccine's role in modulating macrophage activity. Dendritic Cell Population per State. h) Analysis of dendritic cell states provided valuable information on the vaccine's influence on antigen-presenting cells. Cytokine Production i) The simulation illustrated the vaccine's impact on cytokine production, shedding light on its role in immune signaling.

System Immunology: Analysis of HPV-Related Gene Expression in PBMCs

The Gene Expression Omnibus (GEO) was queried from 12-10-2023 to 19-10-2023 using the keyword HPV. This yielded 54,677 results in the Profiles Database and 8,491 records in the DataSets Database. Only studies relevant to humans were considered, resulting in 7,997 papers with human datasets. Papers without CEL supplementary data were excluded, leaving 1,486 papers for further analysis.

CEL files, integral to Affymetrix DNA microarray analysis, store intensity calculations crucial for understanding gene expression patterns. After refining the search with the keyword "cervical," 184 papers remained. Abstracts from these papers were scrutinized for relevance to gene expression in PBMCs of patients with HPV and cervical cancers post-HPV vaccination. The selection process is illustrated in the Figure 26.

One study meeting all criteria was included, focusing on gene expression in PBMCs from a double-blind, randomized, placebo-controlled phase II trial of a monovalent HPV-16 L1 VLP vaccine. Participants, low-risk for HPV16 exposure, received three intramuscular doses. Microarray data from 27 subjects (20 vaccine, 7 placebo) at months 0 and 2 were considered (GEO: GSE13587).

GEO2R was employed to select groups from pre and postvaccinated participants, categorized by media or HPV VLP vaccine. Normalized data was downloaded for analysis Figure 27, resulting in 8,793 genes. After filtering for significance and p-value, 635 genes showed significant expression changes, including 19 with over 2.5-fold changes based on LogFC. Among these, three genes exhibited a 2.5-fold increase, while 16 showed a negative expression.

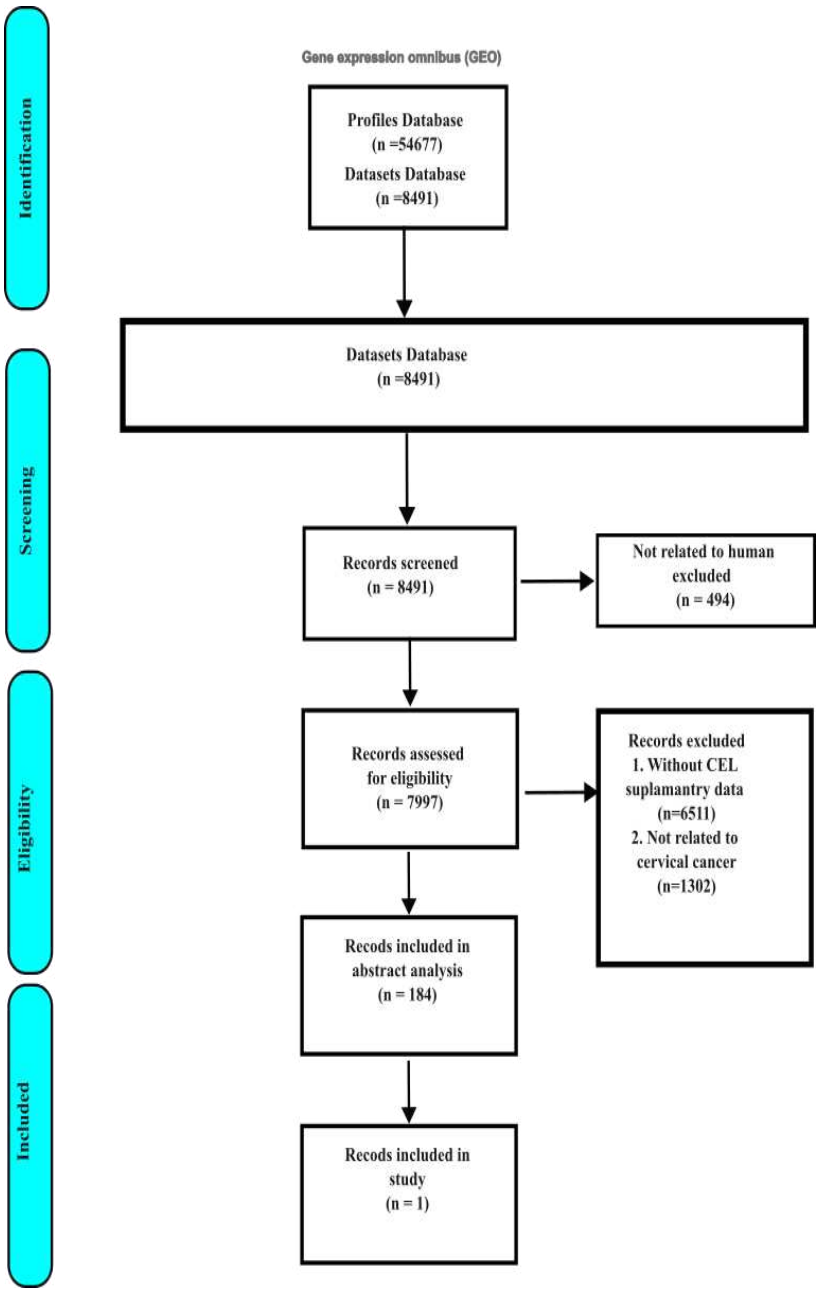


Figure 26 Outlines the process of paper selection for gene expression analysis. The diagram illustrates the sequential steps involved in choosing relevant papers for the analysis of gene expression data. This visual aid serves as a guide to the selection criteria and decision-making process applied in the context of gene expression studies.

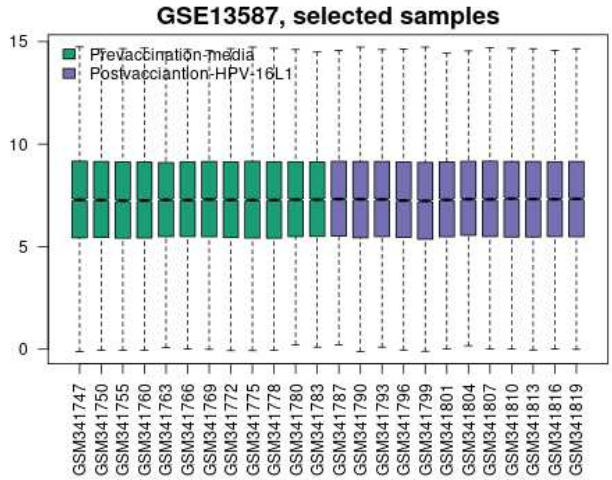


Figure 27 The normalized samples are visually represented for both groups: those who received a placebo are depicted in green, while individuals receiving HPV16 virus-like particles are represented in purple.

Network Analysis

To conduct the network analysis, all genes were initially subjected to the STRING server (STRING) based on their p-value significance. Subsequently, a tab-separated value (TSV) file was generated and imported into Cytoscape. The network in Cytoscape underwent enrichment analysis utilizing betweenness centrality, degree, and closeness centrality. Specifically, 100 hub genes and 1858 edges were selected based on their high values in degree (43-186), betweenness centrality (0.002-0.132), and closeness centrality (0.202-0.566). These selected hub genes were then reintroduced into the STRING server. The process is demonstrated in Figure 28.

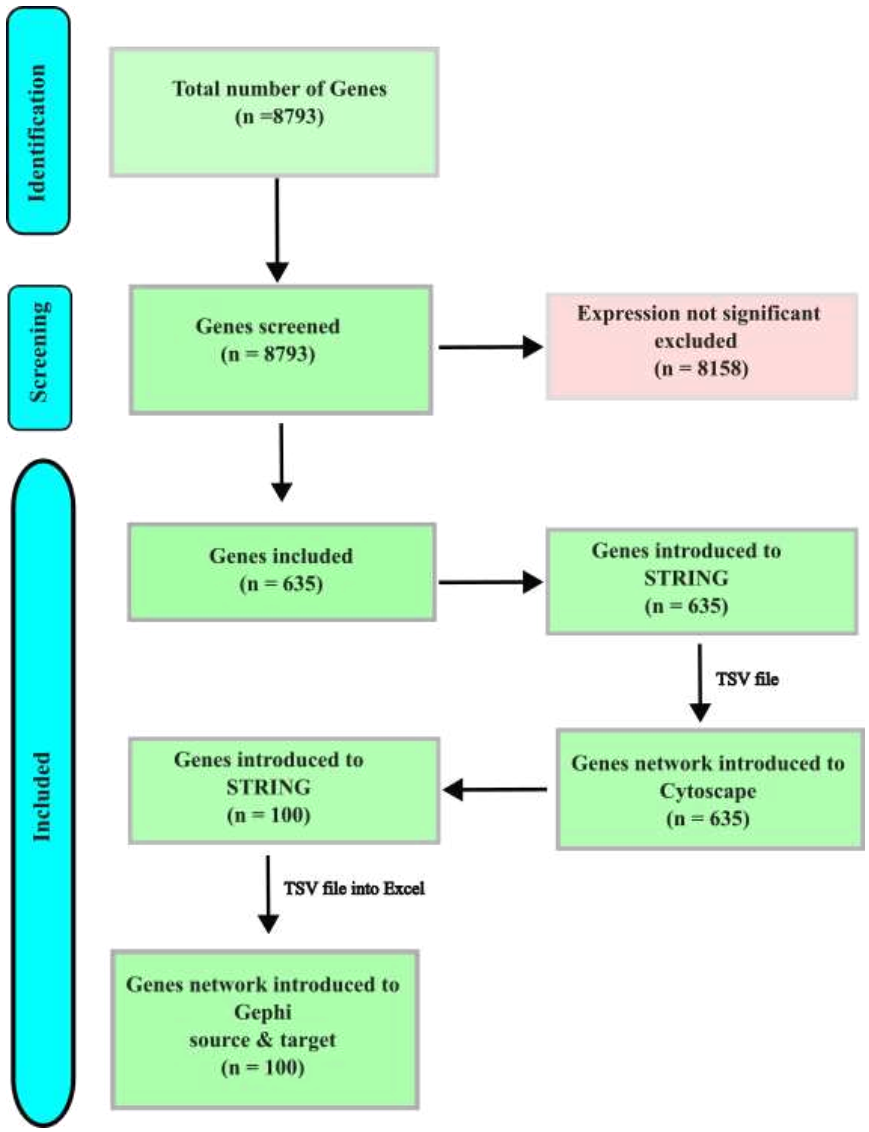


Figure 28 A total of 8793 genes exhibiting expression variation between participants receiving a placebo and those receiving the HPV virus-like particle vaccine were identified. Subsequently, 635 genes demonstrated significant variability between the two groups. The selected genes underwent further analysis by introducing them to the STRING website to explore potential networks among them. To identify the top 100 hub genes, a TSV file extracted from STRING was imported into Cytoscape. The selection of these hub genes was based on criteria such as Degree, Betweenness Centrality, and Closeness Centrality. These 100 hub genes were then exported into a CSV file for additional analysis within the STRING platform. Another gene network was constructed in STRING, utilizing the exported hub genes.

Following the creation of a new network in STRING, a TSV file was obtained for further analysis in Gephi software. The Gephi analysis involved editing the TSV file based on source and target parameters. The gene network was further scrutinized using eigenvector centrality and modularity analysis, as illustrated in Figure 29. This Gephi analysis revealed the identification of five distinct modules within the gene network, enhancing our understanding of the underlying biological relationships, Figure 30.

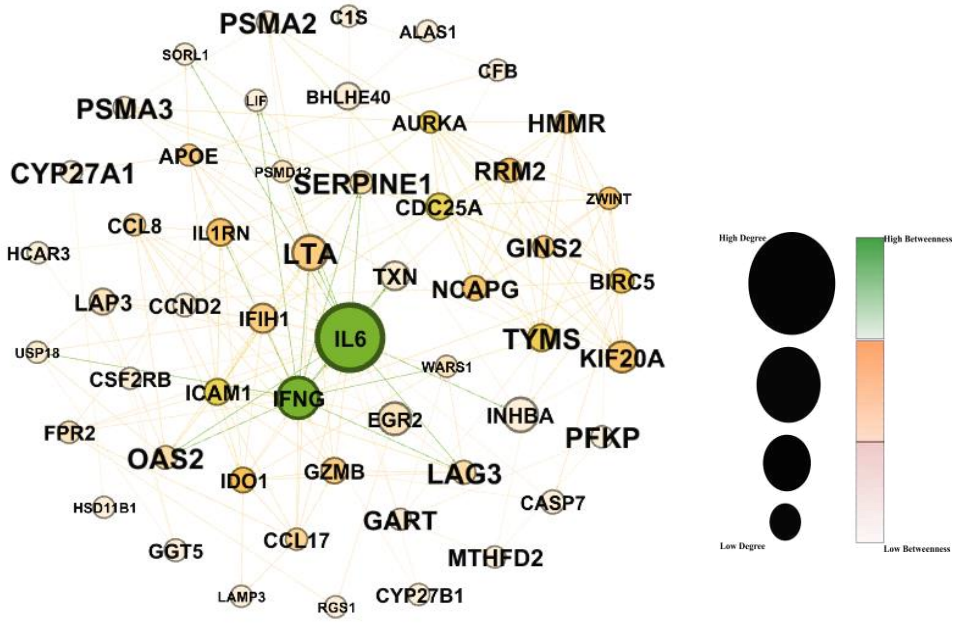


Figure 29 The Gephi network analysis of 100 hub genes reveals the pivotal role of IL6 in the network. IL6 occupies a critical position, as indicated by its prominence in the network. Following closely on the second rank is IFN-γ. *The size of the nodes corresponds to the Betweenness centrality, highlighting the nodes' significance in information flow. Additionally, the color of the nodes signifies the Degree, with a more greenish hue indicating a larger degree. The backbone of this network is constructed from 53 nodes and 185 edges, providing a structural foundation for the interactions among the identified hub genes. This analysis offers valuable insights into the key players and their connectivity within the biological network.*

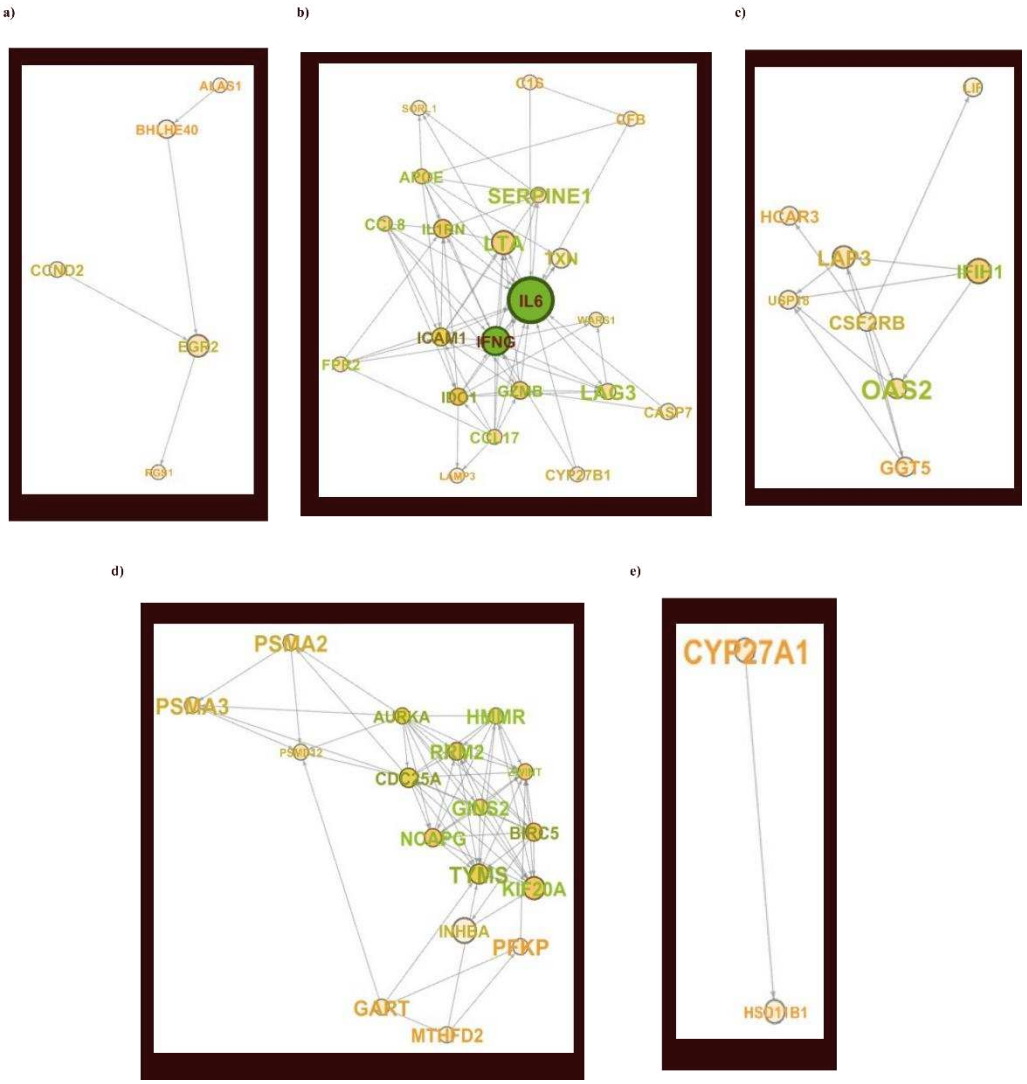


Figure 30 Utilizing Gephi for modulatory analysis, the network exhibits a modular organization with five distinct modules: a) First Module: Consisting of a network of genes with 5 members. b) Second Module: Comprising 21 genes. c) Third Module: Involving 8 genes. d) Fourth Module: Encompassing 17 genes. e) Fifth Module: Comprising 2 members. This modular breakdown reveals the intricate composition of the network, with each module representing a subset of genes that likely collaborate in specific biological processes or pathways.

Importance of Hub Genes in Biological Processes

Utilizing the systems biology server Enricher (Enricher) [189] our analysis reveals the pivotal role of hub genes in orchestrating the regulation of cytokine and chemokine production (See Figure 31). This comprehensive examination underscores the

significance of these genes in influencing key biological processes associated with immune responses and signaling pathways.

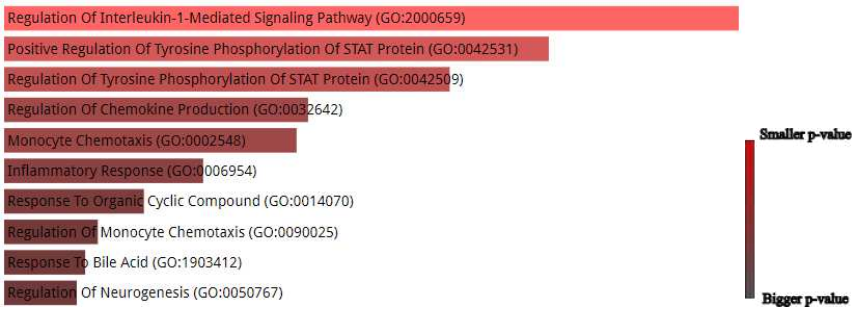


Figure 31 The majority of genes participate in biological processes were related to cytokine and chemokine production, as well as the regulation of inflammatory responses. A pivotal role of STAT proteins is the regulation of cytokine production [221]. *The hierarchical arrangement of the chart is based on p-values.*

The VirusMINT website [222], in connection with Enricher, highlights genes implicated in viral infections like HIV and HPV16, as illustrated in Figure 32.

Figure 32 reveals a network of hub genes intricately linked to viral infection. Notably, the presence of HPV16 emerges as a key factor significantly influencing the expression patterns of these genes..

Our investigation involved examining the DisGeNET database (DisGeNET) [223] and cross-referencing the same genes using the Enricher platform. The results indicated a significant association of these genes with autoimmune diseases, as depicted in the accompanying Figure 33. This comprehensive analysis enhances our understanding of the genetic links to autoimmune conditions.

Figure 33 DisGeNET shows the high relationship between the hub genes and autoimmune diseases such as Rheumatoid Arthritis, Graves, Arteriosclerosis, osteoporosis.

Checking the physicochemical charachteristics of human IL6

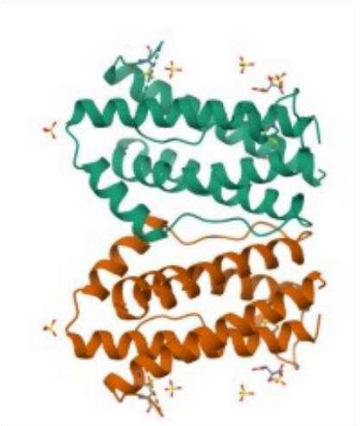
Initially, we identified the Uniprot ID for IL6 as P05231 to ascertain the total number of amino acids. Subsequently, the structural details of IL6 were extracted from the PDB entry 1ALU, revealing a resolution of 1.9 Å and a respectable R-Value work of 0.21. The comprehensive physicochemical structure of IL6 is presented in the accompanying Table 13.

Table 13 Physicochemical structure of the IL6

Protein	**Molecular weight**	**Theoretical pI**	**Instability index**	**Estimated half-life** (mammalian reticulocytes)	**Aliphatic index**	**GRAVY**
IL6	23718.22	6.17	57.70	30 hours	87.50	-0.271

Utilizing Jpred4 (Jpred4)[224] for protein secondary structure prediction revealed a predominant helical structure in most parts of the protein, with a smaller segment adopting a β-sheet conformation, albeit not with high confidence. The Ramachandran analysis exhibited that 95.2% of the protein resides in the most favored regions, with the remaining 4.8% distributed in additional allowed regions and no presence in generally allowed or disallowed regions. Figure 34 illustrates the PDB 2D structure of IL6, accompanied by a secondary structure analysis. This comprehensive analysis provides insights into the prevalent structural motifs and conformational preferences within the protein.

a)

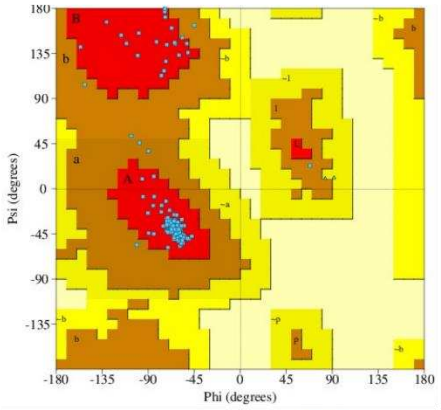


b)

SOPMA :

Alpha helix	(Hh) :	126 is	59.43%
3_{10} helix	(Gg) :	0 is	0.00%
Pi helix	(Ii) :	0 is	0.00%
Beta bridge	(Bb) :	0 is	0.00%
Extended strand	(Ee) :	3 is	1.42%
Beta turn	(Tt) :	3 is	1.42%
Bend region	(Ss) :	0 is	0.00%
Random coil	(Cc) :	80 is	37.74%
Ambiguous states (?)	:	0 is	0.00%
Other states	:	0 is	0.00%

c)



		No. of residues	%-tage
Most favoured regions	[A,B,L]	139	95.2%
Additional allowed regions	[a,b,l,p]	7	4.8%
Generously allowed regions	[~a,~b,~l,~p]	0	0.0%
Disallowed regions	[XX]	0	0.0%
Non-glycine and non-proline residues		146	100.0%
End-residues (excl. Gly and Pro)		4	
Glycine residues		4	
Proline residues		3	
Total number of residues		157	

Figure 34 a)3D structure of the IL-6, b) secondary structure analysis of IL6 using SOPMA server, c) Ramachandran plot analysis of IL6

Identifying Druggable Pockets in IL6 Structures

We utilized the PockDrug website (PockDrug) to predict potential antagonist pockets on the protein structure. Our analysis revealed nine probable sites within the protein structure suitable for drug or ligand interaction. Additionally, we employed Computed Atlas of Surface Topography of Proteins (CASTp) to explore the presence of other pockets.

Figure 35 illustrates four pockets with a high probability of interacting with a drug, as identified by PockDrug, along with the pocket predicted by CASTp.

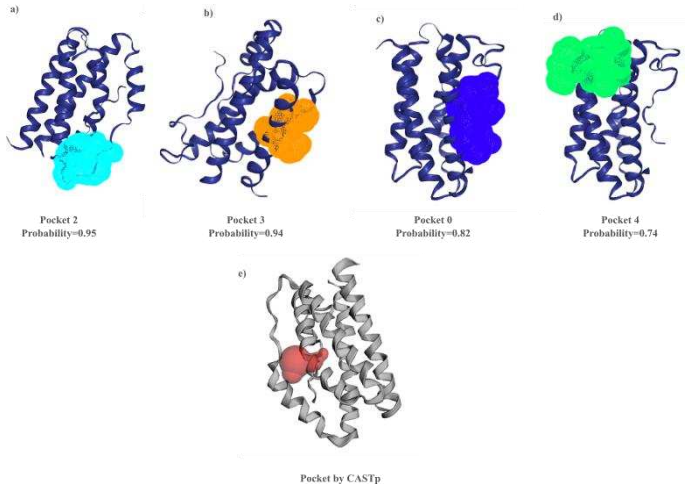


Figure 35 PockDrug and CASTp were employed for predictions, yielding distinct pockets denoted as a, b, c, d, and e. a) Represents pocket 2, exhibiting the highest drugability probability among the predicted pockets. b) Corresponds to the third pocket with a drugability probability of 0.94. c) Represents pocket number 0, demonstrating a drugability probability of 0.82. d) Denotes pocket 4 with a drugability probability of 0.74. e) Represents the pocket predicted by CASTp. Values closer to 1 indicating higher probability.

Exploring Drug Availability for IL6 Targets

To investigate the availability of drugs targeting IL6, we conducted a search on the DrugBank website (DrugBank) [225] focusing on the Target sequence associated with IL6. Our inquiry revealed 12 drugs with potential antigenic, inhibitory, and agonistic effects on the IL6 sequence. A comprehensive presentation of these findings can be found in the accompanying Table 14. Cheking if some drugs are available for IL6.

Table 14 IL6 agonist and antagonist drug

Drug	Drug group	Pharmacological action	Actions	Pubchem ID
Ginseng	investigational, nutraceutical	unknown	antagonist	46508934
YSIL6	investigational	unknown		347909893
VX-702	investigational	unknown		347827732
CRx-139	investigational	unknown		347910206
Andrographolide	investigational	unknown		347827741
Atiprimod	investigational	unknown		-
Siltuximab	approved, investigational	yes	Antagonist antibody	347910394
Dilmapimod	investigational	unknown		347828437
Polaprezinc	experimental	unknown	inhibitor	310265128
Foreskin fibroblast (neonatal)	approved	unknown	agonist	-
Foreskin keratinocyte (neonatal)	approved	yes	agonist	-
Olokizumab	investigational	yes	antagonist	347911425

Next, Binding Data Bank (Binding DB) [194] was mined for substances with Half-maximal inhibitory concentration (IC50) capable of inhibiting human IL6. A dataset of 41 inhibitory substances was extracted in both SDF and TSV formats. These molecules were preserved for molecular computing and descriptor calculation using PaDEL software, contributing to the creation of a robust QSAR model. Further expanding the search, the KEGG database,(KEGG) was consulted to identify novel drug candidates for IL6. Pirfenidone(PubChem ID: 7848646), pomalidomide (PubChem ID: 96025659), and

sirukumab (PubChem ID: 135626800) were discovered through KEGG, and their SMILE structures were generated using Marvin. Chimera was utilized to construct the 3D structures, and OpenBabel facilitated the conversion of these structures into SDF format. To uncover additional ligands for IL6, a quest was undertaken in the ZINC20 database. This search aimed to identify lead-like compounds with descriptors or fingerprints akin to those of known inhibitors from Binding DB or KEGG. By comparing descriptors of known inhibitory compounds with unknown molecules, potential IL6-targeting molecules in our dataset were predicted. The search in ZINC20 yielded a file of lead-like compounds (283 molecules, 1.98 MB in volume) downloaded in SDF format.

This comprehensive investigation leverages multiple databases to identify a spectrum of potential drug candidates for IL6, setting the stage for further analyses in the realm of QSAR modeling.

QSAR Modeling Process

IL6-related SDF files sourced from ZINC, DrugDatabank, and KEGG were processed through PaDEL software for descriptor and fingerprint extraction, resulting in 1443 descriptors. Subsequently, SMLR software was employed to identify high-quality descriptors. Through sparse classification, SMLR identified 36 influential descriptors from the initial pool of 1443, shaping the QSAR model.

This streamlined approach enhances efficiency and reliability, yielding a less complex model. The selected 36 descriptors, along with molecular fingerprints, contribute to the model's effectiveness. The specific descriptors are outlined in the accompanying table 15.

Table 15 Enriched descriptors by SMLR that were used in QSAR model

Type	Descriptor	Description	Descriptor class
2D	apol	Sum of the atomic polarizabilities (including implicit hydrogens)	APol descriptor
2D	ATS0s	Broto-Moreau autocorrelation - lag 0 / weighted by I-state	Auto correlation descripto

Table 15 Enriched descriptors by SMLR that were used in QSAR model

Type	Descriptor	Description	Descriptor class
2D	GATS4e	Geary autocorrelation of lag 4 weighted by Sanderson electronegativity	Auto correlation descriptor
2D	minsssN	Minimum atom-type E-State: >N-	Electrotopological State Atom Type descriptor
2D	nRing	Number of rings	Ring Count descriptor
2D	sumI	Sum of the intrinsic state values I	Electrotopological State Atom Type descriptor
2D	ETA_Eta_B	Branching index EtaB	Extended Topochemical Atom descriptor
2D	AATS2p	Average Broto-Moreau autocorrelation - lag 2 / weighted by polarizabilities	Auto correlation descriptor
2D	AATSC7i	averaged and centered moreau-broto autocorrelation of lag 7 weighted by ionization potential	Auto correlation descriptor
2D	nF8Ring	number of 8-membered fused rings in a molecular structure	Ring Count descriptor
2D	AATSC3i	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by first ionization potential	Auto correlation descriptor
2D	GATS8c	geary coefficient of lag 8 weighted by gasteiger charge	Geary coefficient descriptor
2D	MDEC-44	molecular distance edge between all quaternary carbons	molecular distance edge Descriptor
2D	MLFER_A	Overall or summation solute hydrogen bond acidity	MLFER descriptor
2D	MATS3i	Moran autocorrelation - lag 3 / weighted by first ionization potential	Auto correlation descriptor

Table 15 Enriched descriptors by SMLR that were used in QSAR model

Type	Descriptor	Description	Descriptor class
2D	VE3_Dzp	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by polarizabilities	Barysz Matrix descriptor
2D	minsOm	Minimum atom-type E-State: -O-	Electrotopological State Atom Type descriptor
2D	SpMax3_Bhm	Largest absolute eigenvalue of Burden modified matrix - n 3 / weighted by relative mass	Burden Modified Eigen values descriptor
2D	ALogP	Ghose-Crippen LogKow	ALOGP descriptor
2D	VE3_Dt	Logarithmic coefficient sum of the last eigenvector from detour matrix	Detour Matrix descriptor
2D	GATS2m	Geary autocorrelation - lag 2 / weighted by mass	Auto correlation descriptor
2D	AATS4s	Average Broto-Moreau autocorrelation - lag 4 / weighted by I-state	Auto correlation descriptor
2D	SaasC	Sum of atom-type E-State: :C:-	Electrotopological State Atom Type descriptor
2D	AATSC3p	Average centered Broto-Moreau autocorrelation - lag 3 / weighted by polarizabilities	Auto correlation descriptor
2D	VE3_DzZ	Logarithmic coefficient sum of the last eigenvector from Barysz matrix / weighted by atomic number	Barysz Matrix descriptor
2D	minHCsatu	Minimum atom-type H E-State: HÂ on C sp3 bonded to unsaturated C	Electrotopological State Atom Type descriptor
2D	nwHBd	Count of E-States for weak Hydrogen Bond donors	Electrotopological State Atom Type descriptor

Table 15 Enriched descriptors by SMLR that were used in QSAR model

Type	Descriptor	Description	Descriptor class
2D	GATS3e	Geary autocorrelation - lag 3 / weighted by Sanderson electronegativities	Auto correlation descriptor
2D	nHBint7	Count of E-State descriptors of strength for potential Hydrogen Bonds of path length 7	Electrotopological State Atom Type descriptor
2D	ATSC7c	centered Broto-Moreau autocorrelation of lag 7 weighted by gasteiger charge	Auto correlation descriptor
2D	ATSC2e	Centred Broto-Moreau autocorrelation of lag 2 weighted by Sanderson electronegativity	Auto correlation descriptor
2D	ASP.4	The Average Simple Path of order 4	PaDEL ChiPath descriptor
2D	VR1_Dze	Randic-like eigenvector-based index from Barysz matrix / weighted by Sanderson electronegativities	Barysz Matrix descriptor
2D	Zagreb	Sum of the squares of atom degree over all heavy atoms i	Zagreb Index descriptor
2D	VR1_Dt	Randic-like eigenvector-based index from detour matrix	Detour Matrix descriptor
2D	AATSC2v	Average centered Broto-Moreau autocorrelation - lag 2 / weighted by van der Waals volumes	Auto correlation descriptor

Running QSAR model in Chemoface

In the supervised machine learning process, Chemoface was trained using enriched descriptors generated by SMLR (x-axis) and IC50 values (y-axis) from molecules sourced from the Binding DB. Eight samples were chosen for the training process. The model, depicted in [Figure], predicted the IC50 values for molecules from ZINC, database. The model was developed using the known descriptors (x), generated by SMLR, and IC50 values (y) from the Binding DB dataset.

The data underwent auto-scale pretreatment and were analyzed using the partial least squares (PLS) statistical method, with a maximum of five latent variables (LV=5). The Calibration plot fitted the model based on trained and test data (No. LV=5). The

Prediction plot extended the model to molecules from ZINC database, lacking IC50 values. Linear regression established a correlation line between molecules with known IC50 and predicted values, resulting in the equation y = 0.8086 – 0.2826 and R2 = 0.81.

The results, including the linear regression equation and R2, are provided in Figure 36. Among the molecules, those with IC50 values between 0 and 10 were selected for pharmacophore modeling.

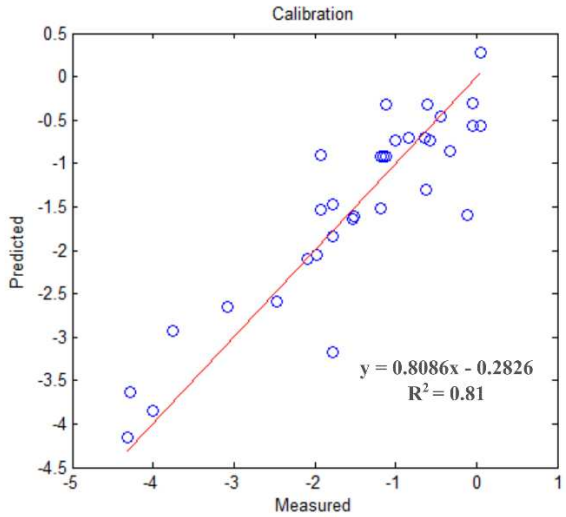


Figure 36 QSAR model shows the correlation between molecular structure and IC50 values for molecules obtained from ZINC20, with a commendable R2 value of 0.81.

Pharmacophore modeling of IL6 and its ligands ligands

The optimization of the structure matching the real ligand of IL6 was achieved using Schrodinger-Maestro. Initially, a reference model structure was constructed based on well-known IL-6 ligands sourced from the Binding DB. In the ensuing pharmacophore model development, three pivotal features were identified: two Rings: The pharmacophore model integrates two rings as essential structural components, enhancing its representation of IL6 ligand interactions. Acceptor: A hydrogen acceptor feature is incorporated into the pharmacophore model, playing a vital role in mediating interactions with other molecular entities. Hydrophobic Part: the pharmacophore model encompasses a hydrophobic feature crucial for shaping the spatial arrangement and understanding molecular interactions within the structure. The spatial orientation of these features is visually depicted in Figure 37 an accompanying figure associated with the

pharmacophore model. It is noteworthy that the matching tolerance for each feature in the model is set at 2 Ångströms (A0).

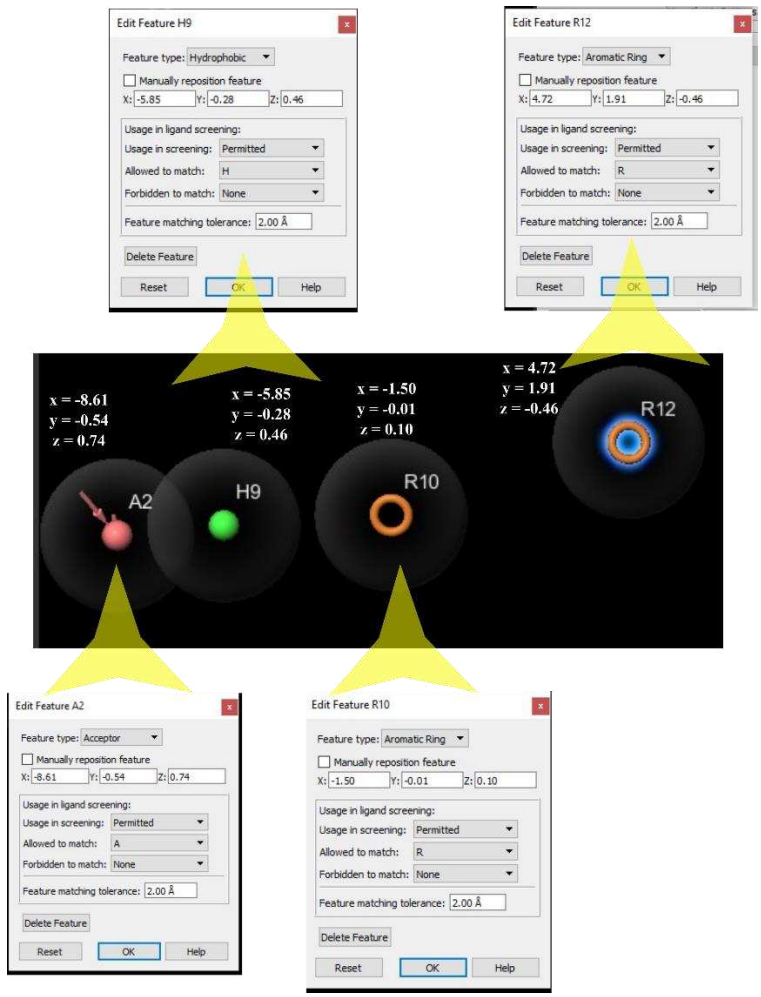


Figure 37 In the developed pharmacophore model, three essential features were identified. Firstly, the model incorporates two rings, serving as vital structural components. Additionally, a hydrogen acceptor feature is integral to the pharmacophore, playing a crucial role in interactions with other molecular entities. Moreover, the pharmacophore model includes a hydrophobic feature that contributes to the spatial arrangement, crucial for understanding molecular interactions. The spatial orientation of these features is visually depicted in an associated picture, and it's noteworthy that the matching tolerance for each feature is set at 2 Å.

In the subsequent analysis, molecules sourced from ZINC20, underwent scrutiny with Maestro to ascertain their alignment with the pharmacophore models. Out of 217 lead-like molecules extracted from ZINC, a singular match was identified with the model. Notably, the matched structure corresponds to 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate, with PubChem CID 816391, as illustrated in the accompanying Figure 38.

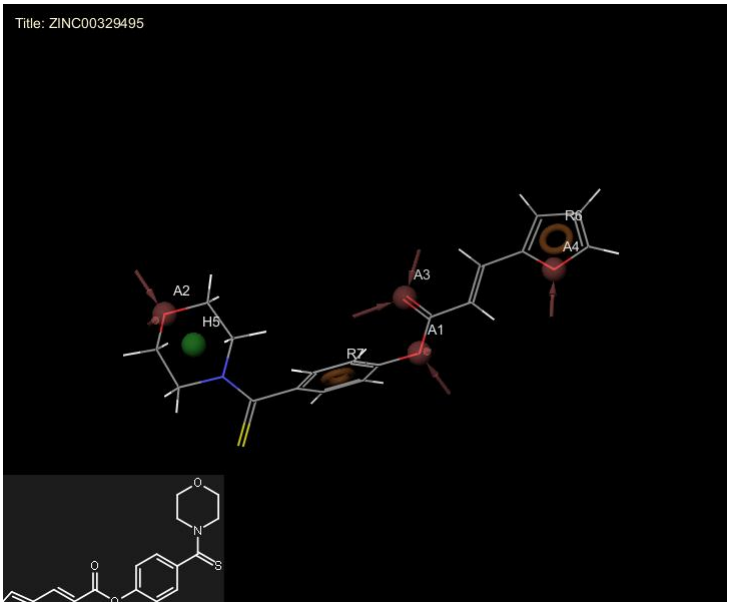


Figure 38 Employing pharmacophore modeling revealed a singular match within the ZINC database, specifically identified as ZINC ID=00329495. This unique structure corresponds to 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate.

Docking results for 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate

To elucidate the interaction between 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate and IL6, the molecular docking process was executed using PyRx software. PyRx identified nine potential binding sites conducive to protein-ligand interaction. A detailed summary of these outcomes is presented in the accompanying Table 17. The table presents various conformations of ligand-protein interactions, with a focus on evaluating the reliability of different models. Best Model Selection Criteria: the first model, characterized by low energy in binding affinity and RMSD, is identified as the best. Criteria for model assessment include RMSD/up, RMSD/down, both equating to zero in the optimal model. Root Mean Square Deviation (RMSD): RMSD measures differences in atomic positions across different protein conformations. Higher RMSD values indicate more significant errors in the model. Reliability Indicator: The reliability of a model is inferred from the difference between upper and lower RMSD. If the RMSD difference is approximately 2, the model is considered reliable. This approach provides a systematic method for selecting the most accurate ligand-protein interaction model based on energy, binding affinity, and RMSD metrics.

Table 17 Different Conformations of Ligand-Protein Interactions.

Ligand-Protein conformations	Binding Affinity	RMSD/up	RMSD/down
1	-5.8	0	0
2	-5.6	24.711	21.925
3	-5.5	8.565	6.041
4	-5.5	33.61	30.616
5	-5.4	25.04	23.407
6	-5.3	32.734	29.238
7	-5.2	8.539	5.023
8	-5.2	10.79	7.6
9	-5.2	32.026	29.342

To unveil the initial conformation of ligand-protein interactions with the lowest affinity energy level and lowest RMSDs, a 2D-QSAR model was employed. Utilizing BIOVIA Studio Visualizer, the interactions were analyzed and visualized (refer to Figure). Interactions Between IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate: Aspartic Acid (ASP) Residue at Position 160 (A160): Forms a Pi Anion bond, indicating interaction between the pi system in the ligand and an anion (ASP). This enhances binding specificity. Lysine (LYS) Residue at Position 46 (A46): Participates in a Pi Alkyl interaction, involving the interaction between the pi system in the ligand and an alkyl group in the amino acid side chain. This contributes to ligand stabilization within the binding site. Tryptophan (TRP) Residue at Position 157 (A157): Forms a PiPi-T shaped interaction, characterized by the stacking of two pi systems, resembling a T shape. This interaction significantly contributes to the overall stability of the ligand-protein complex.

These findings provide a detailed understanding of key interactions, shedding light on the molecular dynamics governing IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2- furyl) acrylate binding, Figure 39.

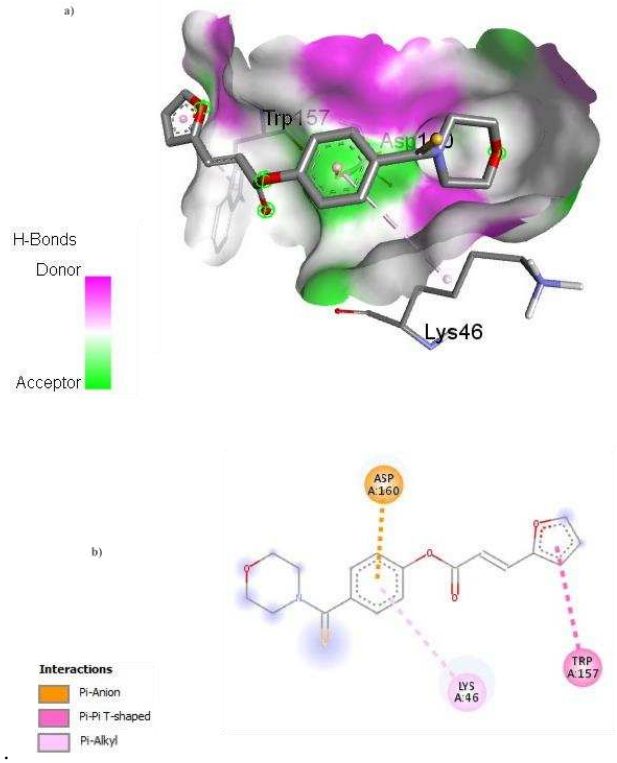


Figure 39 Molecular Interaction Overview, a) 3D Diagram of IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate Interactions Illustrates the spatial arrangement of IL6 and the ligand, showcasing hydrogen donors and involved amino acids in the interaction. b) 2D Diagram of the Same Interaction; Represents a two-dimensional projection of the IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate interaction. Interaction Details: Three interactions occur between amino acids in the IL6 backbone and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate, specifically involving aspartic acid and lysine at R10. The interactions with R10 are facilitated through Pi-Anion and Pi-Alkyl bonds based on our pharmacophore model. Tryptophan exhibits a Pi-Pi-shaped interaction with the ligand at R12. These findings provide a detailed insight into the molecular interactions between IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate, emphasizing key amino acids and bonding patterns.

Molecular Dynamics Analysis

To gain a comprehensive understanding of the docking phenomenon, molecular dynamics simulations were conducted. The Root Mean Square Deviation (RMSD) value for the protein remained stable throughout the molecular dynamics, as depicted in the figure. Although there were fluctuations initially, the system eventually stabilized Figure 40a.

The Root Mean Square Fluctuation (RMSF) exhibited an increasing trend over time, indicating the flexible movement of 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate during the simulation, as illustrated in the Figure 40b.

Hydrogen bonds emerged as pivotal contributors to the interaction between IL6 and 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate. The figure demonstrates that a maximum of two hydrogen bonds formed during the molecular dynamics simulation Figure 40c.

The Radius of Gyration (Rg) analysis was employed to assess the compactness of IL6 in the presence of 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate. The results, presented in the figure, provide insights into the structural compactness of IL6 during the molecular dynamics simulation Figure 40d.

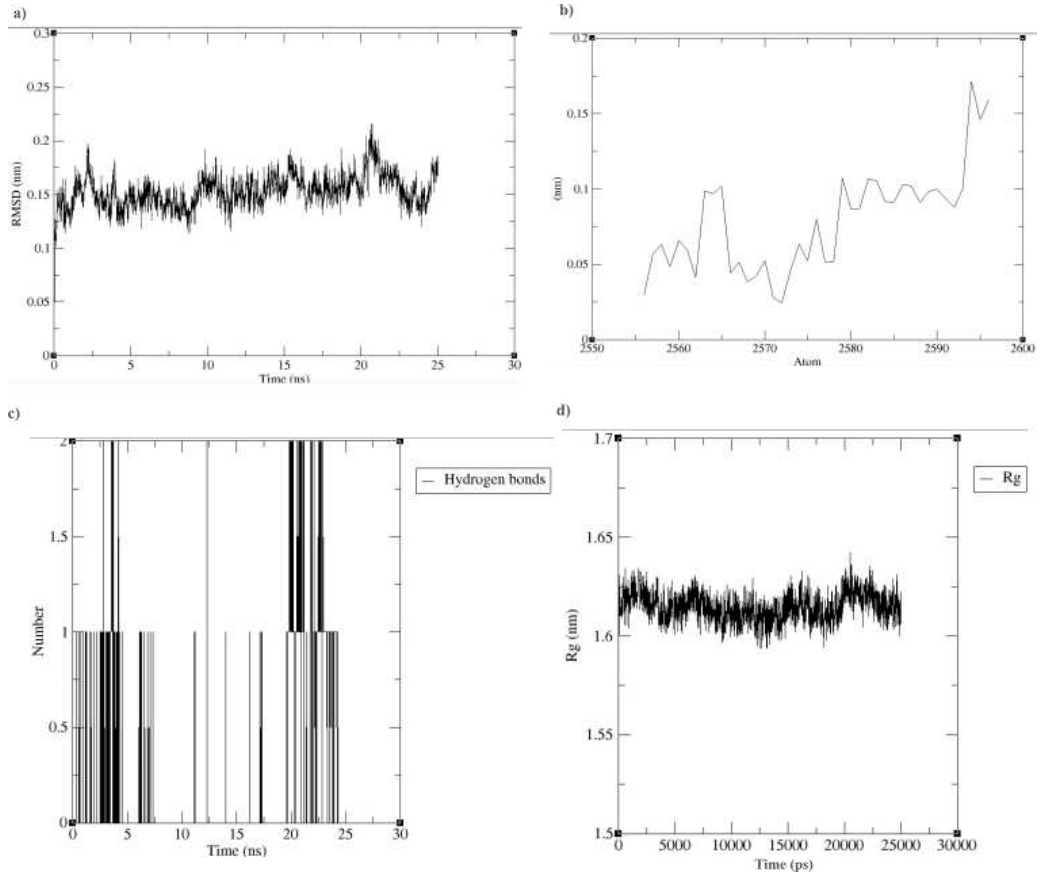


Figure 1 Molecular Dynamics Analysis of IL6 Interaction with 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate. a) RMSD of IL6 Backbone: The graph illustrates the Root Mean Square Deviation (RMSD) of the IL6 backbone while interacting with 4-(4-morpholinylcarbothioyl) phenyl 3-(2-furyl) acrylate. b) Ligand RMSF and Flexibility: Representing the fluctuation (RMSF) of the ligand during the interaction, indicating its flexible movement during Molecular Dynamics (MD) simulation. c) Hydrogen Bonds Dynamics: The graph displays the number of hydrogen bonds formed over time in nanoseconds (ns) during the MD simulation. d) Radius of Gyration (Rg) of IL6:Illustrating the Rg of IL6 around its axis versus time in picoseconds (ps) during the MD simulation. The stability of IL6 Rg is depicted in this graph.

Assessing Drug Likeness of 4-(4-Morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate Using SwissADME

The physicochemical profile of 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate was systematically evaluated through SwissADME (http://www.swissadme.ch/) utilizing its SMILES representation. SwissADME, a web-based computational tool for small molecule analysis in drug discovery [226]. encompasses diverse functionalities for assessing physicochemical attributes, predicting pharmacokinetic parameters, gauging drug-likeness, and scrutinizing other pivotal factors in drug development. Upon navigating to the SwissADME platform, 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate was submitted via the Simplified Molecular Input Line Entry System (SMILES) notation, or the option to upload a file containing its chemical structure. The computational process was initiated by clicking the "Run" button, leading to the generation of results as illustrated in Figure 41.

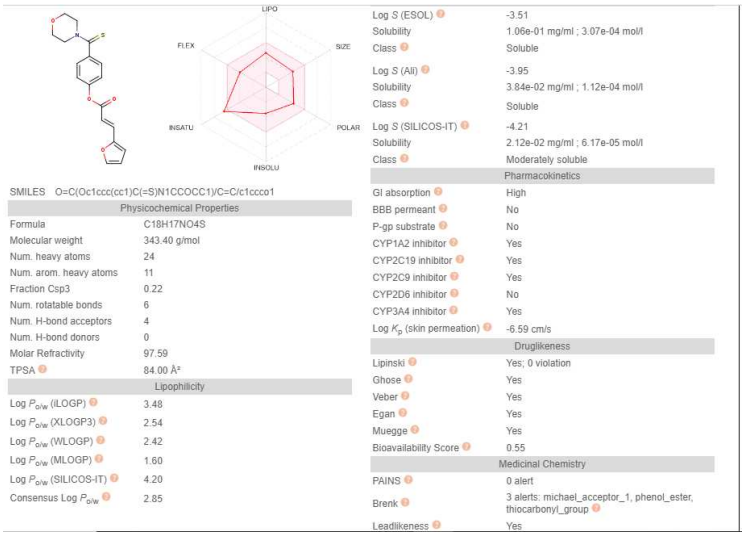


Figure 41 Physicochemical Characteristics of 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate Analyzed with SwissADME.

Chapter 7: Discussion

Discussion

Hypothesis

How robust is the T cell immune response, encompassing activation and proliferation, specifically targeting HPV 16 antigens (E6, E7, and L1) within both CD8 and CD4 T cell subsets? Furthermore, is there evidence of cross-reactivity in patients with different HPV species, beyond HPV16, where T cells can be activated or proliferated by HPV16 and HPV18 antigens? Assessing the effectiveness of conventional vaccines, namely the bivalent Cervarix and the quadrivalent/nine-valent Gardasil, is a key aspect. Additionally, this study aims to explore avenues for vaccine enhancement.

The investigation involves a detailed analysis of T cell activation and proliferation using AIM (Activation Induced Marker) and LPA (Lymphocyte Proliferation Assay) assays among both HPV-vaccinated, non-vaccinated participants, and patients with different stages of HPV induced cervical cancer. Specifically, the study delves into a distinct subset of CD4 T cells, namely follicular T cells expressing CXCR5, recognized for their critical role in B cell antibody production and maturation. The primary functions of T cells expressing CXCR5, known as T follicular helper (Tfh) cells, include providing essential help to B cells, leading to class-switching, secretion of antigen-specific antibodies, and the formation of immunological memory [227].

The research incorporates experimental analyses coupled with immunoinformatics to design a multi-epitope vaccine for HPV. A comprehensive examination of system immunology involves assessing gene expression in peripheral blood mononuclear cells (PBMCs) before and after vaccination. The final objective is to employ network analysis to identify a key gene that could potentially enhance vaccine efficacy. Moreover, if feasible, the study aims to discover or design a drug, utilizing QSAR (Quantitative Structure-Activity Relationship) and pharmacophore modeling, to modulate the identified key gene—either by blocking or stimulating it.

This multifaceted approach seeks to advance our understanding of HPV immunity, optimize current vaccines, and explore novel strategies for more effective preventive measures.

Objectives:

1. Assess CD8 and CD4 T Cell Responses: Analyze the activation and proliferation of CD8 and CD4 T cells against HPV 16 antigens (E6, E7, and L1) using AIM and LPA assays in HPV-vaccinated, non-vaccinated, and HPV related cervical cancer individuals.

2. Explore Follicular T Cell Dynamics: Investigate the role of follicular T cells expressing CXCR5 in the immune response against HPV antigens, considering their influence on B cell antibody production and maturation.

3. Analyze HPV Clade Differences: Utilize sequence alignment and phylogenetic tree analysis to determine the genetic distances between different HPV clades.

4. Design Immunoinformatics-Based Vaccine: Explore the feasibility of designing a multiple epitope vaccine using immunoinformatics, capable of providing broad-spectrum immunity against diverse HPV infections, both pre- and post-infection.

5. Examine Gene Expression Changes: Utilize systems immunology and systems biology to investigate differences in gene expression levels in PBMCs before and after vaccination, identifying genes with altered expression.

6. Assess Drugability of Target Gene: Investigate the drugability of the target gene and protein associated with the immune response against HPV, followed by screening potential chemical compounds using ZINC, Binding DB, and Drug Databank.

7. Utilize Computational Models for Drug Design: Apply pharmacophore modeling, QSAR, 2D-QSAR, molecular dynamics, and docking to evaluate the binding efficacy of identified ligands to the target protein.

This structured approach aims to comprehensively understand T cell immunity against HPV, explore vaccine design possibilities, and assess potential therapeutic interventions at the genetic and molecular levels.

Key findings

1. **Vaccination Regimen Optimization:** Our study underscores the significance of optimizing vaccination effectiveness by considering the number of injections, shedding light on the crucial role of dosing in achieving optimal immunity against HPV.

2. **Antigen Selection Rationale:** The exclusive use of the L1 antigen in common HPV vaccines (e.g., Cervarix and Gardasil) is justified by our identification of diverse phylogenetic variations across different HPV clades, emphasizing the need for a universally effective vaccine targeting the common denominator, L1.

3. **Representative Role of HPV 16 and 18:** Through phylogenetic analysis, we established that HPV 16 and 18 act as representatives of larger groups of high-risk HPVs, streamlining research efforts toward these pivotal strains for broader implications.

4. **Antigenic Focus Refinement:** Contrary to conventional wisdom, we found moderate L1 antigenicity for HPV16 and HPV18. Our discovery of higher antigenicity in E6, E7, and L2 suggests a refined focus for vaccine development, optimizing immune responses.

5. **Multi-Epitope Vaccine Development:** Introducing a multiple epitope vaccine targeting E6, E7, and L1 antigens specific to HPV16 and HPV18 demonstrates a sophisticated approach toward enhancing vaccine efficacy.

6. **Broad Immune Stimulation:** Achieving over 95% population coverage, our vaccine not only stimulates CD4 and CD8 T cells but also induces B cell activation, resulting in antibody production and a comprehensive immune response.

7. **Immune System Simulation Insights:** Our simulations reveal an increase in antiviral cytokines such as IFN-γ and IL2, accompanied by population growth in T cells, B cells, and Macrophages, providing a dynamic understanding of the immune response.

8. **Gene Expression Dynamics:** Studying gene expressions in PBMCs unveils a nuanced picture, indicating an increase in anti-viral cytokine IFN-γ alongside a rise in proinflammatory cytokine IL6 post-vaccination, highlighting the complexity of immune modulation.

9. **Compound Screening for Immunomodulation:** Employing advanced screening procedures, we identified 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate as a potential attenuator of IL6-induced adverse effects, offering a promising avenue for immunomodulation.

10. **Molecular Dynamics Insights:** Utilizing molecular dynamics with Gromacs, we elucidated the intricate interactions between IL6 and 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate, providing molecular-level insights into the mechanism of attenuation.

These findings collectively advance our understanding of HPV vaccination strategies, paving the way for tailored, effective interventions against HPV-associated diseases.

AIM Assay Results for CD4/CD8 T Cells and HPV-16 Antigens

CD4/CD8 T Cell Activation for HPV16-E6:

No significant difference in CD4 and CD8 T cell activation between vaccinated cases and non-vaccinated controls for HPV-E6.

CD4/CD8 T Cell Activation for HPV16-E7:

Only CD4 T cells exhibit significant activation for HPV16-E7, emphasizing a selective response. Statistical analysis: Th-E7 p-value: 0.039058, CI95%: [-0.26, -0.01], Cohen-d: 2.248314, BF10: 2.96.

CD4/CD8 T Cell Activation for HPV16-L1:

No significant difference in CD4 and CD8 T cell activation for HPV`6-L1.

Findings underscore selective CD4 T cell activation in response to HPV16-E7, indicating nuanced T cell responses to specific antigens. The observed T cell activation for HPV-16 in vaccinated participants compared to non-vaccinated individuals despite the availability of only the L1 antigen in the vaccine may be attributed to cross-reactivity or bystander effects. Several factors could contribute to this phenomenon. Cross-reactivity occurs when T cells, primed by the L1 antigen in the vaccine, exhibit responses to other antigens, potentially including non-L1 antigens from HPV-16. This cross-reactivity might lead to T cell activation even against antigens not directly targeted by the vaccine [228]. Additionally, bystander activation refers to the non-specific activation of T cells in response to inflammatory signals. The vaccine-induced immune response could create an inflammatory environment, leading to the activation of T cells against various HPV-16 antigens, irrespective of the specific antigen composition in the vaccine [229]. Also, memory T cells generated by the vaccine may possess a broader specificity, allowing them to recognize and respond to a wider range of antigens, including those not included in the vaccine formulation [230]. Furthermore, based on our homology analysis, activation of T cells in HPV-vaccinated participants might be attributed to the homology between HPV16-E7 and *Listeria monocytogenes* proteins. In this case, studies have utilized *L. monocytogenes* as a vaccine vector expressing HPV16-E7, inducing T cell responses [231–233].

Lymphocyte proliferation Assay Results for CD4/CD8 T Cells and HPV-16 Antigens

CD4 T Cell Activity Against HPV16-E6 Antigen

The study shifted focus to Lymphocyte Proliferation Assay (LPA) analysis to investigate CD4 T cell activity against the HPV16-E6 antigen in a diverse participant pool, including vaccinated, non-vaccinated individuals, and patients in early cervical cancer stages.

In the examination of CD4 T cell proliferation among participant groups, the absence of statistical significance in both ANOVA and Tukey HSD implies a lack of discernible differences. Although modest to moderate effect sizes suggest potential trends, these remain insufficient to achieve statistical significance. In the context of statistical analysis, "effect size" refers to the magnitude of the difference observed between groups. When

we say "modest to moderate effect sizes suggest potential trends," it means there are observable differences between groups, but these differences are not large. They fall within a moderate range. However, the statement continues to mention that these effect sizes are "insufficient to achieve statistical significance." This implies that, despite the presence of trends or differences, the observed effects are not strong enough to be considered statistically significant. In statistical terms, achieving significance means demonstrating that the observed results are unlikely to occur by random chance.

So, in simpler terms, even though there are some trends or differences that we can see (modest to moderate effect sizes), these differences are not strong or consistent enough to confidently say that they are not due to random variability. Statistical significance requires a higher level of certainty in the observed patterns, and in this case, the observed trends fall short of reaching that level of certainty [234–236].

The interpretation of meaningful changes is guided by Minimal Detectable Change (MDC) values, underscoring the necessity of acknowledging practical significance. The study's low statistical power accentuates limitations in detecting differences, necessitating caution in result interpretation. These collective findings emphasize the importance of a nuanced approach, wherein both statistical and practical significance are considered within the context of the study's inherent limitations. Minimal Detectable Change (MDC) is a statistical measure used to determine the smallest change in a variable that can be considered meaningful or significant. In the context of CD4 proliferation against the E6 antigen: MDC values serve as a benchmark for identifying changes that go beyond random variability [237]. In fact, it represents the minimum change needed in CD4 proliferation to confidently say that the observed difference is not due to measurement error alone. The MDC for CD4 proliferation-E6 falls within the range of 9.22 to 13.94. This range indicates the minimum magnitude of change in CD4 proliferation that can be considered significant. A change in CD4 proliferation below 9.22 may not be distinguishable from measurement variability and might be considered within the normal fluctuation. Changes between 9.22 and 13.94 suggest a potential meaningful shift in CD4 proliferatin. Values beyond 13.94 provide stronger evidence of a substantial and noteworthy change. Researchers and practitioners should consider MDC values when interpreting study results. Values within or beyond the MDC range influence decisions about the practical significance of observed changes in CD4 proliferation.

CD4 T Cell Proliferation Against HPV16-E7 Antigen

In this section of the Discussion chapter, the analysis of CD4 T cell proliferation against the E7 antigen is presented. The study aimed to explore potential differences among different participant groups, including non-vaccinated individuals, vaccinated participants, and patients in the first stage of HPV-related cervical cancer. Although the *p-value* is greater than the conventional significance level (0.05), suggesting no statistically significant difference among groups, further analyses were conducted. Also, the *p-adj values* indicate that none of the pairwise comparisons reached statistical significance. To understand the practical significance of the observed differences,

standardized mean differences (SMD) were calculated along with 95% confidence intervals (CI). These effect size measures suggest minimal to moderate effects without reaching statistical significance.

The study's statistical power was calculated at 0.072, indicating a relatively low power to detect differences. The required sample size for adequate power was estimated to be 900 participants. This suggests that the study may have been underpowered to detect significant differences, potentially influencing the outcomes. Further considerations regarding sample size and study design are warranted for future research in this domain.

In the context of the provided information, "Overall MDC" refers to the Overall Minimal Detectable Change. This value, measured for different CD4 T cell activities, indicates the smallest change in the measured variable that is considered significant. The range of MDC was 13-18.7, so CD4 T cell proliferation against HPV16-E7 fall below the Overall MDC, it suggests that these changes may not be practically significant or distinguishable from random variability. On the other hand, if changes exceed the Overall MDC, it implies that the observed differences are likely meaningful and not merely due to chance.

CD4 T Cell proliferation Against HPV16-L1 Antigen

The presented data describes an investigation into CD4 T cell proliferation against the HPV16-L1 antigen in different participant groups, including non-vaccinated individuals (non-vacs), vaccinated participants (vacs), and patients with the first stage of HPV-related cervical cancer.

The ANOVA results suggest that there is no statistically significant difference among the groups, as the *p-value* exceeds the conventional significance level of 0.05. However, further analyses were conducted to explore specific pairwise comparisons using the Tukey HSD test. None of the pairwise comparisons (non-vacs vs. patient, non-vacs vs. vacs, patient vs. vacs) reached statistical significance, as indicated by the *p-adj values*.

However, in the case of SMD and Cohen’s d interpretation ranges, CD4 proliferation-L1-vacs-nonvacs was -1.48 (large Effect): This value indicates the extent of the difference in CD4 T cell proliferation between vaccinated and non-vaccinated individuals concerning the L1 antigen. The negative sign implies a decrease in CD4 proliferation in non-vaccinated participants compared to vaccinated, suggesting a large effect size. SMD for CD4 proliferation-L1-vacs-patients indicates moderate effect size (-0.63), that implies increase CD4 proliferation in vaccinated participants compared to patients. A reduced in CD4 proliferation in non-vaccinated individuals compared to patients, with a moderate effect size (0.53).

The MDC values (ranging from 24.63 to 30.51) provide insights into the magnitude of differences that are practically significant. These values represent the amount of change needed to exceed measurement error based on a predetermined confidence level. In this context, they offer a threshold for determining whether observed changes in CD4 T cell proliferation for HPV16-L1 are practically meaningful.

CD8 T Cell proliferation Against HPV16-E6 Antigen

The presented results indicate the following findings in the comparative analysis of CD8 T cell proliferatin against the HPV E6 antigen. The non-significant p-value (> 0.05) suggests no statistically significant differences in CD8 T cell proliferatin among participant groups. The effect size (Partial Eta Squared) indicates a small effect, contributing to 9.73% of the observed variance. In the case of pairwise comparison, none reached statistical significance (p-adj > 0.05). This aligns with the ANOVA result, further supporting no significant pairwise differences. A statistical power of 0.074 indicates a low probability of detecting true differences. The study is underpowered, suggesting a limitation in its ability to identify significant effects. The low power underscores the need for caution in interpreting the study findings. A larger sample size is recommended to enhance the study's ability to detect meaningful differences and increase the reliability of the results. To achieve adequate statistical power, the study would require a larger sample size of 831 participants. The current sample size may be insufficient to draw robust conclusions or detect existing differences effectively.

In the case of effect sizes, CD8 T cells proliferation in patient shows moderately more, 0.57 and 0.63, in compare to vaccinated and non-vaccinated respectively. CD8 proliferation-E6-vacs-patients: SMD = 0.5709 (Moderate Effect): An SMD of 0.5709 suggests a moderate effect size. This indicates a more substantial difference in CD8 T cell proliferation against the E6 antigen between vaccinated participants and patients with the first stage of HPV-related cervical cancer. CD8 proliferation-E6-nonvacs-patients: SMD = 0.6317 (Moderate Effect): An SMD of 0.6317 indicates a moderate effect size, similar to the previous case. This implies a moderate difference in CD8 T cell proliferation against the E6 antigen between non-vaccinated participants and patients with HPV-related cervical cancer.

Overall MDC values: Small to moderate differences with practical significance, as indicated by values ranging from 1.96 to 4.71: Interestingly, the Minimal Detectable Change (MDC) values provide a range (1.96 to 4.71) representing small to moderate differences in CD8 T cell proliferation against the E6 antigen. This suggests that observed changes in the study are practically significant and not merely due to measurement variability.

CD8 T Cell proliferation Against HPV16-E7 Antigen

The ANOVA results suggest that there are no statistically significant differences in CD8 T cell proliferation against the HPV E7 antigen among participant groups. The p-value exceeds the typical significance threshold of 0.05, indicating a lack of significant group variation. The Tukey HSD test compared group pairs to explore differences further, and none of the pairwise comparisons reached statistical significance (p-adj > 0.05), supporting the ANOVA findings. The low statistical power (below 0.8) indicates a limitation in the study's ability to detect differences. A larger sample size (1308 participants) is recommended for robust conclusions.

In the context of effect size, CD8 proliferation-E7-vacs-nonvacs: SMD = -0.1043 (Small Effect), CD8 proliferation-E7-vacs-patients: SMD = -0.5999 (Moderate Effect), and CD8 proliferation-E7-nonvacs-patients: SMD = -0.5912 (Moderate Effect) were seen. These values suggest small to moderate differences in CD8 T cell proliferation against the E7 antigen, with the most notable difference observed between vaccinated and non-vaccinated participants. Also, overall MDCs represent the Minimal Detectable Change, indicating the smallest change that can be considered significant. They range from 2.69 to 3.76, suggesting practical significance in observed differences.

CD8 T Cell proliferation Against HPV16-L1 Antigen

The ANOVA results suggest no statistically significant differences in CD8 T cell proliferation against the HPV L1 antigen among participant groups. Although the p-value is slightly above the conventional significance level (0.05), the effect size (np2) indicates a moderate influence of the group variable on CD8 T cell proliferation against the L1 antigen. The Tukey HSD test compared group pairs, none of the pairwise comparisons reached statistical significance (p-adj > 0.05), confirming the ANOVA results. The low-to-moderate, 0.198, statistical power suggests limitations in detecting differences. Increasing the sample size to 143 participants is recommended for improved statistical power.

SMD, quantify the magnitude of differences between groups. The values provide insights into the practical significance of observed effects. In this context: CD8 proliferation-L1-vacs-nonvacs: SMD = -1.2840 (Large Effect): This indicates a substantial and practically significant difference in CD8 T cell proliferation against the L1 antigen between vaccinated (vacs) and non-vaccinated (nonvacs) participants. CD8 proliferation-L1-vacs-patients: SMD = -0.8037 (large Effect): A large effect signifies a meaningful difference in CD8 T cell proliferation between vaccinated participants and those with HPV-related cervical cancer (patients). CD8 proliferation-L1-nonvacs-patients: SMD = 0.5847 (moderate Effect): A small effect suggests a less pronounced but still noteworthy difference in CD8 T cell proliferation between non-vaccinated participants and those with cervical cancer.

These SMD values help interpret the practical significance of the observed variations in CD8 T cell proliferation among different participant groups.

Minimal Detectable Change (MDC) values represent the smallest change in a measurement that is considered significant or meaningful. In this context, CD8 proliferation-L1-vacs-nonvacs, overall MDC: 19.52: This indicates that changes in CD8 T cell proliferation against the L1 antigen in the range of 19.52 are practically significant when comparing vaccinated and non-vaccinated participants. CD8 proliferation-L1-vacs-patients, overall MDC: 18.13: For comparisons between vaccinated participants and those with cervical cancer, changes in CD8 T cell proliferation exceeding 18.13 are considered practically significant. CD8 proliferation-L1-nonvacs-patients, overall MDC: 6.43: Changes in CD8 T cell proliferation beyond 6.43 are considered practically significant when comparing non-vaccinated participants with those having cervical cancer.

Interpretation of overall SMD Confidence Intervals

The wide SMD confidence intervals for CD4 and CD8 activation against HPV16 antigens (E6, E7 and L1) indicate substantial uncertainty in effect estimates, making it challenging to precisely determine the impact of vaccination and patient status on CD4 and CD8 activation. These results may indicate that participants may exhibit diverse responses to HPV vaccination due to individual differences in immune system functioning, genetics, or health status. This variability is reflected in the broad confidence intervals [238]. Also, this can be influenced by factors such as assay sensitivity and specificity. The wide intervals underscore the need for further research to elucidate the factors influencing T cell activation against HPV antigens. Larger sample sizes and more comprehensive studies could enhance precision and reliability in effect size estimation [239].

Comparing the fold increase of T cells proliferation with MDC

The fold increases in T cell proliferation compared to actin, along with the Minimal Detectable Change (MDC) values, provide insights into the practical significance of these changes. In summary, TH1-L1, TFH1-L1, and TFH1-E7 show practically significant changes in T cell proliferation, surpassing their respective MDC values. The other cases might not reach practical significance based on the MDCs. TH1 cells are involved in cellular immune responses, while TFH1 cells play a role in supporting B cell responses, crucial for antibody production. The changes in these subsets likely reflect the adaptive immune system's recognition and response to HPV antigens introduced by the vaccines [240].

The logistic regression analysis provides insights into the impact of vaccination parameters on the fold increase in T cell proliferation. The logistic regression underscores the substantial impact of the intercept, number of vaccine doses, and age on T cell activation. In this scenario, as age increases by one unit, the odds of T cell activation decrease. The other two factors have positive on T cells proliferation. Interestingly, the brand name of the vaccine does not significantly contribute to variations in T cell proliferation fold increase, emphasizing the robustness of T cell response across different vaccine formulations. Research suggests that both age and the number of HPV vaccine doses have a significant impact on T cell proliferation in response to vaccination. The age of the recipient has been shown to significantly influence the generation of HPV-specific T cell responses. Specifically, the age of the recipient can impact the proliferation of T cells in response to HPV vaccination. Younger individuals may exhibit a more robust T cell response, emphasizing the importance of early vaccination to enhance immune memory [241]. Additionally, Studies have demonstrated that the number of doses differentially impacts human B and T cell immune memory responses to HPV vaccination. Adjustments in the vaccination schedule, such as the timing and number of doses, may influence the magnitude and durability of T cell responses [242].

Lack of Observable Changes in T Cells Proliferation against HPV16 Antigens

Thus, the absence of detectable changes in proliferation of CD8 and CD4 T cells against HPV16 E6, E7, and L1 antigens in vaccinated, non-vaccinated, and patients with first-stage HPV-related cervical cancer could be influenced by several factors:

1. **Variability in Immune Responses**: Individuals may exhibit variability in their immune responses to HPV vaccination and infection. Factors such as genetics, overall health, and immune system status contribute to the diversity in T cell reactions [243].

2. **Timing of Analysis**: The time at which T cell responses are assessed is crucial. If the analysis occurs too early or too late after vaccination or infection, the changes in T cell proliferation might not be captured. The kinetics of immune responses vary among individuals[242].

3. **HPV Antigen Specificity**: T cell responses can be antigen-specific. If the antigens assessed (E6, E7, and L1) are not the primary targets of the immune response or if other antigens are more dominant, changes in CD8 and CD4 T cells may not be apparent. E6 and E7 oncoproteins are considered primary targets for T cell responses because they are expressed in HPV-associated cancers. T cells, especially cytotoxic T lymphocytes (CTLs), recognize and eliminate cells expressing these oncoproteins, contributing to anti-tumor immunity. HPV16 E6/E7-specific T cells demonstrate cytotoxic proliferation against antigen-positive targets, indicating their role in directly targeting and eliminating infected or transformed cells [244]. While L1 is a structural protein involved in the formation of viral capsids, it may not be as prominent a target for T cell responses as E6 and E7, which are more directly linked to viral oncogenesis. Studies investigating T cell responses to multiple HPV16 proteins have shown that E6 and E7 are the primary focus, with L1 potentially playing a more limited role in the immune response [44]. Additionally, in the context of antigen presentation, E6 and E7 oncoproteins, being viral oncoproteins associated with HPV-driven cancers, are effectively presented by antigen presenting cells (APCs). Antigen presentation of E6 and E7 by APCs leads to the activation of both CD4+ and CD8+ T cells, contributing to a robust immune response against HPV-infected or transformed cells [245]. In contrast, L1, being a structural protein involved in the formation of viral capsids, may have a role in antigen presentation, but it might not be as prominent as E6 and E7. Unlike E6 and E7, L1 has less direct involvement in HPV-driven oncogenesis. However, it may still contribute to the overall immune response against HPV16 [108]. However, it should be considered that HPV has evolved immune evasion mechanisms, and certain viral proteins like E6 and E7 are known to interfere with immune recognition. If these proteins successfully evade the immune response, it can hinder the activation of T cells specific to these antigens [246].

4. **Individual Immune Status**: Pre-existing immunity or immunosuppression in certain individuals can affect the magnitude of T cell responses. Patients with first-stage HPV-related cervical cancer might have heterogeneous immune profiles.

5. **Methodological Considerations**: The techniques used to measure T cell responses, such as assays or detection methods, can influence the sensitivity and specificity of the results. Inaccuracies in measurement may lead to the failure to observe changes. The sensitivity of LPA can be influenced by the background proliferation in unstimulated cells. Performing a comparative analysis, such as comparing stimulated vs. unstimulated conditions, aids in discerning specific T cell proliferation. Additionally, the choice of mitogens or antigens used to stimulate T cell proliferation is critical. If the selected stimuli do not effectively engage T cells, it may result in reduced sensitivity. Utilizing antigens relevant to the specific T cell population of interest can enhance sensitivity [247].

6. **Sample Size and Population Diversity**: Insufficient sample size or a lack of diversity in the study population may limit the generalizability of findings. Larger and more diverse cohorts may be needed to capture subtle variations [242].

Understanding the intricacies of immune responses to HPV requires comprehensive and context-specific investigations. It may be beneficial to consider a combination of immunological assays, longitudinal studies, and a diverse study population to better elucidate the factors influencing T cell proliferation in response to HPV.

Phylogenetic Analysis of HPV Species

Phylogenetic analysis of E6, E7, and L1 antigens across high-risk, low-risk, probably high-risk, and unknown-risk HPV species reveals extensive diversity. E6 spans 6 clades, E7 encompasses 7, and L1 branches into 9 distinct clades. This complexity poses a challenge for designing a multiple epitope vaccine, as it would necessitate incorporating antigens from numerous clades, resulting in a large and intricate peptide sequence. Phylogenetic diversity refers to the genetic variability and evolutionary relationships among different strains or types of a virus, in this case, Human Papillomavirus (HPV). Additionally, HPV is a diverse group of viruses with various species or types. Each species has distinct genetic characteristics and includes multiple strains [248]. So, the varied genetic composition of these antigens across different species suggests a complex evolutionary history for HPV. Consequently, the diverse antigens may contribute to differences in how the virus interacts with host cells, evades the immune system, and causes diseases [249].

Classifying HPV species based on risk (low, high, probably high, and unknown) and analyzing common ancestry reveals interesting patterns. Lower clade diversity is observed as the number of species decreases. Notably, the unknown-risk category

exhibits the highest diversity, with all species belonging to different clades. In the probably high-risk category, challenges arise in finding a common ancestor for E6, while E7 and L1 show categorization into 4 and 3 clades, respectively. High-risk species, fortunately, exhibit more manageable diversity, with L1 categorized into 2 clades, and E6 and E7 into 3 clades each. The data demonstrates the correlation between how HPV species are classified based on risk (low, high, probably high, and unknown), the genetic diversity within clades, and the total number of HPV species. The unknown-risk category is notable for having high genetic diversity. This means that within this category, different HPV species exhibit significant genetic variations, indicating a diverse evolutionary landscape. So, it is difficult to design a vaccine for this class of HPV. In contrast, high-risk species demonstrate a more organized genetic structure. This suggests that within the high-risk category, there is a discernible pattern or organization in the genetic makeup of the HPV species [250]. Probably high-risk HPV class is a good candidate for designing a multiple epitope vaccine.

The fortunate alignment of HPV16 and HPV18 with clades containing a larger number of species in the high-risk group provides a strategic advantage. Developing a multiple epitope vaccine targeting these specific clades could enhance coverage across a significant portion of high-risk HPV species. The genetic alignment of HPV16 and HPV18 within clades that encompass a higher number of species in the high-risk group is a fortuitous occurrence observed in several studies. For instance, research on the rational design of peptide vaccines against multiple HPV types, including HPV 16 and 18, has identified specific peptides within the L1 capsid protein sequences, highlighting the potential for targeted vaccine development [251]. Additionally, studies focusing on the development of HPV16, 18, 31, and 45 E5 and E7 peptides acknowledge the prevalence of these high-risk HPV types, further emphasizing the importance of strategic vaccine design [252]. In silico and in vivo analyses of high-risk papillomavirus L1 and the use of a polytope DNA vaccine containing multiple T-cell and B-cell epitopes also support the idea of a strategic advantage in targeting specific clades for comprehensive vaccine coverage [253]. Therefore, the alignment of HPV16 and HPV18 with species-rich clades is substantiated by various studies, providing a compelling rationale for the development of a multiple epitope vaccine tailored to these specific genetic configurations.

Despite the lower diversity in L1 compared to E6 and E7, the inherent genetic variability in HPV, driven by factors like point mutations and amino acid substitutions, underscores the challenges in vaccine design. Geographical differences in genetic variations among HR-HPV types may impact biological functions, potentially influencing clinical outcomes. The challenges in vaccine design arise from the need to account for this inherent diversity, especially in critical regions like E6 and E7, to ensure broad efficacy against different viral strains. Furthermore, geographical differences in genetic variations among high-risk HPV types have been identified, indicating potential implications for biological functions that may, in turn, influence clinical outcomes [254]. These findings highlight the intricate interplay between genetic diversity, geographical factors, and the

design of HPV vaccines, emphasizing the need for comprehensive strategies to address the complexities associated with vaccine development.

The evolution of Papillomaviridae suggests that E6 and E7 proteins may not only result from convergent evolution but also share a common evolutionary history. Understanding this diversity is pivotal for comprehending the virus's behavior and holds promise for developing effective preventive and therapeutic strategies. Research highlights the pivotal role of understanding this diversity in comprehending the behavior of the virus. Studies emphasize the importance of unraveling the common evolutionary origins of E6 and E7 proteins, shedding light on their shared ancestry and potential functional interplay [255]. Such insights are crucial for the development of effective preventive and therapeutic strategies. By uncovering the shared evolutionary history of these proteins, researchers aim to exploit this knowledge to design targeted interventions that could mitigate the impact of Papillomaviridae, potentially offering innovative approaches to prevent and treat associated diseases.

Elucidating Protein Characteristics in HPV16 and HPV18

The presented table outlines key characteristics of proteins associated with Human Papillomavirus (HPV) strains 16 and 18. Each protein's molecular weight, theoretical isoelectric point (pI), instability index, estimated half-life in mammalian reticulocytes, aliphatic index, and GRAVY (Grand Average of Hydropathy) score are detailed.

HPV16-L1 and HPV18-L1, crucial components of common HPV vaccines, exhibit similar molecular weights, emphasizing their structural resemblance. However, other proteins, such as HPV16-E6, are substantially larger, suggesting diverse structural complexities among viral proteins. Larger proteins may have unique features or functional domains, contributing to the varied roles they play in the viral life cycle or pathogenesis. This insight into the distinct sizes of HPV proteins informs our understanding of their diverse functions within the virus and underscores the complexity of developing effective vaccines targeting different components of the virus [256].

The pI values indicate the pH at which proteins carry no net electrical charge. Notably, HPV18-E7 has a lower pI compared to HPV16-E7, suggesting differences in charge distribution. These variations may influence protein interactions within the host environment. For instance, the lower pI of HPV-18-E7 suggests that it carries a net positive charge at physiological pH levels. This can influence its interactions with negatively charged proteins or regions of other proteins within the host environment. Potential interacting proteins could include those involved in cell cycle regulation or cellular signaling pathways, as the E7 protein of high-risk HPVs like HPV-18 is known for its role in modulating host cell processes. Additionally, proteins associated with DNA damage response or repair mechanisms might be affected by the interaction with HPV-18-E7 [257].

Higher instability indices, as seen in HPV16-E6, imply increased vulnerability to degradation. This susceptibility may impact the protein's functionality and persistence

within the host, potentially influencing viral pathogenicity. Higher instability may lead to a shorter half-life of HPV16-E6 within host cells, affecting its ability to form stable interactions with cellular proteins [258]. Instability could impact E6's binding to crucial cellular factors, such as interferon regulatory factors [259], influencing immune responses. Also, E6 plays a role in manipulating cellular processes, including cell cycle regulation and apoptosis. Its increased instability might disrupt the protein's normal functioning, potentially altering its capacity to modulate these processes and promoting cellular abnormalities [260]. Additionally, Instability could compromise E6's effectiveness in evading host defenses, potentially reducing the virus's overall pathogenicity.

All proteins exhibit an estimated half-life of 30 hours in mammalian reticulocytes. This uniformity suggests a consistent temporal presence of these viral proteins during the viral life cycle, emphasizing their importance in viral processes. In the context of mammalian reticulocytes, the estimated half-life of proteins, including those originating from viruses, is approximately 30 hours. The reliability of the 30-hour half-life suggests that these proteins play crucial roles and are integral to key viral processes within reticulocytes [261]. Thus, selecting them to produce a multiepitope vaccine against HPV 16;18 is reliable.

The aliphatic index reflects protein stability; higher values indicate greater stability. Notably, HPV18-E7 stands out with a significantly elevated aliphatic index, potentially indicating enhanced structural robustness compared to other proteins. The significance of a high aliphatic index in HPV18-E7 implies that the protein may better withstand environmental changes, chemical stresses, or other factors that could potentially destabilize proteins. This structural robustness can be advantageous for the proper functioning and persistence of HPV18-E7 in cellular environments [262].

GRAVY scores represent hydrophobicity; negative values indicate hydrophilicity. While most proteins exhibit negative scores, highlighting water-solubility, HPV16-E6 stands out with a notably lower GRAVY score, suggesting increased hydrophobicity. Higher hydrophobicity of HPV16-E6 implies a greater tendency for the protein to repel water molecules and prefer interactions with non-polar substances. It also may have biological implications, affecting its interactions with cellular components or influencing its stability in different cellular environments. HPV16-E6 interacts with cellular proteins involved in signaling pathways. Altered hydrophobicity may modulate these interactions, impacting downstream signaling events and cellular responses [263]. Furthermore, Hydrophobicity can influence the binding affinity of HPV16-E6 with other cellular proteins, impacting the formation of complexes essential for viral replication and cellular transformation [264]. Moreover, HPV16-E6 contains zinc-binding domains. Changes in hydrophobicity may affect the stability of these domains, influencing the protein's structural integrity and functional properties [265].

In conclusion, the diverse characteristics among these HPV proteins underscore the intricate interplay between their structures and functions. Understanding these nuances is

vital for elucidating viral pathogenicity and informing targeted therapeutic strategies such as designing a multi-epitope vaccine.

Strategic Considerations in HPV Vaccine Development

Antigenicity assessment through VaxiJen yielded intriguing results, particularly in the context of HPV 16 and HPV 18 E6, E7, and L1 antigens. Notably, the puzzling revelation emerged when comparing the antigenicity of L1, a crucial component in widely used HPV vaccines like Cervarix and Gardasil. The preference for L1 in HPV vaccines, such as Gardasil and Cervarix, despite the higher antigenicity of E6, E7, and L2, stems from strategic considerations. L1 is chosen for its ability to induce a robust immune response against the viral capsid, promoting neutralizing antibodies that prevent initial infection events [266]. The classical view of designing HPV vaccines was primarily based on a prophylaxis point of view. The selection of L1 in vaccines like Gardasil and Cervarix reflects a strategic emphasis on preventing initial HPV infection. L1, when expressed as virus-like particles (VLPs), mimics the natural structure of the virus, inducing a robust immune response that generates neutralizing antibodies against the viral capsid.

This prophylactic approach aims to establish immune protection before any viral infection occurs. By targeting L1, which is integral to the structure of the virus, the vaccines strategically focus on hindering the initial steps of the infection process. The classical design prioritizes the prevention of HPV acquisition, thereby reducing the risk of associated diseases, including cervical cancer [267].

The prophylactic approach in designing HPV vaccines focuses on preventing initial infections. While successful, it has certain limitations. For example, Prophylactic vaccines primarily target HPV prevention and may not effectively treat existing infections or associated diseases. Also, those already exposed to certain HPV types might not benefit fully from prophylactic vaccines, as they mainly prevent infections before exposure[268]. Additionally, Prophylactic vaccines primarily cover specific HPV types included in the vaccine, leaving individuals susceptible to other non-vaccine types. So, they do not eliminate the need for cervical cancer screening, as they may not cover all oncogenic HPV types [269]. Moreover, High-risk populations with increased HPV exposure may require additional strategies beyond prophylaxis for comprehensive protection [270].

Contrary to expectations, the antigenicity of L1 was found to be lower compared to E6 and E7. This finding is unexpected, given the central role of L1 in common HPV vaccines. The discrepancy prompts a closer examination of the antigenic properties of L1 and its implications for vaccine design. The unexpected findings suggest a need to reevaluate the current understanding of L1's role in immunity and may necessitate adjustments in vaccine design strategies to enhance the overall efficacy of HPV vaccines [267].

Phylogenetic analysis of HPV species, however, provided a crucial context. Despite the lower antigenicity, L1 emerges as the most conserved antigen among the analyzed

antigens. This contradicts the antigenicity results but highlights the importance of considering phylogenetic conservation in understanding the immune response to HPV antigens. This contradiction between antigenicity and genetic conservation emphasizes the importance of considering the evolutionary history and genetic makeup of HPV antigens in understanding the immune response. While antigenicity is a measure of how well an antigen can induce an immune response, phylogenetic conservation indicates the degree of genetic similarity or stability over time [271].

The finding underscores that, despite its lower antigenicity, L1's genetic conservation might play a crucial role in the immune response to HPV [272]. This insight suggests that the immune system may recognize and respond to conserved elements in L1, which could have implications for vaccine design and understanding the broader dynamics of HPV infection.

The paradoxical nature of L1 antigenicity poses challenges for vaccine development strategies. While it may not exhibit the highest antigenicity, its conservation suggests a critical role in the viral lifecycle. Balancing the inclusion of antigens with high antigenicity, like E6 and E7, with those that are phylogenetically conserved, like L1, becomes a crucial consideration for comprehensive vaccine coverage[253]. On the other hand, in vaccine design, a balance needs to be struck between including antigens with high antigenicity, such as E6 and E7, which may trigger a strong immune response, and those that are phylogenetically conserved, like L1. This balance is essential to ensure comprehensive vaccine coverage against a diverse range of HPV strains. Including both highly immunogenic antigens and conserved ones enhances the effectiveness of the vaccine, offering protection against a broader spectrum of viral variants [273]. This consideration highlights the complexity of HPV vaccine development, where understanding the interplay between antigenicity and conservation is crucial for creating effective vaccines.

Considering the diminished antigenicity of L1 in comparison to E6 and E7, there arises a concern about its potential to induce a robust immune response. Our investigation uncovered a nuanced interplay encompassing antigenicity, phylogenetic conservation, and immunogenicity--a trifecta crucial for vaccine optimization. This revelation prompted a rational approach to HPV vaccine design. The imperative lies in striking an optimal balance: eliciting a potent immune response while strategically targeting conserved elements. This meticulous strategy aims to ensure a comprehensive and effective defense, encompassing a broad spectrum of HPV strains.

Secondary Structure Analysis of HPV16 and HPV18 Proteins

Secondary structure analysis provides valuable insights into the three-dimensional arrangement of proteins, offering clues about their function and stability. By understanding the secondary structure, researchers can predict the potential functions of proteins. Certain structural motifs are associated with specific biological activities, such as binding sites or catalytic proliferation, and identifying these motifs aids in elucidating the protein's role in cellular processes [274]. Moreover, the secondary structure is

intimately linked to the stability of proteins. Stable structures contribute to a protein's overall resilience against denaturation or unfolding, impacting its functional longevity [275]. Furthermore, the arrangement of secondary structures provides critical information about how different regions of a protein interact with each other and with other molecules. This knowledge is essential for comprehending the structure-function relationship of proteins [276].

Alpha-Helix (α-helix) Content:

HPV18-E6 exhibits a significantly higher alpha-helix content (60.13%), suggesting a more defined and stable secondary structure compared to other proteins. Proteins with a well-defined secondary structure, such as a high alpha-helix content, tend to be more stable. This stability is crucial for the proper functioning and longevity of the protein in its biological context [277].

Extended Strand Formation or β-strands:

HPV16-L1 and HPV18-L1 show notable extended strand percentages, indicative of beta-sheet structures. This structural feature might be crucial for their roles in viral assembly. β-sheets are a common secondary structure in proteins where neighboring strands align and form a sheet held together by hydrogen bonds. The structures often contribute to the stability and rigidity of proteins. In the context of viral assembly, these structures may play a critical role in forming the outer shell or capsid of the virus [278].

Random Coil Abundance:

HPV16-E7 and HPV18-E7 display elevated random coil percentages, signifying structural flexibility. This flexibility could be linked to their diverse functional roles during the viral life cycle. Random coil structures are often adopting a flexible and disordered conformation. So, it indicates the ability of the proteins to adopt various conformations, allowing them to be versatile in their interactions [279]. Also, The flexibility of these proteins may enable them to adapt to different stages of the viral life cycle, contributing to processes such as infection, replication, or evasion of host defenses [280].

Thus, variations in secondary structure among HPV proteins highlight their functional diversity and structural adaptations. These findings contribute to our understanding of HPV proteins behavior and can guide us for constructing a better multiepitope vaccine.

Structural analyzing of HPV16 and HPV18 Antigens

All modeled proteins show a high percentage in the favored region. It is indicating reliable and accurate structural predictions [281, 282].

Differences in Allowed Regions:

HPV16-E6 and HPV18-E6 have a minimal allowed region, suggesting well-defined conformations. The minimal allowed region indicates that the phi (ϕ) and psi (ψ) angles of the amino acid residues in these proteins fall within a restricted, specific range, suggesting a stable and defined structure [283]. While HPV16-E7 and HPV18-E7 display a slightly larger allowed region, indicating structural flexibility. The larger allowed region indicates a broader range of phi and psi angles for the amino acid residues, suggesting that these proteins have a degree of conformational variability. This structural flexibility can be indicative of regions within the proteins that can adopt different conformations, possibly playing a role in their functional diversity or adaptability [284].

Limited Disallowed Regions:

Generally, the models exhibit minimal disallowed regions, reinforcing the overall reliability of the predicted structures. The majority of the dihedral angles (ϕ and ψ) in the modeled protein structure fall within energetically favorable regions on the Ramachandran plot. These regions correspond to conformations that are sterically allowed and structurally stable. A lower presence in disallowed regions suggests that the modeled protein structure is consistent with known protein structures and conforms well to the expected backbone geometry [285].

These results underscore the quality and reliability of the modeled structures, providing a basis for designing the multiepitope vaccine.

Molecular Docking Analysis of T Cytotoxic and T-Helper

The docking results of T cytotoxic and T-helper epitopes for HPV16 and HPV18 into HLA-A, HLA-BBR, and HLA-DQ are crucial for understanding the potential interactions between these epitopes and major histocompatibility complex (MHC) molecules. The molecular docking of selected epitopes with HLA-A, HLA-DBR, and HLA-DQ alleles provides insights into the potential binding interactions, crucial for eliciting a robust immune response against HPV. The choice of a 9-mer or 15-mer epitope selection strategy adds granularity to the vaccine design, ensuring a focused and tailored immune response against specific regions of the HPV genome [286].

All of the selected epitopes could strongly with low energy, attach to their MHC molecules based on their size. All the 9-mer epitopes docked perfectly into the MHC-I, HLA-A, and all the epitopes with 15-mere sizes attach into MHC-II molecules with low energy and high affinity, with average of -170 kcal/mol with RMSD near zero. The successful docking of epitopes into various HLA molecules supports their potential as candidates for inducing both CD8+ and CD4+ T cell responses [287].

These findings are promising for the development of an effective HPV vaccine. Further experimental validation and functional assays are necessary to confirm these in silico predictions.

Comprehensive Evaluation of Multi-Epitope Vaccine Attributes

The physicochemical profile of the designed vaccine, as presented in Table 1, reveals promising attributes. The molecular weight of 38798.99 KD, theoretical isoelectric point (pI) of 5.93, and an instability index of 20.37 indicate a stable protein. The vaccine appears to have a moderate molecular weight, a slightly acidic isoelectric point, and a stability index suggesting overall stability, which are favorable characteristics for a vaccine design [288]. A moderate molecular weight is often preferred to ensure the vaccine is immunogenic and effectively presented to the immune system. Extreme molecular weights may impact the protein's ability to be processed and presented as antigens, affecting overall vaccine effectiveness [289]. Furthermore, molecular weight influences the formulation of the vaccine, affecting factors such as solubility and ease of administration [290]. In the case of pI, however the protein is stable, but the optimal isoelectric point (pI) for designing a multi-epitope vaccine can vary based on the specific characteristics of the vaccine and the desired immune response. Generally, a pI in the range of 7 to 9 is often considered favorable . A vaccine with a pI within the physiological pH range (around 7.4) are more stable, have better solubility, and eliciting an appropriate immune response, as it influences the interaction of the vaccine with immune cells and receptors [64]. However, The vaccines' features in overall contribute to the potential effectiveness and reliability of the vaccine The estimated half-life of 7.2 hours in mammalian reticulocytes suggests adequate persistence. The optimal estimated half-life of a multi-epitope vaccine in mammalian reticulocytes can vary, but several studies provide insights into potential ranges. It's important to note that the optimal half-life may depend on various factors, including the specific design of the vaccine, the nature of epitopes used, and the overall vaccine formulation. Therefore, considering a range of greater than 20 hours to 30 hours may be indicative of potential stability and effectiveness in mammalian reticulocytes [291, 292]. The aliphatic index of 66.55 signifies a protein structure rich in aliphatic amino acids, contributing to its stability. The optimal aliphatic index for designing a multi-epitope vaccine can vary, but studies suggest that a higher aliphatic index is generally desirable for thermostability, a range between 70 to 90 are found in the literature [64, 293, 294]. Moreover, the negative GRAVY score (-0.107) indicates hydrophilicity, potentially enhancing solubility. The optimal GRAVY (Grand Average of Hydropathicity) score for multiple epitope vaccine designing is context-dependent, but a negative GRAVY score is generally considered favorable [295].

Allergenicity Assessment

AllerCatPro analysis negated allergenic potential, affirming the safety of the vaccine. Experimental tests for gluten, IgE production, and comparison with common allergens and autoimmune-stimulating allergens further support its non-allergenic nature. When AllerCatPro shows no evidence for allergenicity, it suggests a low predicted risk of inducing an allergic response based on the available data [125]. However, careful interpretation and consideration of additional factors are necessary for a comprehensive understanding of the protein's allergenic potential. the complexity of determining allergenic potential and emphasizes the importance of a holistic evaluation, incorporating

multiple factors beyond the predictions [296] made by tools like AllerCatPro. It encourages a comprehensive understanding that goes beyond individual assessments to make informed conclusions about a protein's allergenicity.

Antigenicity Prediction

VaxiJen 2.0 predicted an antigenic score of 0.64, signifying the vaccine's potential to evoke a robust immune response. This supports its candidacy as an effective immunogen. The optimal range of antigenicity in VaxiJen for multiple epitope vaccine design can vary, but a common threshold used in several articles is greater than 0.4 [161, 297, 298].

B Cell Stimulation

In addition to physicochemical attributes, the vaccine was tested for B cell stimulation, revealing 13 linear epitopes and 2 conformational epitopes. These epitopes play a crucial role in eliciting specific immune responses, and antibody production enhancing the vaccine's potential efficacy. The identified multiple epitopes exhibit diversity, meaning they represent various regions of a foreign protein. These epitopes have the potential to bind to B cell receptors, triggering the production of antibodies. A diverse set of epitopes is advantageous as it engages different aspects of the immune system, increasing the likelihood of a robust and effective B cell immune response. This diversity enhances the vaccine's ability to recognize and combat a wide range of pathogens or foreign substances, contributing to a more comprehensive and adaptive immune defense [299].

Ramachandran Plot Analysis

The Ramachandran plot illustrates favorable structural conformations. Notably, 62.5% of residues fall within the most favored regions, while 33.3% are in additional allowed regions. A generous 4.2% in generously allowed regions indicates acceptable flexibility. Importantly, no residues are in disallowed regions, affirming structural integrity. Ramachandran plot suggests a structurally sound and well-folded protein with a majority of residues adopting favorable conformations. The presence of residues in additional allowed and generously allowed regions indicates a degree of flexibility, which can be beneficial for the protein's function [285].

The Z score of -3.88 denotes a significant deviation from the mean, suggesting a unique but acceptable protein fold. A negative Z score suggests that the multiepitope vaccine structure is energetically more favorable than the average of misfolded structures. This is generally a positive sign, indicating that the predicted fold is unique and lower in energy [300]. The magnitude of the Z score is essential. A Z score of -3.88 is considered significant, indicating a substantial deviation from the mean. Such deviations are often associated with distinctive, yet acceptable, protein folds. In fact, the interpretation of a Z score involves comparing it to a distribution of scores from misfolded structures. A Z score of -3.88 suggests that the multiepitope vaccine has a unique but still acceptable protein fold, deviating significantly from random or misfolded conformations [301].

Consequently, the vaccine exhibits favorable physicochemical attributes, safety from allergenic reactions, high antigenicity, and a structurally sound conformation, positioning it as a promising candidate for further immunological exploration.

Docking Insights into TLR9-CpG Motif Interaction

The successful docking of the vaccine with TLR9 suggests potential recognition and binding between the CpG motif and TLR9, a crucial player in innate immunity. The negative energy value further supports the stability of this interaction.

Recognition of the CpG motif by TLR9 is of immunological significance. TLR9 is known to specifically recognize unmethylated CpG motifs, triggering innate immune responses. This interaction can potentially enhance the immunogenicity of the vaccine by activating innate immune pathways [302]. LR-9, primarily expressed in professional innate immune cells like dendritic cells, macrophages, and NK cells. It recognizes specific DNA motifs containing cytosine–phosphate–guanine (CpG) dideoxynucleotides [303]. Upon binding to unmethylated CpG motifs, TLR-9 initiates downstream signaling cascades, activating innate immune responses. Methylation of self-DNA, unavailability of self-DNA [304], and conjunction with other DNA sensors, such as cGAS-STING [305], enhance the discrimination between self and non-self DNA. This activation triggers the production of various pro-inflammatory cytokines, such as interferons and interleukins, and stimulates the maturation and activation of antigen-presenting cells [306]. TLR9 triggers a signaling cascade that involves the recruitment of adapter proteins, such as MyD88 leading to the activation of downstream kinases and transcription factors, including IRAKs and NF-κB [307]. This cascade nduces the release of proinflammatory cytokines, including IL-6, IL-1β, IL-10, IL-17, TNF-α, and type I interferons (IFNs) [308].

The molecular docking analysis suggests that the CpG motif added to the vaccine sequence is likely to be recognized by TLR9, highlighting a potential avenue for enhancing the vaccine's immunogenicity through the activation of innate immune responses.

Framework for Efficient Production of Multi-Epitope HPV Vaccine

E.coli K-12 strain is genetically stable [309] and this stability of the K12 strain ensures consistent and reliable production of the designed multi-epitopes HPV vaccine. Also, the *E. coli* K12 strain offers practicality, cost-effectiveness, safety, and reliable expression capabilities, making it a suitable choice for the production of your HPV vaccine [310]. The selection of the pET28a (+) plasmid from Addgene and its subsequent importation into the Benchling server marked a critical step in the in-silico cloning process. The pET28a(+) plasmid is well-recognized and widely used in molecular biology studies. Its compatibility with *E. coli* K-12 strains facilitates efficient cloning and transformation processes, ensuring successful integration of the vaccine sequence into the bacterial genome [311]. Also, the plasmid offers versatility in terms of the genes it can carry and express. Researchers can insert their gene of interest into the pET28a(+) plasmid, enabling the expression of specific proteins that form the basis of the vaccine [312].

Additionally, pET28a(+) is designed for use with the T7 RNA polymerase-based expression system. This system allows controlled and inducible protein expression, providing researchers with the ability to regulate the production of the vaccine in E. coli K12 efficiently . Also, it accommodates codon-optimized gene sequences effectively. This is crucial for enhancing protein expression levels and ensuring the efficient translation of the vaccine's genetic code in E. coli K-12 [313]. The termination codon TGA and the Kozak sequence GCCGCCACCAUGG were purposefully integrated for proper termination and initiation of translation. The Kozak sequence enhances the efficiency of translation initiation by providing a favorable context for the start codon (AUG). This sequence, positioned around the start codon, optimally engages with the eukaryotic ribosomal machinery, promoting the accurate initiation of translation in E. coli K12 strains [314, 315].

This meticulous in silico cloning strategy establishes a solid foundation for subsequent laboratory-based experiments and holds promise for the efficient production of the designed multiepitope HPV vaccine.

In Silico Insights into Multi-Injection Vaccination

The in silico immune simulations following a three-injection regimen of the designed vaccine unveiled a complex interplay of immune responses, providing valuable insights into the vaccine's efficacy. These simulations provide insights into factors such as protein collective motion, deformability, and epitope interactions, contributing to the design and validation of multi-epitope vaccines [316]. Stochastic agent-based immune simulation platforms have been employed to evaluate two-dose and three-dose vaccination protocols, offering a deeper understanding of adenoviral COVID-19 vaccination strategies [317]. Furthermore, immunoinformatics and immune simulation analyses have been utilized in designing multi-epitope vaccines for viruses like the lassa virus, showcasing the broad applicability of in silico approaches in vaccine development [318].

To facilitate the future design of COVID-19 vaccines, a computational approach integrating delta binding free energy has been proposed, demonstrating the power of in silico prediction in identifying immune-escaping hot spots [319]. Overall, in silico immune simulations provide a valuable tool for comprehending the complex interplay of immune responses, aiding in the development and optimization of effective vaccines.

Humoral Immune Responses:

The primary injection triggered a substantial elevation in IgM levels, indicative of the initial humoral response. IgM antibodies are produced early in the humoral immune response to viral infections and provide fast protective immunity [320]. After vaccination, the timeline typically involves the detection of IgM antibodies by one week postvaccination, peaking around two weeks postvaccination, and then gradually declining over several months [321]. It can be an advantage of our vaccine as after HPV vaccination, it is not commonly reported to see IgM antibodies. Studies on HPV vaccination primarily focus on the measurement of

IgG antibodies [322]. While some studies mention the detection of IgM-HPV antibodies in a small percentage of cases, such as in maternal HPV-antibodies, it is not a predominant or expected immune response after classical HPV vaccination [323].

Subsequent injections led to a shift towards a more diverse humoral response, with increased IgM + IgG, IgG1 + IgG2, IgG1, and IgG2 populations. This suggests the vaccine's ability to induce a broad spectrum of antibodies. The detection of IgG antibodies is a standard indicator of an effective immune response to HPV vaccination [324]. After HPV vaccination, the induction of high-quality and sustained serum IgG antibody titers against HPV L1 is observed, providing protection against HPV infection [325]. Neutralizing antibodies, predominantly of the IgG class, are perceived as a key determinant of prophylactic HPV vaccine performance [324]. The levels of anti-HPV 16 IgG antibodies in sera and oral fluids have been shown to correlate significantly with vaccination, indicating the systemic and mucosal immune response [326]. IgG1 is the most abundant subclass of IgG in response to HPV. IgG1 antibodies are effective in neutralizing viruses and promoting opsonization, facilitating the elimination of pathogens. Studies have shown a significant presence of IgG1 antibodies in the immune response against HPV [327]. While IgG2 antibodies are generally less abundant, they play a crucial role in complement activation, enhancing the immune response. In the context of HPV, the specific role of IgG2 in eradication might be influenced by its capacity to interact with Fc-receptors, which are essential for driving innate immune effector functions [328]

Cellular Immune Responses:

CTL and HTL T cell populations exhibited heightened responses after each exposure, emphasizing the vaccine's immunogenicity in activating T cell epitopes. CTCs play a crucial role in HPV immunization by recognizing and eliminating cells infected with the virus. They induce cell death in infected cells, preventing the virus from replicating and spreading. HPV-specific CD8+ T cells are known to exhibit lytic capacity, contributing to the clearance of HPV-infected cells [329]. T helper cells, specifically CD4+ T cells, are essential for coordinating the immune response. They support the activation of cytotoxic T cells and B cells. CD4+ T cells with a T helper phenotype are involved in HPV immunization, promoting an effective antiviral response. Additionally, they contribute to the development of a memory immune response for long-term protection [330]. The interplay between CTCs and T helper cells is crucial for mounting a comprehensive and lasting immune defense against HPV.

Consistent NK cell proliferation throughout exposures suggests a stable immune surveillance component. Sustained NK cell activity suggests that the vaccine is effectively stimulating the innate immune system. NK cells play a crucial role in the early defense against infections, and their consistent activity may contribute to an enhanced immune response [331]. Moreover, NK cells are known to

influence and support adaptive immune responses. Their continuous activation may facilitate the activation of other immune cells, such as T cells and B cells, contributing to a robust and adaptive immune defense [332]. NK cells are vital in recognizing and eliminating virus-infected cells. Consistent NK cell activity implies ongoing surveillance and potential clearance of infected cells, enhancing the vaccine's ability to protect against HPV threats [333]. Additionally, NK cells constant activity mean the adjuvant, CpG DNA, may contribute to the observed NK cell activity. Adjuvants play a crucial role in enhancing the NK cells activity and immune response to vaccines [334]. Thus, the observed NK cell activity could be an indicator of the vaccine's efficacy in priming the immune system for a sustained and vigilant response against the targeted pathogens.

Antigen Levels and Memory Response:

The decline in antigen levels after exposures indicates effective clearance, highlighting the vaccine's role in inducing a robust immune memory. The decline in vaccine antigen levels over time is a normal part of the immune response [335]. The initial antigen clearance signifies the resolution of the acute immune response. The gradual decline in vaccine antigen levels, coupled with the establishment of immunological memory, underscores the efficacy of the vaccine. It demonstrates the vaccine's ability to train the immune system for sustained protection and memory recall [336].

Persistent elevation in B and T cell memory throughout exposures underscores the vaccine's potential for sustained immune protection. The prolonged elevation in B cell memory suggests that memory B cells continue to produce antibodies efficiently. These antibodies play a crucial role in recognizing and neutralizing the pathogen upon subsequent encounters, providing long-lasting protection [337]. The sustained elevation in T cell memory signifies the endurance of memory T cells. These cells contribute to a rapid and targeted immune response. Memory T cells are essential for coordinating immune defenses and orchestrating the elimination of infected cells [338].

Cytokine Production and Immune Signaling:

Significant increases in IFN-γ, IL-10, IL-23, and IL-12 levels upon subsequent exposures indicate a pro-inflammatory and regulatory cytokine milieu. This suggests the vaccine's impact on orchestrating a balanced immune response with cellular and humoral components. The simultaneous elevation of pro-inflammatory (IFN-γ, IL-23, IL-12) and anti-inflammatory (IL-10) cytokines creates a balanced cytokine milieu. This balance is crucial for effective immune responses without causing excessive inflammation, preventing potential tissue damage [339]. Furthermore, the observed pattern of cytokine changes upon subsequent exposures may indicate the establishment of immunological memory. This memory enables a quicker and more targeted immune response upon encountering the same or similar pathogens in the future [340]. Also, IL-23 and

IL-12 are associated with the regulation of adaptive immunity, influencing T cell responses. The coordinated increase in these cytokines suggests an orchestrated immune response involving both innate and adaptive components [341].

In silico simulations demonstrate the designed vaccine's ability to elicit a dynamic and multifaceted immune response, emphasizing its potential as a promising candidate for further experimental validation.

Post Vaccination: Gene Expression Insights in PBMCs

The identified genes with significant expression changes in PBMCs offer potential insights into the immune response elicited by the HPV-16 L1 VLP vaccine. Further exploration of these genes and their functional implications may contribute to our understanding of vaccine efficacy and guide future research in HPV vaccination. Gene expression analysis of peripheral blood mononuclear cells (PBMCs) following HPV vaccination reveals crucial adaptive and innate immune responses induced by the vaccine [342]. Furthermore, differentially expressed genes in PBMCs exhibit pathways directly and indirectly linked to inflammation, providing insights into the immune system's response to HPV vaccination [183]. Also, HPV vaccination, specifically with L1 VLP, induces a wide spectrum of cytokines in PBMCs, emphasizing the activation of immune signaling molecules crucial for antiviral responses [343]. In the case of therapeutic purposes, understanding the gene expression profiles in response to HPV vaccination aids in designing more effective vaccines by identifying key immune responses and pathways. Moreover, monitoring local immune responses and HPV gene expression during therapies, can offer insights into treatment resistance and outcomes [344]. Additionally, sequencing T cell receptors (TCRs) of HPV-specific T cells, leading to complete resolution of HPV infections, highlighting a potential avenue for immunotherapeutic interventions [345].

The meticulous selection process and subsequent gene expression analysis provide a foundation for understanding the molecular responses of PBMCs to the HPV vaccines, paving the way for future investigations and potential advancements in HPV vaccination strategies.

Decoding the Network Analysis of Top 100 Hub Genes

The network analysis of the top 100 hub genes conducted in Gephi reveals pivotal insights into the molecular landscape of the studied context. Notably, IL6 emerges as a critical player, occupying the top position in the network, indicative of its central role. IFN-γ follows closely, as visually demonstrated in the accompanying illustration.

The network's structural foundation comprises 53 nodes and 185 edges, forming a robust backbone that underlies the intricate connections among the identified genes. Gephi allows the creation of visually intuitive representations of gene networks, enabling researchers to observe relationships and patterns within complex datasets [188].

Modulatory analysis further refines the network into five distinct modules, each characterized by a specific set of genes, emphasizing the complexity and modularity of the underlying biological system[346].

A substantial proportion of these genes actively participate in biological processes associated with cytokine and chemokine production, as well as the regulation of inflammatory responses. The hierarchal arrangement of the chart is meticulously derived based on *p*-values, establishing a statistical foundation for the observed relationships.

Bile acids and neurogenesis, both linked to gene expression in PBMCs, showcase intriguing biological interconnections. The role of bile acids extends beyond immunomodulation, playing a key role in the dynamic interplay between the microbiota and the mucosal immune system. In the context of neurogenesis, the suggestion to explore a tissue similar to the cervix, like the gut lumen, is insightful. Bile acids, traditionally known for emulsifying dietary lipids, have evolved beyond their classic role. Recent studies unveil their pleiotropic functions as signaling metabolites, orchestrating diverse metabolic and inflammatory pathways across various cells and tissues. Bile acids engage in dynamic interactions with host receptors and the microbiota, actively circulating between the liver and ileum. This evolution in understanding has linked disrupted bile acid circulation and metabolism to the development of cholestatic liver diseases, metabolic syndrome, colon cancer, and inflammatory bowel diseases (IBDs). The three-dimensional interplay between bile acids, the microbiota, and the mucosal immune system is a key focus, elucidating mechanisms governing intestinal homeostasis and inflammation [347]. Additionally, exploring neuroimmunology in a tissue resembling the cervix, such as the gut lumen, could provide valuable insights. The gut-brain axis underscores the bidirectional communication between the gut and the central nervous system, impacting neuroimmune interactions. Neuroepithelial crosstalk plays a pivotal role in gut physiology, particularly in maintaining the mucosal barrier during homeostasis and inflammation. The intricate communication between sensory neurons and epithelial cells has been a subject of limited understanding. Recent research reveals that Nav1.8+ CGRP+ nociceptor neurons interact closely with intestinal goblet cells, triggering mucus secretion for gut protection. Commensal microbes activate nociceptors, influencing homeostatic CGRP release and contributing to gut barrier protection [348].

Furthermore, the hub genes exhibit a noteworthy association with viral infections, particularly highlighting the relevance of HPV16 in influencing the expression patterns of these genes. Persistent HPV infection can induce chronic inflammation. Once established, the infection can alter the immune microenvironment, leading to the recruitment of immune cells and causing inflammation [349]. In addition, Chronic inflammation is a prolonged immune response that persists over time. In the context of HPV infection, the presence of the virus in basal cells may trigger an ongoing immune response, leading to chronic inflammation. So, There is a biologically plausible connection between chronic inflammation and HPV infection [350]. The initial inflammation is critically important for viral eradication. Stealthy viral replication[351] and low viral gene expression [352] are two mechanism that are used by HPV to operate

without alerting the host immune system and initiating the inflammation. However, chronic or persistent HPV infection is considered a key driver of cervical cancer, and inflammation is identified as an important factor in this process [353]. IL-6 is found in the serum and exfoliated cervical cells of individuals with persistent high-risk HPV infections, which are the main cause of cervical cancer and its precursor lesions [354]. Furthermore, polymorphisms in the IL-6 gene have been studied in the context of HPV infections and cervical cancer development. Genetic variations in IL-6 may contribute to susceptibility and outcomes of HPV-associated diseases [355]. Moreover, HPV16 and HPV18, through their E6/E7 and E6* proteins, have been found to upregulate IL-6. This interaction may contribute to the immunomodulatory effects of HPV on the host's immune response [356]. Additionally, DisGeNET underscores a significant correlation between the hub genes such as and autoimmune diseases such as Rheumatoid Arthritis (RA), Graves, Arteriosclerosis, and osteoporosis, providing valuable clinical implications for further investigation. IL-6 dysregulation is implicated in various autoimmune diseases, such as RA, contributing to their pathogenesis [357]. Levels of IL-6 in the serum correlate with the disease activity of RA. Successful treatment with Disease-Modifying Anti-Rheumatic Drugs (DMARDs) or Tumor Necrosis Factor (TNF) inhibitors has been shown to reduce serum IL-6 concentrations [358]. Consequently, diseases characterized by continual synthesis of IL-6, like rheumatoid arthritis, can be treated by targeting IL-6. Inhibition of IL-6 is explored as a therapeutic strategy in autoinflammatory diseases [359].

The comprehensive network and modulatory analyses shed light on the intricate interplay of hub genes, offering a molecular perspective on cytokine regulation, viral interactions, and potential links to autoimmune disorders in the studied context.

Targeting Druggable Pockets in IL-6

IL-6, known for its pro-inflammatory characteristics[360], poses challenges for drug design and dosage due to its structural features, including a high alpha-helix content and an acidic pI charge. Mitigating its impact on HPV-infected tumors necessitates innovative solutions. Additionally, its elevated aliphatic degree enhances stability in high-temperature tissue environments. High alpha-helix content contributes to structural stability, making it resistant to changes. This stability can impede drug interactions as the protein structure may not readily accommodate binding [361]. Also, an acidic pI charge indicates a prevalence of acidic residues. This can affect drug solubility and alter the electrostatic interactions crucial for drug-protein binding. Designing drugs that effectively interact with acidic regions becomes more complex [362]. High aliphatic content enhances the protein's stability in high-temperature environments. While this is advantageous for the protein's natural function, it complicates drug design, as conventional drugs may struggle to compete with the protein's inherent stability. The protein with high aliphatic contents, have lower numbers of binding site and high number of hydrophobic contents. These characters make them difficult to set as drugs target [363].

An analysis of IL-6 unveils potential druggable pockets, deepening our insight into therapeutic opportunities. Predictions from PockDrug [364, 365] and CASTp [366] shed light on the probability of drugability for specific pockets. Pocket 2 (a): Identified as the most promising, boasting the highest drugability probability. This indicates a robust potential for targeted drug interactions, making it a prime candidate for further investigation and drug development . Third Pocket (b): With a high probability of 0.94, this pocket represents another favorable site for drug binding. Its proximity to a probability of 1 underscores a strong likelihood of successful [8] drug targeting, emphasizing its significance in potential therapeutic interventions. Pocket Number 0 (c): While not the highest, a substantial probability of 0.82 suggests significant potential for drug interactions. This adds to the repertoire of potential target sites for drug design and development [367]. Pocket 4 (d): Despite a slightly lower probability of 0.74, this pocket presents a noteworthy opportunity for drug binding, indicating a reasonable likelihood of successful drug targeting within this site. CASTp Predicted Pocket (e): Identified by CASTp, this pocket introduces another dimension to potential druggable sites in IL-6. Further exploration could unveil unique opportunities for therapeutic interventions.

The diverse set of predicted pockets in IL-6, each with varying drugability probabilities, provides a comprehensive landscape for targeted drug design. These findings pave the way for experimental validation and further exploration of specific pockets to develop effective drugs targeting IL-6.

From Drug Databases to a Potent IL-6 Inhibitor

Exploring drug databases like DrugBank provides confidence in drugability, yet most findings are investigational or agonistic. To target IL-6 with inhibitory effects, Binding DB offers crucial insights, revealing compounds predominantly resembling ginseng-like structures. Further exploration in ZINC databases enriches the compound pool. These materials, sourced from ZINC and Binding DB, lay the foundation for identifying the ideal IL-6 inhibitory structure. Leveraging these insights, a robust Quantitative Structure-Activity Relationship (QSAR) model is constructed, boasting an R^2 value of 0.81. This high R^2 value signifies a strong fit for the model, enhancing its reliability in predicting the inhibitory efficacy of compounds against IL-6. The obtained R^2 value of 0.81 in the Quantitative Structure-Activity Relationship (QSAR) model signifies a substantial concordance between the model and the dataset. R^2, the coefficient of determination, quantifies the fraction of variability in the dependent variable (response) that can be anticipated from the independent variables (descriptors) within the model. An R^2 of 0.81 indicates that approximately 81% of the fluctuations in the response variable are elucidated by the model [368].

In the domain of QSAR, a heightened R^2 value denotes a robust alignment between the model's forecasts and the empirically observed values, underscoring the reliability and effectiveness of the model. These findings underscore the triumphant development and validation of a QSAR model, showcasing its prowess in predicting molecular behavior from external databases. The results accentuate the model's efficacy in pinpointing

pharmacophores within a specified biological activity range. Molecules with IC50 values between 0 and 10, chosen from screening, underwent pharmacophore modeling in Schrodinger-Maestro. This approach hones in on compounds sharing a specific structure aligned with the IC50 range. The software, mirroring molecules binding to IL6 with low IC50, identified a structure comprising two rings, an acceptor, and a hydrophobic component. Notably, 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate exclusively met these criteria, solidifying its relevance in this distinct category.

Insights from Docking and 2D-QSAR Analysis

The molecular docking analysis between 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate and IL6 revealed nine potential binding sites, as predicted by PyRx software. The most favorable conformation, indicated by the lowest binding affinity (-5.8) and zero RMSD values, suggests a highly stable interaction. Low binding affinity and RMSD=0 indicate a high level of accuracy and agreement between the predicted and actual binding poses in molecular docking studies. A low binding affinity suggests a strong interaction between the ligand and the receptor, indicating a favorable binding configuration. When combined with RMSD=0, it signifies that the predicted binding pose perfectly aligns with the experimental or native structure[369]. Also, the combination of low binding affinity and RMSD=0 enhances the reliability of molecular docking predictions. It implies that the computational model accurately reproduces the experimentally observed binding mode[370]. So, low binding affinity signifies a strong interaction, and RMSD=0 indicates a perfect alignment of the predicted and actual binding poses. Together, they represent a highly accurate and reliable docking model, essential for drug discovery and design applications. In the context of ligand-protein interactions, especially in tools like PyRex, the affinity values are typically represented as free energy changes (ΔG). In this scenario, a more negative ΔG indicates a stronger binding affinity, suggesting a more stable interaction between the ligand and the protein. A negative ΔG signifies that the reaction is thermodynamically favorable. Moreover, the negative sign emphasizes the release of energy during the binding process, contributing to the stability of the ligand-protein complex [371].

In the 2D-QSAR model using BIOVIA Studio Visualizer, the first conformation demonstrated significant interactions. The Pi Anion bond formed between the ligand and aspartic acid (ASP) residue at position 160 (A160) enhances binding specificity. Additionally, the Pi Alkyl interaction involving the ligand's pi system and lysine (LYS) residue at position 46 (A46) contributes to ligand stabilization within the binding site. The PiPi-T shaped interaction with tryptophan (TRP) residue at position 157 (A157) further enhances the overall stability of the ligand-protein complex. Pi interactions play crucial roles in ligand-protein stability, contributing to the overall binding affinity. Involves the interaction between the pi-electron cloud of an aromatic system and an anion (e.g., sulfate groups). It enhances stability by forming attractive forces between the aromatic amino acids in the protein and anionic ligands [372]. Pi Alkyl interactions occurs when the pi-electron cloud of an aromatic group interacts with the electron cloud

of an alkyl group [373]. This interaction contributes to stability by fostering favorable interactions between aromatic amino acids in the protein and alkyl groups in the ligand.

Qualities of 4-(4-Morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate

Assessing the physicochemical characteristics of 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate through Lipinski rules reveals promising drug-like qualities. With a molecular weight of 343.40, 4 hydrogen bond acceptors (HBA), 0 hydrogen bond donors (HBD), and a LogP of 3.48, it aligns with lead-like structures in medicinal chemistry.

These attributes imply favorable drug-like properties, indicating high solubility and potential medicinal relevance. The low LogP suggests limited penetration of the blood-brain barrier, making it suitable for applications where central nervous system effects are undesirable. Conversely, the high gastrointestinal absorption suggests an efficient route for systemic distribution. The compound complies with essential drug design and safety criteria, positioning it as a promising candidate for drug development. Its lead-like classification, determined by factors such as molecular weight, hydrogen bond characteristics, and lipophilicity, underscores its potential utility in the early stages of drug discovery [374] . The drugability of a compound is often evaluated based on Lipinski's Rule of Five, which consists of the following criteria:

Molecular Mass: The molecular mass should be less than 500 Dalton.

Lipophilicity (LogP): The LogP (partition coefficient) should be less than 5.

Hydrogen Bond Donors: There should be fewer than 5 hydrogen bond donors.

Hydrogen Bond Acceptors: The number of hydrogen bond acceptors should be less than 10.

Polar Surface Area (PSA): The polar surface area should be less than 140 Å².

These rules serve as guidelines for assessing the drug-like properties of a compound, with adherence to these criteria often associated with favorable oral bioavailability and pharmacokinetic profiles. They provide a framework to evaluate key physicochemical properties that influence a compound's potential as a drug candidate [375].

The absence of hydrogen bond donors (HBD) and the presence of hydrogen bond acceptors (HBA) represent pertinent considerations in drug design. These characteristics play a pivotal role in governing the molecule's interactions within biological systems, influencing pharmacokinetics and potential therapeutic effects. Compounds with excessive hydrogen bond donors (HBD > 5) may face challenges in oral absorption, potentially leading to poor bioavailability. In addition, compounds with fewer hydrogen bond donors and acceptors are more likely to exhibit optimal absorption in gastrointestinal tract and distribution characteristics [376] . The balance between the

number of Hydrogen Bond Acceptors (HBA) and Hydrogen Bond Donors (HBD) in a compound is crucial for its drugability. Fewer HBA and HBD contribute to a favorable equilibrium between solubility and lipophilicity. This balance allows the compound to efficiently traverse cell membranes and the gastrointestinal (GI) tract, facilitated by its lipophilic properties. Simultaneously, the compound's solubility characteristics enable widespread distribution throughout the body. The smaller size of compounds with fewer HBD and HBA further enhances their ability to navigate biological membranes with ease.

This intricate interplay between solubility, lipophilicity, and molecular size underscores the significance of HBA and HBD in shaping a compound's pharmacokinetic profile. The compound's ability to navigate biological barriers efficiently, both in terms of size and physicochemical properties, is pivotal for its successful application in drug development. [377].

In medicinal chemistry, molecular filters play a pivotal role by considering parameters like molecular weight to selectively target compounds with desirable drug-like attributes. Molecular weight acts as a crucial determinant, ensuring adherence to the desired drug profile and thereby improving ADMET characteristics (absorption, distribution, metabolism, excretion, and toxicity). A molecular weight less than 500 Da is essential for optimal absorption, enhanced bioavailability, faster entry into the bloodstream, and increased cell permeability, attributed to the smaller size of the compounds[378]. Moreover, a low Polar Surface Area (PSA) below 140 Å² is associated with higher cell permeability. Compounds with lower PSA exhibit favorable characteristics, including increased absorption. Notably, these molecules demonstrate a heightened ability to traverse the Blood-Brain Barrier (BBB) efficiently, emphasizing their potential for faster BBB penetration [379]. This strategic consideration of molecular weight and PSA in drug design enhances the likelihood of developing compounds with superior pharmacokinetic properties and the ability to navigate biological barriers effectively.

However, the presence of alerts in the Brenk analysis of 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate should be carefully considered. A Brenk alert value of 3 typically indicates the presence of three specific structural alerts or fragments in a compound that could be associated with potential toxicity, chemical reactivity, and metabolic instability. In general, a lower number of Brenk alerts is desirable, indicating a molecule with fewer structural features associated with toxicity or other concerns [380]. SwissSimilarity, an invaluable web tool for molecular analysis, facilitates ligand-based virtual screening across a spectrum of small to ultralarge libraries of compounds. Notably, it aids in the rapid assessment of structural similarities and potential issues in chemical structures. In the context of the compound 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate, the interpretation of Brenk alerts highlights the need for nuanced consideration, as the significance may vary based on the specific chemical structure and intended application of the compound.

To optimize the safety and efficacy of the aforementioned compound, structural modifications are imperative, necessitating the exclusion of problematic features and

fragments. SwissSimilarity proves instrumental in this process, allowing for a comprehensive analysis of analogous structures to inform targeted adjustments. For a practical demonstration, the SwissSimilarity website (http://www.swisssimilarity.ch/) can be explored using the compound's SMILES notation, especially those in clinical trials that have passed FDA filters [8]. In our investigation, we identified a structural similarity between 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate and UC-781, a reverse transcriptase inhibitor used in the context of AIDS [381]. Considering this similarity, there arises a potential application of UC-781 in managing the adverse effects of IL-6 during HPV infection.

Limitations of study

1. **Antigen Selection in Experimental Design:**

    - The experimental phase encountered a limitation in the selection of optimal antigens for both the Antigen-Induced Activity (AIM) and lymphocyte proliferation assays. A more focused investigation on E6 and E7 was prioritized over L1 due to their dominance as antigens, crucial for T-cell activation and proliferation.

2. **Validation in Vivo:**

    While vaccine design primarily occurs through computational tools and simulations, it's crucial to acknowledge that in vivo experiments are essential to validate the real-world effects of the designed vaccine. Simulation outcomes must be rigorously tested in living organisms to ensure the predicted immune responses translate effectively into practical vaccine development.

3. **Sample Size and Budget Constraints:**

    - The study faced challenges related to sample size and budget constraints, restricting the scope to conduct only the Lymphocyte Proliferation Assay (LPA) and impeding the continuation of the AIM assay.

4. **pH Levels in Vaccine Formulation:**

    - The pH of the vaccine formulation was measured at 5.9, slightly acidic for an ideal vaccine. Most vaccines fall within the pH range of 7-9, and optimizing the pH within this range is crucial for vaccine stability and effectiveness.

5. **Suboptimal Reticulocyte Incubation Period:**

- The optimal incubation period for reticulocytes in vaccine development is typically 20-30 hours. However, the current vaccine protocol involves a shorter incubation period of around 7 hours, potentially affecting the optimal maturation and response of reticulocytes.

6. **Missing HPV16:18-E7 Epitopes:**

    - The vaccine lacks the HPV16:18-E7 epitopes, indicating a potential limitation in its ability to stimulate immunity against these specific targets. Considering the incorporation of live-attenuated *L. monocytogenes* [382] may be explored as a strategy to address this deficiency.

7. **Brenk Analysis Alerts:**

    - During the Brenk analysis of 4-(4-morpholinylcarbothioyl)phenyl 3-(2-furyl)acrylate, alerts were identified, suggesting potential concerns related to toxicity, chemical reactivity, and metabolic instability. Further investigation and refinement may be needed to ensure the safety and stability of the compound.

These limitations highlight areas for future optimization and refinement to enhance the robustness and effectiveness of the developed vaccine.

Chapter 8: References

References:

1. Chesson HW, Dunne EF, Hariri S, Markowitz LE. The Estimated Lifetime Probability of Acquiring Human Papillomavirus in the United States. Sexually transmitted diseases. 2014;41:660.

2. Arrossi S, Almonte M, Herrero R, Gago J, Sánchez Antelo V, Szwarc L, et al. Psycho-social impact of positive human papillomavirus testing in Jujuy, Argentina results from the Psycho-Estampa study. Preventive Medicine Reports. 2020;18:101070.

3. Handler MZ, Handler NS, Majewski S, Schwartz RA. Human papillomavirus vaccine trials and tribulations: Clinical perspectives. Journal of the American Academy of Dermatology. 2015;73:743–56.

4. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide - PubMed. https://pubmed.ncbi.nlm.nih.gov/10451482/. Accessed 16 Apr 2022.

5. Human papillomavirus vaccines: WHO position paper, May 2017-Recommendations. Vaccine. 2017;35:5753–5.

6. Roman BR, Aragones A. Epidemiology and incidence of HPV-related cancers of the head and neck. Journal of surgical oncology. 2021;124:920–2.

7. Arbyn M, De Sanjosé S, Saraiya M, Sideri M, Palefsky J, Lacey C, et al. EUROGIN 2011 roadmap on prevention and treatment of HPV-related disease. International journal of cancer. 2012;131:1969–82.

8. Li J, Ding J, Zhai K. Detection of Human Papillomavirus DNA in Patients with Breast Tumor in China. PLOS ONE. 2015;10:e0136050.

9. Heng B, Glenn WK, Ye Y, Tran B, Delprado W, Lutze-Mann L, et al. Human papilloma virus is associated with breast cancer. British journal of cancer. 2009;101:1345–50.

10. Sigaroodi A, Nadji SA, Naghshvar F, Nategh R, Emami H, Velayati AA. Human papillomavirus is associated with breast cancer in the north part of Iran. TheScientificWorldJournal. 2012;2012.

11. Bedoya AM, Jaramillo R, Baena A, Castaño J, Olaya N, Zea AH, et al. Location and Density of Immune Cells in Precursor Lesions and Cervical Cancer. Cancer microenvironment : official journal of the International Cancer Microenvironment Society. 2013;6:69–77.

12. Wielgos A, Pietrzak B, Sikora M, Martirosian G, Suchonska B, Gozdowska J, et al. Human Papillomavirus (HPV) DNA Detection Using Self-Sampling Devices in Women Undergoing Long Term Immunosuppressive Therapy. Viruses. 2020;12.

13. Sager R, Frei P, Steiner UC, Fink D, Betschart C. Genital Dysplasia and Immunosuppression: Why Organ-Specific Therapy Is Important. Inflammatory intestinal diseases. 2019;4:154–60.

14. Stelzle D, Tanaka LF, Lee KK, Ibrahim Khalil A, Baussano I, Shah ASV, et al. Estimates of the global burden of cervical cancer associated with HIV. The Lancet Global Health. 2021;9:e161–9.

15. Zur Hausen H. Papillomaviruses and cancer: from basic studies to clinical application. Nature reviews Cancer. 2002;2:342–50.

16. Ding DC, Chiang MH, Lai HC, Hsiung CA, Hsieh CY, Chu TY. Methylation of the long control region of HPV16 is related to the severity of cervical neoplasia. European Journal of Obstetrics and Gynecology and Reproductive Biology. 2009;147:215–20.

17. Forcier M, Musacchio N. An overview of human papillomavirus infection for the dermatologist: disease, diagnosis, management, and prevention. Dermatologic therapy. 2010;23:458–76.

18. Major AG, Pitty LP, Farah CS. Cancer stem cell markers in head and neck squamous cell carcinoma. Stem cells international. 2013;2013.

19. Cubie HA. Diseases associated with human papillomavirus infection. Virology. 2013;445:21–34.

20. Béziat V. Human genetic dissection of papillomavirus-driven diseases: new insight into their pathogenesis. Human genetics. 2020;139:919–39.

21. 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 research. 2017;45:D499–506.

22. Stanley M. Pathology and epidemiology of HPV infection in females. Gynecologic oncology. 2010;117 2 Suppl:S5.

23. Gheit T. Mucosal and Cutaneous Human Papillomavirus Infections and Cancer Biology. Frontiers in oncology. 2019;9 MAY.

24. Dunne EF, Park IU. HPV and HPV-associated diseases. Infectious disease clinics of North America. 2013;27:765–78.

25. Vinzón SE, Rösl F. HPV vaccination for prevention of skin cancer. Human vaccines & immunotherapeutics. 2015;11:353–7.

26. Derkay CS, Bluher AE. Update on Recurrent Respiratory Papillomatosis. Otolaryngologic clinics of North America. 2019;52:669–79.

27. Egawa N, Egawa K, Griffin H, Doorbar J. Human Papillomaviruses; Epithelial Tropisms, and the Development of Neoplasia. Viruses. 2015;7:3863.

28. Yoon CS, Kim KD, Park SN, Cheong SW. α6 Integrin Is the Main Receptor of Human Papillomavirus Type 16 VLP. Biochemical and Biophysical Research Communications. 2001;283:668–73.

29. Bousarghin L, Touzé A, Sizaret P-Y, Coursaget P. Human papillomavirus types 16, 31, and 58 use different endocytosis pathways to enter cells. Journal of virology. 2003;77:3846–50.

30. Schelhaas M, Shah B, Holzer M, Blattmann P, Kühling L, Day PM, et al. Entry of human papillomavirus type 16 by actin-dependent, clathrin- and lipid raft-independent endocytosis. PLoS pathogens. 2012;8.

31. Ferreira AR, Ramalho AC, Marques M, Ribeiro D. The Interplay between Antiviral Signalling and Carcinogenesis in Human Papillomavirus Infections. Cancers. 2020;12.

32. Cosper PF, Bradley S, Luo L, Kimple RJ. Biology of HPV Mediated Carcinogenesis and Tumor Progression. Seminars in radiation oncology. 2021;31:265–73.

33. Lugović-Mihić L, Djaković HCI, Kuna M, Šešerko A. The Influence of Psychological Stress on HPV Infection Manifestations and Carcinogenesis. Cellular physiology and biochemistry : international journal of experimental cellular physiology, biochemistry, and pharmacology. 2021;55:71–88.

34. Moody CA, Laimins LA. Human papillomavirus oncoproteins: pathways to transformation. Nature reviews Cancer. 2010;10:550–60.

35. Maglennon GA, McIntosh P, Doorbar J. Persistence of viral DNA in the epithelial basal layer suggests a model for papillomavirus latency following immune regression. Virology. 2011;414:153–63.

36. Burchell AN, Winer RL, de Sanjosé S, Franco EL. Chapter 6: Epidemiology and transmission dynamics of genital HPV infection. Vaccine. 2006;24 Suppl 3 SUPPL. 3.

37. O'Rourke K. Cutaneous β-HPV may predict cutaneous squamous cell carcinoma. Cancer. 2022;128:646.

38. Egawa N, Doorbar J. The low-risk papillomaviruses. Virus research. 2017;231:119–27.

39. Ryndock EJ, Meyers C. A risk for non-sexual transmission of human papillomavirus? Expert review of anti-infective therapy. 2014;12:1165–70.

40. Scott-Wittenborn N, Fakhry C. Epidemiology of HPV Related Malignancies. Seminars in radiation oncology. 2021;31:286–96.

41. Crosbie EJ, Einstein MH, Franceschi S, Kitchener HC. Human papillomavirus and cervical cancer. Lancet (London, England). 2013;382:889–99.

42. Maucort-Boulch D, Plummer M, Castle PE, Demuth F, Safaeian M, Wheeler CM, et al. Predictors of human papillomavirus persistence among women with equivocal or mildly abnormal cytology. International journal of cancer. 2010;126:684–91.

43. Li K, Yin R, Wang D, Li Q. Human papillomavirus subtypes distribution among 2309 cervical cancer patients in West China. Oncotarget. 2017;8:28502–9.

44. Steele JC, Mann CH, Rookes S, Rollason T, Murphy D, Freeth MG, et al. T-cell responses to human papillomavirus type 16 among women with different grades of cervical neoplasia. British Journal of Cancer 2005 93:2. 2005;93:248–59.

45. Krishna S, Ulrich P, Wilson E, Parikh F, Narang P, Yang S, et al. Human Papilloma Virus Specific Immunogenicity and Dysfunction of CD8 + T Cells in Head and Neck Cancer. Cancer research. 2018;78:6159–70.

46. Tawfeik AM, Mora A, Osman A, Moneer MM, El-Sheikh N, Elrefaei M. Frequency of CD4+ regulatory T cells, CD8+ T cells, and human papilloma virus infection in Egyptian Women with breast cancer. International Journal of Immunopathology and Pharmacology. 2020;34:1–12.

47. Béziat V, Rapaport F, Hu J, Titeux M, Bonnet des Claustres M, Bourgey M, et al. Humans with inherited T cell CD28 deficiency are susceptible to skin papillomaviruses but are otherwise healthy. Cell. 2021;184:3812-3828.e30.

48. Eberhardt CS, Kissick HT, Patel MR, Cardenas MA, Prokhnevska N, Obeng RC, et al. Functional HPV-specific PD-1 + stem-like CD8 T cells in head and neck cancer. Nature. 2021;597:279–84.

49. Hewavisenti R, Ferguson A, Wang K, Jones D, Gebhardt T, Edwards J, et al. CD103+ tumor-resident CD8+ T cell numbers underlie improved patient survival in oropharyngeal squamous cell carcinoma. Journal for ImmunoTherapy of Cancer. 2020;8:e000452.

50. Zhang Y, Wang N, Ding M, Yang Y, Wang Z, Huang L, et al. CD40 Accelerates the Antigen-Specific Stem-Like Memory CD8+ T Cells Formation and Human Papilloma Virus (HPV)-Positive Tumor Eradication. Frontiers in Immunology. 2020;11:1012.

51. Maskey N, Thapa N, Maharjan M, Shrestha G, Maharjan N, Cai H, et al. Infiltrating CD4 and CD8 lymphocytes in HPV infected uterine cervical milieu. Cancer management and research. 2019;11:7647–55.

52. Chen Z, Dong D, Zhu Y, Pang N, Ding J. The role of Tim-3/Galectin-9 pathway in T-cell function and prognosis of patients with human papilloma virus-associated cervical carcinoma. FASEB journal : official publication of the Federation of American Societies for Experimental Biology. 2021;35.

53. Ao C, Zeng K. The role of regulatory T cells in pathogenesis and therapy of human papillomavirus-related diseases, especially in cancer. Infection, genetics and evolution :

journal of molecular epidemiology and evolutionary genetics in infectious diseases. 2018;65:406–13.

54. Raverdeau M, Cunningham SP, Harmon C, Lynch L. γδ T cells in cancer: a small population of lymphocytes with big implications. Clinical & Translational Immunology. 2019;8:e01080.

55. Rawat SS, Keshri AK, Kaur R, Prasad A. Immunoinformatics Approaches for Vaccine Design: A Fast and Secure Strategy for Successful Vaccine Development. Vaccines (Basel). 2023;11:221.

56. Human Papillomavirus Vaccines: An Updated Review - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7565290/. Accessed 3 Dec 2023.

57. Pinto LA, Dillner J, Beddows S, Unger ER. Immunogenicity of HPV prophylactic vaccines: Serology assays and their use in HPV vaccine evaluation and development. Vaccine. 2018;36 32Part A:4792–9.

58. Bhattacharjee R, Kumar L, Dhasmana A, Mitra T, Dey A, Malik S, et al. Governing HPV-related carcinoma using vaccines: Bottlenecks and breakthroughs. Frontiers in Oncology. 2022;12.

59. EMA. Human papillomavirus vaccines - Cervarix, Gardasil, Gardasil 9, Silgard. European Medicines Agency. 2018. https://www.ema.europa.eu/en/medicines/human/referrals/human-papillomavirus-vaccines-cervarix-gardasil-gardasil-9-silgard. Accessed 3 Dec 2023.

60. Vaccines | Free Full-Text | CpG Oligonucleotides as Cancer Vaccine Adjuvants. https://www.mdpi.com/2076-393X/3/2/390. Accessed 3 Dec 2023.

61. Zhang L. Multi-epitope vaccines: a promising strategy against tumors and viral infections. Cell Mol Immunol. 2018;15:182–4.

62. Umitaibatin R, Harisna AH, Jauhar MM, Syaifie PH, Arda AG, Nugroho DW, et al. Immunoinformatics Study: Multi-Epitope Based Vaccine Design from SARS-CoV-2 Spike Glycoprotein. Vaccines (Basel). 2023;11:399.

63. Damas MSF, Mazur FG, Freire CC de M, Cunha AF da, Pranchevicius M-C da S. A Systematic Immuno-Informatic Approach to Design a Multiepitope-Based Vaccine Against Emerging Multiple Drug Resistant Serratia marcescens. Frontiers in Immunology. 2022;13.

64. Mahmud S, Rafi MO, Paul GK, Promi MM, Shimu MSS, Biswas S, et al. Designing a multi-epitope vaccine candidate to combat MERS-CoV by employing an immunoinformatics approach. Sci Rep. 2021;11:15431.

65. A systems immunology study comparing innate and adaptive immune responses in adults to COVID-19 mRNA and adenovirus vectored vaccines - ScienceDirect.

https://www.sciencedirect.com/science/article/pii/S2666379123000769. Accessed 3 Dec 2023.

66. Davis MM, Tato CM, Furman D. Systems immunology: just getting started. Nat Immunol. 2017;18:725–32.

67. Wakabayashi R, Nakahama Y, Nguyen V, Espinoza JL. The Host-Microbe Interplay in Human Papillomavirus-Induced Carcinogenesis. Microorganisms 2019, Vol 7, Page 199. 2019;7:199.

68. Hewavisenti RV, Arena J, Ahlenstiel CL, Sasson SC. Human papillomavirus in the setting of immunodeficiency: Pathogenesis and the emergence of next-generation therapies to reduce the high associated cancer risk. Frontiers in Immunology. 2023;14.

69. Orr MT, Ireton GC, Beebe EA, Huang P-WD, Reese VA, Argilla D, et al. Immune subdominant antigens as vaccine candidates against Mycobacterium tuberculosis. J Immunol. 2014;193:2911–8.

70. A guide to systems-level immunomics | Nature Immunology. https://www.nature.com/articles/s41590-022-01309-9. Accessed 3 Dec 2023.

71. Chu X, Zhang B, Koeken VACM, Gupta MK, Li Y. Multi-Omics Approaches in Immunological Research. Frontiers in Immunology. 2021;12.

72. Hernández-Gea V, Campreciós G, Betancourt F, Pérez-Campuzano V, Seijo S, Díaz A, et al. Co-expression gene network analysis reveals novel regulatory pathways involved in porto-sinusoidal vascular disease. J Hepatol. 2021;75:924–34.

73. Frontiers | Integration of Immunome With Disease-Gene Network Reveals Common Cellular Mechanisms Between IMIDs and Drug Repurposing Strategies. https://www.frontiersin.org/articles/10.3389/fimmu.2021.669400/full. Accessed 3 Dec 2023.

74. Kumar V, Parate S, Danishuddin, Zeb A, Singh P, Lee G, et al. 3D-QSAR-Based Pharmacophore Modeling, Virtual Screening, and Molecular Dynamics Simulations for the Identification of Spleen Tyrosine Kinase Inhibitors. Front Cell Infect Microbiol. 2022;12:909111.

75. Martín-Villamil M, Sanmartín I, Moreno Á, Gallego J. Pharmacophore-Based Discovery of Viral RNA Conformational Modulators. Pharmaceuticals. 2022;15:748.

76. Pharmacophore modeling and its applications. https://www.researchgate.net/publication/362708089_Pharmacophore_modeling_and_its_applications. Accessed 3 Dec 2023.

77. Kaserer T, Beck KR, Akram M, Odermatt A, Schuster D. Pharmacophore Models and Pharmacophore-Based Virtual Screening: Concepts and Applications Exemplified on Hydroxysteroid Dehydrogenases. Molecules. 2015;20:22799–832.

78. Tran Q-H, Nguyen Q-T, Vo N-Q-H, Mai TT, Tran T-T-N, Tran T-D, et al. Structure-based 3D-Pharmacophore modeling to discover novel interleukin 6 inhibitors: An in silico screening, molecular dynamics simulations and binding free energy calculations. PLOS ONE. 2022;17:e0266632.

79. Calvo-Barreiro L, Talagayev V, Pach S, Abdel-Rahman SA, Wolber G, Gabr MT. Discovery of ICOS-Targeted Small Molecules Using Pharmacophore-Based Screening. ChemMedChem. n/a n/a:e202300305.

80. Wang Z, Sun L, Xu Y, Liang P, Xu K, Huang J. Discovery of novel JAK1 inhibitors through combining machine learning, structure-based pharmacophore modeling and bio-evaluation. Journal of Translational Medicine. 2023;21:579.

81. Tian Y-S, Kawashita N, Arai Y, Okamoto K, Takagi T. Pharmacophore Modeling and Molecular Docking Studies of potential inhibitors to E6 PBM–PDZ from Human Papilloma Virus (HPV). Bioinformation. 2015;11:401–6.

82. Tian Y-S, Kawashita N, Arai Y, Okamoto K, Takagi T. Pharmacophore Modeling and Molecular Docking Studies of potential inhibitors to E6 PBM-PDZ from Human Papilloma Virus (HPV). Bioinformation. 2015;11:401–6.

83. Bastikar V, Bastikar A, Gupta P. Quantitative structure–activity relationship-based computational approaches. Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV-2 Infection. 2022;:191–205.

84. Cai Z, Zafferani M, Akande OM, Hargrove AE. Quantitative Structure-Activity Relationship (QSAR) Study Predicts Small-Molecule Binding to RNA Structure. J Med Chem. 2022;65:7262–77.

85. Naik PK, Singh T, Singh H. Quantitative structure-activity relationship (QSAR) for insecticides: development of predictive in vivo insecticide activity models. SAR QSAR Environ Res. 2009;20:551–66.

86. Knowledge gaps in immune response and immunotherapy involving nanomaterials: Databases and artificial intelligence for material design - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0142961220307158. Accessed 2 Dec 2023.

87. Fukunishi Y, Yamasaki S, Yasumatsu I, Takeuchi K, Kurosawa T, Nakamura H. Quantitative Structure-activity Relationship (QSAR) Models for Docking Score Correction. Mol Inform. 2017;36:1600013.

88. Jawarkar RD, Sharma P, Jain N, Gandhi A, Mukerjee N, Al-Mutairi AA, et al. QSAR, Molecular Docking, MD Simulation and MMGBSA Calculations Approaches to Recognize Concealed Pharmacophoric Features Requisite for the Optimization of ALK Tyrosine Kinase Inhibitors as Anticancer Leads. Molecules. 2022;27:4951.

89. Sun L, Liu H, Ye Y, Lei Y, Islam R, Tan S, et al. Smart nanoparticles for cancer therapy. Sig Transduct Target Ther. 2023;8:1–28.

90. Molecules | Free Full-Text | A Multidisciplinary Approach to Coronavirus Disease (COVID-19). https://www.mdpi.com/1420-3049/26/12/3526. Accessed 2 Dec 2023.

91. Theoretical Modeling of HPV: QSAR and Novodesign with Fragment Approach. https://www.researchgate.net/publication/269180681_Theoretical_Modeling_of_HPV_QSAR_and_Novodesign_with_Fragment_Approach. Accessed 3 Dec 2023.

92. Sobral PS, Luz VCC, Almeida JMGCF, Videira PA, Pereira F. Computational Approaches Drive Developments in Immune-Oncology Therapies for PD-1/PD-L1 Immune Checkpoint Inhibitors. Int J Mol Sci. 2023;24:5908.

93. Shahab M, Al-Madhagi H, Zheng G, Zeb A, Alasmari AF, Alharbi M, et al. Structure based virtual screening and molecular simulation study of FDA-approved drugs to inhibit human HDAC6 and VISTA as dual cancer immunotherapy. Sci Rep. 2023;13:14466.

94. Albumin/vaccine nanocomplexes that assemble in vivo for combination cancer immunotherapy | Nature Communications. https://www.nature.com/articles/s41467-017-02191-y. Accessed 2 Dec 2023.

95. Drug screening and biomarker gene investigation in cancer therapy through the human transcriptional regulatory network - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2001037023000521. Accessed 2 Dec 2023.

96. Alabbas AB. Integrativesubtractive proteomics, immunoinformatics, docking, and simulation approaches reveal candidate vaccine against Sin Nombre orthohantavirus. Frontiers in Immunology. 2022;13.

97. Razzaghi-Asl N, Mirzayi S, Mahnam K, Adhami V, Sepehri S. In silico screening and molecular dynamics simulations toward new human papillomavirus 16 type inhibitors. Res Pharm Sci. 2022;17:189–208.

98. Vaccines | Free Full-Text | In-Silico Molecular Modeling Studies to Identify Novel Potential Inhibitors of HPV E6 Protein. https://www.mdpi.com/2076-393X/10/9/1452. Accessed 3 Dec 2023.

99. Screening of small molecular compounds with carcinogenic inhibition function of HPV-16 E6 - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1878535223002216. Accessed 3 Dec 2023.

100. Recent advances in targeting the “undruggable” proteins: from drug discovery to clinical trials | Signal Transduction and Targeted Therapy. https://www.nature.com/articles/s41392-023-01589-z. Accessed 2 Dec 2023.

101. Tan C, Zhu F, Pan P, Wu A, Li C. Development of multi-epitope vaccines against the monkeypox virus based on envelope proteins using immunoinformatics approaches. Frontiers in Immunology. 2023;14.

102. Riandono FD, Istyastono E. MOLECULAR DYNAMICS SIMULATIONS OF THE STK630921 INTERACTIONS TO INTERLEUKIN-17A. International Journal of Applied Pharmaceutics. 2023;:250–5.

103. IJMS | Free Full-Text | Insights into Interactions between Interleukin-6 and Dendritic Polyglycerols. https://www.mdpi.com/1422-0067/22/5/2415. Accessed 2 Dec 2023.

104. Povinelli APR, Zazeri G, Jones AM, Cornélio ML. A Computational–Experimental Investigation of the Molecular Mechanism of Interleukin-6-Piperine Interaction. Int J Mol Sci. 2022;23:7994.

105. Tran Q-H, Nguyen Q-T, Tran T-TN, Tran T-D, Le M-T, Trinh D-TT, et al. Identification of small molecules as potential inhibitors of interleukin 6: a multi-computational investigation. Mol Divers. 2022;:1–16.

106. Gupta M, Ha K, Agarwal R, Quarles LD, Smith JC. Molecular dynamics analysis of the binding of human interleukin-6 with interleukin-6 α-receptor. Proteins: Structure, Function, and Bioinformatics. 2021;89:163–73.

107. Luedemann M, Stadler D, Cheng C-C, Protzer U, Knolle PA, Donakonda S. Montelukast is a dual-purpose inhibitor of SARS-CoV-2 infection and virus-induced IL-6 expression identified by structure-based drug repurposing. Computational and Structural Biotechnology Journal. 2022;20:799–811.

108. 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. Frontiers in Immunology. 2018;9.

109. Molecular Docking and Molecular Dynamics Simulation-Based Identification of Natural Inhibitors against Druggable Human Papilloma Virus Type 16 Target | Trends in Sciences. https://tis.wu.ac.th/index.php/tis/article/view/4891. Accessed 3 Dec 2023.

110. Zhong L, Li Y, Xiong L, Wang W, Wu M, Yuan T, et al. Small molecules in targeted cancer therapy: advances, challenges, and future perspectives. Sig Transduct Target Ther. 2021;6:1–48.

111. Choi SB, Yap B, Choong YS, Wahab H. Molecular Dynamics Simulations in Drug Discovery. 2018. p. 652–65.

112. Xu Y, Qiu Y, Yuan S, Wang H. Prognostic implication of human papillomavirus types in cervical cancer patients: a systematic review and meta-analysis. Infect Agent Cancer. 2020;15:66.

113. Ansari HR, Flower DR, Raghava GPS. AntigenDB: an immunoinformatics database of pathogen antigens. Nucleic Acids Res. 2010;38 Database issue:D847-853.

114. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC1780059/. Accessed 29 Nov 2023.

115. Kolaskar AS, Tongaonkar PC. A semi-empirical method for prediction of antigenic determinants on protein antigens. FEBS Lett. 1990;276:172–4.

116. Singh H, Raghava GPS. ProPred1: prediction of promiscuous MHC Class-I binding sites. Bioinformatics. 2003;19:1009–14.

117. NetTepi: an integrated method for the prediction of T cell epitopes | Immunogenetics. https://link.springer.com/article/10.1007/s00251-014-0779-0. Accessed 2 Dec 2023.

118. Dhanda SK, Mahajan S, Paul S, Yan Z, Kim H, Jespersen MC, et al. IEDB-AR: immune epitope database—analysis resource in 2019. Nucleic Acids Res. 2019;47:W502–6.

119. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 2020;48:W449–54.

120. Bhasin M, Raghava GPS. Prediction of CTL epitopes using QM, SVM and ANN techniques. Vaccine. 2004;22:3195–204.

121. Stranzl T, Larsen MV, Lundegaard C, Nielsen M. NetCTLpan: pan-specific MHC class I pathway epitope predictions. Immunogenetics. 2010;62:357–68.

122. Saha S, Raghava GPS. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins. 2006;65:40–8.

123. Reliable B Cell Epitope Predictions: Impacts of Method Development and Improved Benchmarking | PLOS Computational Biology. https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1002829. Accessed 2 Dec 2023.

124. Sharma N, Patiyal S, Dhall A, Pande A, Arora C, Raghava GPS. AlgPred 2.0: an improved method for predicting allergenic proteins and mapping of IgE epitopes. Brief Bioinform. 2021;22:bbaa294.

125. Maurer-Stroh S, Krutz NL, Kern PS, Gunalan V, Nguyen MN, Limviphuvadh V, et al. AllerCatPro—prediction of protein allergenicity potential from the protein sequence. Bioinformatics. 2019;35:3020–7.

126. Rey J, Murail S, de Vries S, Derreumaux P, Tuffery P. PEP-FOLD4: a pH-dependent force field for peptide structure prediction in aqueous solution. Nucleic Acids Research. 2023;51:W432–7.

127. Kozakov D, Hall DR, Xia B, Porter KA, Padhorny D, Yueh C, et al. The ClusPro web server for protein–protein docking. Nat Protoc. 2017;12:255–78.

128. Cheng J, Randall AZ, Sweredoski MJ, Baldi P. SCRATCH: a protein structure and structural feature prediction server. Nucleic Acids Res. 2005;33 Web Server issue:W72–6.

129. CD137 - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/medicine-and-dentistry/cd137. Accessed 2 Dec 2023.

130. CD25 - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/immunology-and-microbiology/cd25. Accessed 2 Dec 2023.

131. Pathogens | Free Full-Text | Inflammatory and Immune Responses during SARS-CoV-2 Infection in Vaccinated and Non-Vaccinated Pregnant Women and Their Newborns. https://www.mdpi.com/2076-0817/12/5/664. Accessed 1 Dec 2023.

132. Vaccines | Free Full-Text | Effect of a Third Dose of SARS-CoV-2 mRNA BNT162b2 Vaccine on Humoral and Cellular Responses and Serum Anti-HLA Antibodies in Kidney Transplant Recipients. https://www.mdpi.com/2076-393X/10/6/921. Accessed 1 Dec 2023.

133. Fazeli M, Soleimanjahi H, Dadashzadeh S. Further Stimulation of Cellular Immune Responses through Association of HPV-16 E6, E7 and L1 Genes in order to produce more Effective Therapeutic DNA Vaccines in Cervical Cancer Model. Iran J Cancer Prev. 2015;8:18–23.

134. Kadish AS, Ho GYF, Burk RD, Romney SL, Angeletti RH, Wang Y, et al. Lymphoproliferative Responses to Human Papillomavirus (HPV) Type 16 Proteins E6 and E7: Outcome of HPV Infection and Associated Neoplasia. JNCI: Journal of the National Cancer Institute. 1997;89:1285–93.

135. Moser B. CXCR5, the Defining Marker for Follicular B Helper T (TFH) Cells. Front Immunol. 2015;6:296.

136. E. TZ, Shahar T, Leker R, Szalayova I, Bratincsák A, Key S, et al. Sensitive detection of GFP utilizing tyramide signal amplification to overcome gene silencing. Exp Cell Res. 2007;313:1943–50.

137. Polymers | Free Full-Text | Detection of Streptavidin Based on Terminal Protection and Cationic Conjugated Polymer-Mediated Fluorescence Resonance Energy Transfer. https://www.mdpi.com/2073-4360/13/5/725. Accessed 2 Dec 2023.

138. Wang C, Kang SG, Lee J, Sun Z, Kim CH. The roles of CCR6 in migration of Th17 cells and regulation of effector T-cell balance in the gut. Mucosal Immunol. 2009;2:173–83.

139. Coyle AJ, Lehar S, Lloyd C, Tian J, Delaney T, Manning S, et al. The CD28-related molecule ICOS is required for effective T cell-dependent immune responses. Immunity. 2000;13:95–105.

140. Groom JR, Luster AD. CXCR3 in T cell function. Exp Cell Res. 2011;317:620–31.

141. 3.1. Statistics in Python — Scientific Python Lectures. https://lectures.scientific-python.org/packages/statistics/index.html. Accessed 1 Dec 2023.

142. Data Structures for Statistical Computing in Python. https://www.researchgate.net/publication/340177686_Data_Structures_for_Statistical_Computing_in_Python. Accessed 1 Dec 2023.

143. Virtanen P, Gommers R, Oliphant TE, Haberland M, Reddy T, Cournapeau D, et al. SciPy 1.0: fundamental algorithms for scientific computing in Python. Nat Methods. 2020;17:261–72.

144. Harris CR, Millman KJ, van der Walt SJ, Gommers R, Virtanen P, Cournapeau D, et al. Array programming with NumPy. Nature. 2020;585:357–62.

145. Matplotlib: A 2D Graphics Environment | IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/4160265. Accessed 1 Dec 2023.

146. Journal of Open Source Software: seaborn: statistical data visualization. https://joss.theoj.org/papers/10.21105/joss.03021. Accessed 1 Dec 2023.

147. Seabold S, Perktold J. Statsmodels: Econometric and Statistical Modeling with Python. Austin, Texas; 2010. p. 92–6.

148. Journal of Open Source Software: Pingouin: statistics in Python. https://joss.theoj.org/papers/10.21105/joss.01026. Accessed 1 Dec 2023.

149. Dontje ML, Dall PM, Skelton DA, Gill JMR, Chastin SFM. Reliability, minimal detectable change and responsiveness to change: Indicators to select the best method to measure sedentary behaviour in older adults in different study designs. PLoS One. 2018;13:e0195424.

150. Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, Sayers EW. GenBank. Nucleic Acids Res. 2016;44 Database issue:D67–72.

151. The UniProt Consortium. UniProt: the Universal Protein Knowledgebase in 2023. Nucleic Acids Research. 2023;51:D523–31.

152. Clustal Omega for making accurate alignments of many protein sequences - Sievers - 2018 - Protein Science - Wiley Online Library. https://onlinelibrary.wiley.com/doi/10.1002/pro.3290. Accessed 1 Dec 2023.

153. Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation | Nucleic Acids Research | Oxford Academic.

https://academic.oup.com/nar/article/49/W1/W293/6246398?login=true. Accessed 1 Dec 2023.

154. Tamura K, Stecher G, Kumar S. MEGA11: Molecular Evolutionary Genetics Analysis Version 11. Molecular Biology and Evolution. 2021;38:3022–7.

155. Walker JM, editor. The Proteomics Protocols Handbook. Totowa, NJ: Humana Press; 2005.

156. Geourjon C, Deléage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Comput Appl Biosci. 1995;11:681–4.

157. Kelley LA, Mezulis S, Yates CM, Wass MN, Sternberg MJ. The Phyre2 web portal for protein modelling, prediction and analysis. Nat Protoc. 2015;10:845–58.

158. Bhattacharya D, Nowotny J, Cao R, Cheng J. 3Drefine: an interactive web server for efficient protein structure refinement. Nucleic Acids Research. 2016;44:gkw336.

159. Laskowski RA, Thornton JM. PDBsum extras: SARS-CoV-2 and AlphaFold models. Protein Sci. 2022;31:283–9.

160. Laskowski RA, Rullmannn JA, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK-NMR: programs for checking the quality of protein structures solved by NMR. J Biomol NMR. 1996;8:477–86.

161. Oliullah Rafi M, Al-Khafaji K, Takim Sarker M, Taskin-Tok T, Samad Rana A, Shahedur Rahman M. Design of a multi-epitope vaccine against SARS-CoV-2: immunoinformatic and computational methods. RSC Advances. 2022;12:4288–310.

162. Chauhan V, Rungta T, Goyal K, Singh MP. Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach. Sci Rep. 2019;9:2517.

163. Ebrahimi S, Mohabatkar H, Behbahani M. Predicting Promiscuous T Cell Epitopes for Designing a Vaccine Against Streptococcus pyogenes. Appl Biochem Biotechnol. 2019;187:90–100.

164. Identification of Conserved and HLA Promiscuous DENV3 T-Cell Epitopes | PLOS Neglected Tropical Diseases. https://journals.plos.org/plosntds/article?id=10.1371/journal.pntd.0002497. Accessed 1 Dec 2023.

165. Jespersen MC, Peters B, Nielsen M, Marcatili P. BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. Nucleic Acids Res. 2017;45 Web Server issue:W24–9.

166. Ponomarenko J, Bui H-H, 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.

167. Bui H-H, Sidney J, Li W, Fusseder N, Sette A. Development of an epitope conservancy analysis tool to facilitate the design of epitope-based diagnostics and vaccines. BMC Bioinformatics. 2007;8:361.

168. Predicting population coverage of T-cell epitope-based diagnostics and vaccines - PubMed. https://pubmed.ncbi.nlm.nih.gov/16545123/. Accessed 1 Dec 2023.

169. Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. Basic local alignment search tool. J Mol Biol. 1990;215:403–10.

170. Li H, Huang E, Zhang Y, Huang S-Y, Xiao Y. HDOCK update for modeling protein-RNA/DNA complex structures. Protein Science. 2022;31:e4441.

171. HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5793843/. Accessed 29 Nov 2023.

172. SWISS-MODEL: an automated protein homology-modeling server - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC168927/. Accessed 29 Nov 2023.

173. Maleki A, Russo G, Parasiliti Palumbo GA, Pappalardo F. In silico design of recombinant multi-epitope vaccine against influenza A virus. BMC Bioinformatics. 2022;22:617.

174. Wiederstein M, Sippl MJ. ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res. 2007;35 Web Server issue:W407–10.

175. 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:W526-531.

176. Urquiza-García U, Millar AJ. Expanding the bioluminescent reporter toolkit for plant science with NanoLUC. Plant Methods. 2019;15:68.

177. Armour CR, Sovacool KL, Close WL, Topçuoğlu BD, Wiens J, Schloss PD. Machine learning classification by fitting amplicon sequences to existing OTUs. mSphere. 2023;8:e0033623.

178. Khan A, Khan S, Saleem S, Nizam-Uddin N, Mohammad A, Khan T, et al. Immunogenomics guided design of immunomodulatory multi-epitope subunit vaccine against the SARS-CoV-2 new variants, and its validation through in silico cloning and immune simulation. Comput Biol Med. 2021;133:104420.

179. Kumar A, Rathi E, Kini SG. Computational design of a broad-spectrum multi-epitope vaccine candidate against seven strains of human coronaviruses. 3 Biotech. 2022;12:240.

180. Identification and validation of viral antigens sharing sequence and structural homology with tumor-associated antigens (TAAs). | Journal for ImmunoTherapy of Cancer. https://jitc.bmj.com/content/9/5/e002694. Accessed 30 Nov 2023.

181. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, et al. NCBI GEO: archive for functional genomics data sets—update. Nucleic Acids Research. 2013;41:D991–5.

182. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res. 2002;30:207–10.

183. Ndiaye M, Diop G, Derbois C, Spadoni J-L, Noirel J, Medina-Santos R, et al. Gene expression profiling of peripheral blood mononuclear cells from women with cervical lesions reveals new markers of cancer. Oncology Reports. 2023;49:1–9.

184. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al. STRING v10: protein-protein interaction networks, integrated over the tree of life. Nucleic Acids Res. 2015;43 Database issue:D447-452.

185. KEGG as a reference resource for gene and protein annotation | Nucleic Acids Research | Oxford Academic. https://academic.oup.com/nar/article/44/D1/D457/2502600. Accessed 30 Nov 2023.

186. Cline MS, Smoot M, Cerami E, Kuchinsky A, Landys N, Workman C, et al. Integration of biological networks and gene expression data using Cytoscape. Nat Protoc. 2007;2:2366–82.

187. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.

188. Bastian M, Heymann S, Jacomy M. Gephi: An Open Source Software for Exploring and Manipulating Networks. Proceedings of the International AAAI Conference on Web and Social Media. 2009;3:361–2.

189. Chen EY, Tan CM, Kou Y, Duan Q, Wang Z, Meirelles GV, et al. Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. BMC Bioinformatics. 2013;14:128.

190. PDBe: improved findability of macromolecular structure data in the PDB | Nucleic Acids Research | Oxford Academic. https://academic.oup.com/nar/article/48/D1/D335/5613681. Accessed 30 Nov 2023.

191. Lagunas-Rangel FA, Liao S, Williams MJ, Trukhan V, Fredriksson R, Schiöth HB. Drosophila as a Rapid Screening Model to Evaluate the Hypoglycemic Effects of

Dipeptidyl Peptidase 4 (DPP4) Inhibitors: High Evolutionary Conservation of DPP4. Biomedicines. 2023;11:3032.

192. ZINC20-A Free Ultralarge-Scale Chemical Database for Ligand Discovery - PubMed. https://pubmed.ncbi.nlm.nih.gov/33118813/. Accessed 30 Nov 2023.

193. S K, J C, T C, A G, J H, S H, et al. PubChem 2023 update. Nucleic acids research. 2023;51.

194. Liu T, Lin Y, Wen X, Jorissen RN, Gilson MK. BindingDB: a web-accessible database of experimentally determined protein–ligand binding affinities. Nucleic Acids Res. 2007;35 Database issue:D198–201.

195. Marvin Desktop Suite History of Changes | Chemaxon Docs. https://docs.chemaxon.com/display/docs/marvin-desktop-suite-history-of-changes.md. Accessed 30 Nov 2023.

196. Yap CW. PaDEL-Descriptor: An Open Source Software to Calculate Molecular Descriptors and Fingerprints. Journal of computational chemistry. 2011;32:1466–74.

197. Ganji M, Bakhshi S, Shoari A, Ahangari Cohan R. Discovery of potential FGFR3 inhibitors via QSAR, pharmacophore modeling, virtual screening and molecular docking studies against bladder cancer. J Transl Med. 2023;21:111.

198. Krishnapuram B, Carin L, Figueiredo MAT, Hartemink AJ. Sparse multinomial logistic regression: fast algorithms and generalization bounds. IEEE Trans Pattern Anal Mach Intell. 2005;27:957–68.

199. Nunes CA, Freitas MP, Pinheiro ACM, Bastos SC. Chemoface: a novel free user-friendly interface for chemometrics. J Braz Chem Soc. 2012;23:2003–10.

200. Fernandes ÍA, Braga Resende D, Ramalho TC, Kuca K, da Cunha EFF. Theoretical Studies Aimed at Finding FLT3 Inhibitors and a Promising Compound and Molecular Pattern with Dual Aurora B/FLT3 Activity. Molecules. 2020;25:1726.

201. PHASE: a new engine for pharmacophore perception, 3D QSAR model development, and 3D database screening: 1. Methodology and preliminary results | Journal of Computer-Aided Molecular Design. https://link.springer.com/article/10.1007/s10822-006-9087-6. Accessed 30 Nov 2023.

202. Kirubakaran P, Muthusamy K, Singh KHD, Nagamani S. Ligand-based Pharmacophore Modeling; Atom-based 3D-QSAR Analysis and Molecular Docking Studies of Phosphoinositide-Dependent Kinase-1 Inhibitors. Indian Journal of Pharmaceutical Sciences. 2012;74:141.

203. Amengor CDK, Orman E, Danquah CA, Ben IO, Biniyam PD, Harley BK. Pyridine-N-Oxide Alkaloids from Allium stipitatum and Their Synthetic Disulfide Analogs as Potential Drug Candidates against Mycobacterium tuberculosis: A

Molecular Docking, QSBAR, and ADMET Prediction Approach. Biomed Res Int. 2022;2022:6261528.

204. SwissParam 2023: A Modern Web-Based Tool for Efficient Small Molecule Parametrization | Journal of Chemical Information and Modeling. https://pubs.acs.org/doi/10.1021/acs.jcim.3c01053. Accessed 30 Nov 2023.

205. SwissParam: A fast force field generation tool for small organic molecules - Zoete - 2011 - Journal of Computational Chemistry - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1002/jcc.21816. Accessed 30 Nov 2023.

206. Molecular dynamics simulation of a small protein using GROMACS — GROMACS tutorials https://tutorials.gromacs.org documentation. https://tutorials.gromacs.org/docs/md-intro-tutorial.html. Accessed 30 Nov 2023.

207. Wall ME, Calabró G, Bayly CI, Mobley DL, Warren GL. Biomolecular Solvation Structure Revealed by Molecular Dynamics Simulations. J Am Chem Soc. 2019;141:4711–20.

208. NOSÉ S. A molecular dynamics method for simulations in the canonical ensemble. Molecular Physics. 2002;100:191–8.

209. Rao P, Shukla A, Parmar P, Rawal RM, Patel BV, Saraf M, et al. Proposing a fungal metabolite-flaviolin as a potential inhibitor of 3CLpro of novel coronavirus SARS-CoV-2 identified using docking and molecular dynamics. J Biomol Struct Dyn. 2022;40:348–60.

210. Rao P, Shukla A, Parmar P, Rawal RM, Patel B, Saraf M, et al. Reckoning a fungal metabolite, Pyranonigrin A as a potential Main protease (Mpro) inhibitor of novel SARS-CoV-2 virus identified using docking and molecular dynamics simulation. Biophysical Chemistry. 2020;264:106425.

211. McCormack SE, Cruz CRY, Wright KE, Powell AB, Lang H, Trimble C, et al. HPV-specific T-cells can be generated from naïve T cells for use as immunotherapeutic strategy for immunocompromised patients. Cytotherapy. 2018;20:385–93.

212. Keselman HJ, Rogan JC. The Tukey multiple comparison test: 1953–1976. Psychological Bulletin. 1977;84:1050–6.

213. Frontiers | Protein pI and Intracellular Localization. https://www.frontiersin.org/articles/10.3389/fmolb.2021.775736/full. Accessed 29 Nov 2023.

214. (PDF) Applicability of Instability Index for In vitro Protein Stability Prediction. https://www.researchgate.net/publication/331439870_Applicability_of_Instability_Index_for_In_vitro_Protein_Stability_Prediction. Accessed 29 Nov 2023.

215. Sha Z, Zhao J, Goldberg AL. Measuring the overall rate of protein breakdown in cells and the contributions of the ubiquitin-proteasome and autophagy-lysosomal pathways. Methods Mol Biol. 2018;1844:261–76.

216. Thermostability and Aliphatic Index of Globular Proteins. The Journal of Biochemistry. 1980. https://doi.org/10.1093/oxfordjournals.jbchem.a133168.

217. Gravy index score (average hydrophobicity and hydrophilicity) of colon... ResearchGate. https://www.researchgate.net/figure/Gravy-index-score-average-hydrophobicity-and-hydrophilicity-of-colon-proteins-measured_fig3_230755139. Accessed 29 Nov 2023.

218. Larsen MV, Lundegaard C, Lamberth K, Buus S, Lund O, Nielsen M. Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction. BMC Bioinformatics. 2007;8:424.

219. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000;16:276–7.

220. Kamens J. The Addgene repository: an international nonprofit plasmid and data resource. Nucleic Acids Research. 2015;43:D1152–7.

221. Čokić VP, Mitrović-Ajtić O, Beleslin-Čokić BB, Marković D, Buač M, Diklić M, et al. Proinflammatory Cytokine IL-6 and JAK-STAT Signaling Pathway in Myeloproliferative Neoplasms. Mediators Inflamm. 2015;2015:453020.

222. Chatr-aryamontri A, Ceol A, Peluso D, Nardozza A, Panni S, Sacco F, et al. VirusMINT: a viral protein interaction database. Nucleic Acids Res. 2009;37 Database issue:D669–73.

223. Piñero J, Bravo À, Queralt-Rosinach N, Gutiérrez-Sacristán A, Deu-Pons J, Centeno E, et al. DisGeNET: a comprehensive platform integrating information on human disease-associated genes and variants. Nucleic Acids Research. 2017;45 Database issue:D833.

224. Drozdetskiy A, Cole C, Procter J, Barton GJ. JPred4: a protein secondary structure prediction server. Nucleic Acids Res. 2015;43 Web Server issue:W389–94.

225. Wishart DS, Knox C, Guo AC, Shrivastava S, Hassanali M, Stothard P, et al. DrugBank: a comprehensive resource for in silico drug discovery and exploration. Nucleic Acids Res. 2006;34 Database issue:D668–72.

226. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017;7:42717.

227. Matsui K, Adelsberger JW, Kemp TJ, Baseler MW, Ledgerwood JE, Pinto LA. Circulating $CXCR5^{+}CD4^{+}$ T Follicular-Like Helper Cell and Memory B Cell Responses to Human Papillomavirus Vaccines. PLoS One. 2015;10:e0137195.

228. Pinto LA, Castle PE, Roden RB, Harro CD, Lowy DR, Schiller JT, et al. HPV-16 L1 VLP vaccine elicits a broad-spectrum of cytokine responses in whole blood. Vaccine. 2005;23:3555–64.

229. Einstein MH, Baron M, Levin MJ, Chatterjee A, Fox B, Scholar S, et al. Comparison of the immunogenicity of the human papillomavirus (HPV)-16/18 vaccine and the HPV-6/11/16/18 vaccine for oncogenic non-vaccine types HPV-31 and HPV-45 in healthy women aged 18–45 years. Hum Vaccin. 2011;7:1359–73.

230. Pasmans H, Berkowska MA, Diks AM, de Mooij B, Groenland RJ, de Rond L, et al. Characterization of the early cellular immune response induced by HPV vaccines. Front Immunol. 2022;13:863164.

231. Su L, Zhang Y, Zhang X, Liu T, Liu S, Li Y, et al. Combination immunotherapy with two attenuated Listeria strains carrying shuffled HPV-16 E6E7 protein causes tumor regression in a mouse tumor model. Sci Rep. 2021;11:13404.

232. Gunn GR, Zubair A, Peters C, Pan Z-K, Wu T-C, Paterson Y. Two Listeria monocytogenes Vaccine Vectors That Express Different Molecular Forms of Human Papilloma Virus-16 (HPV-16) E7 Induce Qualitatively Different T Cell Immunity That Correlates with Their Ability to Induce Regression of Established Tumors Immortalized by HPV-161. The Journal of Immunology. 2001;167:6471–9.

233. Jia Y, Yin Y, Duan F, Fu H, Hu M, Gao Y, et al. Prophylactic and therapeutic efficacy of an attenuated Listeria monocytogenes-based vaccine delivering HPV16 E7 in a mouse model. International Journal of Molecular Medicine. 2012;30:1335–42.

234. Sullivan GM, Feinn R. Using Effect Size—or Why the P Value Is Not Enough. J Grad Med Educ. 2012;4:279–82.

235. How to calculate, use, and report variance explained effect size indices and not die trying - Iacobucci - 2023 - Journal of Consumer Psychology - Wiley Online Library. https://myscp.onlinelibrary.wiley.com/doi/10.1002/jcpy.1292. Accessed 9 Dec 2023.

236. What Does Effect Size Tell You? 2022. https://www.simplypsychology.org/effect-size.html. Accessed 9 Dec 2023.

237. Kovacs FM, Abraira V, Royuela A, Corcoll J, Alegre L, Tomás M, et al. Minimum detectable and minimal clinically important changes for pain in patients with nonspecific neck pain. BMC Musculoskelet Disord. 2008;9:43.

238. HPV16 E6-specific T cell response and HLA-A alleles are related to the prognosis of patients with cervical cancer - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8447512/. Accessed 10 Dec 2023.

239. Welters MJP, Ma W, Santegoets SJAM, Goedemans R, Ehsan I, Jordanova ES, et al. Intratumoral HPV16-Specific T Cells Constitute a Type I–Oriented Tumor Microenvironment to Improve Survival in HPV16-Driven Oropharyngeal Cancer. Clinical Cancer Research. 2018;24:634–47.

240. Schiller JT, Castellsagué X, Villa LL, Hildesheim A. An Update of Prophylactic Human Papillomavirus L1 Virus-Like Particle Vaccine Clinical Trial Results. Vaccine. 2008;26 Suppl 10:K53–61.

241. Smolen KK, Gelinas L, Franzen L, Dobson S, Dawar M, Ogilvie G, et al. Age of recipient and number of doses differentially impact human B and T cell immune memory responses to HPV vaccination. Vaccine. 2012;30:3572–9.

242. Yang A, Farmer E, Lin J, Wu T-C, Hung C-F. The current state of therapeutic and T cell-based vaccines against human papillomaviruses. Virus Res. 2017;231:148–65.

243. Welters MJP, Kenter GG, Piersma SJ, Vloon APG, Löwik MJG, Berends-van der Meer DMA, et al. Induction of tumor-specific CD4+ and CD8+ T-cell immunity in cervical cancer patients by a human papillomavirus type 16 E6 and E7 long peptides vaccine. Clin Cancer Res. 2008;14:178–87.

244. Ramos CA, Narala N, Vyas GM, Leen AM, Gerdemann U, Sturgis EM, et al. Human Papillomavirus Type 16 E6/E7-Specific Cytotoxic T Lymphocytes for Adoptive Immunotherapy of HPV-Associated Malignancies. J Immunother. 2013;36:66–76.

245. Bhatt KH, Neller MA, Srihari S, Crooks P, Lekieffre L, Aftab BT, et al. Profiling HPV-16–specific T cell responses reveals broad antigen reactivities in oropharyngeal cancer patients. J Exp Med. 2020;217:e20200389.

246. Aggarwal C, Cohen RB, Morrow MP, Kraynyak KA, Sylvester AJ, Cheung J, et al. Immune Therapy Targeting E6/E7 Oncogenes of Human Papillomavirus Type 6 (HPV-6) Reduces or Eliminates the Need for Surgical Intervention in the Treatment of HPV-6 Associated Recurrent Respiratory Papillomatosis. Vaccines (Basel). 2020;8:56.

247. Nikbakht M, Pakbin B, Nikbakht Brujeni G. Evaluation of a new lymphocyte proliferation assay based on cyclic voltammetry; an alternative method. Sci Rep. 2019;9:4503.

248. Ye M, Li S, Luo P, Tang X, Gong Q, Mei B. Genetic variation of E6, E7, and L1 genes of human papillomavirus 51 from central China. J Med Virol. 2022;94:2811–23.

249. Rojas-Cruz AF, Reyes-Bermúdez A. Phylogenetic analysis of Alphapapillomavirus based on L1, E6 and E7 regions suggests that carcinogenicity and tissue tropism have appeared multiple times during viral evolution. Infection, Genetics and Evolution. 2019;67:210–21.

250. Chen Z, Schiffman M, Herrero R, DeSalle R, Anastos K, Segondy M, et al. Classification and Evolution of Human Papillomavirus Genome Variants: Alpha-5 (HPV26, 51, 69, 82), Alpha-6 (HPV30, 53, 56, 66), Alpha-11 (HPV34, 73), Alpha-13 (HPV54) and Alpha-3 (HPV61). Virology. 2018;516:86–101.

251. Dey S, De A, Nandy A. Rational Design of Peptide Vaccines Against Multiple Types of Human Papillomavirus. Cancer Inform. 2016;15 Suppl 1:1–16.

252. Namvar A, Panahi HA, Agi E, Bolhassani A. Development of HPV16,18,31,45 E5 and E7 peptides-based vaccines predicted by immunoinformatics tools. Biotechnol Lett. 2020;42:403–18.

253. 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;9:15225.

254. 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. Frontiers in public health. 2021;8.

255. Van Doorslaer K, Burk RD. Evolution of Human Papillomavirus Carcinogenicity. Adv Virus Res. 2010;77:41–62.

256. Yousefi Z, Aria H, Ghaedrahmati F, Bakhtiari T, Azizi M, Bastan R, et al. An Update on Human Papilloma Virus Vaccines: History, Types, Protection, and Efficacy. Front Immunol. 2022;12:805695.

257. Lee HS, Kim MW, Jin KS, Shin H-C, Kim WK, Lee SC, et al. Molecular Analysis of the Interaction between Human PTPN21 and the Oncoprotein E7 from Human Papillomavirus Genotype 18. Mol Cells. 2021;44:26–37.

258. White EA, Kramer RE, Tan MJA, Hayes SD, Harper JW, Howley PM. Comprehensive Analysis of Host Cellular Interactions with Human Papillomavirus E6 Proteins Identifies New E6 Binding Partners and Reflects Viral Diversity. J Virol. 2012;86:13174–86.

259. Papillomavirus E6 proteins - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0042682208007253. Accessed 27 Nov 2023.

260. Cosper PF, Hrycyniak LCF, Paracha M, Lee DL, Wan J, Jones K, et al. HPV16 E6 induces chromosomal instability due to polar chromosomes caused by E6AP-dependent degradation of the mitotic kinesin CENP-E. Proceedings of the National Academy of Sciences. 2023;120:e2216700120.

261. Wang L, Ye X, Zhao T. The physiological roles of autophagy in the mammalian life cycle. Biol Rev Camb Philos Soc. 2019;94:503–16.

262. Roman A, Munger K. The papillomavirus E7 proteins. Virology. 2013;445:138–68.

263. Illiano E, Demurtas OC, Massa S, Di Bonito P, Consalvi V, Chiaraluce R, et al. Production of functional, stable, unmutated recombinant human papillomavirus E6 oncoprotein: implications for HPV-tumor diagnosis and therapy. Journal of Translational Medicine. 2016;14:224.

264. Xu ML, Kim SC, Kim HJ, Ju W, Kim YH, Kim H-J. Two-step chromatographic purification of glutathione S-transferase-tagged human papillomavirus type 16 E6 protein and its application for serology. Protein Expression and Purification. 2017;132:19–26.

265. Liu Y, Cherry JJ, Dineen JV, Androphy EJ, Baleja JD. Determinants of Stability for the E6 Protein of Papillomavirus Type 16. Journal of Molecular Biology. 2009;386:1123–37.

266. Haghshenas MR, Mousavi T, Kheradmand M, Afshari M, Moosazadeh M. Efficacy of Human Papillomavirus L1 Protein Vaccines (Cervarix and Gardasil) in Reducing the Risk of Cervical Intraepithelial Neoplasia: A Meta-analysis. Int J Prev Med. 2017;8:44.

267. Boilesen DR, Nielsen KN, Holst PJ. Novel Antigenic Targets of HPV Therapeutic Vaccines. Vaccines (Basel). 2021;9:1262.

268. Schiller J, Dillner J. Chapter 5. Evaluation of durability of protection of HPV prophylactic vaccines. In: Primary End-points for Prophylactic HPV Vaccine Trials. International Agency for Research on Cancer; 2014.

269. Han L, Zhang B. Can prophylactic HPV vaccination reduce the recurrence of cervical lesions after surgery? Review and prospect. Infectious Agents and Cancer. 2023;18:66.

270. Mo Y, Ma J, Zhang H, Shen J, Chen J, Hong J, et al. Prophylactic and Therapeutic HPV Vaccines: Current Scenario and Perspectives. Frontiers in Cellular and Infection Microbiology. 2022;12.

271. Chiesa IJ, Perez MS, Nuñez GG, Pirola DA. Genetic variability and phylogeny analysis of partial L1 gene of human papillomavirus variants in Buenos Aires, Argentina. Virusdisease. 2016;27:41–7.

272. Mobini Kesheh M, Shavandi S, Azami J, Esghaei M, Keyvani H. Genetic diversity and bioinformatic analysis in the L1 gene of HPV genotypes 31, 33, and 58 circulating in women with normal cervical cytology. Infectious Agents and Cancer. 2023;18:19.

273. Farmer E, Cheng MA, Hung C-F, Wu T-C. Vaccination Strategies for the Control and Treatment of HPV Infection and HPV-Associated Cancer. Recent Results Cancer Res. 2021;217:157–95.

274. Podtelezhnikov AA, Wild DL. Reconstruction and Stability of Secondary Structure Elements in the Context of Protein Structure Prediction. Biophys J. 2009;96:4399–408.

275. Deller MC, Kong L, Rupp B. Protein stability: a crystallographer's perspective. Acta Crystallogr F Struct Biol Commun. 2016;72 Pt 2:72–95.

276. de Brevern AG. Impact of protein dynamics on secondary structure prediction. Biochimie. 2020;179:14–22.

277. Nominé Y, Masson M, Charbonnier S, Zanier K, Ristriani T, Deryckère F, et al. Structural and Functional Analysis of E6 Oncoprotein: Insights in the Molecular Pathways of Human Papillomavirus-Mediated Pathogenesis. Molecular Cell. 2006;21:665–78.

278. Chen H-S, Conway MJ, Christensen ND, Alam S, Meyers C. Papillomavirus Capsid Proteins Mutually Impact Structure. Virology. 2011;412:378–83.

279. McLaughlin-Drubin ME, Münger K. The Human Papillomavirus E7 Oncoprotein. Virology. 2009;384:335–44.

280. Pal A, Kundu R. Human Papillomavirus E6 and E7: The Cervical Cancer Hallmarks and Targets for Therapy. Frontiers in Microbiology. 2020;10.

281. Chen Y-R, Peng S-L, Tsay Y-W. Protein Secondary Structure Prediction Based on Ramachandran Maps. In: Huang D-S, Wunsch DC, Levine DS, Jo K-H, editors. Advanced Intelligent Computing Theories and Applications. With Aspects of Theoretical and Methodological Issues. Berlin, Heidelberg: Springer; 2008. p. 204–11.

282. Revisiting the Ramachandran plot based on statistical analysis of static and dynamic characteristics of protein structures - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1047847723000023. Accessed 27 Nov 2023.

283. Hollingsworth SA, Karplus PA. A fresh look at the Ramachandran plot and the occurrence of standard structures in proteins. Biomol Concepts. 2010;1:271–83.

284. Ho BK, Brasseur R. The Ramachandran plots of glycine and pre-proline. BMC Struct Biol. 2005;5:14.

285. Ramachandran Plot - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/biochemistry-genetics-and-molecular-biology/ramachandran-plot. Accessed 26 Nov 2023.

286. Sanami S, Rafieian-Kopaei M, Dehkordi KA, Pazoki-Toroudi H, Azadegan-Dehkordi F, Mobini G-R, et al. In silico design of a multi-epitope vaccine against HPV16/18. BMC Bioinformatics. 2022;23:311.

287. Jain S, Baranwal M. Conserved peptide vaccine candidates containing multiple Ebola nucleoprotein epitopes display interactions with diverse HLA molecules. Med Microbiol Immunol. 2019;208:227–38.

288. Sahay A, Piprodhe A, Pise M. In silico analysis and homology modeling of strictosidine synthase involved in alkaloid biosynthesis in catharanthus roseus. J Genet Eng Biotechnol. 2020;18:44.

289. Singh A, Thakur M, Sharma LK, Chandra K. Designing a multi-epitope peptide based vaccine against SARS-CoV-2. Sci Rep. 2020;10:16219.

290. Naz A, Shahid F, Butt TT, Awan FM, Ali A, Malik A. Designing Multi-Epitope Vaccines to Combat Emerging Coronavirus Disease 2019 (COVID-19) by Employing Immuno-Informatics Approach. Frontiers in Immunology. 2020;11.

291. Mahmoodi S, Amirzakaria JZ, Ghasemian A. In silico design and validation of a novel multi-epitope vaccine candidate against structural proteins of Chikungunya virus using comprehensive immunoinformatics analyses. PLoS One. 2023;18:e0285177.

292. Imon RR, Samad A, Alam R, Alsaiari AA, Talukder MdEK, Almehmadi M, et al. Computational formulation of a multiepitope vaccine unveils an exceptional prophylactic candidate against Merkel cell polyomavirus. Frontiers in Immunology. 2023;14.

293. Banerjee S, Majumder K, Gutierrez GJ, Gupta D, Mittal B. Immuno-informatics approach for multi-epitope vaccine designing against SARS-CoV-2. bioRxiv. 2020;:2020.07.23.218529.

294. Samad A, Ahammad F, Nain Z, Alam R, Imon RR, Hasan M, et al. Designing a multi-epitope vaccine against SARS-CoV-2: an immunoinformatics approach. J Biomol Struct Dyn. :1–17.

295. Hossain MdS, Hossan MI, Mizan S, Moin AT, Yasmin F, Akash A-S, et al. Immunoinformatics approach to designing a multi-epitope vaccine against Saint Louis Encephalitis Virus. Informatics in Medicine Unlocked. 2021;22:100500.

296. Westerhout J, Krone T, Snippe A, Babé L, McClain S, Ladics GS, et al. Allergenicity prediction of novel and modified proteins: Not a mission impossible! Development of a Random Forest allergenicity prediction model. Regulatory Toxicology and Pharmacology. 2019;107:104422.

297. Shahab M, Alzahrani AK, Duan X, Aslam M, Abida, Imran M, et al. An Immunoinformatics Approach to Design Novel and Potent Multi-Epitope-Based Vaccine to Target Lumpy Skin Disease. Biomedicines. 2023;11.

298. Peng C, Tang F, Wang J, Cheng P, Wang L, Gong W. Immunoinformatic-Based Multi-Epitope Vaccine Design for Co-Infection of Mycobacterium tuberculosis and SARS-CoV-2. J Pers Med. 2023;13:116.

299. Epitope - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/pharmacology-toxicology-and-pharmaceutical-science/epitope. Accessed 26 Nov 2023.

300. Zhang L, Skolnick J. What should the Z-score of native protein structures be? Protein Sci. 1998;7:1201–7.

301. Impact of C-Score and Z-Score in protein modelling. https://www.biostars.org/p/99924/. Accessed 26 Nov 2023.

302. Klinman DM, Barnhart KM, Conover J. CpG motifs as immune adjuvants. Vaccine. 1999;17:19–25.

303. TLR9 - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/immunology-and-microbiology/tlr9. Accessed 26 Nov 2023.

304. Lamphier MS, Sirois CM, Verma A, Golenbock DT, Latz E. TLR9 and the Recognition of Self and Non-Self Nucleic Acids. Annals of the New York Academy of Sciences. 2006;1082:31–43.

305. Amadio R, Piperno GM, Benvenuti F. Self-DNA Sensing by cGAS-STING and TLR9 in Autoimmunity: Is the Cytoskeleton in Control? Frontiers in Immunology. 2021;12.

306. Toll-like receptor 9, CpG DNA and innate immunity - PubMed. https://pubmed.ncbi.nlm.nih.gov/12243247/. Accessed 26 Nov 2023.

307. Kawasaki T, Kawai T. Toll-Like Receptor Signaling Pathways. Frontiers in Immunology. 2014;5.

308. Xu X, Wang G, Ai L, Shi J, Zhang J, Chen Y-X. Melatonin suppresses TLR9-triggered proinflammatory cytokine production in macrophages by inhibiting ERK1/2 and AKT activation. Sci Rep. 2018;8:15579.

309. Bang HB, Lee YH, Lee YJ, Jeong KJ. High-Level Production of Human Papillomavirus (HPV) Type 16 L1 in Escherichia coli. 2016;26:356–63.

310. Liu Y-Y, Zhang H-J, Shen E-C, Chen D, Wang Y, Fu S-J, et al. Stability of trivalent human papillomavirus (types 16, 18, 58) recombinant vaccine (Escherichia coli). Chin Med J (Engl). 2021;134:3020–2.

311. XU J, BAI X, WANG LB, SHI HN, VAN DER GIESSEN JWB, BOIREAU P, et al. Immune responses in mice vaccinated with a DNA vaccine expressing serine protease-like protein from the new-born larval stage of Trichinella spiralis. Parasitology. 2017;144:712–9.

312. Mahapatra SR, Dey J, Kaur T, Sarangi R, Bajoria AA, Kushwaha GS, et al. Immunoinformatics and molecular docking studies reveal a novel Multi-Epitope peptide vaccine against pneumonia infection. Vaccine. 2021;39:6221–37.

313. Jiang F, Han Y, Liu Y, Xue Y, Cheng P, Xiao L, et al. A comprehensive approach to developing a multi-epitope vaccine against Mycobacterium tuberculosis: from in silico design to in vitro immunization evaluation. Frontiers in Immunology. 2023;14.

314. Wang K, Zhou L, Chen T, Li Q, Li J, Liu L, et al. Engineering for an HPV 9-valent vaccine candidate using genomic constitutive over-expression and low lipopolysaccharide levels in Escherichia coli cells. Microbial Cell Factories. 2021;20:227.

315. Wei M, Wang D, Li Z, Song S, Kong X, Mo X, et al. N-terminal truncations on L1 proteins of human papillomaviruses promote their soluble expression in Escherichia coli and self-assembly in vitro. Emerg Microbes Infect. 2018;7:160.

316. Madanagopal P, Muthusamy S, Pradhan SN, Prince PR. Construction and validation of a multi-epitope in silico vaccine model for lymphatic filariasis by targeting Brugia malayi: a reverse vaccinology approach. Bull Natl Res Cent. 2023;47:47.

317. Frontiers | In-silico evaluation of adenoviral COVID-19 vaccination protocols: Assessment of immunological memory up to 6 months after the third dose. https://www.frontiersin.org/articles/10.3389/fimmu.2022.998262/full. Accessed 26 Nov 2023.

318. In silico vaccine design: A tutorial in immunoinformatics - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S2772442522000156. Accessed 26 Nov 2023.

319. Huang S-H, Chen Y-T, Lin X-Y, Ly Y-Y, Lien S-T, Chen P-H, et al. In silico prediction of immune-escaping hot spots for future COVID-19 vaccine design. Sci Rep. 2023;13:13468.

320. Fraussen J. IgM responses following SARS-CoV-2 vaccination: insights into protective and pre-existing immunity. EBioMedicine. 2022;77:103922.

321. Immunoglobulin M Antibody - an overview | ScienceDirect Topics. https://www.sciencedirect.com/topics/neuroscience/immunoglobulin-m-antibody. Accessed 26 Nov 2023.

322. Toh ZQ, He L, Chen C, Huang A, Russell FM, Garland SM, et al. Measurement of Human Papillomavirus-Specific Antibodies Using a Pseudovirion-Based ELISA Method. Front Immunol. 2020;11:585768.

323. Syrjänen S, Waterboer T, Rintala M, Pawlita M, Syrjänen K, Louvanto K, et al. Maternal HPV-antibodies and seroconversion to HPV in children during the first 3 years of life. Sci Rep. 2022;12:2227.

324. Full article: Review of long-term immunogenicity following HPV vaccination: Gaps in current knowledge. https://www.tandfonline.com/doi/full/10.1080/21645515.2021.1908059. Accessed 26 Nov 2023.

325. Pattyn J, Van Keer S, Tjalma W, Matheeussen V, Van Damme P, Vorsters A. Infection and vaccine-induced HPV-specific antibodies in cervicovaginal secretions. A review of the literature. Papillomavirus Research. 2019;8:100185.

326. Handisurya A, Schellenbacher C, Haitel A, Senger T, Kirnbauer R. Human papillomavirus vaccination induces neutralising antibodies in oral mucosal fluids. Br J Cancer. 2016;114:409–16.

327. Wang Z-H, Kjellberg L, Abdalla H, Wiklund F, Wiklund C, Knekt P, et al. Type Specificity and Significance of Different Isotypes of Serum Antibodies to Human Papillomavirus Capsids. The Journal of Infectious Diseases. 2000;181:456–62.

328. Roy V, Jung W, Linde C, Coates E, Ledgerwood J, Costner P, et al. Differences in HPV-specific antibody Fc-effector functions following Gardasil® and Cervarix® vaccination. npj Vaccines. 2023;8:1–11.

329. Hancock G, Blight J, Lopez-Camacho C, Kopycinski J, Pocock M, Byrne W, et al. A multi-genotype therapeutic human papillomavirus vaccine elicits potent T cell responses to conserved regions of early proteins. Sci Rep. 2019;9:18713.

330. Wang X, Che Y, Chen B, Zhang Y, Nakagawa M, Wang X. Evaluation of immune responses induced by a novel human papillomavirus type 16 E7 peptide-based vaccine with Candida skin test reagent as an adjuvant in C57BL/6 mice. Int Immunopharmacol. 2018;56:249–60.

331. Wagstaffe HR, Mooney JP, Riley EM, Goodier MR. Vaccinating for natural killer cell effector functions. Clin Transl Immunology. 2018;7:e1010.

332. Rydyznski CE, Waggoner SN. Boosting vaccine efficacy the natural (killer) way. Trends in immunology. 2015;36:536.

333. Alrubayyi A, Touizer E, Hameiri-Bowen D, Charlton B, Gea-Mallorquí E, Hussain N, et al. Natural killer cell responses during SARS-CoV-2 infection and vaccination in people living with HIV-1. Sci Rep. 2023;13:18994.

334. Cox A, Cevik H, Feldman HA, Canaday LM, Lakes N, Waggoner SN. Targeting natural killer cells to enhance vaccine responses. Trends in Pharmacological Sciences. 2021;42:789–801.

335. Zimmermann P, Curtis N. Factors That Influence the Immune Response to Vaccination. Clin Microbiol Rev. 2019;32:e00084-18.

336. Pollard AJ, Bijker EM. A guide to vaccinology: from basic principles to new developments. Nat Rev Immunol. 2021;21:83–100.

337. Terreri S, Piano Mortari E, Vinci MR, Russo C, Alteri C, Albano C, et al. Persistent B cell memory after SARS-CoV-2 vaccination is functional during breakthrough infections. Cell Host Microbe. 2022;30:400-408.e4.

338. BNT162b2 vaccination induces durable SARS-CoV-2–specific T cells with a stem cell memory phenotype | Science Immunology. https://www.science.org/doi/10.1126/sciimmunol.abl5344. Accessed 26 Nov 2023.

339. Kak G, Raza M, Tiwari BK. Interferon-gamma (IFN-γ): Exploring its implications in infectious diseases. Biomolecular Concepts. 2018;9:64–79.

340. Mühl H. Pro-Inflammatory Signaling by IL-10 and IL-22: Bad Habit Stirred Up by Interferons? Frontiers in Immunology. 2013;4.

341. Gerosa F, Baldani-Guerra B, Lyakh LA, Batoni G, Esin S, Winkler-Pickett RT, et al. Differential regulation of interleukin 12 and interleukin 23 production in human dendritic cells. J Exp Med. 2008;205:1447–61.

342. García-Piñeres AJ, Hildesheim A, Dodd L, Kemp TJ, Yang J, Fullmer B, et al. Gene Expression Patterns Induced by HPV-16 L1 VLP in Leukocytes from Vaccine Recipients. J Immunol. 2009;182:1706–29.

343. García-Piñeres A, Hildesheim A, Dodd L, Kemp TJ, Williams M, Harro C, et al. Cytokine and Chemokine Profiles following Vaccination with Human Papillomavirus Type 16 L1 Virus-Like Particles. Clin Vaccine Immunol. 2007;14:984–9.

344. Vaccines | Free Full-Text | Novel Antigenic Targets of HPV Therapeutic Vaccines. https://www.mdpi.com/2076-393X/9/11/1262. Accessed 25 Nov 2023.

345. Shibata T, Shah S, Evans T, Coleman H, Lieblong BJ, Spencer HJ, et al. Expansion of Human Papillomavirus-Specific T Cells in Periphery and Cervix in a Therapeutic Vaccine Recipient Whose Cervical High-Grade Squamous Intraepithelial Lesion Regressed. Frontiers in Immunology. 2021;12.

346. Wang J, Sui J, Mao C, Li X, Chen X, Liang C, et al. Identification of Key Pathways and Genes Related to the Development of Hair Follicle Cycle in Cashmere Goats. Genes (Basel). 2021;12:180.

347. Chen ML, Takeda K, Sundrud MS. Emerging roles of bile acids in mucosal immunity and inflammation. Mucosal Immunology. 2019;12:851–61.

348. Yang D, Jacobson A, Meerschaert KA, Sifakis JJ, Wu M, Chen X, et al. Nociceptor neurons direct goblet cells via a CGRP-RAMP1 axis to drive mucus production and gut barrier protection. Cell. 2022;185:4190-4205.e25.

349. FERNANDES JV, DE MEDEIROS FERNANDES TAA, DE AZEVEDO JCV, COBUCCI RNO, DE CARVALHO MGF, ANDRADE VS, et al. Link between chronic inflammation and human papillomavirus-induced carcinogenesis (Review). Oncol Lett. 2015;9:1015–26.

350. Tezal M. Interaction between Chronic Inflammation and Oral HPV Infection in the Etiology of Head and Neck Cancers. Int J Otolaryngol. 2012;2012:575242.

351. Immune evasion mechanisms of human papillomavirus: An update - Steinbach - 2018 - International Journal of Cancer - Wiley Online Library. https://onlinelibrary.wiley.com/doi/full/10.1002/ijc.31027. Accessed 25 Nov 2023.

352. Sasagawa T, Takagi H, Makinoda S. Immune responses against human papillomavirus (HPV) infection and evasion of host defense in cervical cancer. Journal of Infection and Chemotherapy. 2012;18:807–15.

353. Frontiers | From Microbiome to Inflammation: The Key Drivers of Cervical Cancer. https://www.frontiersin.org/articles/10.3389/fmicb.2021.767931/full. Accessed 25 Nov 2023.

354. Bonin-Jacob CM, Almeida-Lugo LZ, Puga MAM, Machado AP, Padovani CTJ, Noceti MC, et al. IL-6 and IL-10 in the serum and exfoliated cervical cells of patients infected with high-risk human papillomavirus. PLoS One. 2021;16:e0248639.

355. Lima Júnior SF de, Tavares MMF, Macedo JL de, Oliveira RS de, Heráclio S de A, Maia M de MD, et al. Influence of *IL-6*, *IL-8*, and *TGF-β1* gene polymorphisms on the risk of human papillomavirus-infection in women from Pernambuco, Brazil. Mem Inst Oswaldo Cruz. 2016;111:663–9.

356. Artaza-Irigaray C, Molina-Pineda A, Aguilar-Lemarroy A, Ortiz-Lazareno P, Limón-Toledo LP, Pereira-Suárez AL, et al. E6/E7 and E6* From HPV16 and HPV18 Upregulate IL-6 Expression Independently of p53 in Keratinocytes. Frontiers in Immunology. 2019;10.

357. HIRANO T. Interleukin 6 in autoimmune and inflammatory diseases: a personal memoir. Proc Jpn Acad Ser B Phys Biol Sci. 2010;86:717–30.

358. Interleukin 6 and Rheumatoid Arthritis - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3913495/. Accessed 25 Nov 2023.

359. Koga T, Kawakami A. Interleukin-6 inhibition in the treatment of autoinflammatory diseases. Frontiers in Immunology. 2022;13.

360. IL-6 in Inflammation, Immunity, and Disease - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4176007/. Accessed 25 Nov 2023.

361. Protein Design: From the Aspect of Water Solubility and Stability - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9523718/. Accessed 25 Nov 2023.

362. Pirtskhalava M, Vishnepolsky B, Grigolava M, Managadze G. Physicochemical Features and Peculiarities of Interaction of AMP with the Membrane. Pharmaceuticals (Basel). 2021;14:471.

363. Zhu Q, Chen Z, Paul PK, Lu Y, Wu W, Qi J. Oral delivery of proteins and peptides: Challenges, status quo and future perspectives. Acta Pharmaceutica Sinica B. 2021;11:2416–48.

364. Hussein HA, Borrel A, Geneix C, Petitjean M, Regad L, Camproux A-C. PockDrug-Server: a new web server for predicting pocket druggability on holo and apo proteins. Nucleic Acids Research. 2015;43:W436–42.

365. Borrel A, Regad L, Xhaard H, Petitjean M, Camproux A-C. PockDrug: A Model for Predicting Pocket Druggability That Overcomes Pocket Estimation Uncertainties. J Chem Inf Model. 2015;55:882–95.

366. Tian W, Chen C, Lei X, Zhao J, Liang J. CASTp 3.0: computed atlas of surface topography of proteins. Nucleic Acids Research. 2018;46:W363–7.

367. Liao J, Wang Q, Wu F, Huang Z. In Silico Methods for Identification of Potential Active Sites of Therapeutic Targets. Molecules. 2022;27:7103.

368. Benchmarks for interpretation of QSAR models | Journal of Cheminformatics | Full Text. https://jcheminf.biomedcentral.com/articles/10.1186/s13321-021-00519-x. Accessed 25 Nov 2023.

369. Castro-Alvarez A, Costa AM, Vilarrasa J. The Performance of Several Docking Programs at Reproducing Protein–Macrolide-Like Crystal Structures. Molecules. 2017;22:136.

370. Zheng L, Meng J, Jiang K, Lan H, Wang Z, Lin M, et al. Improving protein–ligand docking and screening accuracies by incorporating a scoring function correction term. Brief Bioinform. 2022;23:bbac051.

371. Zhang Y, Zhou G, Wei Z, Xu H. Predicting Protein-Ligand Binding Affinity via Joint Global-Local Interaction Modeling. 2022.

372. Stabilizing Role of Water Solvation on Anion−π Interactions in Proteins | ACS Omega. https://pubs.acs.org/doi/10.1021/acsomega.1c03264. Accessed 25 Nov 2023.

373. H−π Complexes of Acetylene−Ethylene: A Matrix Isolation and Computational Study | The Journal of Physical Chemistry A. https://pubs.acs.org/doi/10.1021/jp012457g. Accessed 25 Nov 2023.

374. Herbal concoction Unveiled: A computational analysis of phytochemicals' pharmacokinetic and toxicological profiles using novel approach methodologies (NAMs) - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC10440360/. Accessed 24 Nov 2023.

375. Lipinski CA, Lombardo F, Dominy BW, Feeney PJ. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3–25.1. Advanced Drug Delivery Reviews. 2001;46:3–26.

376. Oral Druggable Space beyond the Rule of 5: Insights from Drugs and Clinical Candidates - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S1074552114002890. Accessed 24 Nov 2023.

377. BDDCS, the Rule of 5 and Drugability - PMC. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4910824/. Accessed 24 Nov 2023.

378. Bickerton GR, Paolini GV, Besnard J, Muresan S, Hopkins AL. Quantifying the chemical beauty of drugs. Nat Chem. 2012;4:90–8.

379. Analysing molecular polar surface descriptors to predict blood-brain barrier permeation - PubMed. https://pubmed.ncbi.nlm.nih.gov/23428480/. Accessed 24 Nov 2023.

380. Nascimento CMC, Moura PG, Pimentel AS. Generating structural alerts from toxicology datasets using the local interpretable model-agnostic explanations method. Digital Discovery. 2023;2:1311–25.

381. Dissolution enhancement of the anti-HIV drug UC 781 by formulation in a ternary solid dispersion with TPGS 1000 and Eudragit E100 - ScienceDirect. https://www.sciencedirect.com/science/article/pii/S0939641108002749. Accessed 24 Nov 2023.

382. Maciag PC, Radulovic S, Rothman J. The first clinical use of a live-attenuated Listeria monocytogenes vaccine: A Phase I safety study of Lm-LLO-E7 in patients with advanced carcinoma of the cervix. Vaccine. 2009;27:3975–83.

References

Acknowledgments

I would like to express my sincere gratitude to the Ministry of Health for their support in the realization of my thesis, which is a vital component of the "HPV Immunological markers of cervical persistent infection and oncogenesis: HPVIMMUNO study." The project, funded under Ricerca Corrente (cod. 08068322), has provided valuable insights into the intricate aspects of human papillomavirus (HPV) and its role in cervical health.

Additionally, I extend my heartfelt appreciation to the University of Pavia for the generous funding of my Ph.D. studies. Their support has not only facilitated my academic pursuits but has also enabled me to immerse myself in the vibrant academic and cultural environment of Pavia, Italy.