UNIVERSIDADE FEDERAL DE MINAS GERAIS
INSTITUTO DE CIÊNCIAS BIOLÓGICAS
PROGRAMA INTERUNIDADES DE PÓS-GRADUAÇÃO
EM BIOINFORMÁTICA

SYSTEMATIC REVIEW OF NON-VIRAL STIS FOR THERAPEUTIC AND PREVENTATIVE STATUS LEADS TO THE DEVELOPMENT OF A MULTI-EPITOPE VACCINE AGAINST *TREPONEMA PALLIDUM*: A Bioinformatics Approach

Discente: **Lucas Gabriel Rodrigues Gomes**

BELO HORIZONTE
2022

LUCAS GABRIEL RODRIGUES GOMES

SYSTEMATIC REVIEW OF NON-VIRAL STIS FOR THERAPEUTIC AND PREVENTATIVE STATUS LEADS TO THE DEVELOPMENT OF A MULTI-EPITOPE VACCINE AGAINST *TREPONEMA PALLIDUM*: A Bioinformatics Approach

Versão Final

Dissertação apresentada como requisito parcial para a obtenção do Grau de Mestre pelo Programa de Pós-graduação em Bioinformática, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais.

Discente: **Lucas Gabriel Rodrigues Gomes**
Orientador: **Prof. Dr. Vasco Ariston de Carvalho Azevedo**
Co-orientador: **Dr. Arun Kumar Jaiswal** e **MSc. Roselane Gonçalves dos Santos**

BELO HORIZONTE
2022

043

Gomes, Lucas Gabriel Rodrigues.
Systematic review of non-viral STIs for therapeutic and preventative status leads to the development of a multi-epitope vaccine against *Treponema pallidum*: a bioinformatics approach [manuscrito] / Lucas Gabriel Rodrigues Gomes. – 2022.
82 f. : il. ; 29,5 cm.

Orientador: Prof. Dr. Vasco Ariston de Carvalho Azevedo. Co-orientadores: Dr. Arun Kumar Jaiswal e MSc. Roselane Gonçalves dos Santos.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa Interunidades de Pós-Graduação em Bioinformática.

1. Bioinformática. 2. Doenças Sexualmente Transmissíveis. 3. *Treponema pallidum*. 4. Epitopos. 5. Vacinas. I. Azevedo, Vasco Ariston de Carvalho. II. Jaiswal, Arun Kumar. III. Santos, Roselane Gonçalves dos. IV. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. V. Título.

CDU: 573:004

Ficha catalográfica elaborada pela bibliotecária Fabiane C M Reis – CRB 6 - 2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa Interunidades de Pós-Graduação em Bioinformática da UFMG

ATA DE DEFESA DE DISSERTAÇÃO

LUCAS GABRIEL RODRIGUES GOMES

Às nove horas do dia **13 de dezembro de 2022**, reuniu-se, através de videoconferência, a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"Systematic Review of Non-Viral STIs for Therapeutic and Preventative Status Leads to the Development of a Multi-Epitope Vaccine Against *Treponema pallidum*: a Bioinformatics Approach"**, requisito para obtenção do grau de Mestre em **Bioinformática.** Abrindo a sessão, o Presidente da Comissão, **Dr. Vasco Ariston de Carvalho Azevedo**, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra ao candidato, para apresentação de seu trabalho. Seguiu-se a arguição pelos Examinadores, com a respectiva defesa do candidato. Logo após, a Comissão se reuniu, sem a presença do candidato e do público, para julgamento e expedição de resultado final. Foram atribuídas as seguintes indicações:

Professor(a)/Pesquisador(a)	**Instituição**	**Indicação**
Dr. Vasco Ariston de Carvalho Azevedo - Orientador	Universidade Federal de Minas Gerais	Aprovado
Dr. Arun Kumar Jaiswal - Coorientador	Universidade Federal de Minas Gerais	Aprovado
Dra. Raquel Cardoso de Melo Minardi	Universidade Federal de Minas Gerais	Aprovado
Dr. Bruno Silva Andrade	Universidade Estadual do Sudoeste da Bahia	Aprovado
Dr. Rodrigo Dias de Oliveira Carvalho	Universidade Federal da Bahia	Aprovado

Pelas indicações, o candidato foi considerado: Aprovado

O resultado final foi comunicado publicamente ao candidato pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 13 de dezembro de 2022.

Documento assinado eletronicamente por **Bruno Silva Andrade**, **Usuário Externo**, em 13/12/2022, às 11:29, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Rodrigo Dias de Oliveira Carvalho**, **Usuário Externo**, em 13/12/2022, às 11:37, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Arun Kumar Jaiswal**, **Usuário Externo**, em 14/12/2022, às 09:58, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo**, **Professor do Magistério Superior**, em 14/12/2022, às 11:38, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Raquel Cardoso de Melo Minardi**, **Professora do Magistério Superior**, em 16/12/2022, às 11:08, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **1952364** e o código CRC **1F702766**.

Referência: Processo nº 23072.272149/2022-64 SEI nº 1952364

AGRADECIMENTOS

Agradeço ao meu orientador, o Professor Vasco Azevedo, por aceitar conduzir o meu trabalho de pesquisa, pelas valiosas contribuições ao longo de todo o processo e pela disponibilidade a fornecer apoio sempre.

Aos meus coorientadores, Arun Jaiswal e Roselane dos Santos, pelo apoio incessável no desenvolvimento e escrita desse projeto.

A todos os colegas do LGCM, que proporcionaram companheirismo e conhecimento em igual medida, enriquecimento minha formação como pessoa e como formando. Agradeço especialmente a Sandeep Tiwari e Thaís Rodrigues, cuja ajuda foi vital no desenvolvimento deste trabalho.

Aos colegas de turma, pelo companheirismo, leveza e apoio durante o mestrado.

Aos colegas e companheiros do Curso de Verão em Bioinformática, pela confiança e oportunidade de participar do nosso curso.

Aos meus pais, Paulo César e Luiza Stella, por todo o apoio e esforço que dedicaram à minha vida e à garantia da minha educação e por sempre acreditarem na minha capacidade de superar todos obstáculos.

À minha tia Eliane Gomes, que queria que eu fosse biólogo desde cedo, pelo apoio constante durante toda a graduação e mestrado, pelos ensinamentos e pela sabedoria compartilhada.

À minha namorada, Clara Peralva, pelo apoio e paciência infinita durante os meses de desenvolvimento deste trabalho, e que foram redobrados durante o processo de escrita.

RESUMO

De acordo com a Organização Mundial da Saúde (OMS), Infecções Sexualmente Transmissíveis (ISTs) são uma preocupação de saúde pública significativa, com mais de um milhão de novos casos a cada dia. Aproximadamente 374 milhões de pessoas por ano contraem uma de quatro ISTs não-virais tratáveis: Sífilis; Gonorréia; Clamídia; e Tricomoníase. Apesar da disponibilidade de testes diagnósticos e tratamentos para essas condições, elas continuam sendo um peso em sistemas de saúde globais, particularmente dentre populações vulneráveis. Um fator agravante associado à disseminação de ISTs bacterianas é a indisponibilidade de vacinas para essas infecções, o desenvolvimento das quais seria vital para combatê-las em escala global. Este trabalho buscou revisar o atual estado da arte na aplicação da bioinformática no desenvolvimento de novas vacinas, drogas e métodos diagnósticos para ISTs não-virais. Nós identificamos áreas em que esforços de pesquisa foram concentrados, como *Neisseria gonorrhoeae* e *Chlamydia trachomatis*, mas também descrevemos que outros patógenos, como *Treponema pallidum* e *Trichomonas vaginalis*, receberam menos atenção. Adicionalmente, este estudo buscou desenvolver uma vacina multi-epítopo para *T. pallidum*, agente causador da sífilis, com base em candidatos vacinais determinados através da vacinologia reversa. Esta vacina foi projetada selecionando epítopos imunogênicos de proteínas selecionadas de *T. pallidum* que são capazes de provocar resposta imune do hospedeiro. Estes epítopos foram concatenados em uma sequência de peptídeo quimérica, que foi sujeita a análises imunológicas. As propriedades físico-químicas e características imunológicas da proteína foram determinadas e modeladas com base na sua sequência e estrutura predita, com intenção de determinar sua segurança e imunogenicidade. Acoplamento e dinâmica moleculares foram realizados para determinar sua capacidade de ligar ao TLR-2 humano e a estabilidade das interações no complexo. Uma simulação imune *in silico* foi conduzida para modelar os efeitos imunogênicos da proteína quimera. Finalmente, a sequência passou por tradução reversa para se obter a sequência respectiva de DNA para expressão em *Escherichia coli* K12, e o gene resultante foi clonado no vetor plasmidial pET28a(+). A vacina multi-epítopo resultante é um promissor candidato vacinal para a sífilis, mas deve primeiro ser testada *in vitro* e *in vivo* para validar essas conclusões.

Palavras-chave: *Treponema pallidum,* IST, vacina multi-epítopo, imunoinformática, vacinologia, triagem virtual de drogas

ABSTRACT

According to the World Health Organization (WHO), Sexually Transmitted Infections (STIs) are a major worldwide public health concern, with over a million new cases of infection every day. Approximately 374 million people per year are infected by one of four curable non-viral STIs: Syphilis, Gonorrhea, Chlamydia, and Trichomoniasis. Despite the availability of diagnostic tests and treatments for these maladies, they continue to burden health systems globally, particularly within vulnerable populations. A compounding factor in the continuing spread of bacterial STIs is the unavailability of vaccines for these infections, the development of which would be vital in combating them on a global scale. This work aimed to review the current state of the literature on the application of bioinformatics to develop new vaccines, drugs, and diagnostic methods for non-viral STIs. We have identified areas in which research efforts have been concentrated, such as *Neisseria gonorrhoeae* and *Chlamydia trachomatis*, but have also found that other pathogens, such as *Treponema pallidum* and *Trichomonas vaginalis*, have received less attention. In addition, this study aimed to develop a multi-epitope vaccine for *T. pallidum*, the causative agent for syphilis, based on vaccine candidates previously determined through reverse vaccinology. This vaccine was designed by selecting immunogenic epitopes from proteins selected from *T. pallidum* that are capable of eliciting an immune response from the host. These epitopes are joined into a chimeric peptide sequence that is then subjected to immunological analyses. The chimeric protein's physico-chemical properties and immunogenic characteristics were determined and then modeled based on its sequence and predicted structure to determine its safety and immunogenicity. Molecular docking and dynamics analyses were performed to determine its capacity to bind to the human TLR-2 and the stability of said receptor-ligand binding interactions. An *in silico* immune simulation was conducted to model the protein's immunogenic effects. Finally, the sequence was reverse-translated into its respective DNA sequence for expression in *Escherichia coli* K12, and the resulting gene was cloned *in silico* into the pET28a(+) plasmid vector. This multi-epitope vaccine shows promise as a potential vaccine candidate for syphilis but must be tested *in vitro* and *in vivo* to validate these findings.

Keywords: *Treponema pallidum,* STI, multi-epitope vaccine, immunoinformatics, vaccinology, drug screening

FIGURE LIST

Chapter 1.

Chapter 2.

TABLE LIST

Chapter 1.

Chapter 2.

ABBREVIATION LIST

IST: Infecção Sexualmente Transmissível;

STI: Sexually transmitted infection;

OMS: Organização Mundial da Saúde;

WHO: World Health Organisation;

MSM: Men who have sex with men;

FSW: Female sex workers;

RV: Reverse Vaccinology;

MHC: Major Histocompatibility Complex;

TLR: Toll-like receptor;

NCBI: National Center for Biotechnology Information;

IEDB: Immune Epitope Database and Analysis Resource;

CTL: Cytotoxic T lymphocyte;

HTL: Helper T lymphocyte;

ANN: Artificial Neural Network;

SMM: Scoring Matrix Method;

SVM: Support Vector Machine;

GRAVY: Grand Average of hydropathicity;

PSSM: Position-Specific Scoring Matrix;

MTT: Multiple Template Threading;

HMM: Hidden Markov Model;

RCSB: Research Collaboratory for Structural Bioinformatics;

PDB: Protein Database;

CAI: Codon Adaptation Index;

II: Instability Index;

RMSD: Root Mean Square Deviation;

RMSF: Root Mean Square Fluctuation;

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses;

OMPs: Outer Membrane Protein;

DTH: Delayed-type hypersensitivity.

SUMMARY

1. INTRODUCTION

According to data from the World Health Organization (WHO), Sexually Transmitted Infections are a major public health factor. An estimated 1 million new STI cases occur daily, and approximately 374 million cases of curable non-viral STIs such as syphilis, gonorrhea, chlamydia, and trichomoniasis every year [1,2]. While these maladies are curable and straightforward to prevent through preservatives, they remain an intense drain on financial and personnel resources for global health systems [3]. This drain is particularly notable in high-risk populations, such as men who have sex with men, female sex workers, and the HIV-positive population [1,4]. Due to the difficulties in controlling these STIs, the WHO Global health sector strategy highlights the need to develop novel tools to combat them [5]. In particular, vaccine development is a critical effort to halt the spread of these infections [3].

Syphilis is an STI caused by the spirochete *Treponema pallidum* subspecies *pallidum*, a hard-to-culture pathogen capable of evading the host's immune system [6]. The disease occurs in three stages, evolving from a primary stage, with rashes that appear near the infected sites, to latent, asymptomatic phases and further to chronic phases where the disease can cause lasting damage to the nervous and cardiovascular systems [6]. The disease can be treated through penicillin, and while there have not been reported cases of strains resistant to it, Azithromycin-resistant strains have already been found [7]. However, despite the availability of treatment and the ongoing efforts by the WHO to halt the spread of the disease, syphilis remains in expansion. According to the organization's progress report on STIs, 2020 saw approximately 7.1 million new syphilis cases (2.4-11.5) [2]. Many factors are associated with the disease's continuing prevalence, including social stigma, the difficulty of cultivating and working with the pathogen, and the unavailability of a vaccine [8,9].

Proof-of-concept for a syphilis vaccine has already been demonstrated in the animal model for the disease, generally studied in rabbits, with partially protective results, but its application in humans was either unfeasible or unsuccessful [10]. Subsequent efforts to identify vaccine candidates have focused on studying the Outer Membrane Proteome (OMPs) of the bacterium but have found difficulties both in cultivating the microorganism and isolating and purifying these antigens [10,11]. While some researchers have been able

to identify and isolate vaccine candidates among the T. pallidum OMPs, single-candidate vaccines have so far been unable to elicit protective immunity, suggesting the need to use more than one candidate for a successful vaccine [10].

The application of computational biology-derived methodologies such as Reverse Vaccinology (RV) and immunoinformatics can identify vaccine targets for hard-to-culture and unculturable bacteria. Such methods have already been applied prior to determining vaccine candidates for *Treponema pallidum* by Jaiswal *et al.* [12]. In addition, it is possible to use vaccine candidates, either determined by traditional or reverse vaccinology, in efforts to design a multi-epitope vaccine. As they are derived from antigenic epitopes from a wide range of vaccine candidates, multi-epitope vaccines have the potential to elicit the immune response in a manner that single antigen vaccines would not be able to [13]. In addition, they are designed to provoke a specific desired response and can be constructed in a manner in which they minimize collateral effects, as well as facilitate production and distribution [14,15].

In Chapter 1 of this work, we have performed a systematic review of the literature on the application of bioinformatics in research regarding Sexually Transmitted Infections. This review aimed to evaluate the current state of the art in the application of immunoinformatics to identify vaccine candidates and drug targets and develop multi-epitope vaccines against many non-viral STIs. In doing so, we have identified areas in which research efforts have been concentrated and fields in which more investment is still required. In addition, we have found concerns regarding the standardization and reproducibility of these methodologies in this context.

In our review of the literature on STIs, we have established a need for more efforts turned to the application of immunoinformatics for the development of a syphilis vaccine. Thus, chapter 2 of this work presents a published research article in which we have designed a multi-epitope vaccine against *T. pallidum*, the causative agent of syphilis, based on targets derived from the literature. Vaccine candidates from which we have predicted epitopes were derived from traditional and reverse vaccinology-based studies, and the resulting chimeric peptide shows promise as a vaccine candidate for syphilis.

2. OBJECTIVES

2.1. General Objective

The objective of this work was to perform a systematic review of the literature concerning the application of immunoinformatics to research Sexually Transmitted Infections and to design a multi-epitope vaccine to control the spread of *Treponema pallidum*.

2.2. Specific Objectives

The specific objectives for the completion of this work were:

- Perform a systematic review of the literature on the application of immunoinformatics to research on STIs;
- Select vaccine candidates for *Treponema pallidum* from the literature for immunoinformatics analyses;
- Predict MHC-I, MHC-II, and B-cell epitopes from the vaccine candidates;
- Construct a chimeric peptide vaccine;
- Evaluate the construct's physico-chemical and immunological properties *in silico*;
- Model the secondary and tertiary structure of the construct;
- Perform molecular docking and dynamics of the construct with the human TLR-2;
- Simulate the immune response to the vaccine *in silico*;
- Perform an *in silico* cloning of the vaccine to an expression plasmid.

3. CHAPTER 1

3.1. Systematic Review

A systematic review of reverse vaccinology and immunoinformatics data for non-viral sexually transmitted infections

Lucas Gabriel Rodrigues Gomes, Joyce da Cruz Ferraz Dutra, Rodrigo Profeta, Mariana Vieira Dias, Glen Jasper Yupanqui García, Diego Lucas Neres Rodrigues, Aristóteles Goés Neto, Flávia Figueira Aburjaile, Sandeep Tiwari, Siomar de Castro Soares, **Vasco Aristón de Azevedo, Arun Kumar Jaiswal**

The article is being prepared for submission.

In this review article, we have broadly evaluated the application of immunoinformatics and reverse vaccinology techniques to studies regarding sexually transmitted infections. We have identified gaps in the research in which more studies are still required and bacteria in which efforts have been concentrated. In addition, we point to concerns regarding the standardization of these methodologies. This review was performed in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines, by a team of 5 independent reviewers.

SYSTEMATIC REVIEW OF REVERSE VACCINOLOGY AND IMMUNOINFORMATICS DATA FOR NON-VIRAL SEXUALLY TRANSMITTED INFECTIONS

Lucas Gabriel Rodrigues Gomes[1], Joyce da Cruz Ferraz Dutra[2], Rodrigo Profeta[1], Mariana Vieira Dias[3], Glen Jasper Yupanqui García[4], Diego Lucas Neres Rodrigues[1 5], Aristóteles Goés Neto[6], Flávia Figueira Aburjaile[1 5], Sandeep Tiwari[1 7 8], Siomar de Castro Soares[9], Vasco Azevedo[1*], Arun Kumar Jaiswal[1*]

1 Laboratory of Cellular and Molecular Genetics (LGCM), PG Program in Bioinformatics, Department of Genetics, Ecology, and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais (UFMG), Belo Horizonte 31270-901, Brazil

2 Federal University of Minas Gerais, Institute of Biological Sciences, Department of Microbiology, Belo Horizonte, Minas Gerais Brazil

3 Federal University of Minas Gerais, Institute of Biological Sciences, Department of Genetics, Ecology, and Evolution, Belo Horizonte, Minas Gerais, Brazil

4 Federal University of Minas Gerais, Institute of Biological Sciences, Department of Bioinformatics, Belo Horizonte, Minas Gerais, Brazil

5 Department of Veterinary Medicine, Veterinary School, Universidade Federal de Minas Gerais, Belo Horizonte 31270-901, Brazil.

6 Laboratório de Biologia Molecular e Computacional de Fungos, Departamento de Microbiologia, Instituto de Ciências Biológicas, Universidade Federal de Minas Gerais (UFMG), Belo Horizonte, Minas Gerais CEP 31270-901, Brazil.

7 Institute of Biology, Federal University of Bahia, Salvador, BA, Brazil,

8 Institute of Health Sciences, Federal University of Bahia, Salvador, BA, Brazil

9 Department of Microbiology, Immunology, and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro (UFTM), Uberaba 38025-180, Brazil

* Corresponding authors: Arun Kumar Jaiswal – arunjaiswal1411@gmail.com & Vasco Azevedo - vascoariston@gmail.com

Abstract

Sexually Transmitted Infections (STIs) are a public health burden rising in developed and developing nations. The World Health Organization estimates nearly 374 million new cases of curable STIs yearly. Global efforts to control their spread have been insufficient in fulfilling their objective. As there is no vaccine for many of these infections, these efforts are focused on education and condom distribution. The development of vaccines for STIs is vital for successfully halting their spread. The field of immunoinformatics is a powerful new tool for vaccine development, allowing for the identification of vaccine candidates within a bacterium's genome and allowing for the design of new genome-based vaccine peptides. The goal of this review was to evaluate the usage of immunoinformatics in research focused on non-viral STIs, identifying fields where research efforts are concentrated. Here we describe gaps in applying these techniques, as in the case of *Treponema pallidum* and *Trichomonas vaginalis*.

Keywords: Sexually Transmitted Infections, Vaccinology, Immunoinformatics, Virtual Drug Screening, Multi-epitope vaccine

1 INTRODUCTION

According to the World Health Organization (WHO), an estimated 1 million new cases of sexually transmitted infections (STIs) occur every day worldwide [1–3]. Every year, nearly 374 million new cases of curable STIs such as syphilis, chlamydia, and trichomoniasis are acquired [1–3]. Due to the unavailability of vaccines for these infections, prophylactic measures for their prevention are generally based on the fostering of sexual education and incentives for using barrier contraceptives [3]. While these methods can be effective if adequately applied, they are faced with low adoption rates in populations with little to no access to protective and educational resources and consistently fail to reach STI-vulnerable populations [3,4]. Thus, the WHO's strategy on STIs highlights the need for developing vaccines against these conditions as essential for their control and eventual eradication [4].

Using bioinformatics through its multidisciplinary fields, such as immunoinformatics, can be a significant boon to efforts to identify and develop STI vaccines and tools for diagnostics and treatment. This recent field seeks to apply the vast and rapidly accumulating repertoire of currently sequenced and understood genomes to develop vaccines, drugs, and diagnostic methods [5,6]. These techniques are made possible by the existence of extensive databases generated by previous immunological studies and genome sequencing efforts [5]. The application of computational techniques, such as machine learning, to these databases, allows for the exploration of a given bacterium's genome for sequences of interest *in silico*.

Using computational techniques may not only aid vaccine development efforts in identifying vaccine candidates in a microorganism's genome [7], but also in rationalizing peptides that may serve as vaccines [8]. This systematic literature review aims to evaluate the current state of the art in the usage of these methodologies in the context of sexually transmitted infections. Specifically, this review has examined the study of curable,

non-viral STI-causing microorganisms, such as *Treponema pallidum*, *Neisseria gonorrhoeae*, *Chlamydia trachomatis*, *Trichomonas vaginalis*, *Mycoplasma genitalium* and *Haemophilus ducreyi,* identifying fields where research is already being performed and where more investment is required. While there are other causative agents for STIs that do not fit our search parameters, like fungi, we have elected to focus on these pathogens as they are regarded by the WHO as significant burdens to be eradicated [1–3]. The research investment in this field may reduce the monetary and labor costs involved in vaccine development, which is slow and costly [8,9].

2 METHODS

2.1 Data collection and filtering

The present systematic review was performed per the benchmark work *Preferred Reporting Items for Systematic Reviews and Meta-Analyses* (PRISMA) [10]. Searches were performed amongst all document types available in the Scopus, Web of Science, and PubMed databases published until the 22nd of March, 2022. Document filtering was performed in three steps: [1] identification of documents including selected keywords in the Scopus, Web of Science, and Pubmed databases; [2] automated document triage; and [3] Manual evaluation of study eligibility.

The keywords used in the search were: (("immunoinformatics" OR "Reverse Vaccinology" OR "bioinformatics vaccine targets" OR "rational vaccine design" OR "rational immunogen design" OR "Potential Universal Vaccine Candidates" OR "In silico vaccine design" OR "in silico immunogen design" OR "epitope-based vaccine" OR "epitope-based immunogen" OR "Structure-based immunogen design" OR "Structure-based antigen design" OR "in silico vaccine targets" OR "potential vaccine candidates" OR "subtractive genomics" OR "computational vaccinology" OR "vaccinology") AND ("STD vaccine" OR "sexually transmitted infections" OR "sexually transmitted" OR "Treponema pallidum" OR "syphilis" OR "Mycoplasma genitalium" OR "Neisseria gonorrhoeae" OR "gonorrhea" OR "Haemophilus ducreyi" OR "chancroid" OR "Chlamydia trachomatis" OR "chlamydia") NOT ("virus"). All keyword combinations and boolean operators are detailed in **Supplementary Material S1**.

Documents were automatically exported and extracted as a CSV file through the format_input.py script (**Supplementary Material S2**), where for each article, the Title, Publication year, Digital Object Identifier (DOI), Document type, Language, and Author(s) were extracted. Documents lacking a DOI were automatically identified, and duplicate titles or DOIs were removed. Records for the three separate databases were unified, and redundancy has been removed through the remove_duplicates.py script (**Supplementary Material S3**). Database information for each article was added to the unified file, labeling each article based on the databases on which it could be found. Files that could be found on more than one database were labeled appropriately.

The Federal University of Minas Gerais' internal library network access permissions were used to download each article in PDF format through its DOI. Inaccessible documents were characterized as *Not Available (NA)*.

Three separate reviewers performed document analysis independently, with a fourth being consulted in case of disagreement. Articles were initially analyzed based on their titles and

abstracts, with only articles relating to immunoinformatics and STIs being selected. In a second analysis, studies were reevaluated based on two sets of eligibility and exclusion criteria:

Table 1. Eligibility and exclusion criteria for the inclusion of articles in the review.

	Eligibility criteria
1	Non-viral Sexually Transmitted Disease/Sexually Transmitted Infection (STD/STI)
2	Method: immunoinformatics, such as the application of Reverse Vaccinology and the development of multiepitope chimeric vaccines.
3	Language: English. Original documents with Digital Object Identifiers;
4	Objective: Identify or propose drug targets and vaccine candidates for STIs.

	Exclusion criteria
1	Articles that performed exclusively *in vivo* and *in vitro* experiments
2	Book chapters, posters, and review articles
3	Studies on microorganisms not related to or involved in sexually transmitted infections

2.2 Bibliographic coupling and word cloud analysis

The VOSviewer software (version 1.6.8) was used for bibliographic coupling between authors and publications [11,12]. This analysis creates a citation network based on the DOIs of the articles, which describes associations between authors and references and, by combining clustering and visualization techniques, allows for better analysis. A word cloud was generated based on the PDF files using the NVivo v.20.5.0 software (QSR International Pty Ltd., 2020).

2.3 Data analysis

Data on the following subjects were extracted from the selected documents: (1) publication date; (2) methodology, (3) drug target prediction; (4) vaccine candidate identification; and (5) multi-epitope vaccine design. The world map containing information on scientific production *per country* and the frequency in which each STI-causing microorganism was studied was generated using the ggplot2 v3.3.6 and scatterpie v0.1.7 R packages. The bar graph used to demonstrate the evolution of scientific production on these subjects over the years was generated through the ggplot2 v3.3.6 R package. A circular visualization generated through the circlize R package was utilized to facilitate interaction data analysis [13]. Analyses of each paper's different content and data were performed using the UpSetR version 1.4.0 R package [14]. Association networks relating drug targets and vaccine candidates to their targets were generated using the Gephi v0.9.2, using the Yifan Hu algorithm to distribute nodes and links.

3 RESULTS

3.1 Study characteristics

A workflow diagram of the study selection process for this systematic review is shown in Figure 1. Out of an initial set of 971 articles, automatic filtering could remove 316 documents due to being duplicated, lacking a Digital Object Identifier, or not being a full research article. From the 655 documents that were manually reviewed, 7 exclusion categories were reported. Reason 3 (Not involving an STI-causing bacteria) was responsible for most exclusions (207), and reason 4 (Article not in English) was responsible for the least exclusions (2). Only three papers were excluded for a reason 6 (Not found). In total, 54 articles fulfilled the inclusion criteria and were considered for further analysis (**Supplementary Material S4**).

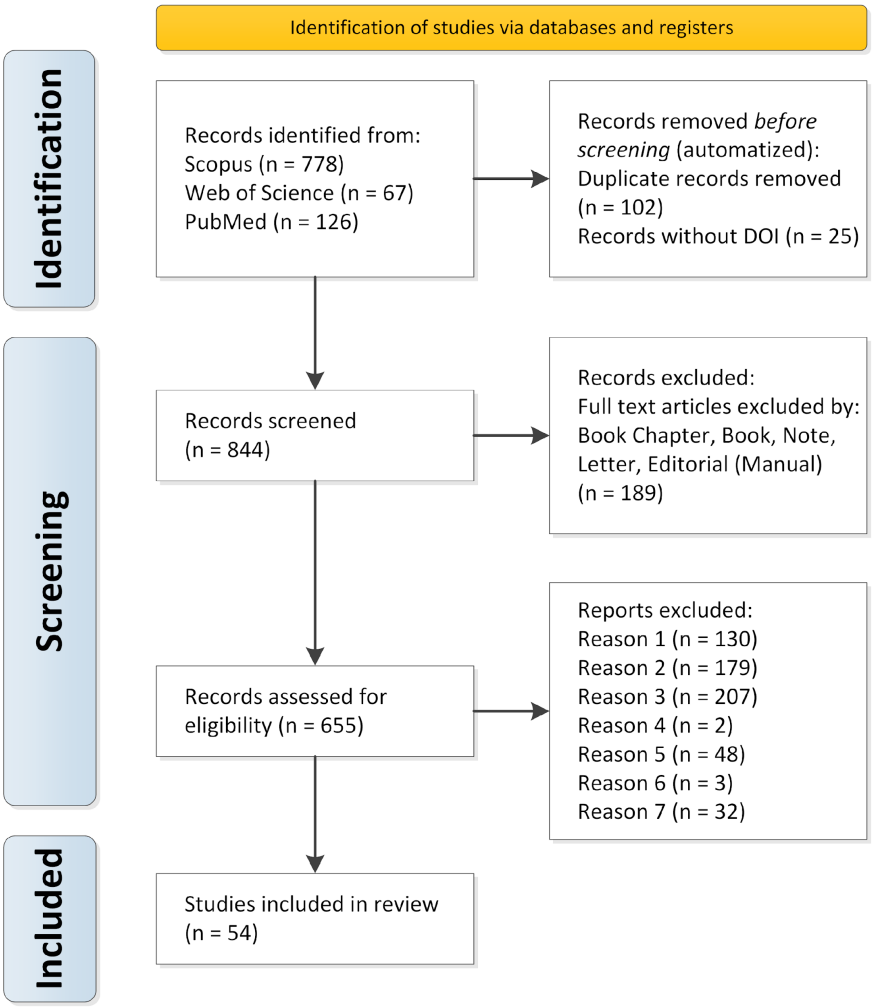


Figure 1. - Prisma Flowchart.

Reason 1= Article did not use immunoinformatics; Reason 2= Unrelated to the topic; Reason 3= Did not involve an STI-causing microorganism; Reason 4= Not in English; Reason 5= Neither an STI-causing microorganism nor immunoinformatics; Reason 6= Not found; Reason 7= Excluded during data extraction.

Bibliographic coupling analysis of the 54 studies was able to show author association based on the citation network (**Figure 2**). Five distinct author groups were determined by this network, each represented by a different node color. The most cited authors were Zielk (2014 and 2016), Semchenko (2019), and Butt (2012). While there was some connectivity between the author groups, this interplay was limited, suggesting there was not much communication between them.

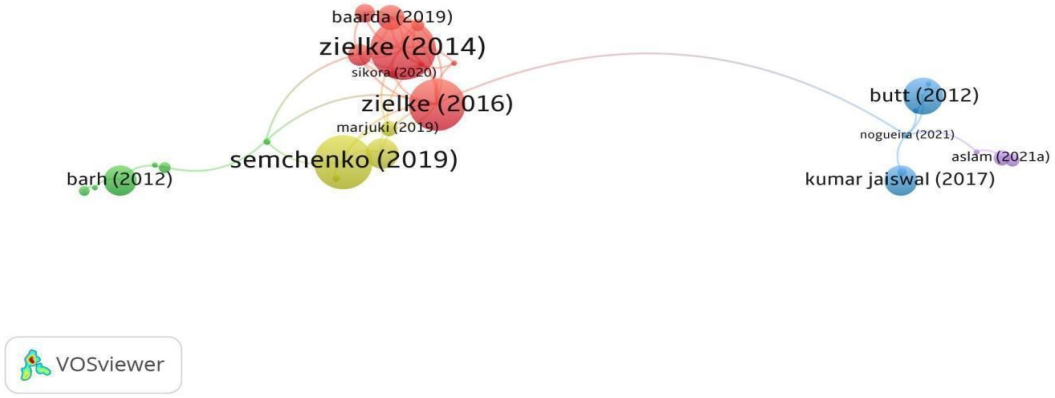


Figure 2. Citation network

Node size correlates to citation numbers, with the larger nodes being more commonly cited works; colors indicate the formation of strongly associated groups, with links representing associations between works.

Word clouds are graphical devices that use textual mining algorithms to determine the most frequently relevant words in a set of documents (Vasconcellos-Silva & Araujo-Jorge, 2019). In this work, the word cloud displayed the 100 most frequent terms in the article set that were included in the keyword set used to perform database searches. The words “protein/proteins” and “vaccine” were the most frequent in the dataset. The words “cell” and “gonorrhoeae” can also be highlighted as frequent, highlighting a greater interest in gonorrhea among immunoinformatics studies (**Figure 3**).

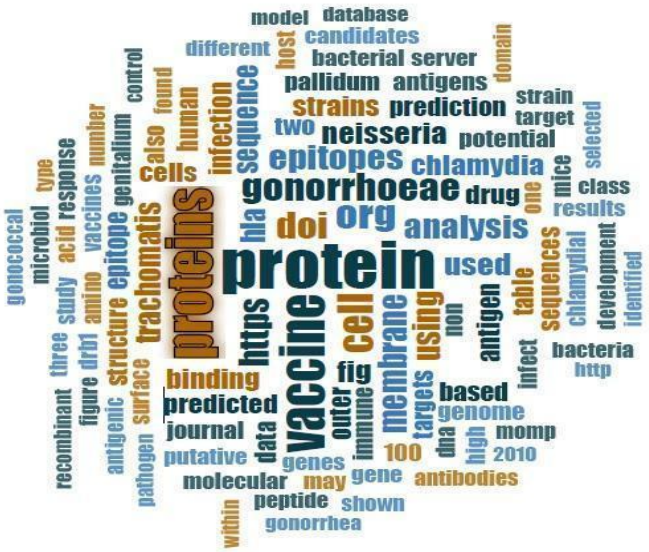


Figure 3. Word cloud constructed based on the selected article set

Top 100 terms in terms of relative frequency in the selected dataset. Word size indicates a higher frequency of the word.

The 54 selected studies were distributed between six continents: North America (United States of America-21 and Mexico-1); South America (Brazil-3 and Argentina-1); Asia (China-5, Pakistan-5, India-7, and Iran-1); Europe (Sweden-1, Germany-1, Denmark-1, and the United Kingdom-1); and Oceania (Australia-4) (**Figure 4**). We identified a concentration of research efforts in the USA, with more research directed at *Neisseria gonorrhoeae*, the causative organism of Gonorrhea, represented in the world map by the color pink. In other countries, a more significant frequency of research was directed at the *Chlamydia trachomatis* pathogen, the causative agent of Chlamydia, represented in yellow. Considering the selected countries with available publications on the area, only Sweden, South Africa, Brazil, and Mexico did not perform any study on Chlamydia immunoinformatics. *Treponema pallidum*, the causative agent of Syphilis, was also represented in multiple countries (USA, Brazil, Sweden, and China), with a more significant number of studies being developed in the US.

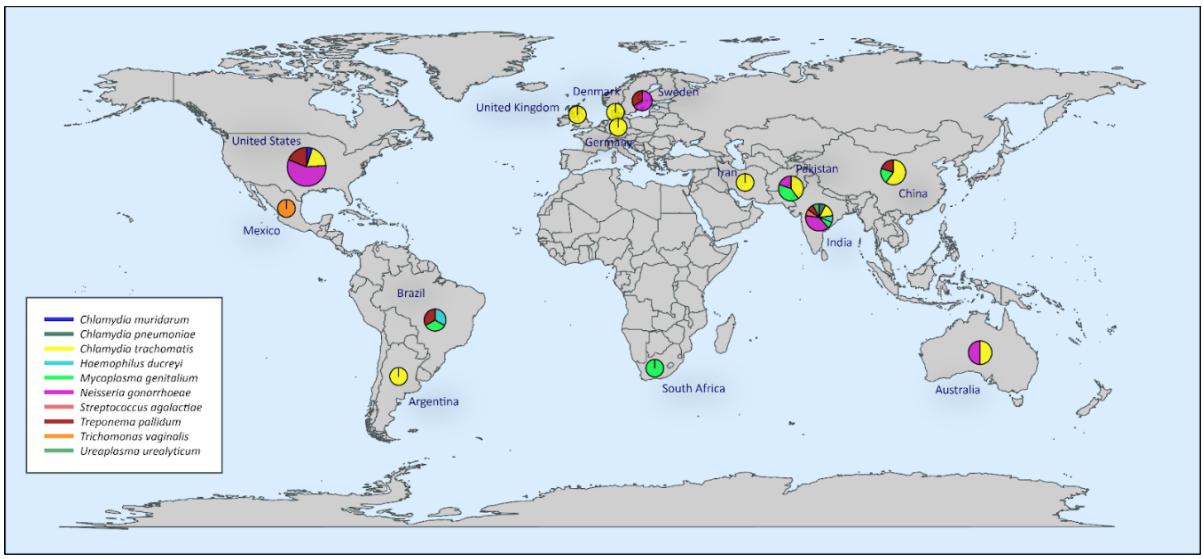


Figure 4. Frequency of STI immunoinformatics studies and their worldwide distribution, as well as the most studied STI-associated bacteria in each country.

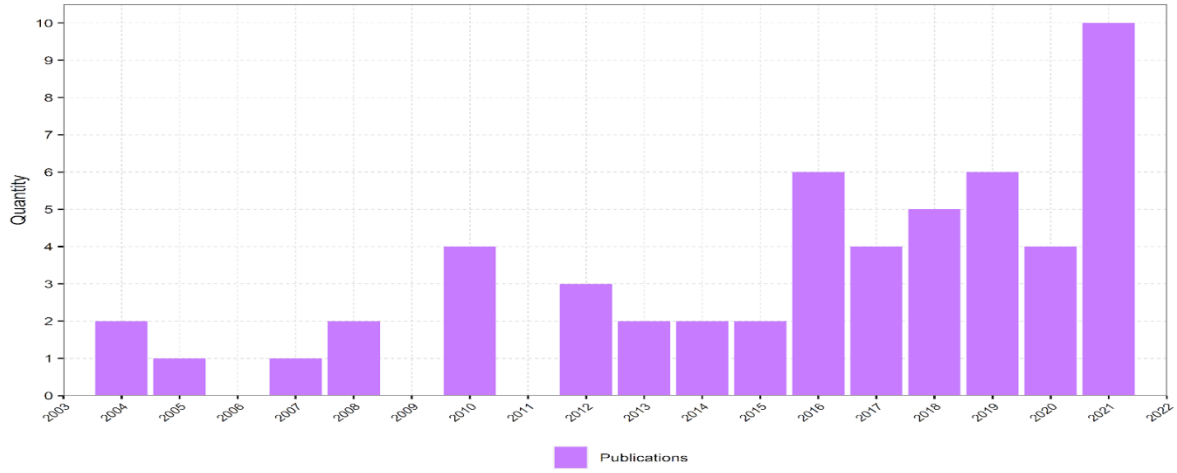


Figure 5 shows the scientific production in the field of STI immunoinformatics over its history. The earliest studies in this field were performed in 2004, with an increase in production in 2010. In 2006, 2009, and 2011 no research was performed in this field, and only in 2016 there was a significant increase compared to 2010. In 2021 there was another significant boom in production in the field.

Figure 5. Evolution of STI-immunoinformatics-associated works over time.

3.2 Methodological characteristics

The usage of immunoinformatics-based tools in support of drug target and vaccine candidate identification studies is an expanding field. Among these tools, Reverse

Vaccinology (R.V.), Virtual Drug Screening (V.D.S.), and the rational design of multi-epitope vaccines (M.V.) are essential highlights.

Of the 54 selected articles, the most commonly used methodology was R.V. (52), followed by M.V.-based studies (13), applied majoritarian to *Neisseria gonorrhoeae* and *Chlamydia trachomatis* research. Virtual Drug Screening analyses were less common (**Figure 6**) [**Supplementary Material S5**].

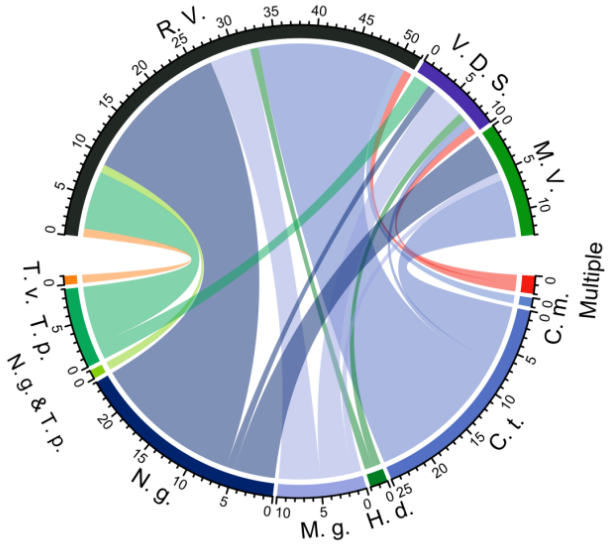


Figure 6. Chord-plot associating methodology usage and the microorganisms they are applied to. R.V.-Reverse Vaccinology; V.D.S.-Virtual Drug Screening; M.V.-Multi-epitope Vaccine; T.v.-*Trichomonas vaginalis*; T.p.-*Treponema pallidum*; N.g.-*Neisseria gonorrhoeae*; M.g.-*Mycoplasma genitalium*; H.d.-*Haemophilus ducreyi*; C.t.-*Chlamydia trichomatis*; C.m.-*Chlamydia muridarum*;

Despite being a widely applied set of methodologies in studying some microorganisms, there are still gaps in knowledge that need to be filled. In STI research, a wide range of tools still needs to be applied in certain areas, such as research on *Haemophilus ducreyi*, *Trichomonas vaginalis*, and *Treponema pallidum*.

3.3 Drug target identification

In this section of the systematic literature review, we assess the current state of the art in using *in silico* techniques to identify drug targets in the context of STIs. Of 54 articles carefully analyzed articles, 14 identified 97 drug targets for STI-causing microorganisms such as - *Treponema pallidum*, *Mycoplasma genitalium*, *Neisseria gonorrhoeae*, *Haemophilus ducreyi*, *Chlamydia trachomatis,* and *Trichomonas vaginalis*.

Out of 14 articles, only 4 provided the information related to the active site, binding affinity (binding score), and information of ligand library (compound list used for docking analysis) identified (**Supplementary Material S6**; **Figure 7**). No standard methodology was observed between articles for identifying drug targets, with various strategies such as –

subtractive genomics, essentiality analysis, and subcellular localization analysis, used for the identification.

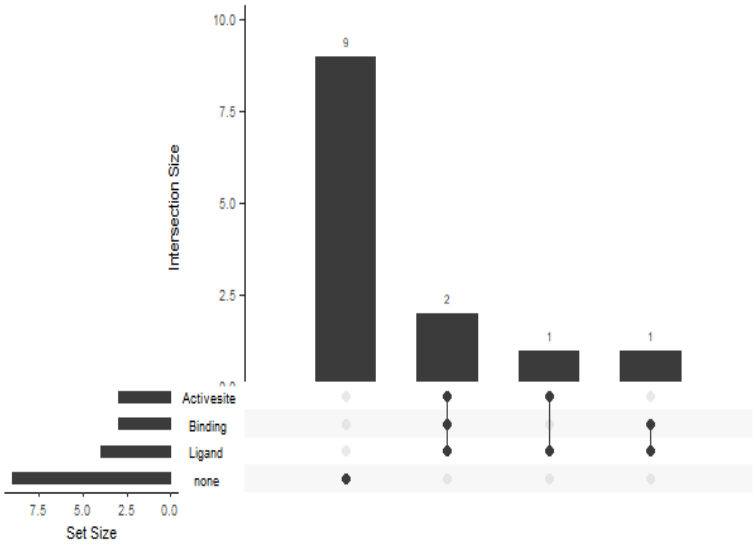


Figure 7. Upset plot showing the availability of information regarding drug targets in the different articles

In comparing all 14 out of 54 selected articles for drug target identification, we identified 97 drug targets reported in different STIs. **Figure 8** shows the complex network of identified drug targets of STIs used in this work. Druggable targets were identified in *C. trachomatis* [15], *T. vaginalis* [16], *H. ducreyi* [17], *M. genitalium* [18–21], *N. gonorrhoeae* [22–25], and *T. pallidum* [26,27]. Of the identified drug targets, two were common to multiple microorganisms, the *ddl* gene was identified as a drug target in both *N. gonorrhoeae* and *T. pallidum*, and the *SecA* gene was identified as drug target for *C. trachomatis*, *T. pallidum* and *M. genitalium*.

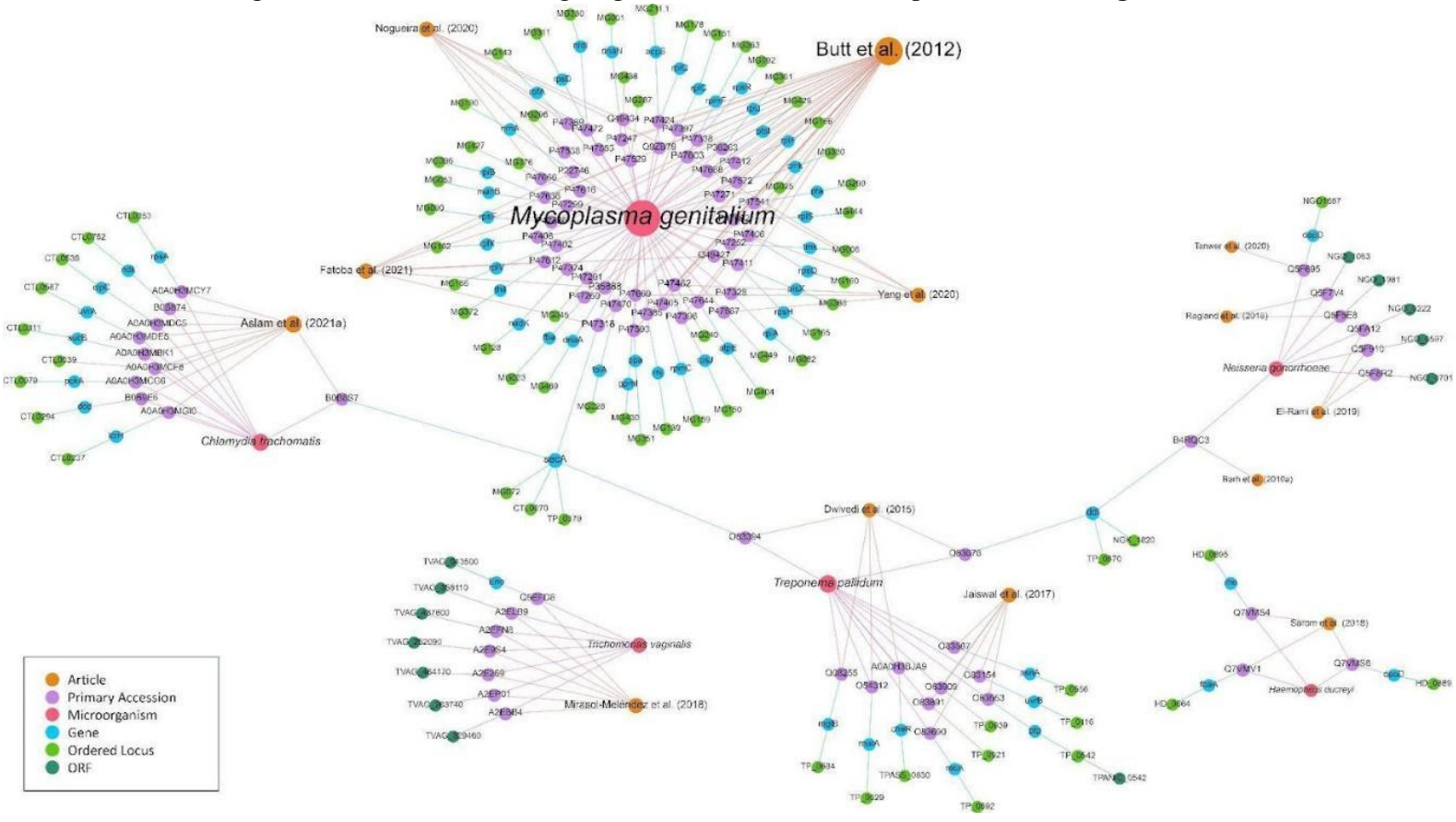


Figure 8. Complex network connecting drug targets to the microorganisms in which they were identified. Node size is proportional to the node's degree. A color legend is embedded within the figure.

3.4 Vaccine candidate identification

Out of 54 articles under analysis, 27 proposed putative vaccine candidates for the STIs being researched. Of those 27, eight articles provided only basic information regarding the candidates and their function, while others were more informative, providing relevant information to the candidate's activity as a vaccinal candidate, such as subcellular localization and adhesion probability. Most articles included information on the candidate's subcellular localization in the pathogen's cell, which is relevant for the proteins' availability to the immune system. Localization predictions were performed through a wide variety of differing tools [**Supplementary Material S7**]. Of those, nine only provided the protein's sublocalization, six also provided the protein's antigenicity as predicted by the Vaxijen tool, and three provided the protein's adhesion probability as predicted by Vaxign. One article provided the protein's antigenicity without providing its subcellular localization (**Figure 9**).

There was no methodological standardization between articles that performed vaccine candidate prediction, with different protocols being performed in the subtractive genomics selection of putative candidates and evaluating their vaccinal potential. Regarding subcellular localization predictions, most articles elected to use PSORTb and CELLO. Regarding the protein's likelihood to act as an antigen, Vaxijen and Vaxign were commonly used, with no intersection between the two tools.

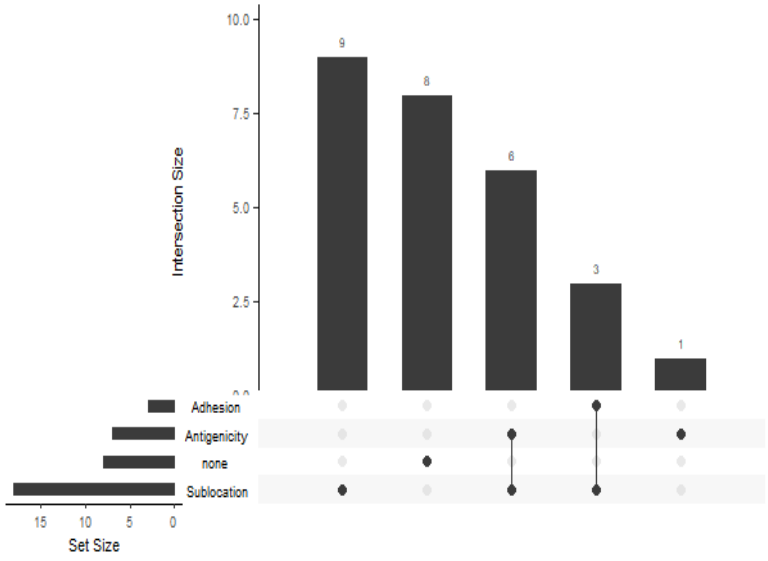
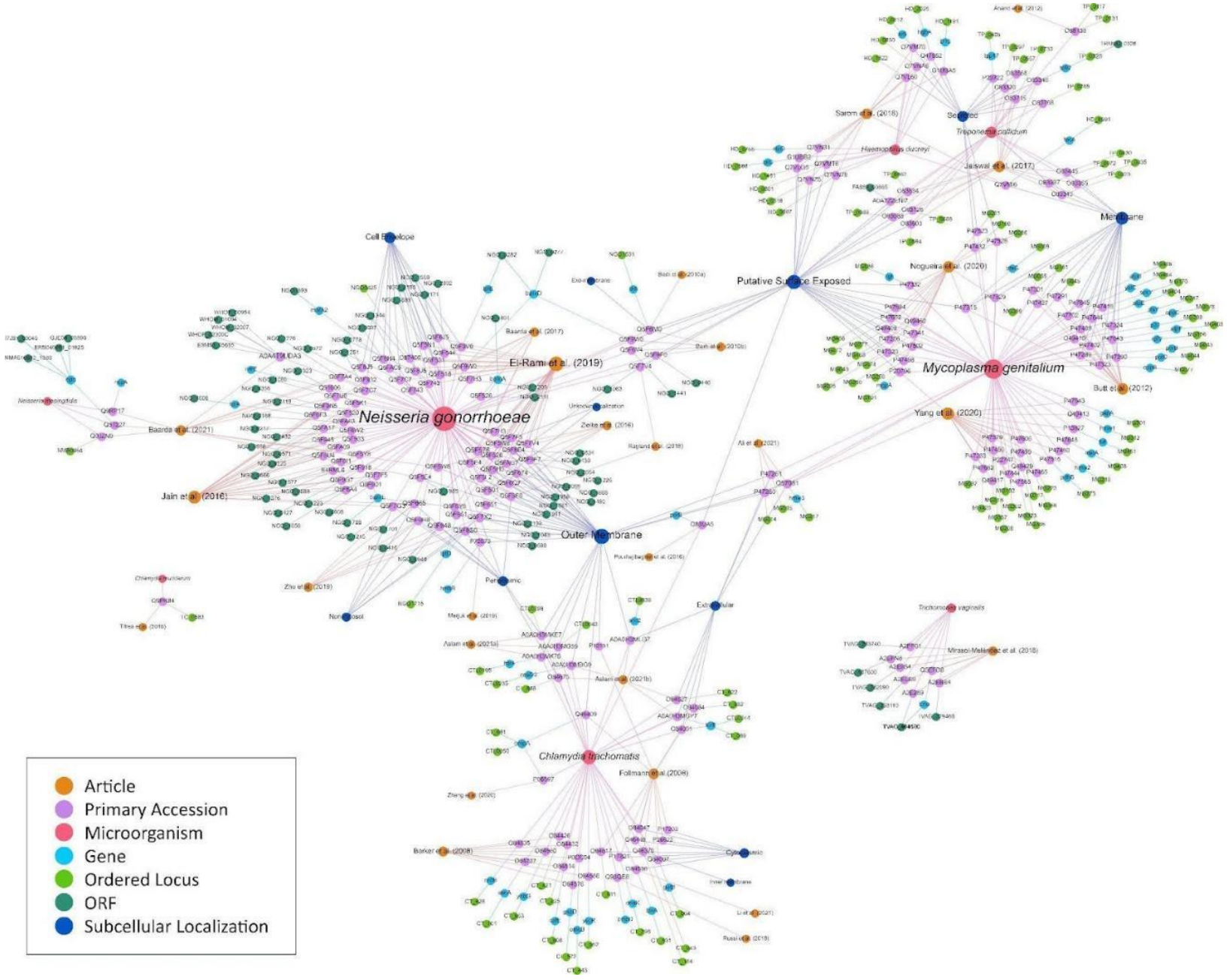


Figure 9. Upset plot showing the availability of information regarding vaccine candidates in the different articles

From the 27 selected articles that proposed vaccine candidates, 274 proteins were proposed as putative vaccine candidates, divided between the different STIs. 4 proteins had inactive

Uniprot entries, and 61 were unavailable in the Uniprot database. The ones that remained in the dataset were subjected to the following analyses. Figure 10 shows a complex network of vaccine candidates and the microorganisms in which they were identified. The most represented microorganisms were *N. gonorrhoeae*, *M. genitalium*, and *C. trachomatis*, with 88, 55 and 30 proposed candidates, respectively. Other bacteria represented within the selected documents had between 1 and 16 proposed proteins. An automated search between the annotated proteins found no typical candidates between the microorganisms. **Figure 10** shows a complex network of vaccine candidates associated with the microorganisms they have been identified in.

Regarding the subcellular localization of vaccine candidates, most proteins studied were either membrane, secreted, or otherwise surface exposed, proteins. There is an absence of standardization in the terms used to describe the proteins' sublocalization, with each article using differing terminology to describe the same meaning and imprecise language in many cases.



using differing terminology to describe the same meaning and imprecise language in many cases.

Figure 10. Complex network connecting vaccine candidates to the microorganisms in which they were identified. Node size is proportional to the node's degree. A color legend is embedded within the figure.

3.5 Epitope prediction and multi-epitope vaccine design

Only 3 selected papers (21, 602, and 609) designed a muti-epitope construct (**Figure 11**). The constructs were designed to be used as a vaccine against *C. trachomatis* infection in the first two and *M. genitalium* infection in the last one. To increase the immunogenic nature of the multi-epitope vaccine, Beta defensin, and Cholera toxin subunit B (CTB) proteins were coupled to the N-terminal of the epitopes in the *C. trachomatis* works, respectively. No adjuvant was indicated in the last paper.

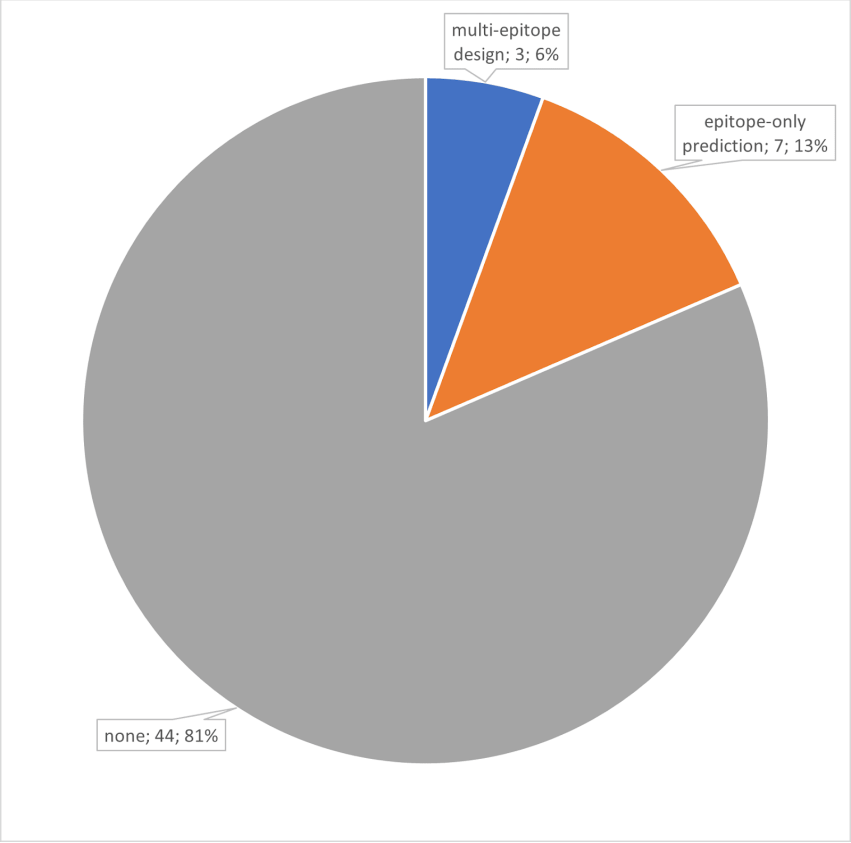


Figure 11. Pie chart of the total articles related to multi-epitope design or epitope-only prediction.

EAAAK, AAY, GPGPG, and KK linkers were used as connectors for epitopes. The primary role of these connectors is to link the immunogenic epitopes of the recombinant fusion protein. They may be flexible, with a certain degree of movement, rigid or cleavable linkers, capable of releasing epitopes in vivo [28]. The rigid EAAAK linker was used in all

three multi-epitope constructs as a separator of epitopes. KK and GPGPG linkers play a role in producing a better conformation of the construct (flexibility) [29]. The AAY linker is the cleavage site for the proteasomes in mammalian cells. Consequently, the epitopes that flank this linker will get disconnected within the cells [30]. The flexible GPGPG linker is reported to increase construct solubility, providing high accessibility and flexibility for adjacent domains [31].

Still, 7 other works (202, 220, 243, 450, 463, 603, and 643) suggested different epitopes as suitable activators of MHC-I, MHC-II, and/or B-cell [**Supplementary Material S8**]. These epitopes were predicted to activate the host immune system effectively. However, they were not included in the design of a multi-epitope construct. Therefore, these studies were discarded in the multi-epitope vaccine design analysis. As an epitope prediction, theses articles correctly indicate their vaccine candidate source. However, with no standardization of what protein ID code (NCBI, Uniprot) is used, which might delay or make it difficult to make quick comparisons between works.

4. Discussion

Sexually Transmitted Infections continue to burden public health, healthcare systems, and vulnerable populations worldwide. Among them, chlamydia, gonorrhoea, syphilis, and trichomoniasis, a set of treatable diseases, are responsible for nearly 374 million new infections every year [1–3]. Despite being treatable and detectable, these infections have been in resurgence in highly developed countries, particularly within the at-risk populations of men who have sex with men (MSM) and female sex workers (FSW), with notable, measured increases, as is the case of syphilis in Europe, for instance [4,32]. The WHO's strategy in response to managing and controlling the spread of these infections points to the importance of developing potentially vital innovations in the fight against STIs. Some vital innovations include point-of-care testing and vaccine development, both of which can be achieved with the help of immunoinformatics [4,8,33]. Immunoinformatics-based approaches can be critical to these developments by allowing for the efficient screening of genome sets and the identification of putative candidates [34,35]. Particularly in the case of *T. pallidum*, where vaccine candidate identification and testing can be made costly by the bacterium's cultivation requirements and cell wall composition, the virtual screening of candidates is vital to vaccine development [36–38]. In addition to vaccine candidate identification, immunoinformatics also allows for the design of immunogenic peptides based on the sequences of existing vaccine candidates, referred to as multi-epitope vaccines [8,9,33]. Other genome and proteome-based approaches may also seek to screen for candidate drug targets in the genome, which may open new avenues for STI treatment [39].

This review aimed to identify fields in which immunoinformatics have already been applied within the larger context of STI research. In identifying areas where investment has already been applied and the ground has already been gained, our goal was to determine fields in which more investment of research time and funding can be beneficial. The field of immunoinformatics is comparatively recent, with the first paper applying it to bacterial STIs being published in 2004, and production was low for the following eleven years. More

recently, since 2016 and especially in 2021, there has been a rise in prominence for the field. Most of this research has been conducted in the North American and Asian continents, with the USA, China, Pakistan, and India being the main hubs in which it is being conducted. The bibliographic coupling-based citation network also points to this concentration of research efforts in tightly knit hubs.

We have observed that most immunoinformatics-based STI research efforts have centered on the pathogens *C. trachomatis* and *N. gonorrhoeae*, the two most common bacterial STIs. *Chlamydia* is estimated to have been responsible for 129 million new STI cases, and gonorrhea for 82 million new STI cases in 2020 [1–3]. However, other pathogens such as *T. pallidum*, *M. genitalium*, *H. ducreyi*, and *T. vaginalis* have received comparatively little attention despite also being responsible for a large number of yearly cases worldwide (7.1 million and 156 million in the cases of *T. pallidum* and *Trichomonas vaginalis*, respectively) [1–3], this represents an important gap in the knowledge we have acquired on these infections. The lack of studies regarding *T. vaginalis* is of particular concern, considering the global burden of trichomoniasis, and may be related to the additional complexity of protozoan genomics.

A significant gap was also seen in the methodologies applied to these microorganisms. While a significant portion of articles performed Reverse Vaccinology analyses to identify vaccine candidates within the genomes of the organisms in question, very few used these candidates to design a multi-epitope vaccine. Fewer articles still performed a Virtual Drug Screening for potential drug targets within the genome. A common thread between articles that performed both R.V. and V.D.S. analyses is a lack of consistency between methodological approaches. There was no standardization in the tools, methods used to determine good drug targets and vaccine candidates, or terminology to define a protein's subcellular localization. As for articles performing epitope prediction and multi-epitope vaccine design, only a minority designed a vaccine construct, while the rest only performed epitope prediction on vaccine candidates. Scientific standardization is vital to the quality and reproducibility of research, particularly in a field such as the Omics sciences, composed of various differing and interacting high-throughput techniques [40].

5. Conclusion

This review sought to evaluate the current literature on the development of vaccines for non-viral Sexually Transmitted Infections, specifically through the lens of Immunoinformatics or the application of bioinformatics and genomics data to a vaccine and drug development. In doing so, it was possible to identify gaps in the current state of the art, where diseases like Chlamydia or Gonorrhea have gotten significantly more attention than the others, which is expected, considering their global burden. However, diseases like Trichomoniasis and Syphilis, responsible for many infections yearly, still require work in vaccine candidate identification and development of multi-epitope vaccines. We have also found that works in the field have not been able to determine common drug and vaccine targets between the microorganisms, which could be used in the development of a multivalent vaccine. In addition, we point to issues in the standardization of procedures in the field of immunoinformatics and the need to determine and develop better quality and more consistent tools and protocols for this field. The development of standardized, reproducible protocols and the review and defining of accurate tools for the purpose would do much for the improvement of the field.

6. References

1. Sexually transmitted infections (STIs) [Internet]. World Health Organization. 2022 [cited 2022 Sep 12]. Available from: https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-(stis)

2. Global progress report on HIV, viral sexually transmitted infections, 2021. World Health Organization [Internet]. 2021 [cited 2022 Sep 12]; Available from: https://www.who.int/publications/i/item/9789240027077

3. Fatima F, Kumar S, Das A. Vaccines against sexually transmitted infections: an update. Clin Exp Dermatol [Internet]. 2022 Aug 1 [cited 2022 Sep 12];47(8):1454–63. Available from: https://onlinelibrary.wiley.com/doi/full/10.1111/ced.15223

4. Global health sector strategy on Sexually Transmitted Infections, 2016-2021 [Internet]. World Health Organization. [cited 2022 Sep 12]. Available from: https://www.who.int/publications/i/item/WHO-RHR-16.09

5. Oli AN, Obialor WO, Ifeanyichukwu MO, Odimegwu DC, Okoyeh JN, Emechebe GO, et al. Immunoinformatics and Vaccine Development: An Overview. Immunotargets Ther [Internet]. 2020 Feb 26 [cited 2022 Nov 5];9:13–30. Available from: https://www.dovepress.com/immunoinformatics-and-vaccine-development-an-overview-peer-reviewed-fulltext-article-ITT

6. Ramana J, Mehla K. Immunoinformatics and Epitope Prediction. Methods Mol Biol [Internet]. 2020 [cited 2022 Nov 5];2131:155–71. Available from: https://pubmed.ncbi.nlm.nih.gov/32162252/

7. Jaiswal AK, Tiwari S, Jamal SB, Barh D, Azevedo V, Soares SC. An In Silico Identification of Common Putative Vaccine Candidates against Treponema pallidum: A Reverse Vaccinology and Subtractive Genomics Based Approach. International Journal of Molecular Sciences 2017, Vol 18, Page 402 [Internet]. 2017 Feb 14 [cited 2022 Nov 5];18(2):402. Available from: https://www.mdpi.com/1422-0067/18/2/402/htm

8. Martinelli DD. In silico vaccine design: A tutorial in immunoinformatics. Healthcare Analytics. 2022 Nov 1;2:100044.

9. Moyle P, Toth I. Modern subunit vaccines: development, components, and research opportunities. ChemMedChem [Internet]. 2013 Mar [cited 2021 Jul 21];8(3):360–76. Available from: https://pubmed.ncbi.nlm.nih.gov/23316023/

10. Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ [Internet]. 2021 Mar 29 [cited 2022 Nov 15];372. Available from: https://www.bmj.com/content/372/bmj.n71

11. van Eck NJ, Waltman L. VOS: A new method for visualizing similarities between objects. Studies in Classification, Data Analysis, and Knowledge Organization [Internet]. 2007 [cited 2022 Nov 15];299–306. Available from: https://link.springer.com/chapter/10.1007/978-3-540-70981-7_34

12. van Eck NJ, Waltman L. Software survey: VOSviewer, a computer program for bibliometric mapping. Scientometrics [Internet]. 2010 Dec 31 [cited 2022 Nov 15];84(2):523–38. Available from: https://link.springer.com/article/10.1007/s11192-009-0146-3

13. Gu Z, Gu L, Eils R, Schlesner M, Brors B. circlize Implements and enhances circular visualization in R. Bioinformatics [Internet]. 2014 Apr 2 [cited 2022 Nov 15];30(19):2811–2. Available from: https://pubmed.ncbi.nlm.nih.gov/24930139/

14. Conway JR, Lex A, Gehlenborg N. UpSetR: an R package for the visualization of intersecting sets and their properties. Bioinformatics [Internet]. 2017 Sep 15 [cited 2022 Nov 15];33(18):2938–40. Available from: https://pubmed.ncbi.nlm.nih.gov/28645171/

15. Aslam M, Shehroz M, Ali F, Zia A, Pervaiz S, Shah M, et al. Chlamydia trachomatis core genome data mining for promising novel drug targets and chimeric vaccine candidates identification. Comput Biol Med. 2021 Sep 1;136:104701.

16. Mirasol-Meléndez E, Brieba LG, Díaz-Quezada C, López-Hidalgo M, Figueroa-Angulo EE, Ávila-González L, et al. Characterization of multiple enolase genes from Trichomonas vaginalis. Potential novel targets for drug and vaccine design. Parasitol Int. 2018 Aug 1;67(4):444–53.

17. Sarom A de, Jaiswal AK, Tiwari S, Oliveira L de C, Barh D, Azevedo V, et al. Putative vaccine candidates and drug targets identified by reverse vaccinology and subtractive genomics approaches to control Haemophilus ducreyi, the causative agent of chancroid. J R Soc Interface [Internet]. 2018 [cited 2022 Nov 15];15(142). Available from: https://royalsocietypublishing.org/doi/10.1098/rsif.2018.0032

18. Nogueira WG, Jaiswal AK, Tiwari S, Ramos RTJ, Ghosh P, Barh D, et al. Computational identification of putative common genomic drug and vaccine targets in Mycoplasma genitalium. Genomics. 2021 Jul 1;113(4):2730–43.

19. Butt AM, Tahir S, Nasrullah I, Idrees M, Lu J, Tong Y. Mycoplasma genitalium: a comparative genomics study of metabolic pathways for the identification of drug and vaccine targets. Infect Genet Evol [Internet]. 2012 Jan [cited 2022 Nov 15];12(1):53–62. Available from: https://pubmed.ncbi.nlm.nih.gov/22057004/

20. Yang Z, Hou J, Mu M, Wu SY. Subtractive proteomics and systems biology analysis revealed novel drug targets in Mycoplasma genitalium strain G37. Microb Pathog. 2020 Aug 1;145:104231.

21. Fatoba AJ, Okpeku M, Adeleke MA. Subtractive genomics approach for identification of novel therapeutic drug targets in Mycoplasma genitalium. Pathogens [Internet]. 2021 Aug 1 [cited 2022 Nov 28];10(8):921. Available from: https://www.mdpi.com/2076-0817/10/8/921/htm

22. Barh D, Narayan Misra A, Kumar A. In Silico Identification of Dual Ability of N. gonorrhoeae ddl for Developing Drug and Vaccine Against Pathogenic Neisseria and Other Human Pathogens. J Proteomics Bioinform [Internet]. 2010 [cited 2022 Nov 15];3(3):82. Available from: http://myhits.isb-sib.ch/cgi-

23. Ragland SA, Humbert M v., Christodoulides M, Criss AK. Neisseria gonorrhoeae employs two protein inhibitors to evade killing by human lysozyme. PLoS Pathog [Internet]. 2018 Jul 1 [cited 2022 Nov 15];14(7):e1007080. Available from: https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1007080

24. El-Rami FE, Zielke RA, Wi T, Sikora AE, Unemo M. Quantitative Proteomics of the 2016 WHO Neisseria gonorrhoeae Reference Strains Surveys Vaccine Candidates and Antimicrobial Resistance Determinants. Mol Cell Proteomics [Internet]. 2019 Jan 1 [cited 2022 Nov 15];18(1):127–50. Available from: https://pubmed.ncbi.nlm.nih.gov/30352803/

25. Tanwer P, Kolora SRR, Babbar A, Saluja D, Chaudhry U. Identification of potential therapeutic targets in Neisseria gonorrhoeae by an in-silico approach. J Theor Biol. 2020 Apr 7;490:110172.

26. Kumar Jaiswal A, Jamal SB, Gabriel L, Gomes R, Profeta R, Sales-Campos H, et al. Neuroinformatics Insights towards Multiple Neurosyphilis Complications. Venereology 2022, Vol 1, Pages 135-160 [Internet]. 2022 Jun 6 [cited 2022 Nov 15];1(1):135–60. Available from: https://www.mdpi.com/2674-0710/1/1/10/htm

27. Dwivedi UN, Tiwari S, Singh P, Singh S, Awasthi M, Pandey VP. Treponema pallidum Putative Novel Drug Target Identification and Validation: Rethinking Syphilis Therapeutics with Plant-Derived Terpenoids. https://home.liebertpub.com/omi [Internet]. 2015 Feb 14 [cited 2022 Nov 15];19(2):104–14. Available from: https://www.liebertpub.com/doi/10.1089/omi.2014.0154

28. Chen X, Zaro JL, Shen WC. Fusion protein linkers: Property, design and functionality. Adv Drug Deliv Rev. 2013 Oct 15;65(10):1357–69.

29. Dong R, Chu Z, Yu F, Zha Y. Contriving Multi-Epitope Subunit of Vaccine for COVID-19: Immunoinformatics Approaches. Front Immunol. 2020 Jul 28;11:1784.

30. Ayyagari VS, Venkateswarulu TC, Abraham Peele K, Srirama K. Design of a multi-epitope-based vaccine targeting M-protein of SARS-CoV2: an immunoinformatics approach. https://doi.org/101080/0739110220201850357 [Internet]. 2020 [cited 2022 Nov 15];40(7):2963–77. Available from: https://www.tandfonline.com/doi/abs/10.1080/07391102.2020.1850357

31. Tarrahimofrad H, Rahimnahal S, Zamani J, Jahangirian E, Aminzadeh S. Designing a multi-epitope vaccine to provoke the robust immune response against influenza A H7N9. Scientific Reports 2021 11:1 [Internet]. 2021 Dec 29 [cited 2022 Nov 15];11(1):1–22. Available from: https://www.nature.com/articles/s41598-021-03932-2

32. Spiteri G, Unemo M, Mårdh O, Amato-Gauci AJ. The resurgence of syphilis in high-income countries in the 2000s: a focus on Europe. Epidemiol Infect [Internet]. 2019 [cited 2022 Sep 12];147:e143. Available from:

https://www.cambridge.org/core/journals/epidemiology-and-infection/article/resurgence-of-syphilis-in-highincome-countries-in-the-2000s-a-focus-on-europe/DF9411C5F899819985C34B97FF404217

33. Jameie F, Dalimi A, Pirestani M, Mohebali M. Development of a Multi-Epitope Recombinant Protein for the Diagnosis of Human Visceral Leishmaniasis. Iran J Parasitol [Internet]. 2021 Feb 16 [cited 2022 Mar 14];16(1):1–10. Available from: https://ijpa.tums.ac.ir/index.php/ijpa/article/view/1869

34. Jaiswal A, Tiwari S, Jamal S, Barh D, Azevedo V, Soares S. An In Silico Identification of Common Putative Vaccine Candidates against Treponema pallidum: A Reverse Vaccinology and Subtractive Genomics Based Approach. Int J Mol Sci [Internet]. 2017 Feb 14 [cited 2021 Jul 21];18(2). Available from: https://pubmed.ncbi.nlm.nih.gov/28216574/

35. Pizza M, Scarlato V, Masignani V, Giuliani MM, Aricò B, Comanducci M, et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. Science [Internet]. 2000 Mar 10 [cited 2022 Nov 15];287(5459):1816–20. Available from: https://pubmed.ncbi.nlm.nih.gov/10710308/

36. Hook EW. Syphilis. The Lancet [Internet]. 2017 Apr 15 [cited 2021 Jul 21];389(10078):1550–7. Available from: http://www.thelancet.com/article/S0140673616324114/fulltext

37. Singh T, Otero CE, Li K, Valencia SM, Nelson AN, Permar SR. Vaccines for Perinatal and Congenital Infections—How Close Are We? Front Pediatr. 2020 Dec 15;0:569.

38. Cameron CE, Lukehart SA. Current Status of Syphilis Vaccine Development: Need, Challenges, Prospects. Vaccine [Internet]. 2014 Mar 3 [cited 2022 Jun 19];32(14):1602. Available from: /pmc/articles/PMC3951677/

39. Barh D, Tiwari S, Jain N, Ali A, Santos AR, Misra AN, et al. In silico subtractive genomics for target identification in human bacterial pathogens. Drug Dev Res [Internet]. 2011 Mar 1 [cited 2022 Nov 15];72(2):162–77. Available from: https://onlinelibrary.wiley.com/doi/full/10.1002/ddr.20413

40. Holmes C, McDonald F, Jones M, Ozdemir V, Graham JE. Standardization and Omics Science: Technical and Social Dimensions Are Inseparable and Demand Symmetrical Study. https://home.liebertpub.com/omi [Internet]. 2010 Jun 22 [cited 2022 Nov 15];14(3):327–32. Available from: https://www.liebertpub.com/doi/10.1089/omi.2010.0022

SUPPLEMENTARY MATERIAL

PubMed

(

("immunoinformatics" OR "Reverse Vaccinology" OR "bioinformatics vaccine targets" OR "rational vaccine design" OR "rational immunogen design" OR "Potential Universal Vaccine Candidates" OR "In silico vaccine design" OR "in silico immunogen design" OR "epitope-based vaccine" OR "epitope-based immunogen" OR "Structure-based immunogen design" OR "Structure-based antigen design" OR "in silico vaccine targets" OR "potential vaccine candidates" OR "subtractive genomics" OR "computational vaccinology" OR "vaccinology")

AND

(

("STD vaccine" OR "sexually transmitted infections" OR "sexually transmitted" OR "Treponema pallidum" OR "syphilis" OR "Mycoplasma genitalium" OR "Neisseria gonorrhoeae" OR "gonorrhea" OR "Haemophilus ducreyi" OR "chancroid" OR "Chlamydia trachomatis" OR "chlamydia")

NOT ("virus")

)

)

126

Scopus

(

ALL("immunoinformatics" OR "Reverse Vaccinology" OR "bioinformatics vaccine targets" OR "rational vaccine design" OR "rational immunogen design" OR "Potential Universal Vaccine Candidates" OR "In silico vaccine design" OR "in silico immunogen design" OR "epitope-based vaccine" OR "epitope-based immunogen" OR "Structure-based immunogen design" OR "Structure-based antigen design" OR "in silico vaccine targets" OR "potential vaccine candidates" OR "subtractive genomics" OR "computational vaccinology" OR "vaccinology")

AND

(

ALL("STD vaccine" OR "sexually transmitted infections" OR "sexually transmitted" OR "Treponema pallidum" OR "syphilis" OR "Mycoplasma genitalium" OR "Neisseria gonorrhoeae" OR "gonorrhea" OR "Haemophilus ducreyi" OR "chancroid" OR "Chlamydia trachomatis" OR "chlamydia")

AND

NOT ALL("virus")

)

)

778

Web of Science

```
# ----------------

ALL=(

        ("immunoinformatics" OR "Reverse Vaccinology" OR "bioinformatics vaccine targets" OR "rational vaccine design" OR "rational immunogen design" OR "Potential Universal Vaccine Candidates" OR "In silico vaccine design" OR "in silico immunogen design" OR "epitope-based vaccine" OR "epitope-based immunogen" OR "Structure-based immunogen design" OR "Structure-based antigen design" OR "in silico vaccine targets" OR "potential vaccine candidates" OR "subtractive genomics" OR "computational vaccinology" OR "vaccinology")

        AND

        (

        ("STD vaccine" OR "sexually transmitted infections" OR "sexually transmitted" OR "Treponema pallidum" OR "syphilis" OR "Mycoplasma genitalium" OR "Neisseria gonorrhoeae" OR "gonorrhea" OR "Haemophilus ducreyi" OR "chancroid" OR "Chlamydia trachomatis" OR "chlamydia")

        NOT ("virus")

        )

)

# 67
```

Supplementary Material S1. Keyword combinations and Boolean operators for article search.

https://github.com/lbmcf/format-input

Supplementary Material S2. Github repository containing the script used to extract basic information from the downloaded articles and save them as a CSV file.

https://github.com/lbmcf/remove-duplicates

Supplementary Material S3. The Github repository contains the script to unify records between all searched databases, removing redundancy.

https://docs.google.com/spreadsheets/d/1vWmoiyEsrHeq85d_kGLF3t71BZUIWym9/edit?usp=share_link&ouid=101463300592951365557&rtpof=true&sd=true

Supplementary Materials S4-S8 have been provided in a separate file:

Supplementary Material S4. List of articles that fit all inclusion criteria. Basic information for each article has been extracted.

Supplementary Material S5. Methodologies applied in each reviewed article.

Supplementary Material S6. Drug targets are identified in each reviewed article.

Supplementary Material S7. Vaccine candidates are identified in each reviewed article.

Supplementary Material S8. Epitope prediction and multi-epitope vaccine construction in each reviewed article.

3.2. Discussion

This systematic review of the literature is presented as chapter 1 of this dissertation, as both an introduction and review of concepts in STI and immunoinformatics research, and as justification for the need of work centered on less-researched microorganisms. In our review we have identified a concentration of efforts in research regarding *Neisseria gonorrhoeae* and *Chlamydia trachomatis*, while others such as *Treponema pallidum* and *Trichomonas vaginalis* have had less efforts directed towards them. This distribution of research initiatives is in line with the relative burden of the diseases, as gonorrhea and chlamydia are responsible for significantly more cases per year than syphilis.

Nonetheless, we have identified that additional research efforts to combat the spread of syphilis are necessary. There is, as of yet, no functioning vaccine protective against *T. pallidum* infection, and none are in trial currently. Attempts to develop one have, so far, met with numerous limitations in regards to the pathogen's difficulty to culture and hard-to-purify outer membrane. Therefore, the application of computational techniques to efforts in finding vaccine candidates and developing an effective immunization against the disease makes itself necessary. In the following chapter 2, we have made use of previously identified vaccine candidates for *T. pallidum* in an effort to develop a multi-epitope vaccine for syphilis.

4. CHAPTER 2

4.1. Published Research Article

***In silico* designed multi-epitope immunogen "Tpme-VAC/LGCM-2022" may induce both cellular and humoral immunity against *Treponema pallidum* infection**

Lucas Gabriel Rodrigues Gomes, Thaís Cristina Vilela Rodrigues, **Arun Kumar Jaiswal, Roselane Gonçalves Santos**, Rodrigo Bentes Kato, Debmalya Barh, Khalid J. Alzahrani, Hamsa Jameel Banjer, Siomar de Castro Soares, **Vasco Azevedo**, Sandeep Tiwari

Published in: *Vaccines* **2022**, *10*(7), 1019; https://doi.org/10.3390/vaccines10071019

In this research article, we have designed a multi-epitope vaccine based on some promising vaccine candidates for *Treponema pallidum*, the causative agent of syphilis. Vaccine candidates were derived from the literature encompassing both traditional and reverse vaccinology, then scanned for immunogenic epitopes, which were screened and selected to compose a chimeric vaccine construct. Several features and traits of this construct were modeled from its sequence and point to it being an adequately immunogenic, safe, and synthesizable peptide.

vaccines

MDPI

Article

In Silico Designed Multi-Epitope Immunogen "Tpme-VAC/LGCM-2022" May Induce Both Cellular and Humoral Immunity against *Treponema pallidum* Infection

Lucas Gabriel Rodrigues Gomes [1,†], Thaís Cristina Vilela Rodrigues [1,†], Arun Kumar Jaiswal [1,†], Roselane Gonçalves Santos [1], Rodrigo Bentes Kato [1], Debmalya Barh [1,2], Khalid J. Alzahrani [3], Hamsa Jameel Banjer [3], Siomar de Castro Soares [4], Vasco Azevedo [1,*] and Sandeep Tiwari [1,*]

1 Laboratory of Cellular and Molecular Genetics (LGCM), PG Program in Bioinformatics, Department of Genetics, Ecology, and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais (UFMG), Belo Horizonte 31270-901, Brazil; lucasgabrielrgomes@ufmg.br (L.G.R.G.); thaiscristinavr@gmail.com (T.C.V.R.); arunjaiswal1411@gmail.com (A.K.J.); roselanegr@gmail.com (R.G.S.); rbkato@gmail.com (R.B.K.); dr.barh@gmail.com (D.B.)
2 Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur 721172, West Bengal, India
3 Department of Clinical Laboratories Sciences, College of Applied Medical Sciences, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia; ak.jamaan@tu.edu.sa (K.J.A.); h.banjer@tu.edu.sa (H.J.B.)
4 Department of Immunology, Microbiology, and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro (UFTM), Uberaba 38025-180, Brazil; siomars@gmail.com
* Correspondence: vascoariston@gmail.com (V.A.); sandip_sbtbi@yahoo.com (S.T.)
† These authors contributed equally to this work.

check for updates

Citation: Gomes, L.G.R.; Rodrigues, T.C.V.; Jaiswal, A.K.; Santos, R.G.; Kato, R.B.; Barh, D.; Alzahrani, K.J.; Banjer, H.J.; Soares, S.d.C.; Azevedo, V.; et al. In Silico Designed Multi-Epitope Immunogen "Tpme-VAC/LGCM-2022" May Induce Both Cellular and Humoral Immunity against *Treponema pallidum* Infection. *Vaccines* **2022**, *10*, 1019. https://doi.org/10.3390/vaccines10071019

Academic Editor: Daniel Dory

Received: 24 May 2022
Accepted: 23 June 2022
Published: 25 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Abstract: Syphilis, a sexually transmitted infection caused by the spirochete *Treponema pallidum*, has seen a resurgence over the past years. *T. pallidum* is capable of early dissemination and immune evasion, and the disease continues to be a global healthcare burden. The purpose of this study was to design a multi-epitope immunogen through an immunoinformatics-based approach. Multi-epitope immunogens constitute carefully selected epitopes belonging to conserved and essential bacterial proteins. Several physico-chemical characteristics, such as antigenicity, allergenicity, and stability, were determined. Further, molecular docking and dynamics simulations were performed, ensuring binding affinity and stability between the immunogen and TLR-2. An in silico cloning was performed using the pET-28a(+) vector and codon adaptation for *E. coli*. Finally, an in silico immune simulation was performed. The in silico predictions obtained in this work indicate that this construct would be capable of inducing the requisite immune response to elicit protection against *T. pallidum*. Through this methodology we have designed a promising potential vaccine candidate for syphilis, namely Tpme-VAC/LGCM-2022. However, it is necessary to validate these findings in in vitro and in vivo assays.

Keywords: *Treponema pallidum*; sexually transmitted infection; syphilis; chimeric multi-epitope vaccine; immunoinformatics

1. Introduction

Syphilis is a sexually transmitted infection (STI) caused by *Treponema pallidum* subspecies *pallidum*, which belongs to the genus *Treponema*. Other pathogens from the genus cause various other non-venereal infections, such as endemic syphilis (*T. pallidum* subsp. *endemicum*), yaws (*T. pallidum* subsp. *pertenue*), and pinta (*T. carateum*) [1]. *T. pallidum* subsp. *pallidum* is a spiral-shaped, slow-growing, obligate human pathogen with a genome size of approximately 1.14 Mb, a protein count of 967, and a GC% content of 52.8%. Long term culture methods for cultivating *T. pallidum* have only recently been developed, and require co-culture with Sf1Ep cottontail rabbit epithelial cells [2]. The disease occurs in three stages:

primary (First two to three weeks, chancres at site of infection), secondary (lesions and rashes throughout body, can take months to resolve), and tertiary (long term neural and cardiovascular complications), alternated with a latent stage (asymptomatic), which can occur at any point of the infection [3]. *Treponema pallidum* subsp. *pallidum* can also cause congenital syphilis in infants through vertical transmission, known as mother-to-child transmission (MTCT). The infection occurs during gestation and can lead to fatal infections in neonates [3].

Syphilis continues to be a worldwide healthcare burden. Even though the infection is easily identifiable and treatable, it is endemic in developing countries and on the rise within selected populations, specifically, men who have sex with men (MSM) and female sex workers (FSW), particularly within the HIV-positive population in developed countries [4]. According to the World Health Organization (WHO), six million people are infected with syphilis each year, and 2.5 million cases of yaws, bejel, and pinta are reported each year [3,5].

Currently, Africa, the Americas, and the Western Pacific are the regions most afflicted by incident cases [5]. Of those, an exceptional burden is on populations of pregnant women. Untreated syphilis in pregnancy is a leading cause of fetal and antenatal morbidity and mortality, resulting in high numbers of stillbirths, preterm infants, and cases of congenital syphilis [4]. According to the WHO, syphilis causes about 200,000 fetal and neonatal deaths each year, and about 215,000 infants are placed at risk of early death [5]. Gay men and sex workers are two other populations at-risk, with a high prevalence of STIs such as syphilis [5].

While syphilis can still be treated with penicillin with no occurrence of penicillin-resistant strains, there have been cases of macrolide-resistant strains, such as Azithromycin [6]. The disease's prevalence in spite of the pathogen's sensitivity to penicillin indicates the pathogen is unlikely to be controlled through screening and treatment alone [7]. Despite concerted efforts from the WHO to contain congenital syphilis and joint efforts to halt the spread of sexually transmitted syphilis, it remains a challenging disease to tackle as a result of various limiting factors [8]. Among said limiting factors, the pathogen's difficulty to cultivate and study, the social stigma associated with the disease, its comorbidity with HIV, and the absence of a vaccine are of note [8].

Proof-of-principle for successful syphilis vaccination in the rabbit model was established in 1973, indicating development of a *T. pallidum* vaccine could be viable, but the immunization procedure demonstrated then was untenable for human application [9]. Attempts to develop a syphilis vaccine have since focused on the targeting of the bacterium's outer membrane proteins (OMPs), of which there are few, and the few that are known are difficult to isolate [7,9]. Certain OMPs had limited potential in eliciting a protective response, suggesting that no single protein will confer full protective immunity against *T. pallidum* [10]. In this study, we sought to tackle the limitations to syphilis vaccine design through the use of computational biology and using an immunoinformatics-based approach.

Modern computational biology techniques, such as reverse vaccinology (RV) and immunoinformatics, have proven to be powerful tools for reducing resources and minimising the time typically spent in developing vaccines [11]. The usage of RV allows us to identify potential vaccine targets of interest in the pathogen's genome while ensuring that no host-homologous sequences are in use. Immunoinformatics allows us to filter epitopes of the target protein according to their capacity to induce an immune response in the host [12]. It has been used to identify vaccine targets and drive vaccine development for a number of bacteria, viruses, parasites, fungi, as well as for cancers [13].

2. Materials and Methods

2.1. Selection of Target Antigenic Proteins

In order to search for and determine epitopes capable of eliciting an immune response from the host, 15 proteins, previously determined to be potential vaccine targets through reverse vaccinology, were selected from the work of Jaiswal et al. 2017 [14]. Their analysis

of the Pan-genome of all available *T. pallidum* genomes identified potential vaccine targets in the core genome of *T. pallidum* sequences. Since these are core proteins, which are present in all *T. pallidum* strains, and were predicted to be surface-exposed, they are likely to have high expression levels across multiple strains and high immunogenic potential [14]. In addition, the proteins are non-homologous to the host, likely essential to the pathogen, and were predicted to be potential antigens [14]. Three other proteins considered to be immunogenic in previous in vivo studies using the outbred rabbit model were also added [15–17]. The amino-acid sequences for each protein, used as the database for this search, were retrieved from the National Center for Biotechnology Information (NCBI) Database. Eighteen proteins were selected for analysis (Table S1).

2.2. Prediction of MHC-I Allele Binding CTL Epitopes

For the cytotoxic T lymphocyte (CTL) epitope prediction, in an effort to improve the confidence in the selected epitopes, two different platforms were used, both of which were used to predict nine amino-acid residue long sequences. The Immune Epitope Database and Analysis Resource (IEDB-AR) is robust and has multiple epitope prediction tools [18]. The major histocompatibility complex I (MHC-I) binding epitope prediction tool was used to identify CTL epitopes in the target proteins. In order to design an immunogen capable of inciting a response in a wide range of population worldwide, a reference set of 27 MHC-I alleles, which has a high frequency in the global population, was selected for epitope binding [19]. The IEDB-AR recommended 2020.09 (NetMHCpan EL 4.1) prediction method was used. Only epitopes having a percentile rank of <1% and an IC50 of <500 nM were selected.

We also applied a tool to assess antigenic processing and transportation. This tool, NETCTL-1.2, uses the ANN and SMM methods to perform predictions. All of the allele supertypes (A1, A2, A3, A24, A26, B7, B8, B44, and B58) [20], which were used in the previous step and are available in this platform, were also used in this step.

2.3. Prediction of MHC-II Allele Binding HTL Epitopes

To predict MHC-II epitopes for binding to Helper T Lymphocytes (HTL), we used two high-quality predictors [21] to identify the higher confidence epitopes, with a standard sequence length of fifteen amino acid residues. The IEDB-AR MHC-II binding epitope [18] and NETMHCII-2.3 [22] tools were used in combination. In this step, only IEDB-AR epitopes with a percentile rank < 3% and IC50 < 1000 nM were kept in the study. The platform NETMHCII-2.3 used an ANN with diverse epitope databases to perform the predictions. High-frequency alleles described in Greenbaum et al., 2011 [19] were used in the MHC-II prediction, both for the IEDB-AR and the NETMHCII 2.3 predictions.

2.4. Prediction of B-Cell Epitopes

We have used the ABCpred tool, an ANN-based tool for the prediction of linear B-cell epitopes, using the default parameters [23].

2.5. Filtering Best Epitopes from Each Protein

We run an in-house python script to determine shared epitopes between the two MHC-I and MHC-II prediction methods and select the predicted epitopes with high confidence between the prediction methods. Finally, we run the script to select overlapping epitopes between MHC-I and B epitopes and MHC-II and B epitopes, that is, epitopes capable of inducing cellular and humoral immune responses. In this case, the window was 2–9 residues for CTL epitopes and 2–15 residues for HTL epitopes. Only the epitopes that showed overlaps in the two cases were selected for further analysis. In order to filter down the number of epitopes and carefully define the final structure, further filtering steps were performed: MHC-II epitopes were filtered based on IC-50, keeping only epitopes with IC50 of up to 50 nM [24]. MHC-I epitopes were filtered using IEDB-AR's immunogenicity tool with a cut-off of 0.1, indicating a higher probability of selecting immunogenic epitopes [25].

Only overlapping CTL and HTL epitopes were selected. The remaining epitopes that differed by only one or two residues were filtered, selecting only those with lower IC50 or higher immunogenicity. The final epitopes were used to construct two different chimeric proteins, which were compared in regards to the overall population coverage of the alleles used in their construction [26], as well as their physico-chemical properties.

2.6. Construction of Multi-Epitope Immunogen Sequence

To determine the final sequence of the chimeric protein, the final epitopes were merged using appropriate linker peptides, the purpose of which is to assist in protein folding and processing. CTL epitopes were linked by AAY linkers, and MHC-II epitopes were joined by GPGPG linkers [11]. The selected adjuvant was the cholera enterotoxin B-subunit (ctxB) [27], linked to the rest of the sequence by the peptide linker EAAAK. Two separate immunogens were constructed, and the immunogen with the higher overall population coverage was selected.

2.7. Prediction of Antigenicity, IFN-γ Induction, Toxicity, and Allergenicity of the Multi-Epitope Immunogen

The final chimeric protein structure was subjected to several analyses to answer important questions regarding its induction of immune response, allergic and toxic potential, and physico-chemical properties. First, VaxiJen was used to assess the antigenic capacity of the amino acid sequence through the automatic cross-covariance method, evaluating the physico-chemical properties of the protein and predicting its immunogenicity without performing alignments [28]. Epitopes capable of inducing IFN-γ production with consequent TCD4+ lymphocyte activation were identified with the IFNepitope predictor, which uses an SVM hybrid method based on protein motifs to perform prediction [29]. The protein sequence was then evaluated for toxic potential by submitting it to Toxinpred [30]. In addition, Allertop v.2.0 was used to evaluate the protein's allergenic propensity based on the amino acid chain structure [31].

2.8. Physico-Chemical Properties and Host and Microbiota Homology Analyses

The molecular mass, theoretical pI, extinction coefficient, aliphatic index, grand average of hydropathicity (GRAVY), estimated half-life for three model organisms (Escherichia coli, yeast, and mammal cells), and instability index of the final protein sequence were evaluated through the ProtParam tool [32]. The solubility index was measured through the Protein-Sol tool [33], which evaluates the protein based on E. coli expression data. The Pipeline Builder for Identification of drug targets for infectious diseases (PBIT) tool [34] was used to search for homology between the chimeric protein and the proteome of the host as well as the gut microbiota of the host.

2.9. Secondary Structure Prediction

The PSIPRED prediction tool was used to determine the secondary structure of the final protein. This tool makes use of a complex ANN and Position-Specific Scoring Matrix (PSSM) based approach to predict the structure and generate pictures [35]. The RaptorX tool was used to provide the ratios of β-strands, α-helixes, and coils.

2.10. Tertiary Structure and Refinement

We used Phyre2 intensive model, RaptorX, and I-TASSER servers to select the best structure for tertiary structure prediction. The Phyre2 intensive method consists of multiple alignments of the sequence of interest with homologous sequences, followed by secondary structure prediction with PSIPRED. Information from these two steps was combined to determine a hidden Markov model. A search for this model was performed in an HMM database of proteins with known structures, and the model with the best score was used to determine modelling and correct errors [36]. RaptorX, which uses multiple-template threading (MTT) and scoring methods to indicate the quality of the models [37]. Finally,

I-TASSER constructed the model with an iterative method based on templates according to fragment assembly simulations with further refinement [38]. Methods for the refinement of amino acid side chains using light and aggressive relaxation were applied through the GalaxyRefiner tool to improve model quality by enhancing the local and global structure of the chimeric protein [39]. To check the quality of the refined structure, we used the PROCHECK tool, available in the SAVES server V6.0, to generate the Ramachandran plot, comparing the structure of the chimeric protein with the geometry of amino acid residues resulting from high-quality structures [40].

2.11. Prediction of Conformational B Cell Epitopes

Conformational epitopes are indispensable in stimulating immune responses. The refined structure was submitted to ElliPro to predict these discontinuous epitopes [41].

2.12. Molecular Docking between the Chimeric Protein and the TLR-2 Recepto

The Toll-like receptor 2 was identified as a vital receptor in detecting T. pallidum infection and assembling an effective immune response against the pathogen [42]. We retrieved the Toll-like receptor-2 (TLR-2) structure from RCSB: (PDB ID: 2z7x) database to determine the interactions of the chimeric protein with this receptor. The structure was edited with the Chimera visualization software, removing water molecules, ligands and side chains [43]. In order to verify the interactions between the chimeric protein and the TLR-2, molecular docking was performed using the Swarmdock server. The proteins were subjected to blind docking, attempting to find the lowest energy conformations across the whole protein. [44]. Hydrogen bonds and hydrophobic interactions were evaluated using the LigPlot^{+} program [45]. The PDBePISA tool was used to calculate the solvation free energy gain ($\Delta^{i}G$) of the final selected complex [46].

2.13. Molecular Dynamics Simulation of the Receptor-Ligand Complex

We performed the molecular dynamics simulation using the Gromacs v5.0 program [47] to enhance understanding of the microscopic structural properties of the interaction between the chimeric protein and the Toll-like receptor. To set simulation parameters, we prepared the software in the following manner: To construct protein topology and information about bonded and non-bonded characteristics, pdb2gmx will be used. The structure will be solvated in a cubic box of TIP3P water molecules. The complete system simulation was performed with the GROMOS96 43A1 force field, and a concentration of 150 mM sodium chloride (NaCl) ions were introduced to neutralize the system. Energy minimization was executed to ensure the quality of the system's geometry and the absence of steric clashes. For this purpose, the steepest descent algorithm was applied. Simulation time was 90 ns.

2.14. In Silico Cloning

In silico cloning was performed to verify the capacity of cloning and expression of the protein in an appropriate expression vector. For this, the codon usage of our peptide sequence was adapted according to the codon usage of the *E. coli* expression system. For this purpose, the JCat tool was used for reverse translation. From that cDNA sequence, the codon optimisation for *E. coli* k12 was performed, returning the Codon Adaptation Index (CAI), which must have a score higher than 0.8, and the GC content should be between 30–70% [48]. Simulation and visualization of the in silico cloning were performed through the SnapGene® software (from Insightful Science, available at snapgene.com) accessed on 1 October 2021, where the sequence of the chimeric protein was inserted in the pET-28a(+) plasmid with the help of Blpi and BamHI restriction enzymes.

2.15. Immune Simulation of Multi-Epitope Immunogen

We used the C-ImmSim server to run an immunological simulation to enhance the description of the immune response outlined by the chimeric protein [49]. The in silico

method uses PSSM for epitope prediction and machine learning to assess interactions. The model also simulates the anatomical regions where crucial events of immunity occur: the bone marrow, where the lymphoid and myeloid cells are produced; the thymus, where the autoreactivity process happens; and the tertiary lymphatic organ, where antigenic presentation occurs, which describes the immunogenic profile. Three injections containing 1000 immunogen proteins each were given at four-week intervals for the simulation. Time steps were set at 1, 84, and 168 (each step representing eight real-life hours and time step 1 being injection time = 0). The total steps were modified to 1050, and other parameters were kept at the default. To check the effectiveness of the selected epitopes, we used the C-ImmSim tool again to simulate injections for only the adjuvant sequence while maintaining parameters, as described above.

3. Results

3.1. Predicted CTL Epitopes

Epitopes were predicted using the IEDB-AR database, yielding epitopes that can be recognized by MHC-I alleles with high frequency in the global population for all 18 proteins under analysis. We also submitted these proteins to the NETCTL 1.2 server to improve the confidence of the chosen epitopes. These epitopes had the size of nine amino acid residues.

3.2. Predicted HTL and B-Cell Epitopes

MHC-II epitopes were predicted through the IEDB-AR and netMHCII 2.3 tools. IEDB-AR epitopes were predicted to bind to the most common MHC-II alleles, according to the IEDB-AR database. These epitopes had a size of 15 amino acid residues.

The protein sequences were submitted to the ABCpred tool, and 16-mer epitopes, with scores higher than 0.51, were predicted for all 18 proteins according to the ability to interact with B lymphocyte receptors.

3.3. Overlapping Epitopes for Both Humoral and Cellular Responses

High confidence epitopes were selected based on their overlap between the two methods for each category, MHC-I and MHC-II, for each protein at a time. Finally, epitopes capable of inducing both humoral and cellular responses were selected according to the overlap between each category and B epitopes. The number of overlapping MHC-I and B epitopes was 729. The number of overlapping MHC-II and B epitopes was 521. These epitopes were subjected to the next screening methods. After the IC-50 screening, 112 MHC-II/B epitopes fit the new threshold. When the immunogenicity screening was applied, 269 MHC-I/B epitopes had scores higher than 0.1. The remaining epitopes were subjected to a search for overlapping MHC-I and MHC-II epitopes, keeping 37 MHC-I epitopes and presenting sequences overlapping with 32 MHC-II epitopes. These epitopes were used to construct two different chimeric immunogens, which were then compared in regards to their overall population coverage and physico-chemical properties. The chimeric protein that was selected was composed of 11 MHC-I epitopes and 15 MHC-II epitopes belonging to ten different proteins (Table 1). The final version of Tpme-VAC/LGCM-2022 had an overall population coverage of 99.93% of HLA Alleles (Figure S1).

3.4. Constructed Multi-Epitope Vaccine Sequence (Tpme-VAC/LGCM-2022), and Host and Microbiota Homology

The Treponema Pallidum Multi Epitope-Vaccine/Laboratory of Cellular and Molecular Genetics-2022 (Tpme-VAC/LGCM-2022) is composed of the following sequences: the cholera enterotoxin B subunit (ctxB) followed by the linker peptide EAAAK. Then, CTL epitopes linked by AAY linker peptides, followed by HTL epitopes linked by GPGPG linker peptides (Figure 1A). The chimeric protein was found to be non-host and non-gut-microbiota homologous.

Table 1. Final epitopes selected for immunogen construct. Highlighted in bold is the overlap between MHC-I and MHC-II epitopes.

	GENE ID/NAME	MHC	EPITOPE	PERCENTILE RANK
1	TP_0049	I	HLRTFLAAV	0.12
2		II	CPSVC**HLRTFLAAV**R	0.9
3		I	SVCGPDFLY	0.22
4	TP_0323	I	ASVALFYAY	0.1
5		II	VGMAVA**ASVALFYAY**	1.1
6		II	IELFSALPYALTVVV	0.6
7		II	EGLMMFGAFSTATVT	0.7
8	TP_0335	I	AAAVTEYAF	0.14
9		II	VLHA**AAAVTEYAF**VL	0.8
10		I	AVHALWNAY	0.05
11		I	HALWNAYAI	0.21
12		II	V**HALWNAYAI**AAAAR	0.25
13		I	TLFAGAAGA	0.07
14		II	RPAGSA**TLFAGAAGA**	0.9
15	TP_0430/ntpK	I	AAAAGADAL	0.59
16		II	GR**AAAAGADAL**AETG	0.25
17		I	GMFGAAAVL	0.15
18		II	**GMFGAAAVL**GISAVG	0.4
19	TP_0435/nlpE	I	YMGAPGAGK	0.11
20	TP_0557	I	RAVRTLLII	0.72
21		II	KRMW**RAVRTLLII**CA	0.5
22	TP_0733	II	GGGGFHLGYEYFFTK	0.3
23	TP_0972/ftr1	II	VGVFVAIRFLSVRLP	0.12
24	TP_0326/BamA	II	GIVSFDFFFDAAMVY	0.12
25		II	GQKWTYELYLEILQK	0.03
26	tprK	II	DYAQARAPAAGAKVS	1.1

3.5. Secondary and Tertiary Structural Properties of Tpme-VAC/LGCM-2022

The result of PSIPRED showed that, among the 558 residues of the sequence, there was an arrangement of 48% α-helices, 15% β-strands, and 37% coil formation (Figure S2).

Three different prediction methods were applied to find the model with the best structural quality. The highest quality model was constructed by the RaptorX server, the Ramachandran plot of this model showing 92.9% of residues in most favored regions (Figure S3). After structural refinement through the GalaxyRefiner tool, the highest quality model had 94.9% amino acid residues in the most favored regions and only 0.2% in a disallowed region (Figure 1B,C).

3.6. Antigenicity, IFN-γ Production, and Conformational B-Cell Epitopes in Tpme-VAC/LGCM-2022

The chimeric protein sequence was determined as a probable antigen according to the VaxiJen tool, with a score of 0.6852. The IFNepitope tool, using the SVM method, predicted 230 epitopes with positive and negative scores. Among these, 69 epitopes had a score greater than 1, considered more capable of inducing the production of this cytokine

(Table S2). The final refined structure of the protein was submitted to ElliPro, and five conformational epitopes with scores above 0.7 were predicted (Figure S4, Table S3).

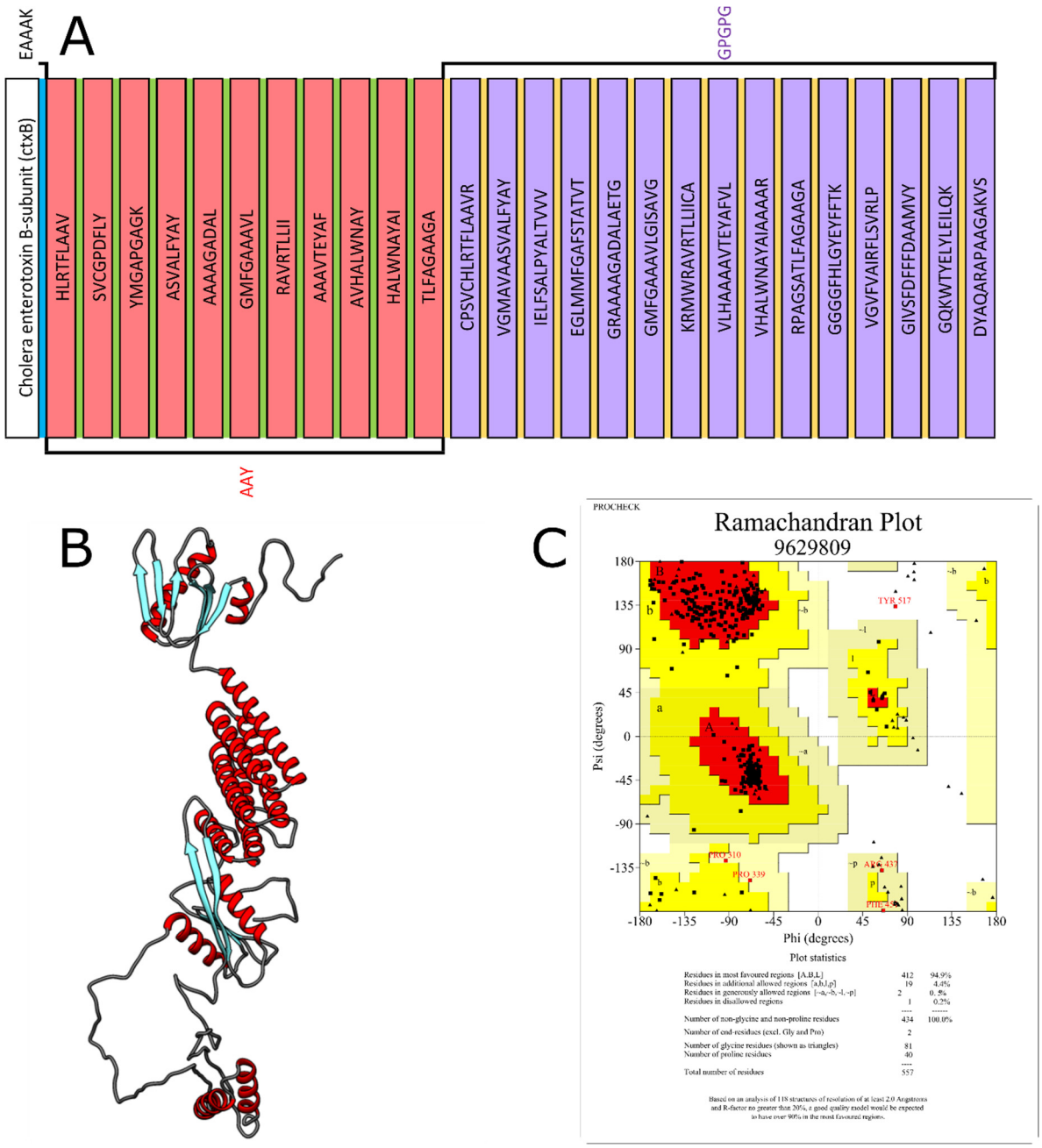


Figure 1. (**A**) Multi-epitope immunogen (Tpme-VAC/LGCM-2022) construct with highlighted peptide linkers and epitopes. Sequence length is 558 amino acid residues. (**B**) Three-dimensional structure modelling of the chimeric protein after refinement by GalaxyRefiner. (**C**) Ramachandran plot for the model after refinement, showing 94.9% residues in most favored regions and 0.2% in disallowed regions.

3.7. Physico-Chemical Properties, Toxicity, and Allergenicity of Tpme-VAC/LGCM-2022

According to the ProtParam web tool, the theoretical molecular mass of the protein is 56,596.15 (56.59 Kd), and its isoelectric point (pI) is 9.03, indicating activity in a basic environment. The Instability Index (II), which is related to the stability of the protein, is 29.92, characterizing it as stable. Its estimated half-life in in vitro mammalian reticulocytes is 30 h, >20 h in in vivo yeast, and >10 h in in vivo *E. coli*. The protein's aliphatic index,

which is associated with stability in the face of temperature changes, was 83.89, the high indexes indicating greater stability. The grand average of hydropathicity (GRAVY) is 0.398, with positive values indicating hydrophobicity. According to Protein-Sol, the predicted scaled solubility was 0.378, a score that is lower than the solubility threshold, which is 0.45, related to solubility in E. coli, scores above it having higher solubility than the average. The pI according to this predictor was 9.520. According to the AllerTOP and ToxinPred tools, the chimeric protein sequence of the chimeric protein has shown no prospect of being allergic or toxic to humans.

3.8. Tpme-VAC/LGCM-2022docks with the TLR2 Receptor

Docking results were evaluated according to binding energy, number of members in a cluster, number of hydrogen bonds, and hydrophobic interactions. The binding energy between the chimeric protein with the TLR-2 receptor was −65.97, comprising ten hydrogen bonds. Three of the Tpme-VAC/LGCM-2022 residues involved in hydrogen bonds (Gln24, Thr22, His20P) belong to the adjuvant, while the remainder of the residues (Asp147, Tyr150, Ala185, Tyr189, Arg202) belong to the selected epitopes. These bonds were formed between the vaccine construct and the extracellular portion of the TLR-2 chain. Moreover, 21 residues were involved in hydrophobic interactions, of which 8 (Met1, Ile2, Lys5, Phe6, Gly7, Val8, Phe9 and Gly21) belonged to the adjuvant, while the remainder (Ala128, Leu135, Ala139, Val143, Leu149 Ala188, Phe192, Ala195, Ala199, Ala200, Ala203, Trp240 and Pro261) belong to the selected epitopes (Figure 2). PDBePISA calculated the $\Delta^{i}G$ of the complex to be −20.6 kcal/mol. A negative $\Delta^{i}G$ value corresponds to hydrophobic interfaces, or positive protein affinity.

3.9. Tpme-VAC/LGCM-2022-TLR2 Complex Is Stable in Molecular Dynamics Simulation

The stability of the interaction between the best-docked complex was evaluated by molecular dynamics using GROMACS 5.0. In MD simulation protocol, the energy minimized structure was carried out in phases: equilibration under a constant number of particles, volume, and temperature (NVT) at 300 K, and a constant number of particles, pressure, and temperature (NPT) at 1 bar, during which the protein atoms and the solvent molecules were allowed to equilibrate around the protein molecule for 1 ns. The system was analyzed using root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF). The pressure plot indicated a fluctuation around 0.5 bar within the 1000 ps stabilization phase (Figure 3A). The temperature plot indicates the system maintained a temperature of 300 K during the same interval (Figure 3B).

The interaction was analyzed by RMSD, which reflects the complex's structural stability. The RMSD plot shows a fluctuation ranging from 0.13 nm to 1.6 nm after an 80 ns time interval. This mild fluctuation indicates the stability of the complex during the tested time interval. To reflect the fluctuation of amino acid side chains, RMSF was analyzed. The RMSF plot shows mild fluctuations of RMSF values around 0.5, indicating uninterrupted interactions between receptor and ligand, and higher peaks with RMSF values around 2.0, indicating highly flexible loop regions in the complex (Figure 3C,D).

3.10. Codon Adaptation and in silico Cloning of Tpme-VAC/LGCM-2022

Using the JCat tool to perform *E. coli* k12 codon adaptation of the protein sequence, we obtained the reverse translated sequence for the protein. The GC content of the sequence was 56.45%, within the optimum range of 30–70%. The CAI index was 1.0 and within the allowed range. Using the SnapGene tool, we created *BlpI* and *BamHI* restriction sites to insert our Tpme-VAC/LGCM-2022 sequence into the pET-28a(+) plasmid vector. The complete length of the insert was 1680 bp (Figure 4).

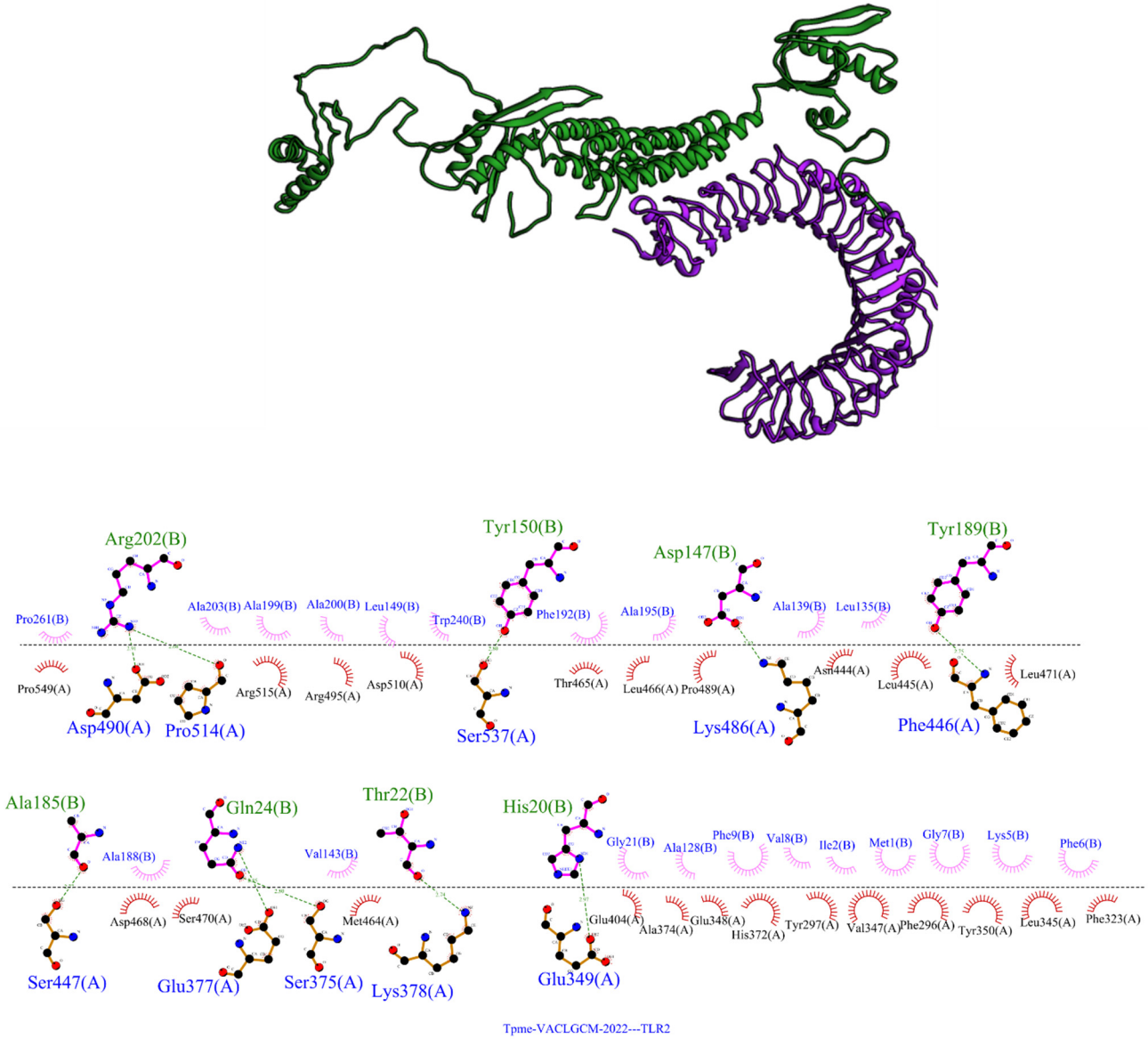


Figure 2. Molecular docking of the Tpme-VAC/LGCM-2022 with the TLR-2 receptor structure representing the interaction with the lowest energy score. 3D structure of the complex. In green, the chimeric protein, and purple is the TLR-2. 3D structure of the complex, highlighting 2D representation of the interactions between TLR 2 and the chimeric protein. The figure represents the residues of the chimeric protein (Chain-B) and TLR2 receptor (Chain-A) with hydrogen bonds (green dotted lines). The residues involved in hydrophobic interactions are shown (Red).

3.11. Tpme-VAC/LGCM-2022 Could Simulate Immune Response

The results provided by the C-ImmSim tool for immune response simulation were compatible with and indicative of the development of immunity. Regarding the B cells population, the simulation predicted an increase in memory cells throughout immunogen injection points with a strong differentiation to the production of IgG and IgM and a decrease in memory B cells. Further, a significant production of the IgM + IgG, IgG1, and IgG1 + IgG2 immunoglobulins is noted over the injections, which are relevant to complement fixation, induction of innate response, and Th1 cell activation (Figure 5A–C).

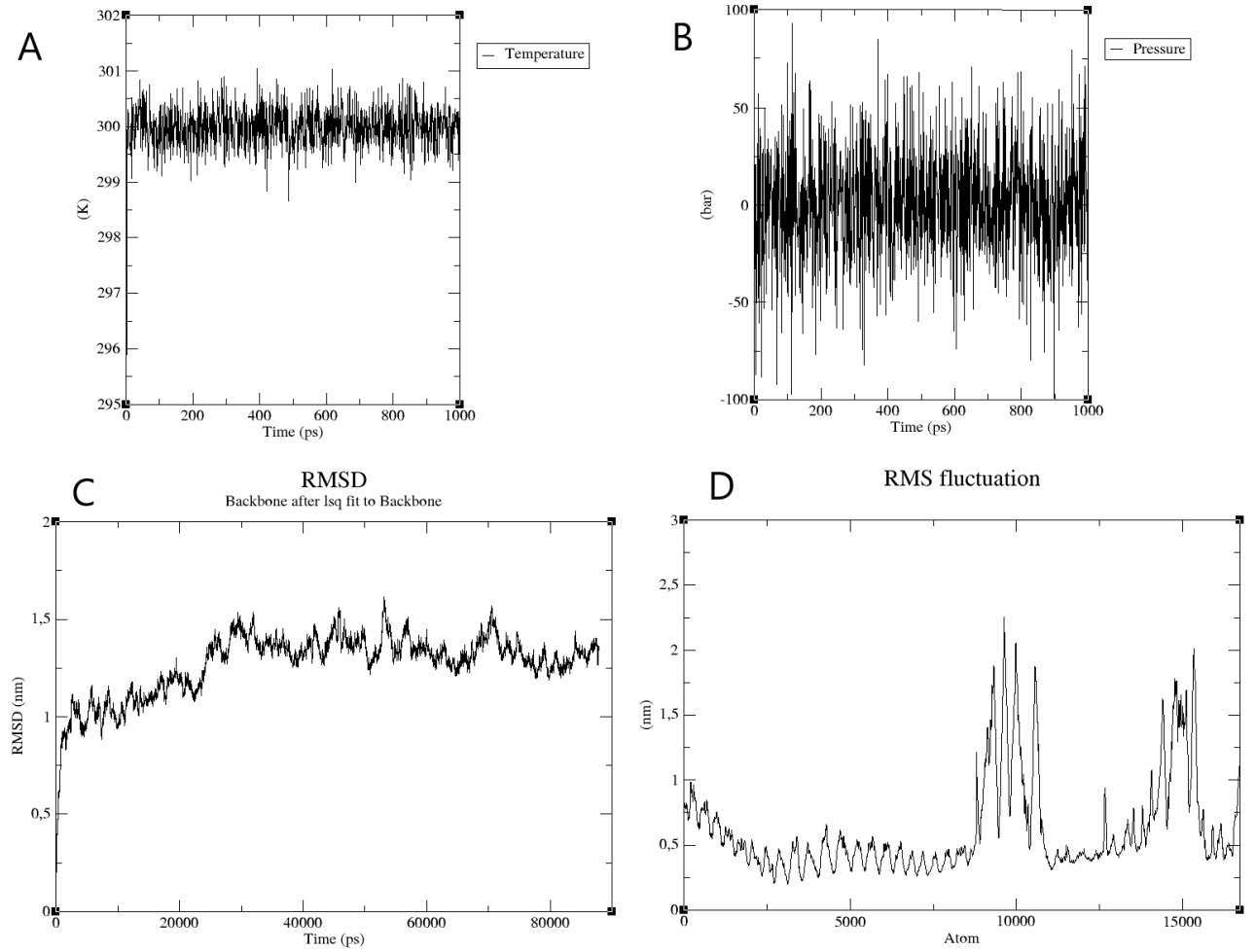


Figure 3. Molecular dynamics simulation plots for the Tpme-VAC/LGCM-2022-TLR2 complex. (**A**) The temperature plot shows that the temperature of the system reaches over 300 K, and fluctuates around 300 K throughout the equilibration phase (1000 ps). (**B**) The pressure plot displays pressure fluctuation during the equilibration phase (1000 ps), with an average pressure of 0.5 bar. (**C**) RMSD plot of the receptor-ligand complex shows no significant deviation, indicating a stable interaction. (**D**) RMSF plot shows mild fluctuations of about 0.5 nm and higher peaks with an RMSF value of 2, due to the highly flexible loop regions in the complex.

In regards to T lymphocytes, an increase in T-helper populations with a strong Th1 differentiation was observed. Th1 lymphocytes are major inducers of cytotoxic T lymphocyte proliferation. They increase their cytotoxic capacity and stimulate the production of IFN-γ. The results also indicate a growth in the active cytotoxic cell population, with a decrease in the number of resting cells (Figure 5D–G). In regards to the innate immune system, the results followed the expected response patterns for the activation and proliferation of natural killer cells and macrophages.

In general, the immune simulation showed an increase, mainly during the secondary and tertiary response in cell types and cytokines that are vital in sustaining effective immune responses with clear peaks on injection days. There is also a clear decrease in the level of active cells a few days after the third injection, which, together with the evidence of the probable induction of the IL-10 and TGF-β cytokines, indicates an attempt by the immune system to control the response and prevent exacerbation (Figure 5H).

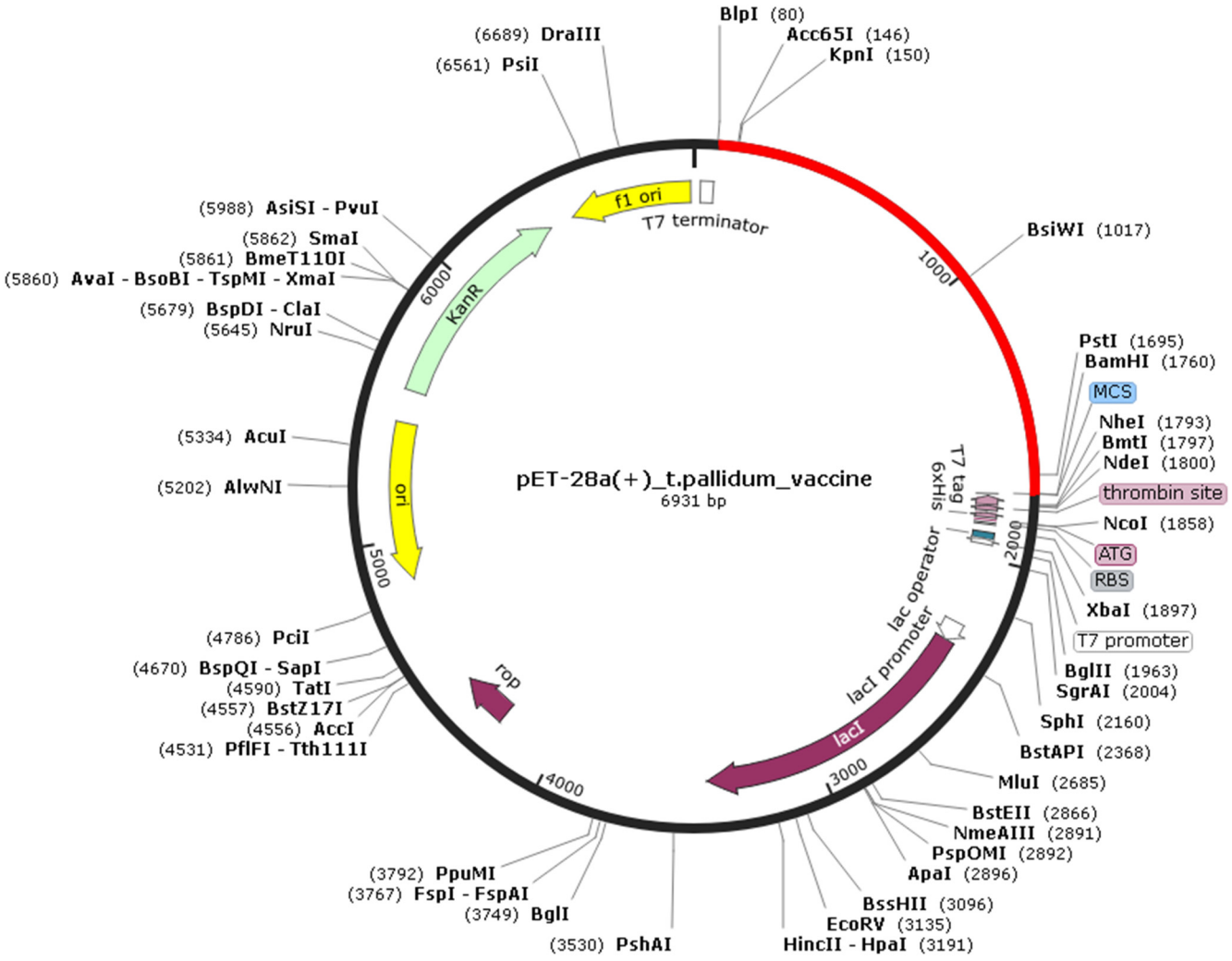


Figure 4. In silico cloning. Reverse translated sequence of the Tpme-VAC/LGCM-2022 protein insert represented in red between the *BlpI* and *BamHI* restriction sites. Vector is represented in black.

We compared the results from the simulation performed with those from our chimeric protein simulation using only the adjuvant. The comparison revealed a less intense proliferation of B cells populations on the adjuvant simulation, indicating the effect of the selected epitopes on the induction of B lymphocytes in silico. This pattern was not maintained in relation to immunoglobulin production, indicating a better activation of the humoral system by the adjuvant than the immunogen. The comparison of the simulations also showed similar patterns for the induction of cytotoxic T cells and cytokine production (Figures S5–S8). A simulation performed with only the epitopes showed similar patterns of immune activation to the full immunogen simulation (Figures S9–S12), indicating that the epitopes are responsible for the immune response elicited by the multi-epitope immunogen.

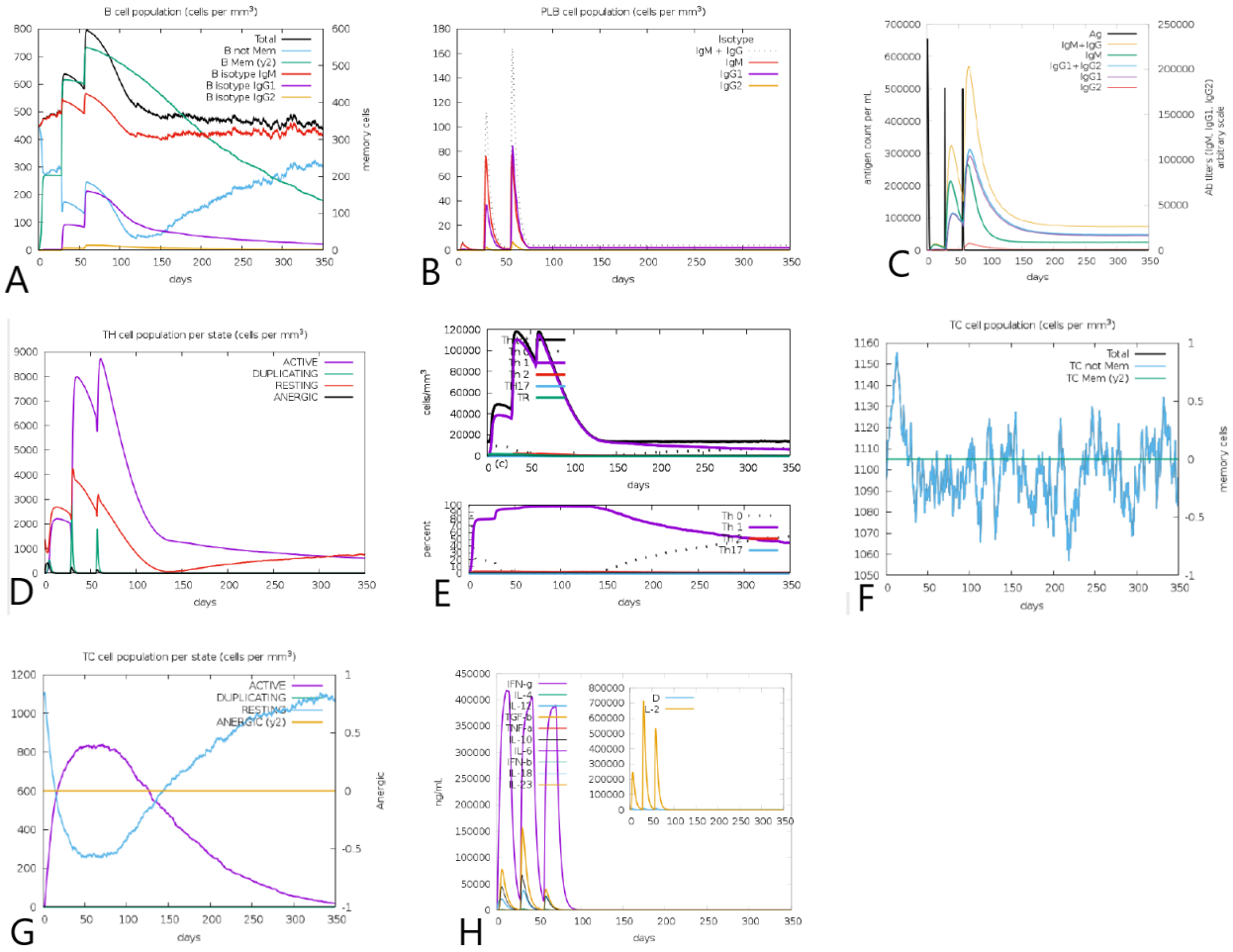


Figure 5. Immuno simulation results of the Tpme-VAC/LGCM-2022 regarding: (**A**) B cells population per mm^3 (**B**) PLB cell population per mm^3. (**C**) Immunoglobulin production. (**D**) Helper T-cell population per state. (**E**) I Helper T-cell differentiation. (**F**) Cytotoxic T-cell population. (**G**) Cytotoxic T-cell population per state. (**H**) Cytokine production.

4. Discussion

The application of current immunoinformatics approaches in the development of multi-epitope immunogens has assisted in the acceleration of the slow and costly vaccine development process, as it allows for the efficient screening of genome sets and identification of candidates [13,50]. This class of vaccine is designed to trigger both the innate and adaptive immune systems, generate protective memory, and reduce the chance of side effects, as well as spontaneous reversions that occur in attenuated vaccines [50]. Peptide-based vaccines are also cheaper and easier to produce on a large scale, can be easily stored and transported due to freeze-drying, and do not require the cultivation of the infectious bacterium [50]. The limitations associated with *T. pallidum* research have meant that there is currently no clinical vaccine for syphilis. Some putative vaccine targets have been identified, but so far have not been successful in eliciting an adequate protective response in vivo [7].

The constructed immunogen may be used as a peptide-based subunit vaccine, and our in silico predictions indicate that it may be a strong candidate for protection against *T. pallidum*. In addition, the immunogen may also be used in the research and development of new diagnostic methods for syphilis [51].

The immunoinformatics strategy adopted in this study aimed to provide a multi-epitope immunogen consisting of epitopes selected from the core genome of the pathogen, thus ensuring its coverage of various strains and the scope of the displayed antigens. These epitopes were derived from eighteen *T. pallidum* vaccine targets. The *T. pallidum*

Outer Membrane Proteome (OMPome) are generally regarded as promising candidates for vaccine targets in syphilis research [52]. Among the vaccine candidates selected for the study, four proteins (TP0897, TP0126, TP0326 and TP0733) are part of the *T. pallidum* OMPome [52].

During syphilis infection, the resolution of primary and secondary syphilitic lesions has generally been associated with cellular infiltrates of predominantly T-cells and macrophages. Clearance of primary lesions in humans is associated with a strong $CD4^+$ T-cell response, as well as the presence of macrophages and NK cells, whilst clearance of secondary legions is associated with a higher abundance of $CD8^+$ T-cells. The induction of delayed-type hypersensitivity (DTH) and Th1-mediated opsonophagocitosis, which is induced mainly by $CD4^+$ T-cells, is thought to be the main mechanism of clearance in syphilitic lesions. While this evidence points strongly to the importance of cell-mediated immune response in the response to *T. pallidum* infection, there is also evidence pointing that the humoral immune response is also essential. *T. pallidum* opsonophagocitosis has been shown to be dependent on the presence of Immune Patient Serum, highlighting the importance of the production of opsonic antibodies in the response to syphilis [10].

Aiming to design an immunogen capable of eliciting the desired response against the bacterium, epitopes were selected for their ability to induce CTL, HTL, and B-cell responses in various HLA supertypes, to stimulate both cellular and humoral immunity in a broad range of the global population. This careful selection of epitopes also allowed for the exclusion of potentially deleterious sequences for the construct, resulting in a multi-epitope immunogen that is non-allergenic, non-toxic, and non-homologous to any host proteins. One disadvantage of this approach is that, in contrast with whole pathogen vaccines, multi-epitope immunogens lack some antigenic determinants, requiring the addition of an enhancer as an adjuvant with strong antigenic properties [27]. The cholera enterotoxin B-subunit (ctxB) protein was selected for this purpose.

The chimeric protein's physico-chemical properties were determined to be within the acceptable parameters for recombinant protein production and application as a sub-unit vaccine. The sequence was optimized for expression in the *E. coli* K12 expression model and cloned into an *E. coli* expression plasmid. Through modelling, molecular docking, and molecular dynamics simulations, the multi-epitope immunogen was determined to be able to form a strong, considerably stable binding to the TLR-2. TLR-2 is one of the main Toll-like receptors involved in *T. pallidum* detection by the host, indicating its capacity to activate the receptor. To enhance our understanding of the protein's immunogenic capacity, an immune simulation was performed. The results indicated it to be a putative inducer of both the innate and adaptive immune systems, with wide differentiation of B-cell population, antibody production, and high activation of Th1 Helper T-cells. Both types of responses are vital to the clearance of active treponemes from syphilis wounds and limit the bacterium's ability to spread within the host [7]. The immunogen also induced both the differentiation of immune memory-associated cells and the production of important cytokines to balance the immune response.

5. Conclusions

In this study, a novel multi-epitope immunogen (Tpme-VAC/LGCM-2022), comprising high-ranked epitopes from eighteen *Treponema pallidum* proteins, was constructed using an immunoinformatics-based approach. Several criteria were applied to select epitopes capable of inducing strong humoral and cellular responses while assessing the protein's allergenicity and toxicity to construct a safe immunogen. The designed immunogen has suitable structural, physiochemical, and immunological properties that can successfully elicit humoral and cellular immune responses against *T. pallidum*. Finally, the in silico immune simulation enhanced our understanding of the capacity of the immunogen to elicit an immune response, highlighting that it is capable of inducing both humoral and cellular immune responses. This feature is vital in the development of a syphilis vaccine, as both the clearance of active treponemes in infectious sites and prevention of its ability

to spread within the host are key for effective syphilis prevention. Therefore, a promising immunogen candidate for *T. pallidum* was designed, notwithstanding, to ensure immunologic efficiency and memory development. Experimental validation must, however, be performed in vitro and in vivo. Further validation of Tpme-VAC/LGCM-2022 may be performed through the recombinant synthesis of the immunogen construct in a vector expression system (pET-28a(+), for expression in *E. coli*). Once synthesized and purified, the construct can then be used in in vitro and in vivo immunization assays in the rabbit model before progressing to human studies.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/vaccines10071019/s1, Table S1. List of Treponema pallidum proteins selected for analysis. Table S2. Predicted IFN-gamma inducing epitopes. 69 MHC-II epitopes capable to induce IFN-gamma (positive) were predicted using the IFNepitope server hybrid method (MERCI and SVM) default parameters. Table S3: Predicted conformational B-cell epitopes. The ElliPro server was used to predict the conformational B cell binding epitopes via the 3D structure of the Mycoplasma pneumoniae multi-epitope vaccine. Figure S1. Population coverage of the alleles used in the construction of the final version of Tpme-VAC/LGCM-2022, with a total population coverage of 99.93%. Figure S2. Representation of the secondary structure of the chimeric protein. The results showed an arrangement of alpha helices (48.0%), β-strand (15.0%), and coil formation (37.0%). Figure S3. Three-dimensional structure modelling of the chimeric protein. (A) Model predicted by RaptorX. (B) Ramachandran plot for the initial model, showing 92,9% residues in most favoured regions and 0,7% in disallowed regions. Figure S4. Representation of B-lymphocyte conformational epitopes with a score above 0.7 present in the chimeric protein sequence. (Red) 6 residue epitopes, score: 0.969. (Green) 25 residue epitopes, score: 0.898. (Blue) 3 residue epitopes, score: 0.859. (Cyan) 85 residue epitopes, score: 0.745. (Purple) 107 residue epitopes, score: 0.741. Figure S5. Immuno simulation results of only the adjuvant regarding B cell population. (A) B cell population per mm3 (B) PLB cell population per mm3. Figure S6. Immuno simulation results of only the adjuvant regarding immunoglobulin production. Figure S7. Immuno simulation results of only the adjuvant regarding T lymphocyte populations. (A) Helper T-Cell population per state. (B) Helper T-Cell differentiation. (C) Cytotoxic T-Cell population. (D) Cytotoxic T-Cell population per state. Figure S8. Immuno simulation results of the chimeric protein regarding cytokine production. Figure S9. Immuno simulation results of only the epitopes regarding B cell population. (A) B cell population per mm3 (B) PLB cell population per mm3. Figure S10. Immuno simulation results of only the epitopes regarding immunoglobulin production. Figure S11. Immuno simulation results of only the epitopes regarding T lymphocyte populations. (A) Helper T-Cell population per state. (B) Helper T-Cell differentiation. (C) Cytotoxic T-Cell population. (D) Cytotoxic T-Cell population per state. Figure S12. Immuno simulation results of the chimeric protein regarding cytokine production.

Author Contributions: L.G.R.G., T.C.V.R., S.T. and A.K.J.: conceived, designed the protocol, L.G.R.G., T.C.V.R., S.T., A.K.J., R.B.K. and R.G.S. collected and analyzed initial data, wrote the paper; S.T., S.d.C.S. and V.A.: coordinated and led the entire project; S.T., S.d.C.S. and V.A.: cross-checked all data and re-analysis; V.A., D.B., S.d.C.S., K.J.A. and H.J.B.: edited the article. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/protein/ accessed on 1 May 2021.

Acknowledgments: We acknowledge the collaboration and assistance of all team members, the laboratories associated with the Omics Science Network (RECOM), and the Brazilian funding agencies CNPq (Conselho Nacional de Desenvolvimento Científico eTecnológico), CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior, Brasil), and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais). VA, and ST acknowledges support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, Número do Processo: 402817/2019-2. A.K.J. acknowledges support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Brazil, Número do Processo:164608/2020-6. K.J.A. is supported by Taif University Researchers Supporting Program (project number: TURSP-2020/128), Taif University, Saudi Arabia.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Radolf, J.; Deka, R.; Anand, A.; Šmajs, D.; Norgard, M.; Yang, X. Treponema Pallidum, the Syphilis Spirochete: Making a Living as a Stealth Pathogen. *Nat. Rev. Microbiol.* **2016**, *14*, 744–759. [CrossRef] [PubMed]
2. Edmondson, D.G.; Hu, B.; Norris, S.J. Long-Term In Vitro Culture of the Syphilis Spirochete Treponema Pallidum Subsp. Pallidum. *mBio* **2018**, *9*, e01153-18. [CrossRef] [PubMed]
3. Jaiswal, A.K.; Tiwari, S.; Jamal, S.B.; Oliveira, L.d.C.; Alves, L.G.; Azevedo, V.; Ghosh, P.; Oliveira, C.J.F.; Soares, S.C. The Pan-Genome of Treponema Pallidum Reveals Differences in Genome Plasticity between Subspecies Related to Venereal and Non-Venereal Syphilis. *BMC Genom.* **2020**, *21*, 1–16. [CrossRef] [PubMed]
4. Spiteri, G.; Unemo, M.; Mårdh, O.; Amato-Gauci, A. The Resurgence of Syphilis in High-Income Countries in the 2000s: A Focus on Europe. *Epidemiol. Infect.* **2019**, *147*, e134. [CrossRef]
5. WHO Data on Syphilis. Available online: https://www.who.int/data/gho/data/themes/topics/topic-details/GHO/data-on-syphilis (accessed on 21 July 2021).
6. Li, Y.; Li, J.; Hu, W.; Luo, H.; Zhou, J.; Li, C.; Chen, C. Gene Subtype Analysis of Treponema Pallidum for Drug Resistance to Azithromycin. *Exp. Ther. Med.* **2018**, *16*, 1009–1013. [CrossRef]
7. Singh, T.; Otero, C.E.; Li, K.; Valencia, S.M.; Nelson, A.N.; Permar, S.R. Vaccines for Perinatal and Congenital Infections—How Close Are We? *Front. Pediatrics* **2020**, *8*, 569. [CrossRef]
8. Hook, E.W. Syphilis. *Lancet* **2017**, *389*, 1550–1557. [CrossRef]
9. Cameron, C.E.; Lukehart, S.A. Current Status of Syphilis Vaccine Development: Need, Challenges, Prospects. *Vaccine* **2014**, *32*, 1602. [CrossRef]
10. Lithgow, K.V.; Cameron, C.E. Vaccine Development for Syphilis. *Expert Rev. Vaccines* **2016**, *16*, 37–44. [CrossRef]
11. Chauhan, V.; Rungta, T.; Goyal, K.; Singh, M.P. Designing a Multi-Epitope Based Vaccine to Combat Kaposi Sarcoma Utilizing Immunoinformatics Approach. *Sci. Rep.* **2019**, *9*, 1–15. [CrossRef]
12. Palatnik-de-Sousa, C.B.; Soares, I.d.S.; Rosa, D.S. Editorial: Epitope Discovery and Synthetic Vaccine Design. *Front. Immunol.* **2018**, *9*, 826. [CrossRef] [PubMed]
13. Martinelli, D.D. In Silico Vaccine Design: A Tutorial in Immunoinformatics. *Healthc. Anal.* **2022**, *2*, 100044. [CrossRef]
14. Jaiswal, A.; Tiwari, S.; Jamal, S.; Barh, D.; Azevedo, V.; Soares, S. An In Silico Identification of Common Putative Vaccine Candidates against Treponema Pallidum: A Reverse Vaccinology and Subtractive Genomics Based Approach. *Int. J. Mol. Sci.* **2017**, *18*, 402. [CrossRef]
15. Haynes, A.M.; Godornes, C.; Ke, W.; Giacani, L. Evaluation of the Protective Ability of the Treponema Pallidum Subsp. Pallidum Tp0126 OmpW Homolog in the Rabbit Model of Syphilis. *Infect. Immun.* **2019**, *87*, e00323-19. [CrossRef] [PubMed]
16. Lithgow, K.V.; Hof, R.; Wetherell, C.; Phillips, D.; Houston, S.; Cameron, C.E. A Defined Syphilis Vaccine Candidate Inhibits Dissemination of Treponema Pallidum Subspecies Pallidum. *Nat. Commun.* **2017**, *8*, 1–10. [CrossRef] [PubMed]
17. Morgan, C.; Lukehart, S.; van Voorhis, W. Protection against Syphilis Correlates with Specificity of Antibodies to the Variable Regions of Treponema Pallidum Repeat Protein K. *Infect. Immun.* **2003**, *71*, 5605–5612. [CrossRef] [PubMed]
18. Martini, S.; Nielsen, M.; Peters, B.; Sette, A. The Immune Epitope Database and Analysis Resource Program 2003-2018: Reflections and Outlook. *Immunogenetics* **2020**, *72*, 57–76. [CrossRef]
19. Greenbaum, J.; Sidney, J.; Chung, J.; Brander, C.; Peters, B.; Sette, A. Functional Classification of Class II Human Leukocyte Antigen (HLA) Molecules Reveals Seven Different Supertypes and a Surprising Degree of Repertoire Sharing across Supertypes. *Immunogenetics* **2011**, *63*, 325–335. [CrossRef]
20. Larsen, M.; Lundegaard, C.; Lamberth, K.; Buus, S.; Lund, O.; Nielsen, M. Large-Scale Validation of Methods for Cytotoxic T-Lymphocyte Epitope Prediction. *BMC Bioinform.* **2007**, *8*, 424. [CrossRef]
21. Lin, H.H.; Zhang, G.L.; Tongchusak, S.; Reinherz, E.L.; Brusic, V. Evaluation of MHC-II Peptide Binding Prediction Servers: Applications for Vaccine Research. *BMC Bioinform.* **2008**, *9*, S22. [CrossRef]
22. Jensen, K.K.; Andreatta, M.; Marcatili, P.; Buus, S.; Greenbaum, J.A.; Yan, Z.; Sette, A.; Peters, B.; Nielsen, M. Improved Methods for Predicting Peptide Binding Affinity to MHC Class II Molecules. *Immunology* **2018**, *154*, 394–406. [CrossRef] [PubMed]

23. Galanis, K.A.; Nastou, K.C.; Papandreou, N.C.; Petichakis, G.N.; Pigis, D.G.; Iconomidou, V.A. Linear B-Cell Epitope Prediction for In Silico Vaccine Design: A Performance Review of Methods Available via Command-Line Interface. *Int. J. Mol. Sci.* **2021**, *22*, 3210. [CrossRef]
24. Zhao, W.; Sher, X. Systematically Benchmarking Peptide-MHC Binding Predictors: From Synthetic to Naturally Processed Epitopes. *PLoS Comput. Biol.* **2018**, *14*, e1006457. [CrossRef] [PubMed]
25. Calis, J.J.A.; Maybeno, M.; Greenbaum, J.A.; Weiskopf, D.; De Silva, A.D.; Sette, A.; Keşmir, C.; Peters, B. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. *PLoS Comput. Biol.* **2013**, *9*, e1003266. [CrossRef] [PubMed]
26. Bui, H.-H.; Sidney, J.; Dinh, K.; Southwood, S.; Newman, M.J.; Sette, A. Predicting Population Coverage of T-Cell Epitope-Based Diagnostics and Vaccines. *BMC Bioinform.* **2006**, *7*, 1–5. [CrossRef] [PubMed]
27. Nosrati, M.; Hajizade, A.; Nazarian, S.; Amani, J.; Namvar Vansofla, A.; Tarverdizadeh, Y. Designing a Multi-Epitope Vaccine for Cross-Protection against Shigella Spp: An Immunoinformatics and Structural Vaccinology Study. *Mol. Immunol.* **2019**, *116*, 106–116. [CrossRef]
28. Doytchinova, I.A.; Flower, D.R. VaxiJen: A Server for Prediction of Protective Antigens, Tumour Antigens and Subunit Vaccines. *BMC Bioinform.* **2007**, *8*, 1–7. [CrossRef]
29. Dhanda, S.K.; Vir, P.; Raghava, G.P. Designing of Interferon-Gamma Inducing MHC Class-II Binders. *Biol. Direct* **2013**, *8*, 1–15. [CrossRef]
30. Gupta, S.; Kapoor, P.; Chaudhary, K.; Gautam, A.; Kumar, R.; Consortium, O.S.D.D.; Raghava, G.P.S. In Silico Approach for Predicting Toxicity of Peptides and Proteins. *PLoS ONE* **2013**, *8*, e73957. [CrossRef]
31. Dimitrov, I.; Bangov, I.; Flower, D.; Doytchinova, I. AllerTOP v.2—A Server for in Silico Prediction of Allergens. *J. Mol. Model* **2014**, *20*, 1–6. [CrossRef]
32. Walker, J. *The Proteomics Protocols Handbook*; Springer Science and Business Media LLC: Totowa, NJ, USA, 2005. [CrossRef]
33. Hebditch, M.; Carballo-Amador, M.A.; Charonis, S.; Curtis, R.; Warwicker, J. Protein–Sol: A Web Tool for Predicting Protein Solubility from Sequence. *Bioinformatics* **2017**, *33*, 3098–3100. [CrossRef] [PubMed]
34. Shende, G.; Haldankar, H.; Barai, R.; Bharmal, M.; Shetty, V.; Idicula-Thomas, S. PBIT: Pipeline Builder for Identification of Drug Targets for Infectious Diseases. *Bioinformatics* **2017**, *33*, 929–931. [CrossRef] [PubMed]
35. Jones, D. Protein Secondary Structure Prediction Based on Position-Specific Scoring Matrices. *J. Mol. Biol.* **1999**, *292*, 195–202. [CrossRef]
36. Kelley, L.; Mezulis, S.; Yates, C.; Wass, M.; Sternberg, M. The Phyre2 Web Portal for Protein Modeling, Prediction and Analysis. *Nat. Protoc.* **2015**, *10*, 845–858. [CrossRef]
37. Källberg, M.; Wang, H.; Wang, S.; Peng, J.; Wang, Z.; Lu, H.; Xu, J. Template-Based Protein Structure Modeling Using the RaptorX Web Server. *Nat. Protoc.* **2012**, *7*, 1511–1522. [CrossRef] [PubMed]
38. Yang, J.; Zhang, Y. Protein Structure and Function Prediction Using I-TASSER. *Curr. Protoc. Bioinform.* **2015**, *52*, 5.8.1–5.8.15. [CrossRef]
39. Ko, J.; Park, H.; Heo, L.; Seok, C. GalaxyWEB Server for Protein Structure Prediction and Refinement. *Nucleic Acids Res.* **2012**, *40*, W294–W297. [CrossRef]
40. Laskowski, R.A.; MacArthur, M.W.; Moss, D.S.; Thornton, J.M. PROCHECK: A Program to Check the Stereochemical Quality of Protein Structures. *J. Appl. Crystallogr.* **1993**, *26*, 283–291. [CrossRef]
41. Ponomarenko, J.; Bui, H.; Li, W.; Fusseder, N.; Bourne, P.; Sette, A.; Peters, B. ElliPro: A New Structure-Based Tool for the Prediction of Antibody Epitopes. *BMC Bioinform.* **2008**, *9*, 514. [CrossRef]
42. Marra, C.M.; Sahi, S.K.; Tantalo, L.C.; Ho, E.L.; Dunaway, S.B.; Jones, T.; Hawn, T.R. Toll-like Receptor Polymorphisms Are Associated with Increased Neurosyphilis Risk. *Sex Transm. Dis.* **2014**, *41*, 440. [CrossRef]
43. Pettersen, E.; Goddard, T.; Huang, C.; Couch, G.; Greenblatt, D.; Meng, E.; Ferrin, T. UCSF Chimera–a Visualization System for Exploratory Research and Analysis. *J. Comput. Chem.* **2004**, *25*, 1605–1612. [CrossRef] [PubMed]
44. Torchala, M.; Moal, I.; Chaleil, R.; Fernandez-Recio, J.; Bates, P. SwarmDock: A Server for Flexible Protein-Protein Docking. *Bioinformatics* **2013**, *29*, 807–809. [CrossRef] [PubMed]
45. Laskowski, R.; Swindells, M. LigPlot+: Multiple Ligand-Protein Interaction Diagrams for Drug Discovery. *J. Chem. Inf. Model* **2011**, *51*, 2778–2786. [CrossRef]
46. Krissinel, E.; Henrick, K. Inference of Macromolecular Assemblies from Crystalline State. *J. Mol. Biol.* **2007**, *372*, 774–797. [CrossRef] [PubMed]
47. Abraham, M.J.; Murtola, T.; Schulz, R.; Páll, S.; Smith, J.C.; Hess, B.; Lindah, E. GROMACS: High Performance Molecular Simulations through Multi-Level Parallelism from Laptops to Supercomputers. *SoftwareX* **2015**, *1–2*, 19–25. [CrossRef]
48. Grote, A.; Hiller, K.; Scheer, M.; Münch, R.; Nörtemann, B.; Hempel, D.; Jahn, D. JCat: A Novel Tool to Adapt Codon Usage of a Target Gene to Its Potential Expression Host. *Nucleic Acids Res.* **2005**, *33*, W526–W531. [CrossRef]
49. Rapin, N.; Lund, O.; Bernaschi, M.; Castiglione, F. Computational Immunology Meets Bioinformatics: The Use of Prediction Tools for Molecular Binding in the Simulation of the Immune System. *PLoS ONE* **2010**, *5*, e9862. [CrossRef]
50. Moyle, P.; Toth, I. Modern Subunit Vaccines: Development, Components, and Research Opportunities. *Chem. Med. Chem.* **2013**, *8*, 360–376. [CrossRef]

51. Jameie, F.; Dalimi, A.; Pirestani, M.; Mohebali, M. Development of a Multi-Epitope Recombinant Protein for the Diagnosis of Human Visceral Leishmaniasis. *Iran. J. Parasitol.* **2021**, *16*, 1–10. [CrossRef]
52. Radolf, J.D.; Kumar, S. The Treponema Pallidum Outer Membrane. *Curr. Top. Microbiol. Immunol.* **2017**, *415*, 1–38. [CrossRef]

Supporting information

Supplementary Table S1. List of *Treponema pallidum* proteins selected for analysis.

	NCBI ID	GENE ID	PREDICTED LOCATION	PROTEIN NAME	PREDICTED FUNCTION
1	WP_010882178.1	TP_0733	SEC	Outer membrane beta-barrel protein	OprG/OmpW-like ion-channel involved in transport of small hydrophobic molecules
2	WP_014342713.1	FA889_00665	PSE	Hypothetical protein	Outer membrane protein/outer membrane enzyme PagP, beta-barrel domain
3	WP_010881878.1	TP_0430 / ntpK	MEM	ATP synthase subunit K	Proton transmembrane transporter activity
4	WP_010882306.1	TP_0862 / slyD	PSE	FKBP-type peptidyl-prolyl cis-trans isomerase	Peptidyl-prolyl cis-trans isomerase activity
5	WP_010881883.1	TP_0435 / nlpE	SEC	Copper resistance protein NlpE	Copper resistance associated lipoprotein
6	WP_010882040.1	TP_0594	PSE	DUF2147 domain-containing protein	Hypothetical conserved protein
7	WP_010882416.1	TP_0972 / ftr1	MEM	FTR1 family iron permease	Iron ion transmembrane transporter activity
8	WP_010881498.1	TP_0049	PSE	M23 family metallopeptidase	M23 domain-containing metallopeptidase
9	WP_010882004.1	TP_0557	SEC	DUF1007 family protein	ABC-type uncharacterized transport system, periplasmic component
10	WP_010881746.1	TP_0297	SEC	SPOR domain-containing protein	Hypothetical protein containing a peptidoglycan binding domain
11	WP_010882234.1	TP_0789	SEC	Outer membrane lipoprotein-sorting protein	Outer membrane lipoprotein-sorting
12	WP_014342788.1	TP_0326 / BamA	SEC	Outer membrane protein assembly factor BamA	Outer membrane beta-barrel assembly factor
13	WP_010881537.1	TP_0088	PSE	Hypothetical protein	Hypothetical protein
14	WP_010881771.1	TP_0323	MEM	ABC transporter permease	Ribose/Galactose transmembrane transporter activity
15	WP_010881783.1	TP_0335	MEM	CPBP family intramembrane metalloprotease	Metalloendopeptidase activity
16	AAC65118.1	TP_0126	MEM	Predicted coding region TP0126	OmpW-like ion-channel involved in transport of small hydrophobic molecules
17	WP_010882196.1	TP_0751	PSE	Vascular adhesion/metalloproteas e pallilysin	Vascular adhesin/metalloprotease
18	AAF45140.1	tprK	CYT	Tpr protein K	*T. pallidum* repeat protein

Supplementary Table S2. Predicted IFN-gamma inducing epitopes. 69 MHC-II epitopes capable to induce IFN-gamma (positive) were predicted using the IFNepitope server hybrid method (MERCI and SVM) default parameters.

Epitope sequence	Method	Result	Score
GTPQNITDLCAEYHN	MERCI	POSITIVE	1
TPQNITDLCAEYHNT	MERCI	POSITIVE	1
PQNITDLCAEYHNTQ	MERCI	POSITIVE	1
QNITDLCAEYHNTQI	MERCI	POSITIVE	1
NITDLCAEYHNTQIH	MERCI	POSITIVE	1
ITDLCAEYHNTQIHT	MERCI	POSITIVE	1
TDLCAEYHNTQIHTL	MERCI	POSITIVE	1
NDKIFSYTESLAGKR	MERCI	POSITIVE	1
DKIFSYTESLAGKRE	MERCI	POSITIVE	1
KIFSYTESLAGKREM	MERCI	POSITIVE	1
IFSYTESLAGKREMA	MERCI	POSITIVE	1
FSYTESLAGKREMAI	MERCI	POSITIVE	1
SYTESLAGKREMAII	MERCI	POSITIVE	1
YTESLAGKREMAIIT	MERCI	POSITIVE	1
EYAFAAYAVHALWNA	MERCI	POSITIVE	1
FVLGPGPGVHALWNA	MERCI	POSITIVE	1
PGPGGIVSFDFFFDA	MERCI	POSITIVE	1
FFFDAAMVYGPGPGG	MERCI	POSITIVE	1
AAMVYGPGPGGQKWT	MERCI	POSITIVE	1
AMVYGPGPGGQKWTY	MERCI	POSITIVE	1
MVYGPGPGGQKWTYE	MERCI	POSITIVE	1
VYGPGPGGQKWTYEL	MERCI	POSITIVE	1
YGPGPGGQKWTYELY	MERCI	POSITIVE	1
GPGPGGQKWTYELYL	MERCI	POSITIVE	1
PGPGGQKWTYELYLE	MERCI	POSITIVE	1
PGGQKWTYELYLEIL	MERCI	POSITIVE	1
GGQKWTYELYLEILQ	MERCI	POSITIVE	1

GQKWTYELYLEILQK	MERCI	POSITIVE	1
QKWTYELYLEILQKG	MERCI	POSITIVE	1
TLFAGAAGAGPGPGG	SVM	POSITIVE	1,0027957
AAVLAAYRAVRTLLI	SVM	POSITIVE	1,021854
SVALFYAYAAYAAAA	SVM	POSITIVE	1,0643286
AAYAAAVTEYAFAAY	SVM	POSITIVE	1,084482
HAIAAISMANEAAAK	SVM	POSITIVE	1,0846252
LFAGAAGAGPGPGGG	SVM	POSITIVE	1,1717537
IAAYAAAVTEYAFAA	SVM	POSITIVE	1,2066697
AYAAAVTEYAFAAYA	SVM	POSITIVE	1,2083039
GAGPGPGGGGGFHLG	SVM	POSITIVE	1,2582538
LIIAAYAAAVTEYAF	SVM	POSITIVE	1,2673741
AAGAGPGPGGGGGFH	SVM	POSITIVE	1,3224505
YAAAVTEYAFAAYAV	SVM	POSITIVE	1,3276033
AGAGPGPGGGGGFHL	SVM	POSITIVE	1,4281056
GAAGAGPGPGGGGGF	SVM	POSITIVE	1,5152442
FAGAAGAGPGPGGGG	SVM	POSITIVE	1,5171639
AGAAGAG-PGPGGGGG	SVM	POSITIVE	1,6719466
YAFAAYAVHALWNAY	MERCI	POSITIVE	2
AFAAYAVHALWNAYA	MERCI	POSITIVE	2
FAAYAVHALWNAYAA	MERCI	POSITIVE	2
AAYAVHALWNAYAAY	MERCI	POSITIVE	2
AYAVHALWNAYAAYH	MERCI	POSITIVE	2
YAVHALWNAYAAYHA	MERCI	POSITIVE	2
AVHALWNAYAAYHAL	MERCI	POSITIVE	2
VHALWNA-YAAYHALW	MERCI	POSITIVE	2
VLGPGPGVHALWNAY	MERCI	POSITIVE	2
LGPGPGVHALWNAYA	MERCI	POSITIVE	2
GPGPGVHALWNAYAI	MERCI	POSITIVE	2
PGPGVHALWNAYAIA	MERCI	POSITIVE	2
GPGVHALWNAYAIAA	MERCI	POSITIVE	2

PGVHALWNAYAIAAA	MERCI	POSITIVE	2
GPGGIVSFDFFFDAA	MERCI	POSITIVE	2
PGGIVSFDFFFDAAM	MERCI	POSITIVE	2
GGIVSFDFFFDAAMV	MERCI	POSITIVE	2
GIVSFDFFFDAAMVY	MERCI	POSITIVE	2
IVSFDFFFDAAMVYG	MERCI	POSITIVE	2
VSFDFFFDAAMVYGP	MERCI	POSITIVE	2
SFDFFFDAAMVYGPG	MERCI	POSITIVE	2
FDFFFDAAMVYGPGP	MERCI	POSITIVE	2
DFFFDAAMVYGPGPG	MERCI	POSITIVE	2
GPGGQKWTYELYLEI	MERCI	POSITIVE	2

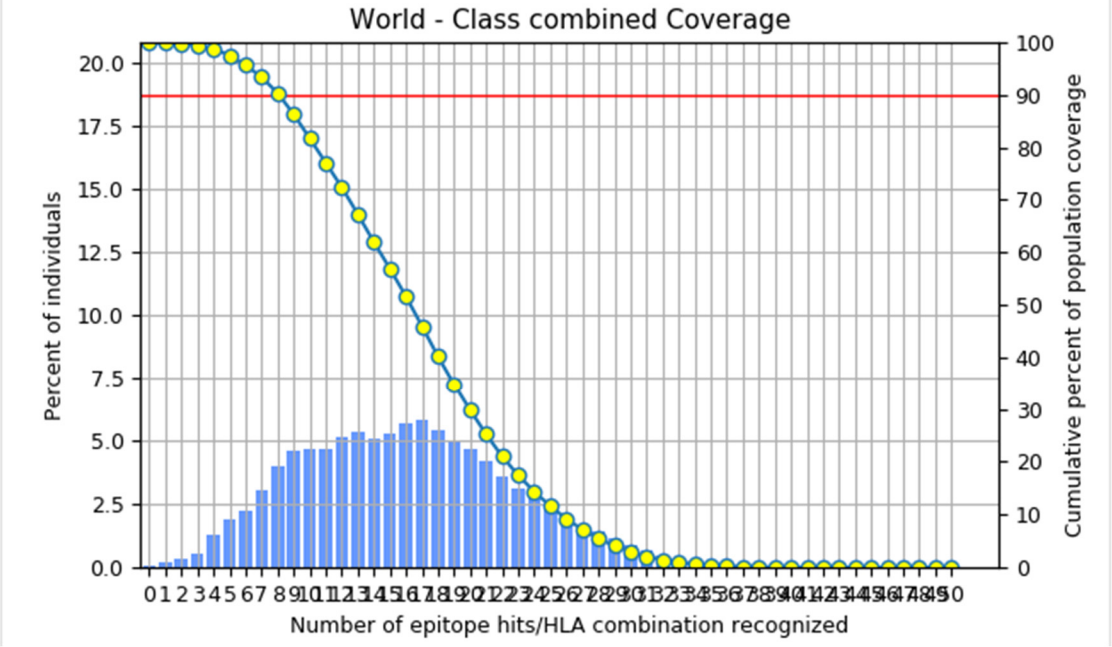
Supplementary Table S3: Predicted conformational B-cell epitopes. The ElliPro server was used to predict the conformational B cell binding epitopes via the 3D structure of the *Mycoplasma pneumoniae* multi-epitope vaccine.

Conformational epitope	Residues and position	Number of residues	Score
1	M1, I2, K3, L4, K5, F6	6	0.969
2	P359, G360, P361, G362, G363, M364, F365, G366, A367, A368, A369, V370, L371, G372, I373, S374, A375, V376, G377, G378, P379, G380, P381, G382, K383	25	0.898
3	G7, V8, F9	3	0.859
4	R387, A388, V389, R390, T391, L392, L393, I394, I395, C396, A397, G398, P399, G400, P401, G402, V403, L404, H405, A406, A407, A408, A409, V410, T411, E412, Y413, A414, F415, V416, L417, G418, P419, G420, P421, G422, V423, H424, A425, L426, W427, N428, A429, Y430, A431, I432, A433, A434, A435, A436, R437, G438, P439, G440, P441, G442, R443, P444, A445, G446, S447, A448, T449, L450, F451, A452, G453, A454, A455, G456, A457, G458, P459, G460, P461, G462, G463, G464, G465, G466, F467, H468, L469, G470, Y471	85	0.745
5	F10, T11, V12, L13, L14, S15, S16, A17, Y18, A19, H20, G21, T22, P23, Q24, N25, D28, L29, C30, A31, E32, Y33, H34, N35, T36, Q37, I38, H39, L41, D43, K44, I45, F46, S47, Y48, T49, E50, S51, L52, A53, G54, K55, R56, E57, M58, A59, I60, I61, T62, F63, K64, N65, G66, A67, T68, F69, Q70, V71, E72, V73, P74,	107	0.741

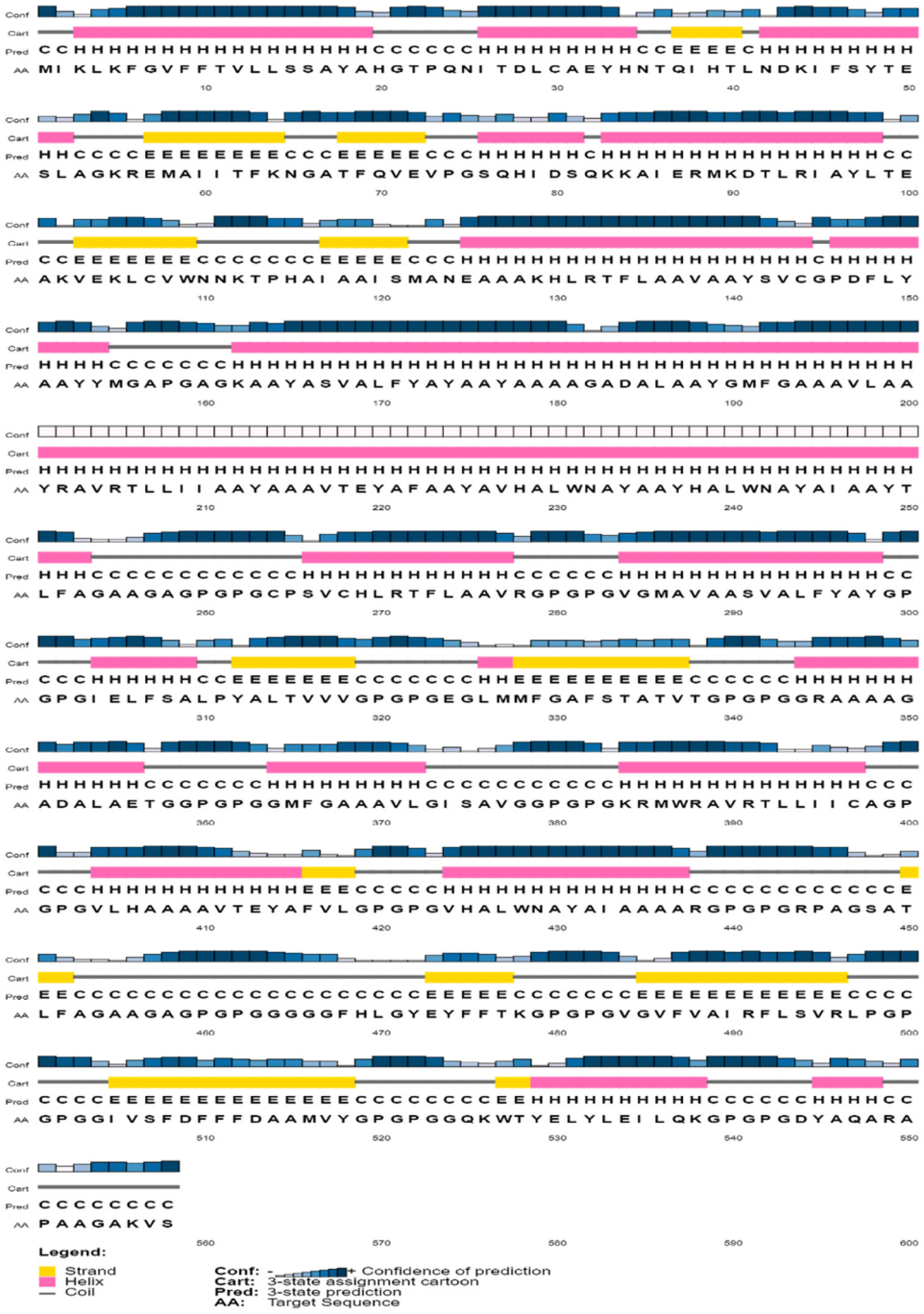
G75, S76, Q77, H78, I79, D80, S81, Q82, K83, K84, A85, I86, E87, R88, M89, K90, D91, T92, L93, R94, I95, A96, Y97, L98, T99, E100, A101, V103, L106, C107, V108, W109, N110, N111, K112, T113, P114, H115, A116, I117, A118, A119, I120, S121, M122, A123

Supplementary Figure S1. Population coverage of the alleles used in the construction of the final version of Tpme-VAC/LGCM-2022, with a total population coverage of 99.93%

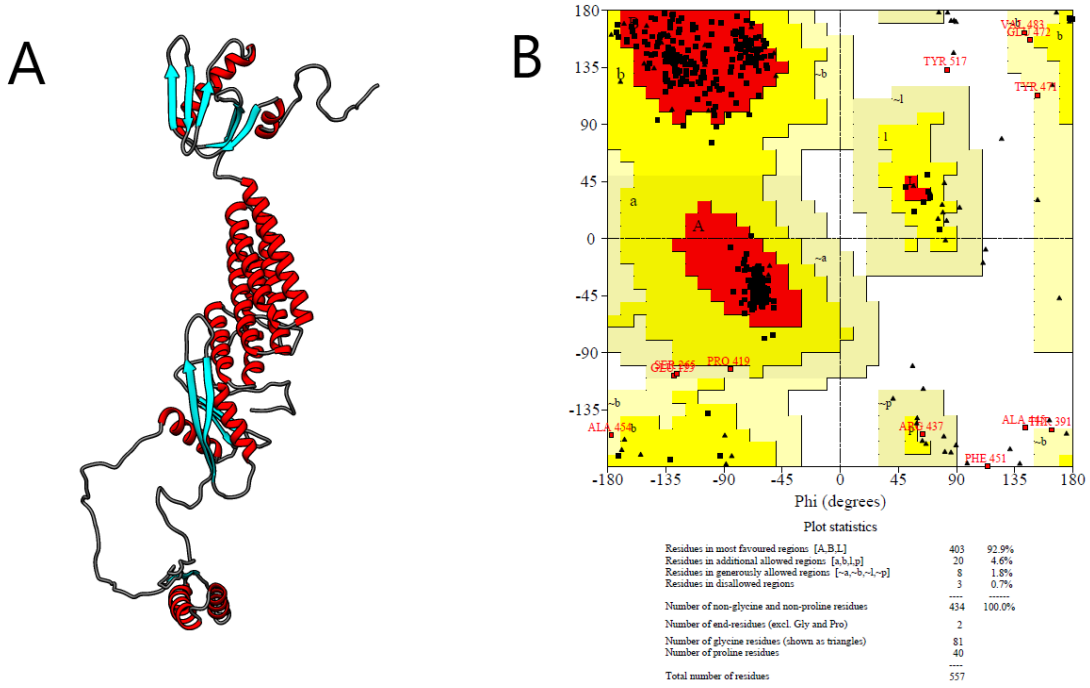
MHC class	Coverage	Average hit	PC90
combined	99.93%	16.05	8.08



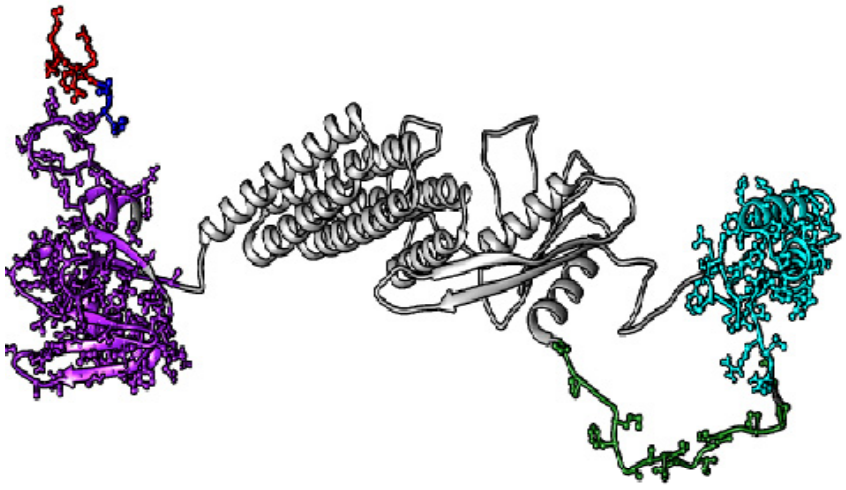
Supplementary Figure S2. Representation of the secondary structure of the chimeric protein. The results showed an arrangement of alpha helices (48.0%), β-strand (15.0%), and coil formation (37.0%).



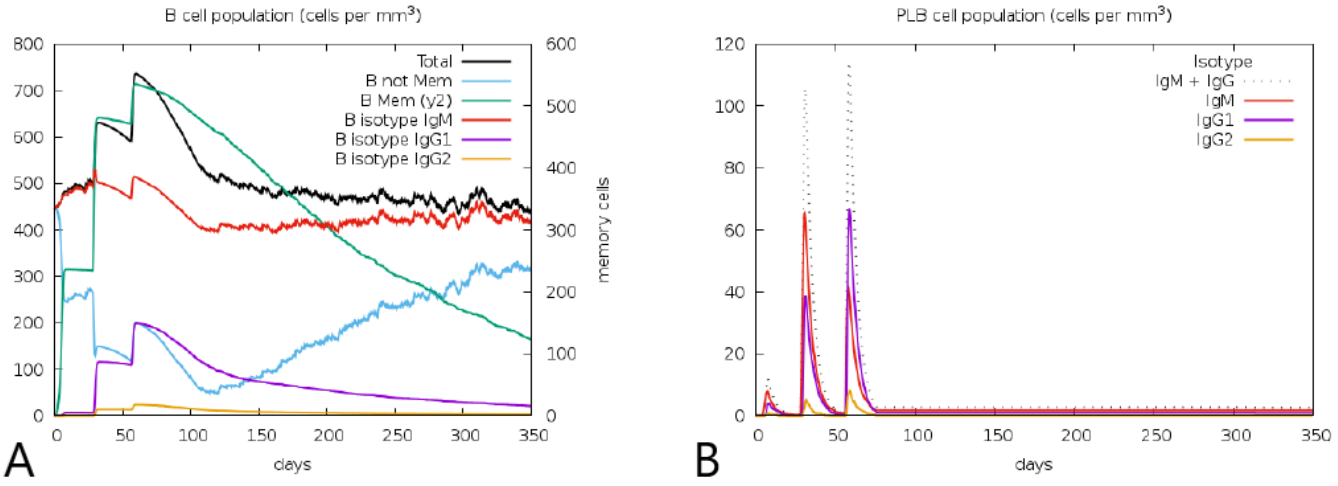
Supplementary Figure S3. Three-dimensional structure modelling of the chimeric protein. (A) Model predicted by RaptorX. (B) Ramachandran plot for the initial model, showing 92,9% residues in most favoured regions and 0,7% in disallowed regions.



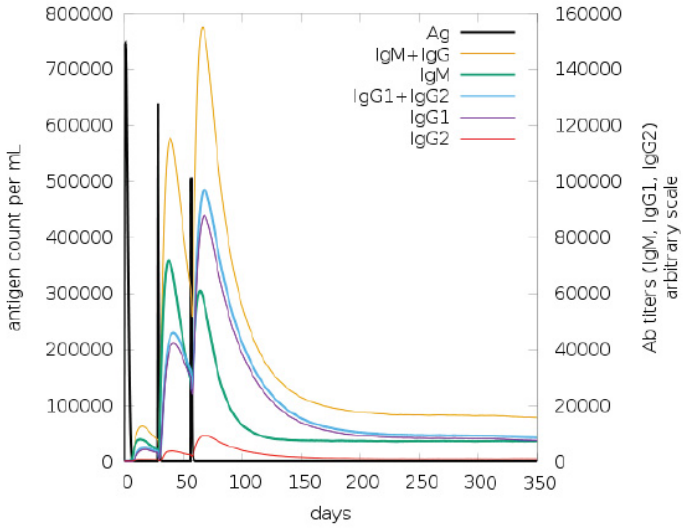
Supplementary Figure S4. Representation of B-lymphocyte conformational epitopes with a score above 0.7 present in the chimeric protein sequence. (Red) 6 residue epitopes, score: 0.969. (Green) 25 residue epitopes, score: 0.898. (Blue) 3 residue epitopes, score: 0.859. (Cyan) 85 residue epitopes, score: 0.745. (Purple) 107 residue epitopes, score: 0.741.



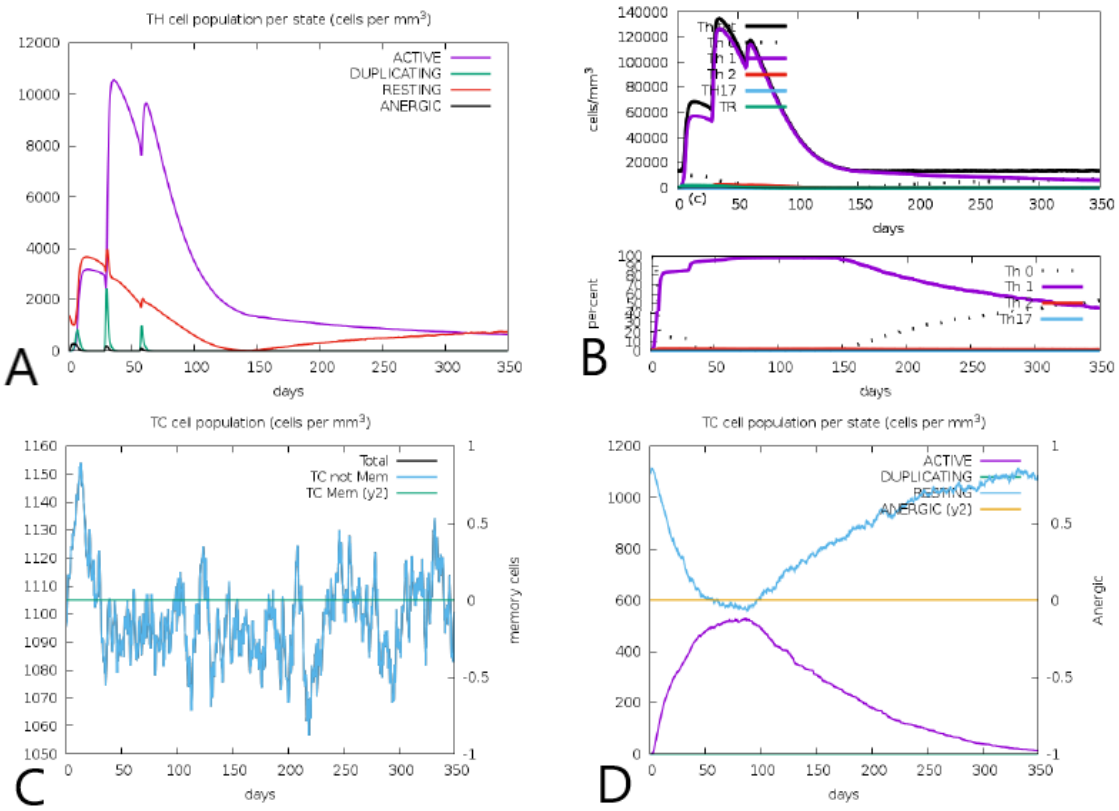
Supplementary Figure S5. Immuno simulation results of only the adjuvant regarding B cell population. (A) B cell population per mm^3 (B) PLB cell population per mm^3.



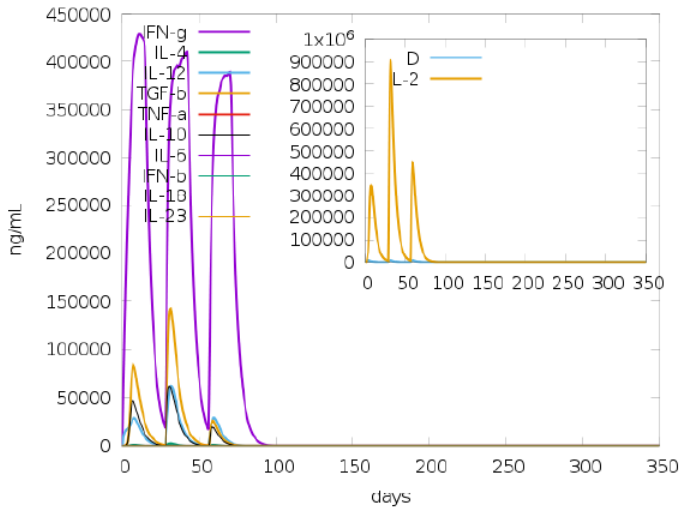
Supplementary Figure S6. Immuno simulation results of only the adjuvant regarding immunoglobulin production.



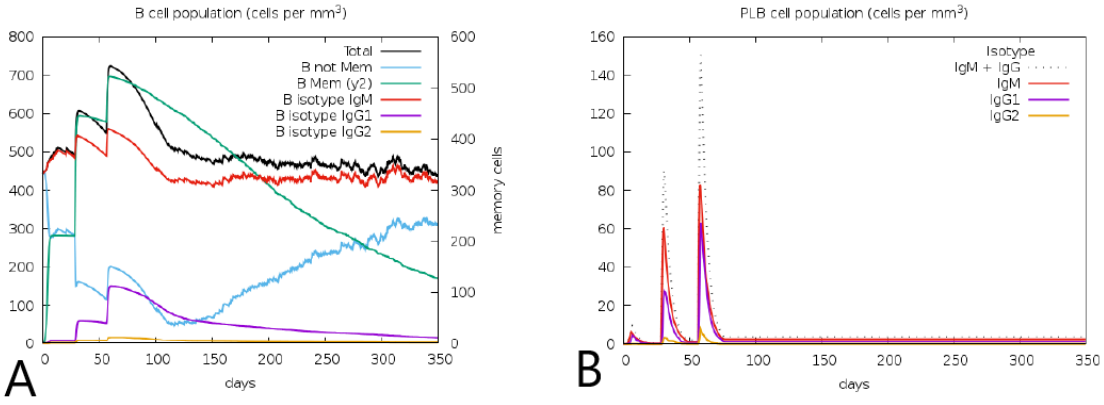
Supplementary Figure S7. Immuno simulation results of only the adjuvant regarding T lymphocyte populations. (A) Helper T-Cell population per state. (B) Helper T-Cell differentiation. (C) Cytotoxic T-Cell population. (D) Cytotoxic T-Cell population per state.



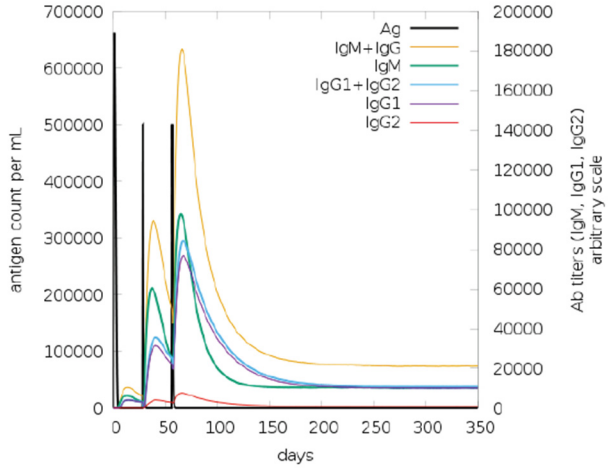
Supplementary Figure S8. Immuno simulation results of the chimeric protein regarding cytokine production.



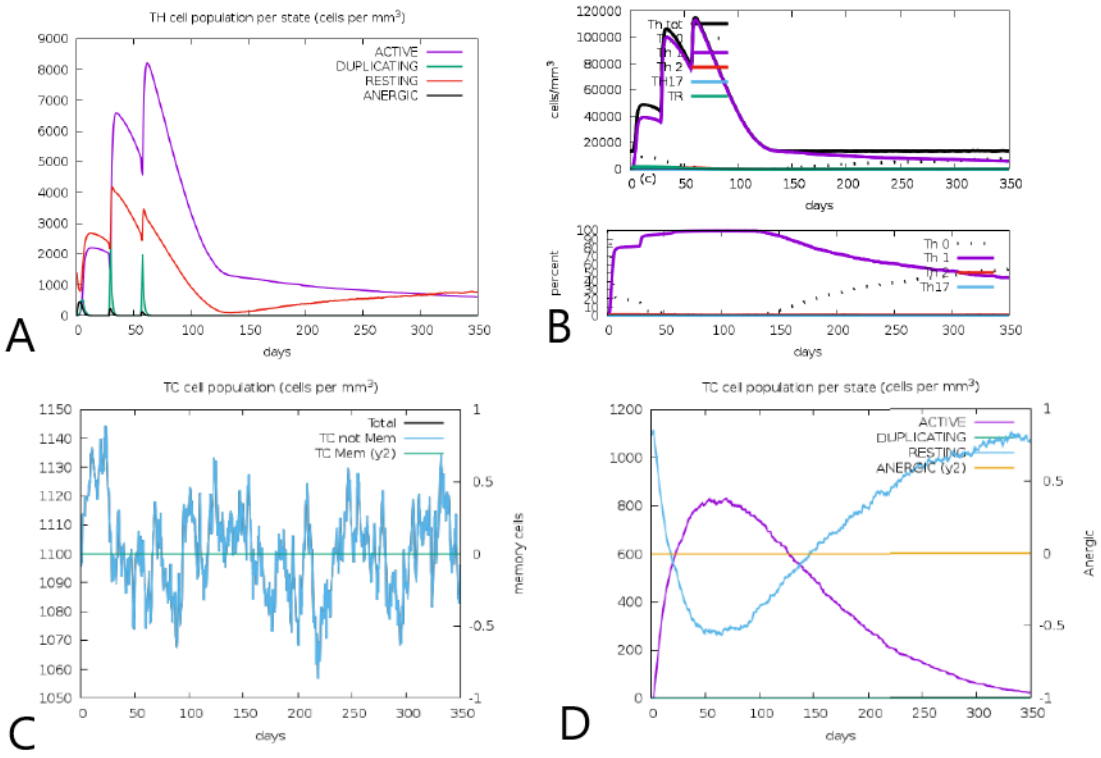
Supplementary Figure S9. Immuno simulation results of only the epitopes regarding B cell population. (A) B cell population per mm^3 (B) PLB cell population per mm^3.



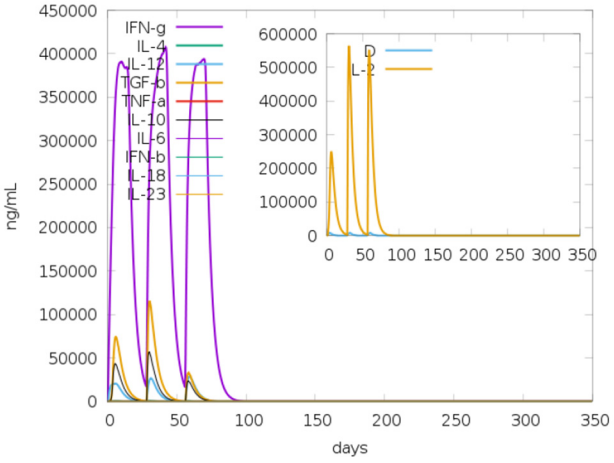
Supplementary Figure S10. Immuno simulation results of only the epitopes regarding immunoglobulin production.



Supplementary Figure S11. Immuno simulation results of only the epitopes regarding T lymphocyte populations. (A) Helper T-Cell population per state. (B) Helper T-Cell differentiation. (C) Cytotoxic T-Cell population. (D) Cytotoxic T-Cell population per state.



Supplementary Figure S12. Immuno simulation results of the chimeric protein regarding cytokine production.



4.2. Discussion

This research article is presented as chapter 2 of this dissertation, and the main research work performed in it. In this article, we have designed a promising multi-epitope vaccine candidate for syphilis, a disease for which there is currently no vaccine either available or in development. As established in chapter 1, prior to this work there were no published works applying the multi-epitope vaccine methodology to this disease, although there have been works focused on finding vaccine candidates for it using Reverse Vaccinology.

5. GENERAL CONCLUSION

Applying computational biology methods may prove vital in developing new tools to combat the spread of Sexually Transmitted Infections. In this work, we sought to evaluate the current state of the art in using these methodologies in research regarding the most common non-viral STIs. In doing so, we have identified that research is concentrated on the more common bacterial STIs of Chlamydia and Gonorrhea, while diseases like Trichomoniasis and Syphilis, which are also responsible for a large number of cases, are more neglected.

More research is required to develop a novel vaccine against Treponema pallidum, for which there are no available vaccines either in development or on the market; this highlights the importance of the research performed at the core of this work, in which we have designed a promising vaccine candidate for syphilis in the form of a multi-epitope vaccine. This vaccine was designed based on immunogenic epitopes derived from promising *T. pallidum* vaccine candidates and evaluated based on their capacity to elicit humoral and cellular immune responses. The construct has adequate structural and physico-chemical properties for its synthesis and application as a vaccine and has been predicted to elicit a desirable immune response.

In conclusion, in this work, we have established a need for research applying immunoinformatics to develop a syphilis vaccine, as there is currently no vaccine available for the disease and very few efforts in applying computational techniques to the field. In response to the need for the development of one such immunogen, we have developed a multi-epitope chimeric peptide based on immunogenic fragments of *T. pallidum* vaccine candidates. The vaccine candidates used in this work were derived from publically available data on the *T. pallidum* genome, and were established both by traditional and reverse vaccinology. This peptide is a promising vaccine candidate for syphilis but will require further in vivo and in vitro testing to validate our findings.

6. PERSPECTIVES

The multi-epitope vaccine peptide was designed and evaluated purely through the application of *in silico* models in this work. While a promising candidate for a syphilis vaccine, further steps in developing a functioning vaccine based on this study will require the recombinant synthesis of this peptide for *in vitro* and *in vivo* tests. The synthetic protein will then be evaluated regarding its physico-chemical properties, its structure may be resolved, and its immunological properties may be confirmed in a murine model.

In addition, our literature review has shown that, while there has been working applied in developing vaccines for other sexually transmitted infections, such as *N. gonorrhoeae* and *C. trachomatis*, no common vaccine candidates for all STIs have been found. Applying Reverse Vaccinology to search for common vaccine candidates for multiple bacterial STIs and developing a multi-epitope vaccine based on those proteins may be a step in the direction of a multivalent STI vaccine.

7. REFERENCES

[1] Sexually transmitted infections (STIs), World Health Organization. (2022). https://www.who.int/news-room/fact-sheets/detail/sexually-transmitted-infections-(stis) (accessed September 12, 2022).

[2] Global progress report on HIV, viral sexually transmitted infections, 2021, World Health Organization. (2021). https://www.who.int/publications/i/item/9789240027077 (accessed September 12, 2022).

[3] F. Fatima, S. Kumar, A. Das, Vaccines against sexually transmitted infections: an update, Clin Exp Dermatol. 47 (2022) 1454–1463. https://doi.org/10.1111/CED.15223.

[4] G. Spiteri, M. Unemo, O. Mårdh, A.J. Amato-Gauci, The resurgence of syphilis in high-income countries in the 2000s: a focus on Europe, Epidemiol Infect. 147 (2019) e143. https://doi.org/10.1017/S0950268819000281.

[5] Global health sector strategy on Sexually Transmitted Infections, 2016-2021, World Health Organization. (n.d.). https://www.who.int/publications/i/item/WHO-RHR-16.09 (accessed September 12, 2022).

[6] J.D. Radolf, R.K. Deka, A. Anand, D. Šmajs, M. v. Norgard, X.F. Yang, Treponema pallidum, the syphilis spirochete: making a living as a stealth pathogen, Nature Reviews Microbiology 2016 14:12. 14 (2016) 744–759. https://doi.org/10.1038/nrmicro.2016.141.

[7] Y. Li, J. Li, W. Hu, H. Luo, J. Zhou, C. Li, C. Chen, Gene subtype analysis of treponema pallidum for drug resistance to azithromycin, Exp Ther Med. 16 (2018) 1009–1013. https://doi.org/10.3892/ETM.2018.6241/HTML.

[8] E.W. Hook, Syphilis, The Lancet. 389 (2017) 1550–1557. https://doi.org/10.1016/S0140-6736(16)32411-4.

[9] T. Singh, C.E. Otero, K. Li, S.M. Valencia, A.N. Nelson, S.R. Permar, Vaccines for Perinatal and Congenital Infections—How Close Are We?, Front Pediatr. 8 (2020) 569. https://doi.org/10.3389/FPED.2020.00569/XML/NLM.

[10] C.E. Cameron, S.A. Lukehart, Current status of syphilis vaccine development: Need, challenges, prospects, Vaccine. 32 (2014) 1602–1609. https://doi.org/10.1016/J.VACCINE.2013.09.053.

[11] J.D. Radolf, S. Kumar, The treponema pallidum outer membrane, Curr Top Microbiol Immunol. 415 (2018) 1–38. https://doi.org/10.1007/82_2017_44/COVER.

[12] A.K. Jaiswal, S. Tiwari, S.B. Jamal, D. Barh, V. Azevedo, S.C. Soares, An In Silico Identification of Common Putative Vaccine Candidates against Treponema pallidum: A Reverse Vaccinology and Subtractive Genomics Based Approach, International Journal of Molecular Sciences 2017, Vol. 18, Page 402. 18 (2017) 402. https://doi.org/10.3390/IJMS18020402.

[13] D.D. Martinelli, In silico vaccine design: A tutorial in immunoinformatics, Healthcare Analytics. 2 (2022) 100044. https://doi.org/10.1016/J.HEALTH.2022.100044.

[14] V. Chauhan, T. Rungta, K. Goyal, M.P. Singh, Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach, Scientific Reports 2019 9:1. 9 (2019) 1–15. https://doi.org/10.1038/s41598-019-39299-8.

[15] P. Moyle, I. Toth, Modern subunit vaccines: development, components, and research opportunities, ChemMedChem. 8 (2013) 360–376. https://doi.org/10.1002/CMDC.201200487.

8. ANNEX

This section includes four additional published authors as co-author and certificates for participation in events, both as attendees and organizers.

venereology

MDPI

Review

Neuroinformatics Insights towards Multiple Neurosyphilis Complications

Arun Kumar Jaiswal [1,*], Syed Babar Jamal [2], Lucas Gabriel Rodrigues Gomes [1], Rodrigo Profeta [3], Helioswilton Sales-Campos [4], Carlo Jose Freire Oliveira [5], Flávia Figueira Aburjaile [6], Sandeep Tiwari [1], Debmalya Barh [7], Marcos Vinicius da Silva [5], Siomar de Castro Soares [5,*] and Vasco Azevedo [1,*]

1 Post-Graduate Program in Bioinformatics, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; lucasgabriel388@gmail.com (L.G.R.G.); sandip_sbtbi@yahoo.com (S.T.)
2 Department of Biological Sciences, National University of Medical Sciences, Rawalpindi, Rawalpindi 46000, Pakistan; babar.jamal@numspak.edu.pk
3 Department of Genetics, Ecology, and Evolution (GEE), Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; profeta.biotec@gmail.com
4 Institute of Tropical Pathology and Public Health, Federal University of Goias (UFG), Goiânia 74690-900, Brazil; tonsales@ufg.br
5 Laboratory of Immunology and Bioinformatics, Department of Microbiology, Immunology, and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro, Uberaba 38025-180, Brazil; carlo.oliveira@uftm.edu.br (C.J.F.O.); marcosuftm@gmail.com (M.V.d.S.)
6 Department of Preventative Veterinary Medicine, School of Veterinary Medicine, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; faburjaile@gmail.com
7 Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Purba Medinipur 721172, India; dr.barh@gmail.com
* Correspondence: arunjaiswal1411@gmail.com (A.K.J.); siomars@gmail.com (S.d.C.S.); vascoariston@gmail.com (V.A.)

Abstract: *Treponema pallidum* subspecies pallidum causes syphilis, a sexually transmitted disease that infects more than 2.1 million pregnant women every year. Due to its maximum death rates and augmented risk of human immunodeficiency virus (HIV) infection, the disease is still a matter of debate in many low- and high-income countries. The infection has three stages that lead to several complications if left untreated and can lead to many tertiary complications in the brain, eyes, ears, heart, and pregnancy. Neurosyphilis is also known as the clinical result of infection of the central nervous system by *Treponema pallidum* subspecies pallidum. It can evolve at any time and from any stage of syphilis exposure. This review briefly explains the severe and multiple neurosyphilitic complications and recently identified cases related to neurosyphilis. We also explained computational neuroscience, neuroinformatics, and in silico models and techniques based on artificial intelligence and other computational and mathematical methods. These techniques have already been applied to several neurological and psychological brain complications and can be applied to neurosyphilis to better understand the persistence of the disease related to the brain that causes neurosyphilis.

Keywords: *Treponema pallidum*; blood–brain barrier; neurosyphilitic meningitis; neurocognitive; cognitive deficits; computational neuroscience

check for updates

Citation: Jaiswal, A.K.; Jamal, S.B.; Gabriel Rodrigues Gomes, L.; Profeta, R.; Sales-Campos, H.; Oliveira, C.J.F.; Figueira Aburjaile, F.; Tiwari, S.; Barh, D.; Silva, M.V.d.; et al. Neuroinformatics Insights towards Multiple Neurosyphilis Complications. *Venereology* **2022**, *1*, 135–160. https://doi.org/10.3390/venereology1010010

Academic Editor: Pasquale Ferrante

Received: 31 March 2022
Accepted: 7 May 2022
Published: 6 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

Syphilis is a sexually-transmitted disease caused by *Treponema pallidum* (*Tp*) subspecies pallidum infection that infects more than 2.1 million pregnant women every year. Due to its maximum death rates of neonates, augmented risk of human immunodeficiency virus (HIV) infection, and continued morbidity particularly in low-income countries [1,2] as well as in high-income countries [3,4], such as Japan, where the rate of cases is increasing at an alarming level in heterosexual men and women, syphilis is a disease of worldwide concern [5]. Principally, the infection is transmitted through sexual contact, exceptionally with blood transfusion and blood products, and transmits vertically from mother to

Inflammation
https://doi.org/10.1007/s10753-022-01734-w

ORIGINAL ARTICLE

SARS-CoV-2 Variants Show a Gradual Declining Pathogenicity and Pro-Inflammatory Cytokine Stimulation, an Increasing Antigenic and Anti-Inflammatory Cytokine Induction, and Rising Structural Protein Instability: A Minimal Number Genome-Based Approach

Debmalya Barh[1,2,14] , Sandeep Tiwari[2] , Lucas Gabriel Rodrigues Gomes[2] , Cecília Horta Ramalho Pinto[3] , Bruno Silva Andrade[4] , Shaban Ahmad[5] , Alaa A. A. Aljabali[6] , Khalid J. Alzahrani[7] , Hamsa Jameel Banjer[7] , Sk. Sarif Hassan[8] , Elrashdy M. Redwan[9] , Khalid Raza[5] , Aristóteles Góes-Neto[2] , Robinson Sabino-Silva[10] , Kenneth Lundstrom[11] , Vladimir N. Uversky[12] , Vasco Azevedo[2] and Murtaza M. Tambuwala[13]

Received 19 July 2022; accepted 23 August 2022

Abstract— Hyper-transmissibility with decreased disease severity is a typical characteristic of the SARS-CoV-2 Omicron variant. To understand this phenomenon, we used various bioinformatics approaches to analyze randomly selected genome sequences (one each) of the Gamma, Delta, and Omicron variants submitted to NCBI from December 15 to 31, 2021. We report that the pathogenicity of SARS-CoV-2 variants decreases in the order of Wuhan > Gamma > Delta > Omicron; however, the antigenic property follows the order of Omicron > Gamma > Wuhan > Delta. The Omicron spike RBD shows lower pathogenicity but higher antigenicity than other variants. The reported decreased disease severity by the Omicron variant may be due to its decreased pro-inflammatory and IL-6 stimulation and increased IFN-γ and IL-4 induction efficacy. The mutations in the N protein are probably associated with this decreased IL-6 induction and human DDX21-mediated increased IL-4 production for Omicron. Due to the mutations, the stability of S, M, N, and E proteins

Highlights
- The pathogenicity of SARS-CoV-2 variants decreases in the order: Wuhan > Gamma > Delta > Omicron
- Omicron spike RBD has lower pathogenicity but higher antigenicity than other variants
- Omicron shows low severity due to its low pro-inflammatory and IL-6 stimulation and increased IFN-γ and IL-4 induction efficacy
- Stronger spike RBD-*h*ACE2 binding of Omicron is associated with increased transmissibility
- The low stability of Omicron spike protein is associated with low systemic infection and severe disease

Extended author information available on the last page of the article

Published online: 10 October 2022

viruses

MDPI

Review

Associations and Disease–Disease Interactions of COVID-19 with Congenital and Genetic Disorders: A Comprehensive Review

Altijana Hromić-Jahjefendić [1,*,†], Debmalya Barh [2,3,*,†], Cecília Horta Ramalho Pinto [4,‡], Lucas Gabriel Rodrigues Gomes [3,‡], Jéssica Lígia Picanço Machado [5,‡], Oladapo Olawale Afolabi [6,‡], Sandeep Tiwari [3], Alaa A. A. Aljabali [7,*], Murtaza M. Tambuwala [8], Ángel Serrano-Aroca [9], Elrashdy M. Redwan [10,11], Vladimir N. Uversky [12] and Kenneth Lundstrom [13,*]

1 Department of Genetics and Bioengineering, Faculty of Engineering and Natural Sciences, International University of Sarajevo, Hrasnicka Cesta 15, 71000 Sarajevo, Bosnia and Herzegovina
2 Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur 721172, India
3 Department of Genetics, Ecology and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; lucasgabriel388@gmail.com (L.G.R.G.); sandip_sbtbi@yahoo.com (S.T.)
4 Department of Biochemistry and Immunology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; ceciliahrp@gmail.com
5 Department of Bioinformatics, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; jessicaligia.pm@gmail.com
6 Department of Physiology and Biophysics, Pharmacology, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; afolabi.oladapo9@gmail.com
7 Department of Pharmaceutics and Pharmaceutical Technology, Faculty of Pharmacy, Yarmouk University, P.O. Box 566, Irbid 21163, Jordan
8 School of Pharmacy and Pharmaceutical Science, Ulster University, Coleraine BT52 1SA, UK; m.tambuwala@ulster.ac.uk
9 Biomaterials and Bioengineering Laboratory, Centro de Investigación Traslacional San Alberto Magno, Universidad Católica de Valencia San Vicente Mártir, c/Guillem de Castro 94, 46001 Valencia, Spain; angel.serrano@ucv.es
10 Department of Biological Science, Faculty of Science, King Abdulaziz University, Jeddah 21589, Saudi Arabia; rredwan@gmail.com
11 Therapeutic and Protective Proteins Laboratory, Protein Research Department, Genetic Engineering and Biotechnology Research Institute, City for Scientific Research and Technology Applications, New Borg EL-Arab 21934, Alexandria, Egypt
12 Department of Molecular Medicine and USF Health Byrd Alzheimer's Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; vuversky@usf.edu
13 PanTherapeutics, Route de Lavaux 49, CH1095 Lutry, Switzerland
* Correspondence: ahromic@ius.edu.ba (A.H.-J.); dr.barh@gmail.com (D.B.); alaaj@yu.edu.jo (A.A.A.A.); lundstromkenneth@gmail.com (K.L.)
† These authors contributed equally to this work.
‡ Are students of Postgraduate Transversal topic III, course "Personal Genome and Precision Health" (NAP 802) 2021 of ICB/UFMG.

check for updates

Citation: Hromić-Jahjefendić, A.; Barh, D.; Ramalho Pinto, C.H.; Gabriel Rodrigues Gomes, L.; Picanço Machado, J.L.; Afolabi, O.O.; Tiwari, S.; Aljabali, A.A.A.; Tambuwala, M.M.; Serrano-Aroca, Á.; et al. Associations and Disease–Disease Interactions of COVID-19 with Congenital and Genetic Disorders: A Comprehensive Review. *Viruses* **2022**, *14*, 910. https://doi.org/10.3390/v14050910

Academic Editor: Manuela Sironi

Received: 30 March 2022
Accepted: 25 April 2022
Published: 27 April 2022

Abstract: Since December 2019, the COVID-19 pandemic, which originated in Wuhan, China, has resulted in over six million deaths worldwide. Millions of people who survived this SARS-CoV-2 infection show a number of post-COVID complications. Although, the comorbid conditions and post-COVID complexities are to some extent well reviewed and known, the impact of COVID-19 on pre-existing congenital anomalies and genetic diseases are only documented in isolated case reports and case series, so far. In the present review, we analyzed the PubMed indexed literature published between December 2019 and January 2022 to understand this relationship from various points of view, such as susceptibility, severity and heritability. Based on our knowledge, this is the first comprehensive review on COVID-19 and its associations with various congenital anomalies and genetic diseases. According to reported studies, some congenital disorders present high-risk for developing severe COVID-19 since these disorders already include some comorbidities related to the structure and function of the respiratory and cardiovascular systems, leading to severe pneumonia. Other congenital disorders rather cause psychological burdens to patients and are not considered high-risk for the development of severe COVID-19 infection.

viruses

MDPI

Article

Potential Molecular Mechanisms of Rare Anti-Tumor Immune Response by SARS-CoV-2 in Isolated Cases of Lymphomas

Debmalya Barh [1,2,*,†], Sandeep Tiwari [2,†], Lucas Gabriel Rodrigues Gomes [2], Marianna E. Weener [3], Khalid J. Alzahrani [4], Khalaf F. Alsharif [4], Alaa A. A. Aljabali [5], Murtaza M. Tambuwala [6], Kenneth Lundstrom [7], Sk. Sarif Hassan [8], Ángel Serrano-Aroca [9], Kazuo Takayama [10], Preetam Ghosh [11], Elrashdy M. Redwan [12], Bruno Silva Andrade [13], Siomar de Castro Soares [14], Vasco Azevedo [2] and Vladimir N. Uversky [15,16]

1 Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur 721172, West Bengal, India
2 Department of Genetics, Ecology and Evolution, Institute of Biological Sciences, Federal University of Minas Gerais, Belo Horizonte 31270-901, Brazil; sandip_sbtbi@yahoo.com (S.T.); lucasgabriel388@gmail.com (L.G.R.G.); vascoariston@gmail.com (V.A.)
3 Clinical Research Center, Oftalmic, CRO, 119334 Bardina Str. 22/4, 119991 Moscow, Russia; marianna.weener@gmail.com
4 Department of Clinical Laboratories Sciences, College of Applied Medical Sciences, Taif University, Taif 21944, Saudi Arabia; ak.jamaan@tu.edu.sa (K.J.A.); alsharif@tu.edu.sa (K.F.A.)
5 Department of Pharmaceutics and Pharmaceutical Technology, Faculty of Pharmacy, Yarmouk University, Irbid 21163, Jordan; alaaj@yu.edu.jo
6 School of Pharmacy and Pharmaceutical Science, Ulster University, Coleraine BT52 1SA, UK; m.tambuwala@ulster.ac.uk
7 PanTherapeutics, CH 1095 Lutry, Switzerland; lundstromkenneth@gmail.com
8 Department of Mathematics, Pingla Thana Mahavidyalaya, Maligram, Paschim Medinipur 721140, West Bengal, India; sksarifhassan@pinglacollege.ac.in
9 Biomaterials and Bioengineering Lab, Centro de Investigación Traslacional San Alberto Magno, Universidad Católica de Valencia San Vicente Mártir, 46001 Valencia, Spain; angel.serrano@ucv.es
10 Center for iPS Cell Research and Application (CiRA), Kyoto University, Kyoto 606-8507, Japan; kazuo.takayama@cira.kyoto-u.ac.jp
11 Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA; preetam.ghosh@gmail.com
12 Department of Biological Science, Faculty of Science, King Abdulazizi University, Jeddah 21589, Saudi Arabia; rredwan@gmail.com
13 Laboratory of Bioinformatics and Computational Chemistry, Department of Biological Sciences, State University of Southwest Bahia (UESB), Jequié 45206-190, Brazil; bandrade@uesb.edu.br
14 Department of Immunology, Microbiology and Parasitology, Institute of Biological and Natural Sciences, Federal University of Triângulo Mineiro (UFTM), Uberaba 38025-180, Brazil; siomars@gmail.com
15 Department of Molecular Medicine and USF Health Byrd Alzheimer's Institute, Morsani College of Medicine, University of South Florida, Tampa, FL 33612, USA; vuversky@usf.edu
16 Research Center for Molecular Mechanisms of Aging and Age-Related Diseases, Moscow Institute of Physics and Technology, Institutskiy pereulok, 9, 141700 Dolgoprudny, Russia
* Correspondence: dr.barh@gmail.com
† These authors contributed equally.

Citation: Barh, D.; Tiwari, S.; Gabriel Rodrigues Gomes, L.; Weener, M.E.; Alzahrani, K.J.; Alsharif, K.F.; Aljabali, A.A.A.; Tambuwala, M.M.; Lundstrom, K.; Hassan, S.S.; et al. Potential Molecular Mechanisms of Rare Anti-Tumor Immune Response by SARS-CoV-2 in Isolated Cases of Lymphomas. *Viruses* **2021**, *13*, 1927. https://doi.org/10.3390/v13101927

Academic Editor: Elmostafa Bahraoui

Received: 27 July 2021
Accepted: 22 September 2021
Published: 25 September 2021

Abstract: Recently, two cases of complete remission of classical Hodgkin lymphoma (cHL) and follicular lymphoma (FL) after SARS-CoV-2 infection were reported. However, the precise molecular mechanism of this rare event is yet to be understood. Here, we hypothesize a potential anti-tumor immune response of SARS-CoV-2 and based on a computational approach show that: (i) SARS-CoV-2 Spike-RBD may bind to the extracellular domains of CD15, CD27, CD45, and CD152 receptors of cHL or FL and may directly inhibit cell proliferation. (ii) Alternately, upon internalization after binding to these CD molecules, the SARS-CoV-2 membrane (M) protein and ORF3a may bind to gamma-tubulin complex component 3 (GCP3) at its tubulin gamma-1 chain (TUBG1) binding site. (iii) The M protein may also interact with TUBG1, blocking its binding to GCP3. (iv) Both the M and ORF3a proteins may render the GCP2-GCP3 lateral binding where the M protein possibly interacts with GCP2 at its GCP3 binding site and the ORF3a protein to GCP3 at its GCP2 interacting residues. (v) Interactions of the M

CVB22

CURSO DE VERÃO EM BIOINFORMÁTICA

04 a 08 de julho de 2022

Organização e apoio

Certificamos que

Lucas Gabriel Rodrigues Gomes

fez parte da **equipe de organização** do VI Curso de Verão em Bioinformática da UFMG, realizado pelo Programa Interunidades de Pós-Graduação em Bioinformática da Universidade Federal de Minas Gerais, no período de 04/07/2022 a 08/07/2022.

Glória Regina Franco
Coordenadora do Projeto da Escola de Verão em Bioinformática da UFMG

Aristóteles Góes Neto
Coordenador do Programa Interunidades de Pós-Graduação em Bioinformática da UFMG

3RD INTERNATIONAL ASSOCIATED LABORATORIES MEETING

LIA 2022 BACT-INFLAM: CLOSING SEMINARS

We hereby certificate that **Lucas Gabriel Rodrigues Gomes** participated as a member of the organizing committee at the 3rd **International Associated Laboratoires Meeting - LIA 2022 BACT-INFLAM: CLOSING SEMINARS** at the Federal University of Minas Gerais (UFMG), Belo Horizonte, Minas Gerais, Brazil, during October 17th to October 21st.

Prof. Dr. Vasco Ariston de Carvalho Azevedo

Profa. Dra. Flávia Figueira Aburjaile

Support: INRAE UFMG icbufmg Escola de Veterinária UFMG Microbiologia Since 1969 UFMG Pós-Graduação Genética UFMG

Future Learn

Issued 2nd July 2021. futurelearn.com/certificates/dcbhh1f

Certificate of Achievement

Lucas Gabriel Rodrigues Gomes

has completed the following course:

ENGLISH AS A MEDIUM OF INSTRUCTION FOR ACADEMICS
UNIVERSITY OF SOUTHAMPTON

This online course explored issues around using English as a medium of instruction in universities. The course covered topics such as practice and research in EMI, how to use the voice effectively, appropriate language for EMI settings and the importance of intercultural awareness in EMI classrooms.

4 weeks, 4 hours per week

Mary Page
Senior Teaching Fellow in English
University of Southampton

Robert Baird
Teaching Fellow in English
University of Southampton

UNIVERSITY OF Southampton

The person named on this certificate has completed the activities in the attached transcript. For more information about Certificates of Achievement and the effort required to become eligible, visit futurelearn.com/proof-of-learning/certificate-of-achievement.

This certificate represents proof of learning. It is not a formal qualification, degree, or part of a degree.

TRANSCRIPT

UNIVERSITY OF Southampton

Lucas Gabriel Rodrigues Gomes

has completed the following course:

ENGLISH AS A MEDIUM OF INSTRUCTION FOR ACADEMICS
UNIVERSITY OF SOUTHAMPTON

This online course explored issues around using English as a medium of instruction in universities. The course covered topics such as practice and research in EMI, global contexts of EMI use, how to use the voice effectively when lecturing and teaching, appropriate language for EMI settings and the importance of intercultural awareness in EMI classrooms. The course worked to build a community of practice around EMI use and to develop confidence in using English in academic contexts.

STUDY REQUIREMENT
4 weeks, 4 hours per week

LEARNING OUTCOMES
- Explore aspects of research and practice in contemporary EMI across the world
- Develop confidence in using English as a medium of instruction and become part of a community of practice in EMI
- Compare the different contexts of use for EMI and reflect upon how this understanding fits with your own teaching context
- Identify challenges in facilitating effective intercultural communication in a variety of contexts (lectures, seminars, writing etc) and explore how these may be addressed
- Explore the role of the voice in effective intercultural communication
- Identify appropriate language to facilitate effective communication in English and discuss the role of language in intercultural communication
- Investigate how to promote successful communication in educational interactions

SYLLABUS
- What is EMI?
- Different contexts of EMI use and different interpretations/meanings of the term including internationalisation 'at home'
- Latest research findings and discussions about how ELF research has influenced EMI
- The role of language in effective intercultural communication
- Useful language for presenting ideas in English (e.g. signposting language)
- The use of the voice in effective intercultural communication, e.g. clarity, speed of delivery, intonation etc
- How far is accuracy in language important?
- Issues and challenges in lecturing for a multilingual and multicultural audience
- Facilitating and managing effective communication in small groups or seminars
- Recognising cultural differences in the international classroom and avoiding stereotypes
- Practical tips for addressing intercultural challenges

Future Learn