frontiers in Immunology

ORIGINAL RESEARCH
published: 29 March 2022
doi: 10.3389/fimmu.2022.862491

B-Cell Epitope Mapping of TprC and TprD Variants of *Treponema pallidum* Subspecies Informs Vaccine Development for Human Treponematoses

Barbara Molini[1†], Mark C. Fernandez[1†], Charmie Godornes[1], Anastassia Vorobieva[2,3], Sheila A. Lukehart[1,4] and Lorenzo Giacani[1,4]*

[1] Department of Medicine, University of Washington, Seattle, WA, United States, [2] VIB-VUB Center for Structural Biology, VIB, Brussels, Belgium, [3] Structural Biology Brussels, Vrije Universiteit Brussel, Brussels, Belgium, [4] Department of Global Health, University of Washington, Seattle, WA, United States

OPEN ACCESS

Edited by:
Pingyu Zhou,
Tongji University, China

Reviewed by:
Steven J. Norris,
University of Texas Health Science Center at Houston, United States
Yi-Pin Lin,
Wadsworth Center, United States
Feijun Zhao,
University of South China, China

***Correspondence:**
Lorenzo Giacani
giacal@u.washington.edu

[†]These authors have contributed equally to this work

Specialty section:
This article was submitted to Microbial Immunology, a section of the journal Frontiers in Immunology

Received: *26 January 2022*
Accepted: *07 March 2022*
Published: *29 March 2022*

Citation:
Molini B, Fernandez MC, Godornes C, Vorobieva A, Lukehart SA and Giacani L (2022) B-Cell Epitope Mapping of TprC and TprD Variants of Treponema pallidum Subspecies Informs Vaccine Development for Human Treponematoses. Front. Immunol. 13:862491. doi: 10.3389/fimmu.2022.862491

Several recent studies have focused on the identification, functional analysis, and structural characterization of outer membrane proteins (OMPs) of *Treponema pallidum* (*Tp*). The *Tp* species encompasses the highly related *pallidum*, *pertenue*, and *endemicum* subspecies of this pathogen, known to be the causative agents of syphilis, yaws, and bejel, respectively. These studies highlighted the importance of identifying surface-exposed OMP regions and the identification of B-cell epitopes that could be protective and used in vaccine development efforts. We previously reported that the TprC and TprD OMPs of *Tp* are predicted to contain external loops scattered throughout the entire length of the proteins, several of which show a low degree of sequence variability among strains and subspecies. In this study, these models were corroborated using AlphaFold2, a state-of-the-art protein structure modeling software. Here, we identified B-cell epitopes across the full-length TprC and TprD variants using the Geysan pepscan mapping approach with antisera from rabbits infected with syphilis, yaws, and bejel strains and from animals immunized with refolded recombinant TprC proteins from three syphilis strains. Our results show that the humoral response is primarily directed to sequences predicted to be on surface-exposed loops of TprC and TprD proteins, and that the magnitude of the humoral response to individual epitopes differs among animals infected with various syphilis strains and *Tp* subspecies. Rather than exhibiting strain-specificity, antisera showed various degrees of cross-reactivity with variant sequences from other strains. The data support the further exploration of TprC and TprD as vaccine candidates.

Keywords: *Treponema pallidum*, syphilis, Tpr proteins, B-cell epitope mapping, vaccine development

INTRODUCTION

The human treponematoses (syphilis, yaws, and bejel) are caused by a group of highly related pathogens classified as subspecies of the spirochete bacterium *Treponema pallidum* (*Tp*). Classically, the *pallidum* subspecies is said to causes syphilis, while the *pertenue* and *endemicum* subspecies are regarded as the causes of yaws and bejel, respectively (1), although the modes of transmission and the clinical manifestations may overlap among subspecies. These diseases are still a concern for public and global health, as they continue to result in substantial morbidity and mortality worldwide. According to the World Health Organization, the global prevalence of syphilis is ~20 million cases, with an incidence of ~6.3 million new cases every year (2). Although most of these infections occur in low- and middle-income countries, syphilis has resurged also in industrialized nations (3–7). If left untreated, syphilis can progress to affect the cardiovascular and central nervous systems of patients, potentially leading to death (8). Additionally, vertical transmission of syphilis is estimated to account for ~1/3 of stillbirths in sub-Saharan Africa (9, 10). Past public health initiatives to eliminate syphilis and congenital syphilis promoted by the CDC and WHO (11, 12) have significantly aided in reducing syphilis incidence and in generating awareness of this disease, but have not achieved their intended elimination goals. Compared to syphilis, less accurate epidemiological data are available on yaws and bejel (13). While the ongoing yaws elimination campaign in Asia and Africa using mass administration of azithromycin has demonstrated promising results (14), such efforts could be undermined by the spreading of macrolide resistant *Tp* subsp. *pertenue*, as recently demonstrated in Papua New Guinea (15). Foci of bejel have been reported in the last two decades, mostly in the Near East and Sahelian Africa (16–19), and bejel strains have recently reported to be transmitted sexually (20).

The chance of success of current and future control campaigns for all treponematoses would significantly increase if effective vaccines were available (21, 22). The most rational approach to vaccine development for these infections requires a clear understanding of the type of immune response that is protective and the identification of suitable candidate antigens to be tested in a pre-clinical animal model (21, 22). Furthermore, because there is very limited or no cross-immunity between subspecies of *Tp* and only sporadic cross-immunity between syphilis strains (23, 24), the identification of antigenic differences in potential vaccine candidates among subspecies and strains is of pivotal importance, as such differences could be key to devising a broadly protective vaccine (22). There is consensus that vaccine candidates are most likely to be found among these spirochetes' surface-exposed antigens, such as (but not limited to) integral outer membrane proteins (OMPs). Integral *Tp* OMPs will necessarily contain a membrane-embedded β-barrel domain composed of antiparallel β-strands joined together by loops that alternatively protrude toward the extracellular environment or the periplasm (25). Because *Tp* clearance from early lesions is dependent on opsonophagocytosis of *Tp* cells by activated macrophages (26, 27), the identification of surface-exposed epitopes that can be targeted by immunization to induce opsonic antibodies and promote macrophage activation is key to vaccine development. Such tasks, however, have been historically challenging due to the inability to steadily propagate the *Tp* subspecies *in vitro*, which was only recently achieved (28), and also because of the uncommon fragility and limited protein content of these spirochetes' OM (29, 30). These limitations have been partially overcome by the ability to predict *in silico* OMP-encoding genes and the structure of their encoded proteins, enabling investigation using structural and functional experimental approaches (31, 32).

Among *Tp* putative OMPs identified to date, there are several members of the *T. pallidum* repeat (Tpr) family of paralogous proteins, including TprC and TprD (encoded by the *tp0117* and *tp0131* genes in the reference Nichols strain, respectively) (33); these are reported to have OM localization and porin activity (34, 35). These paralogs have been identified as vaccine candidates by past studies in which it was demonstrated that the N-terminal conserved region of these antigens elicited strong antibody and T-cell responses during infection, and immunization with this region attenuates syphilitic lesion development upon infectious challenge (36). In this study, we examine the protein sequence variation in TprC and TprD among *T. pallidum* strains and subspecies, and predict, then confirm, the locations of B cell epitopes using antisera from infected and immunized rabbits. Variant specificity and cross-reactivity are analyzed so that epitopes with broad coverage among strains and subspecies can be identified for future evaluation as vaccine antigens.

RESULTS

Sequence Analysis of TprC and TprD Variants

Although the TprC and TprD proteins are identical in the Nichols, Chicago, and Bal73-1 strains, allelic variants of TprC and TprD exist among syphilis strains and the three *Tp* subspecies (35, 37). Among the treponemal strains used in this study (**Figure 1**), four alleles were found at the *tprD* locus, which include the reference *tprD* allele (found in the syphilis Nichols, Chicago, and Bal73-1 strains), and the *tprD2* allele (found in the syphilis strains MexicoA, Sea81-4, Bal3, and UW249) which encodes the $TprD_2$ protein (35). Also the subsp. *pertenue* SamoaD strain and subsp. *endemicum* IraqB strains harbor a $tprD_2$ allele in the *tprD* locus, but their $TprD_2$ amino acid sequences differ from the subsp. *pallidum* $TprD_2$ sequence due to five amino acid substitutions scattered throughout the length of the protein (**Figure 1**) (35). TprD2 has four unique regions that differentiate it from the reference TprD sequence. These include a large central region of 110 amino acids and three smaller regions toward the COOH-terminal end of the protein (**Figure 1**) (35). As previously reported, the *tprC* locus of MexicoA, Sea81-4, and Bal3 encodes a TprC variant with a limited number of amino acid (aa) changes (15 aa for MexicoA, 9 aa for Sea81-4, and 9 aa for Bal3) compared to the reference TprC found in Nichols, Chicago, and Bal73-1 strains (**Figure 1**) (35). The TprC protein of the *pertenue* and

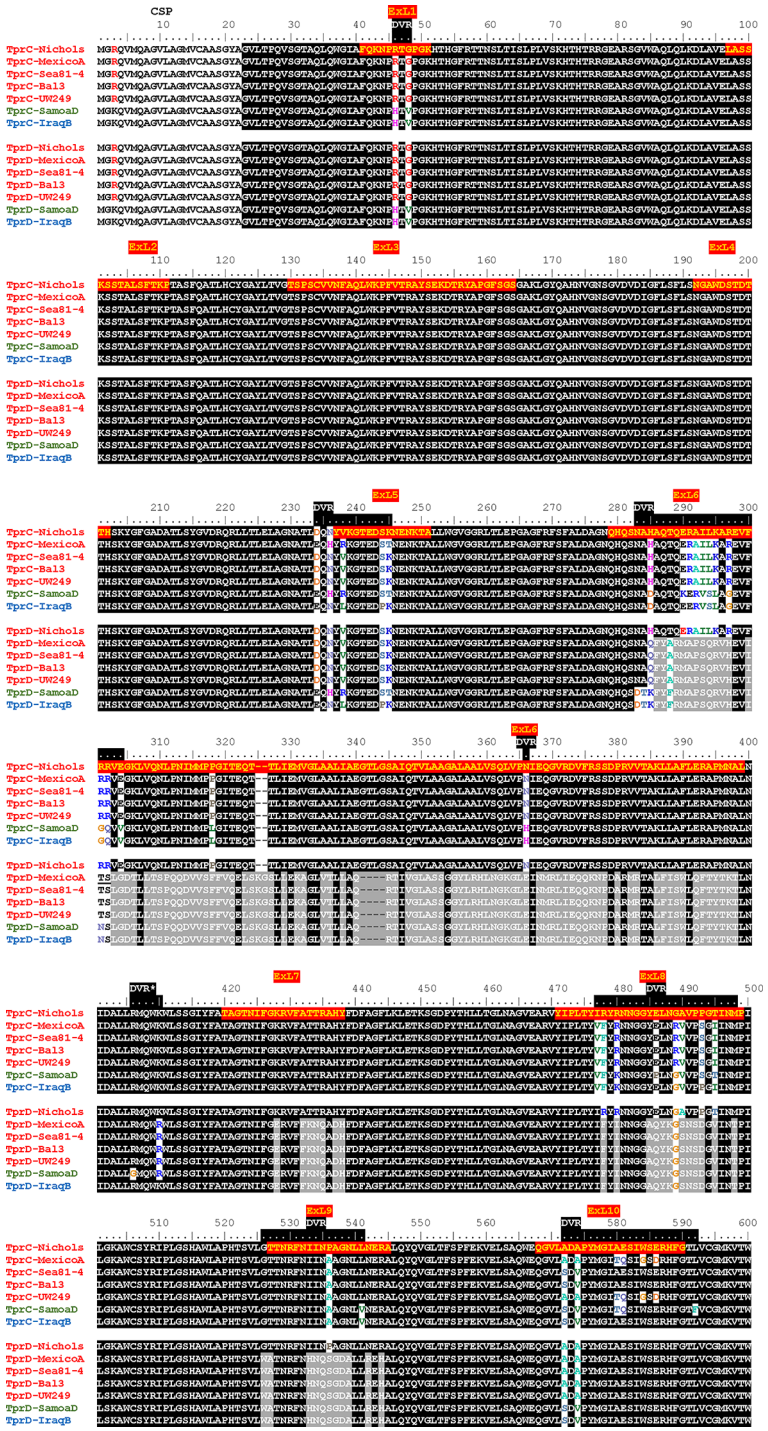

FIGURE 1 | Alignment of amino acid sequences of the TprC and TprD/D_2 variants. *Tp.* subsp. *pallidum* strains (Nichols, MexicoA, Sea81-4, Bal3, and UW249) are indicated in red font on the left of the sequence. The *Tp* subsp. *pertenue* strain (SamoaD) is in green font, and the *Tp* subsp. *endemicum* (IraqB) strain is in blue font. The Chicago and Bal73-1 TprC and TprD sequences (not shown) are identical to the Nichols strain. The MexicoA, Sea8-14, Bal3, UW249, SamoaD, and IraqB strains harbor a $TprD_2$ variant within the *tprD* locus. CSP, predicted cleavable signal peptide; ExL, External Loop. Amino acids encompassing the ExLs predicted by AlphaFold2 are highlighted in red with yellow text only in the top sequence. DVR, Discrete Variable Region. DVRs are highlighted in black along the ruler. *Indicates a DVR found in $TprD_2$ but not TprC and TprD variants.

endemicum strains studied here also shows limited amino acid changes compared to the reference TprC (31 aa for SamoaD and 26 aa for IraqB; **Figure 1**), albeit higher compared to the subsp. *pallidum* strain (35). We previously reported that TprC and TprD/D_2 sequence variation does not occur randomly, but rather is localized in discrete variable regions (DVRs; **Figure 1**) (35). When TprC and TprD variants are compared (with the exclusion of $TprD_2$), seven DVRs are found throughout the protein sequence, while 8 DVRs can be identified within the $TprD_2$ sequences (**Figure 1**). To obtain predictions of TprC and $TprD_2$ structures from their amino acid sequences (**Figure 1**) and map the DVRs on these models, we used the recently developed AlphaFold2 software (https://AlphaFold.ebi.ac.uk/) (38). These new models revealed remarkably similar structures for TprC and $TprD_2$ and identified these proteins as relatively large β-barrel integral OMPs of 20 β-strands connected by ten external loops (ExLs, protruding toward the extracellular milieu), and nine periplasmic loops (**Figure 2**). Elevated structural identity of the transmembrane region was also supported by a 0.75 backbone root-mean-square deviation (RMSD) score (**Figure 2A**). The local model quality, indicated by the Predicted Local Distance Difference Test (pLDDT) value was high in the transmembrane and periplasmic loop regions, and slightly lower in the predicted ExLs, suggesting conformational flexibility (**Figure 2B**). Except for two substitutions (aa 407 and 410 mapping to a periplasmic β-turn), all DVRs localized within a subset of the surface-exposed ExLs (**Figure 2C**). More specifically, DVRs were located in ExL1, ExL5-6 and ExL8-10 of the TprC and $TprD_2$ models; while ExL2-4 harbored conserved loops. ExL6 is also conserved between $TprD_2$ sequences from various isolates, although its shows only 60% of sequence identity to the ExL6 of other TprC and TprD variants (**Figure 1**). DALI software (39) and PDB analyses to identify structurally similar porins (**Table S1**) showed that the highest-scoring structures did not contain the exact number of β-strands predicted by AlphaFold2 for TprC and $TprD_2$ β-barrels, but slightly higher or slightly lower, but well within the models of integral OMPs with no large periplasmic domains. These results suggest that these Tpr proteins belong to a new family of porins not yet represented in the PDB database.

In silico prediction of B-cell epitopes using BepiPred2.0 (http://www.cbs.dtu.dk/services/BepiPred/), IEDB (https://www.iedb.org/), and BCpreds (http://ailab-projects1.ist.psu.edu:8080/bcpred/data.html) (**Tables S2–S5** and **Figure S1**) showed that the putative TprC and TprD ExLs were also enriched in immunogenic epitopes. Therefore, it is possible that the antigenic variability in the ExLs regions has functional significance in immunity to the *T. pallidum* subspecies.

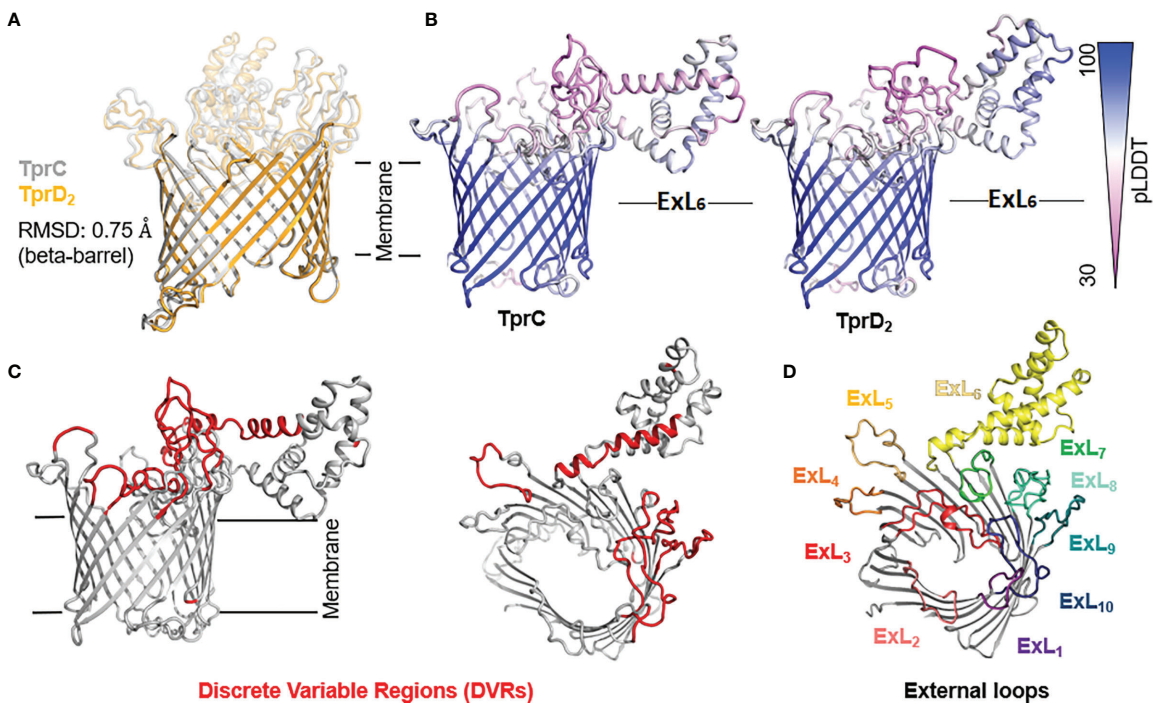

FIGURE 2 | Predicted structures of TprC and $TprD_2$ using AlphaFold2. TprC and $TprD_2$ sequences were obtained from the Sea81-4 genome (NZ_CP003679.1/CP003679.1). **(A)** Both TprC and $TprD_2$ are predicted to fold into similar 20-strands beta-barrels structures. The predicted structures are nearly identical in the transmembrane region (backbone RMSD = 0.75). **(B)** Compared to the scaffolding, slightly lower confidence in the structural models for TprC and $TprD_2$ are seen in the external loop regions based on the per-residue model confidence (pLDDT) mapped on the predicted TprC and $TprD_2$ structures, suggesting structural flexibility for these regions. Noteworthy, however, is that the largest loop (ExL7) is predicted to have a similar helical-rich structure in both TprC and $TprD_2$. Full pLDDT analysis for TprC and $TprD_2$ is reported in **Figure S3**. **(C)** The DVRs (in red on the TprC model structure) identified by aligning TprC and $TprD_2$ sequences of different strains (**Figure 1**) mostly localize in the predicted extracellular loops. **(D)** AlphaFold identified ten surface-exposed external loops (ExLs). Discrete variable regions (DVRs) based on **Figure 1** alignments are predicted to be in ExL1, ExL5-6 and ExL8-10.

To validate the B-cell epitope prediction and evaluate the cross-reactivity of these epitopes across species and strains, we performed experimental B-cell epitope mapping of the TprC, and TprD/D_2 proteins with a Geysan pepscan approach based on overlapping synthetic peptides (40) using sera from animals infected with *Tp* subsp. *pallidum*, *Tp* subsp. *pertenue*, and *Tp* subsp. *endemicum* strains. Furthermore, we compared antibody reactivity in sera from infected rabbits with that of sera from rabbits immunized with a subset of full-length refolded recombinant TprC proteins.

Humoral Responses to Homologous TprC and TprD/D_2 Peptides in Experimentally Infected Rabbits

Groups of three laboratory rabbits were infected intratesticularly (IT) with one of seven syphilis strains (Nichols, Chicago, Bal73-1, MexicoA, Sea81-4, Bal3, and UW249), one yaws strain (SamoaD), and one bejel strain (IraqB). From these animals, serum samples were obtained at day 30, 60, and 90 post-infection. Pooled sera from animals in each infection group/time point were tested in ELISA to assess reactivity to homologous overlapping synthetic peptides (20-mers overlapping by 10 amino acids) representing the TprC and TprD/D_2 variants previously identified in each strain. The full list of synthetic peptides used in this study, with amino acids encompassing predicted ExLs highlighted in red with yellow text, and percentage amino acid homology among peptides across strains is shown in **Table S6**. Peptide nomenclature is explained in **Table S6** footnote. Cumulative absorbance data from the three timepoints (sum of the mean absorbance values for day 30, 60, 90 values for each infected rabbit group) are reported in **Figures 3A–C**. Epitope mapping studies of the NH_2-terminal portion of the protein resulted in the identification of six highly reactive peptide regions (**Figure 3A**) representing sequences shared by all TprC and TprD genes in the studied subspecies *pallidum* strains: C1-C3, C6, C13-C14, C18, C20, and C25-C29. Based on AlphaFold2 structural predictions, 9 of these 13 peptides had at least 7 amino acids mapping to the predicted external loops of the protein, while only four reside in predicted transmembrane scaffolding and periplasmic loop regions (C1, C6, C20 and C25; **Figure 1A** and **Table S7**). It is noteworthy that all three B cell epitope prediction programs uniformly predicted all six of the experimentally determined epitope-containing regions of the NH_2-terminal portion of the subspecies *pallidum* TprC and D proteins (**Tables S2–S5** and **Figure S1**).

Several epitopes were also identified in the COOH-terminal region of these proteins, and corresponded to peptides the same regions in *pallidum* and non-*pallidum* subspecies: C46 and C47 homologs from Nichols (**Figure 3A** and **Table S7**), SamoaD (S-C46, S-C47; **Figure 3B** and **Table S7**) and IraqB (I-C46, I-C47; **Figure 3B** and **Table S7**), C51 homologs from Nichols (N-C51), SamoaD and iraqB (S/I-C51) (**Figures 3A, B** and **Table S7**); and C53-C55 homologs from Nichols, Bal3/Sea81-4 (**Figure 3A** and **Table S7**), and IraqB (**Figure 3B** and **Table S7**). Similarly, the C43, and D45-D47 (ExL8) peptides, mapping to the TprD$_2$ COOH-terminus were found to contain B-cell epitope(s) (**Figure 3C** and **Table S7**). Additional TprD$_2$ peptides found to be reactive were D33-D35 (ExL6), I-C39, D40-41 (ExL7), C49, and D51 (ExL9). In our 3D models of these proteins, all the reactive peptides in the COOH-terminus fall within predicted ExLs (**Tables S6–S7** and **Figure 1**), except for C43, most of I-C39 (75%), C49, and C53, which are predicted transmembrane scaffolding sequences. Of these "scaffold epitopes", only one (C43) was predicted by a B-cell prediction program (**Tables S2–S5** and **Figure S1**).

The percentage of immune sera that showed reactivity to many of the peptides was variable. For example, peptides C3, and C13 were recognized by rabbits infected with 28% of the *Tp* subsp. *pallidum* strains; peptides C1, and C27 were recognized by rabbits infected with 42% of the strains; peptides C14 and C28 were recognized by rabbits infected with 57% of the strains; C6 was recognized by rabbits infected with 71% of the strains; and peptides C2 and C18 were recognized by rabbits infected with 85% of the *Tp* subsp. *pallidum* strains (**Figure 1**). Overall, based upon the AlphaFold2 models, these results show that humoral reactivity elicited to these Tpr antigens during experimental infection is directed primarily to predicted surface-exposed regions of the TprC/D and TprD$_2$ proteins.

Reactivity to Non-Homologous TprC and TprD$_2$ Peptides

Epitope mapping using short peptides based on the TprC and TprD/D_2 sequences from multiple *Tp* strains and sera from infected animals also allowed us to investigate cross-reactivity to non-homologous peptides to determine the fine specificity of the antibody response to these antigens. Such analyses focused on peptides mapping to the proteins` COOH-terminal regions, due to the higher sequence variability in this region, compared to the more conserved NH_2-terminal region (**Figure 1**). Major variable regions include peptides C46 - C47 (mapping to the predicted ExL8), C51 (ExL9), and C55 (ExL10) (**Figure 1** and **Tables S6–S7**). Four distinct variants of each of the C46, C47, and C55 peptides, and three variants of C51, representing all sequences found in the strains studied here, were tested against all nine pools of immune sera obtained at day-90 post-experimental infection.

As shown in **Figures 4A–D**, very few sera were reactive only to their homologous peptide. For example, the Bal73-1and SamoaD antisera were primarily reactive only to their own C46 sequences (**Figure 4A**), although the Bal73-1 antiserum showed a very modest reactivity to the IraqB peptide variant (**Figure 4A**). When reactivity against the C47 peptide was analyzed, the Chicago, Bal73-1 sera reacted only to their own peptide, the IraqB antiserum reacted to all variants but the SamoaD peptide, and the Sea81-4 serum only saw the Nichols C46 variant, but not its homologous peptide. (**Figure 4B**). Only Chicago, Bal73-1, and SamoaD sera showed complete strain-specificity for the C51 peptides (**Figure 4C**), while none of the sera reactive to C55 showed complete strain-specificity (**Figure 4D**). When tested against TprD2 peptides, most antisera did not show any reactivity. There were, however, two exceptions, as the Chicago sera cumulatively showed reactivity to the D34 and D47 peptides,

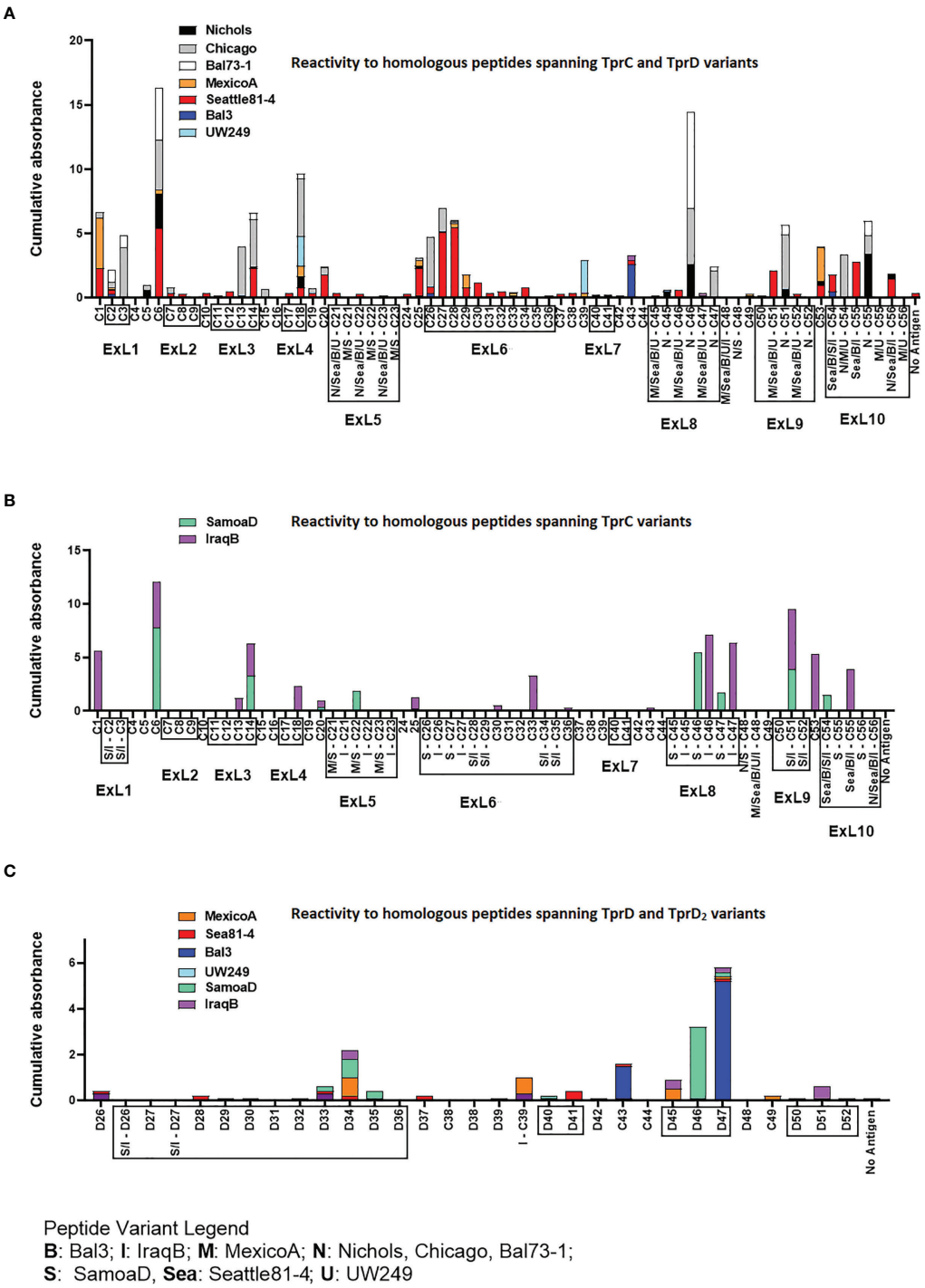

FIGURE 3 | Reactivity of sera from experimentally infected animals to homologous peptides representing the TprC, TprD and $TprD_2$ variants. **(A)** Reactivity to homologous peptides spanning TprC and TprD proteins of sera from rabbits infected with *Tp* subsp. *pallidum* (Nichols, Chicago, Bal73-1, MexicoA, Sea81-4, Bal3, and UW249B) collected at day 30, 60, and 90 post-infection. Nichols, Chicago, and Bal73-1 sequences are identical. **(B)** Reactivity to homologous peptides spanning TprC variants of immune sera from groups of rabbits infected with *Tp* subsp. *pertenue* (SamoaD) or *Tp* subsp. *endemicum* (IraqB) strains collected at day 30, 60, and 90 post-infection. **(C)** Reactivity to homologous peptides spanning TprD and $TprD_2$ variants of sera collected at day 30, 60, and 90 post-infection from all $TprD_2$-containing *Tp* subspecies and strains studied here. Cumulative Absorbance values are the sum of the mean OD values obtained from all animals in the infection group at all three time points. Boxed peptides contain at least seven amino acids (35% of the peptide length) belonging to a predicted ExL. Strain names on x axis are abbreviated as follows: N, Nichols; M, MexicoA; Sea, Sea81-4; B, Bal3; U, UW249; S, SamoaD; I, IraqB.

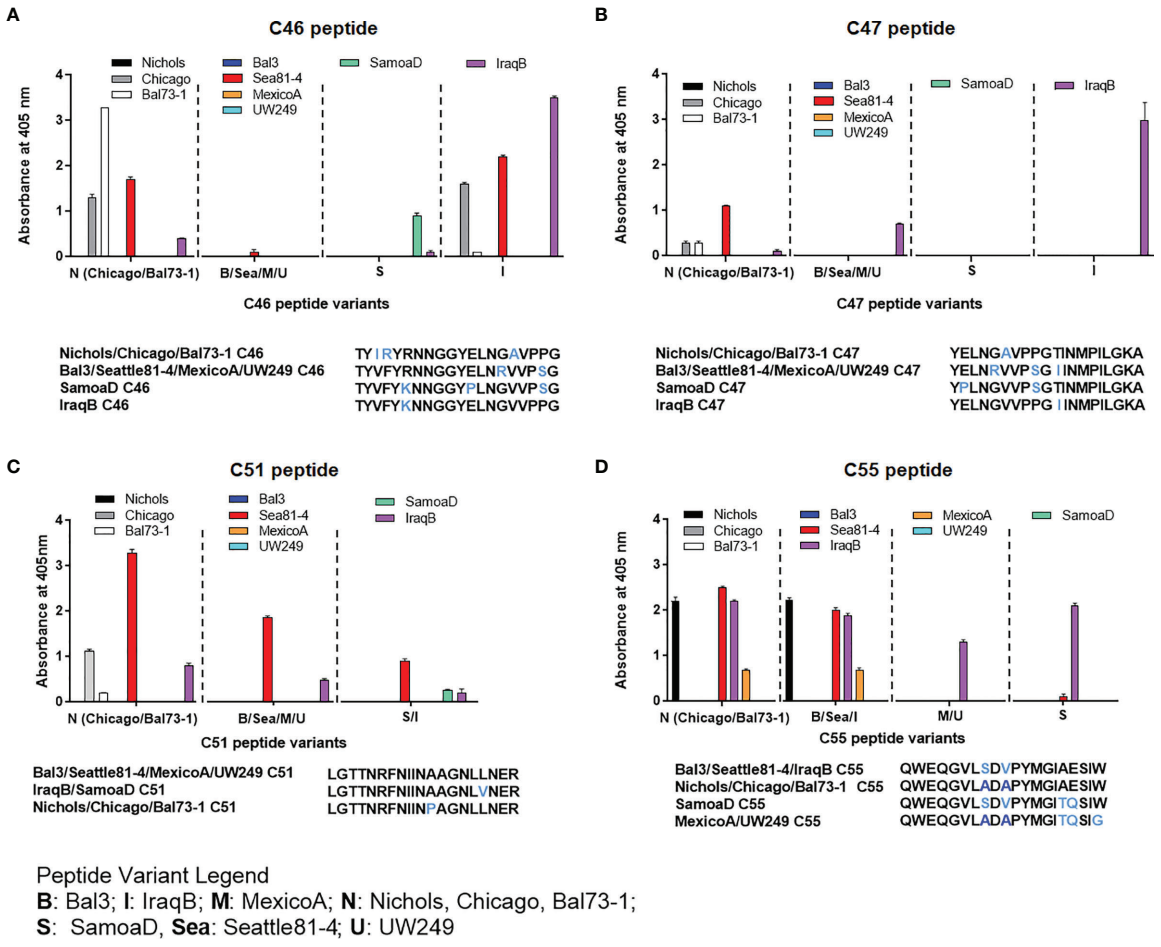

FIGURE 4 | Reactivity of sera from experimentally infected animals to homologous and non-homologous peptides C46, C47, C51, and C55. Humoral reactivity of day-90 sera from experimentally infected animals to homologous and non-homologous TprC peptides. **(A-D)** reactivity to C46, C47, C51, and C55 variants. Strain names on x axes are abbreviated as follows: N, Nichols; M, MexicoA; Sea, Sea81-4; B, Bal3; U, UW249; S, SamoaD; I, IraqB.

with OD values of 3.6 and 6.6, respectively. However, only the D47 peptide was consistently recognized at all time points, while D34 was recognized only at day 60. Overall, these data indicate that cross reactivity is possible, and perhaps suggest that immunization with a given sequence might generate cross-reactive antibodies able to overcome the obstacle of sequence diversity among TprC epitopes, a feature that is desirable in vaccine development as they may be broadly opsonic or neutralizing.

Humoral Response to TprC Peptides Following Rabbit Immunization With Full-Length Refolded Antigens

Refolded antigens, analyzed using circular dichroism (CD), were found to have a β-pleated sheet component of about 48% for all three antigen variants. Random coil was also found to be 48% of the protein structure, while only 4% was identified as alpha helices. Epitopes recognized following immunization with any of three recombinant full-length TprC variants from *Tp* subsp. *pallidum* strains (Nichols/Chicago/Bal73-1, Sea81-4/Bal3, and MexicoA) were also identified to evaluate differences with infection-induced antibody responses. Results showed that sera from animals immunized with the Nichols TprC sequence were highly reactive to peptides C1-C3, C6 and C47, and moderately reactive to peptides C5, C9, C16-18, C28, C32, C34-C35, C53 and C55 (**Figure 5A**). Of these 16 peptides, six mapped almost exclusively to putative surface-exposed loop regions (C28, C32, C34, C35, C47, and C55), five (C1, C5-C6, C16, and C53) mapped to predicted transmembrane scaffolding sequences, while five peptides (C2-C3, C9 and C17-C18) contained both surface-exposed loops and scaffold regions. Sequences of these peptides and location in the predicted protein models are reported in **Table S8**. When tested against non-homologous peptides (**Figure 5B**), the Nichols TprC-immunized sera strongly recognized the SamoaD/IraqB C2-C3 variants, and all three heterologous C47 variants (SamoaD, Iraq B, and Sea81-4), while modest reactivity was seen towards the MexicoA/UW249 C55 peptide variant, the SamoaD/IraqB C34, and both C26

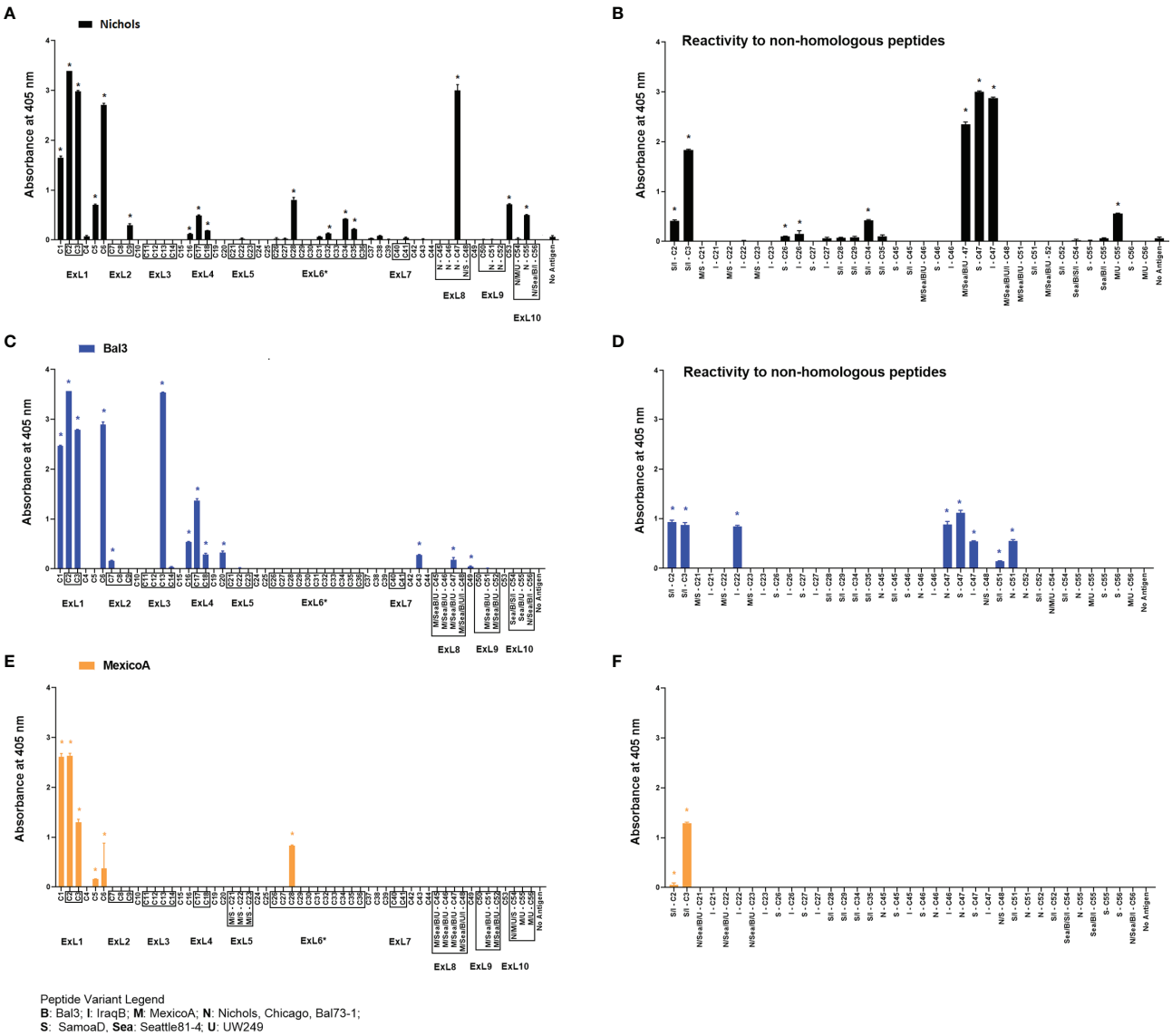

FIGURE 5 | Humoral reactivity to TprC peptides following immunization with refolded recombinant full-length TprC antigens. Reactivity to TprC homologous (left panels) and non-homologous peptides (right panels) in sera from rabbits immunized with Nichols **(A, B)**, Bal3 **(C, D)**, and Mexico A **(E, F)** variants of TprC. Asterisk (*) indicates significant reactivity compared to no antigen control. Peptides encompassing sequences predicted to be within ExLs are boxed. Peptide sequence and homology among strains are reported in **Table S6**. Strain names on x axes are abbreviated as follows: N, Nichols; M, MexicoA; Sea, Sea81-4; B, Bal3; U, UW249; S, SamoaD; I, IraqB.

variants from SamoaD and IraqB (**Figure 5B**). Immunization with the Bal3 variant of TprC elicited high reactivity to peptides C1-3, C6, and C13, and moderate reactivity to peptides C7, C16-18, C20, C43, C47, and C49 (**Figure 5C** and **Table S8**). Of these thirteen peptides, six (C2-C3, C13, C17-C18, and C47) mapped predominantly to ExLs, and seven (C1, C6, C7, C16, C20, C43, and C49) predominantly to the protein transmembrane scaffolding (**Table S8**). Cross-reactivity to non-homologous peptides was seen predominantly to the SamoaD/IraqB C2 and C3, IraqB C22, and all variants of C47 and C51 (**Figure 5D**). Antisera from rabbits immunized with the MexicoA TprC variants primarily recognized homologous peptides C1-3 and, secondarily, C5, C6, and C28 (**Figure 5E**). Of these six, one peptide mapped to the predicted ExL6 (C28), three mapped only to the transmembrane scaffolding (C1, C5-C6) and two mapped to a peptide predicted to contain portions of both (C2-C3) (**Table S8**). Cross-reactivity to the non-homologous SamoaD/IraqB C2 and C3 was also detected (**Figure 5F**). Overall, these data show that, as seen in infection-induced antibody responses, the humoral response following immunization with full-length TprC variants is mainly elicited by predicted surface-exposed sequences, rather than sequences mapping to the β-barrel

transmembrane scaffolding, and that cross-reactivity to non-homologous peptides is possible.

A side-by-side comparison of the infection- vs. immunization-induced humoral response to peptides is shown in **Figure 6**. For this comparison, the mean value of the cumulative reactivity seen in sera at day 30, 60, and 90 sera post-experimental infection is shown for each peptide. Sera from immunized animals were obtained three weeks after the last immunization. All sera were tested at the same dilution. In general, immunization-induced reactivity to most peptides appeared to be higher than that elicited by experimental infection; specific examples include C1-C3, C5, C9, C16-C17, C28, C32, C34-C35, N-C47, and C53 peptides (**Figure 6A**). For Nichols-clade *T. pallidum* strains (**Figure 6A**), which contain identical *tprC* and *tprD* loci, this was most noticeable for epitopes located in the NH_2- and COOH-terminal regions of the protein. In contrast, infection-induced antibody responses to epitopes in the central part of the protein were comparable to or higher than those induced by immunization. For TprD2-containing subsp. *pallidum* strains (**Figures 6B, C**), immunization-induced responses were limited to the NH_2-terminal portion of the protein (including ExL1-3) and virtually no immunization-

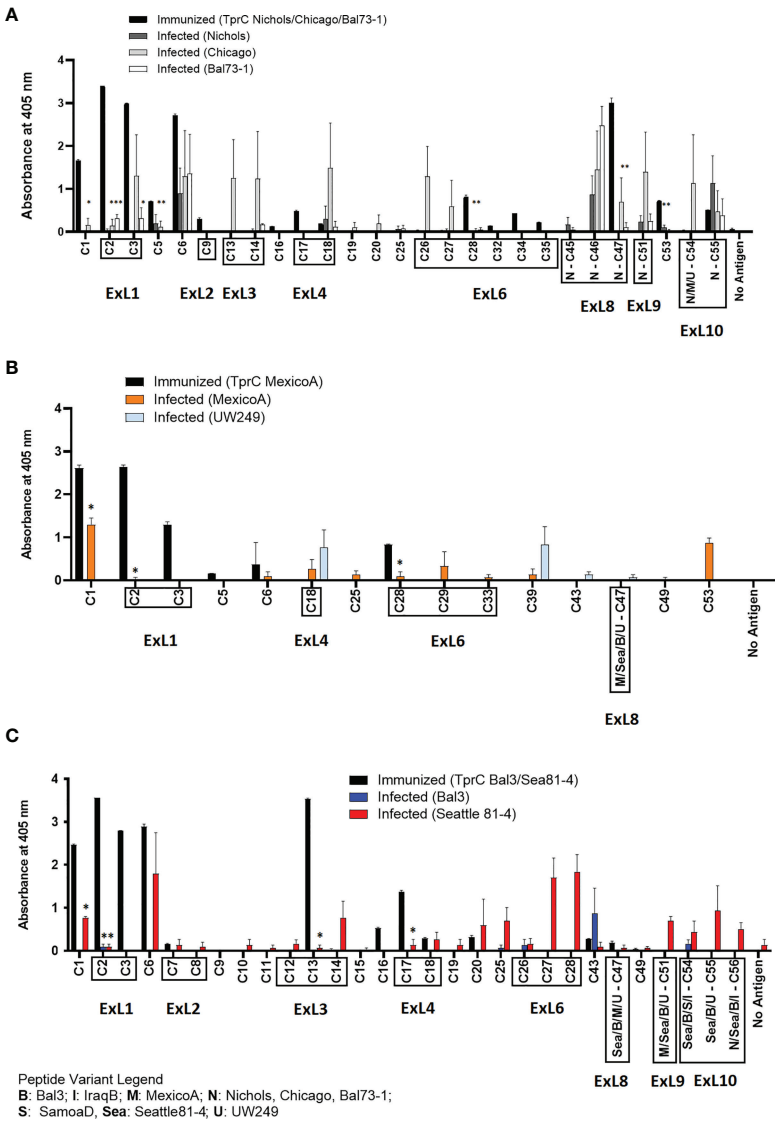

FIGURE 6 | Comparison of reactivity of sera from infected animals vs. immunized animals. **(A-C)** Reactivity to peptides following immunization with TprC variants compared to experimental infection. Data shown are means +/- SEM of 3 rabbits per group: 3 weeks post final boost (immunized) and mean +/- SEM of values for days 30, 60, 90 post-infection (infected). Asterisk (*) indicates a significant difference in reactivity compared to the reactivity value following immunization. Peptides encompassing sequences predicted to be within ExLs are boxed. Strain names on x axes are abbreviated as follows: N, Nichols; M, MexicoA; Sea, Sea81-4; B, Bal3; U, UW249; S, SamoaD; I, IraqB.

induced antibodies were detected for epitopes in the central and COOH-terminal regions, although these were recognized by infection-induced responses. Overall, these data support that, in most cases, immunization elicits a higher reactivity to TprC B-cell epitopes compared to experimental infection, particularly for those epitopes located in the NH_2-terminal portion of the protein. These data support the preferential use of the amino portion of TprC, which contains multiple conserved ExLs, for vaccine studies.

DISCUSSION

The continuing prevalence of syphilis, in the face of highly effective therapy and active control programs, highlights the need for a protective vaccine. The development of such a vaccine calls for a deeper understanding of the mechanisms of protective immunity and the antigens and adjuvants that induce protection. Our laboratories have been examining these issues for many years (22, 31, 36, 41–48). Much of that work has focused on the Tpr antigens of *T. pallidum*. In this current study, B-cell epitope mapping studies of the TprC/D and $TprD_2$ proteins of *Tp* reveal that antibodies arising during experimental infection recognize sequences predicted, using state-of-the-art modeling systems, to fall largely in the proteins' surface-exposed loops. Because opsonic antibodies are required for efficient ingestion and killing of *T. pallidum* by macrophages, surface-exposed epitopes are attractive targets as vaccine candidate antigens.

A broadly protective vaccine would need to be effective against most strains of *T. pallidum*, optimally including the agents of syphilis as well as the endemic treponematoses yaws and bejel. Because some of the external loops of Tpr C/D and $TprD_2$ demonstrate sequence heterogeneity among strains and subspecies of *T. pallidum*, we expected that these epitopes might be strain-specific, similar to the specificity demonstrated for the variable regions of TprK (43, 45, 49). For this reason, we included seven strains of *Tp* subsp. *pallidum* as well as strains from the subspecies *pertenue* and *endemicum* in our work. Unexpectedly, we saw cross reactivity of antibodies toward the variant peptides (**Figure 4**). These findings support the use of TprC/D as at least one component of broadly effective candidate vaccine. It was intriguing that in two instances (Sea81-4 serum for C47; and MexicoA serum for C55) antisera only recognized non-homologous peptides (**Figures 4B, D**). In the case of the Sea81-4 serum, reactivity was detected at day 30 post-infection, but not the at later time points. For the MexicoA serum, one could hypothesize that masking of key epitope residues occurred for the non-recognized homologous peptide during absorption to the ELISA plate which, in turn, could have reduced assay sensitivity, as discussed for peptide arrays by Cretich et al. (50).

The AlphaFold2 structural predictions for TprC/D and $TprD_2$, as well as our CD analyses of purified refolded recombinant TprC variants, support our model (35) that these Tprs are membrane-localized 20-stranded β-barrel proteins containing numerous surface-exposed loops. Very similar models for TprC were previously obtained using I-TASSER (51) (https://zhanggroup.org/I-TASSER/) (35). Interestingly, when AlphaFold2 and I-TASSER results are compared, the only difference is that I-TASSER splits ExL6 (**Figure 1**) into two external loops separated by a β-hairpin, so that I-TASSER predictions harbor 11 external loops instead of 10. AlphaFold2, on the contrary, predicts a significantly larger ExL6, mapping approximately to the proteins' central domains. AlphaFold2 is the new standard for *ab-initio* structural prediction, and in the 2020 Critical Assessment of protein Structure Prediction (CASP) global challenge, it outperformed any other structure prediction algorithm, including I-TASSER (https://predictioncenter.org/casp14/zscores_final.cgi). Furthermore, in a recent preprint (52), AlphaFold2 was shown to work well on structural prediction for membrane proteins, although the exercise focused mostly on alpha-helical membrane proteins, and additional analyses are necessary to establish the same benchmark for β-barrel proteins.

In previous work by Anand et al. (34, 53) significantly different models for the TprC/D proteins were reported, compared to those provided here. These models, however, are not supported by AlphaFold2, which finds the structure of all Subfamily I and Subfamily II Tpr family members to be very similar to the structures for TprC/D and $TprD_2$ in **Figure 2A**. Although there is not unanimous agreement on the structure of these antigens within our scientific community, our epitope mapping data support our AlphaFold2 models, predicting a predominantly β-barrel structure for TprC and TprD/D2 (34, 53). Further studies and integration of all the structural, functional, and immunological data are needed to establish a consensus on the structure of these antigens until crystallographic (or equally reliable) data become available.

This study also provides evidence that infection with different strains might lead to differences in the breadth and intensity of the humoral response against the same epitope, as reported previously for responses to longer portions of the Tpr proteins (54). It is the case, for example, of rabbits infected with the Sea81-4 strain of *Tp* that overall recognize more TprC/D peptides compared to other *Tp* subsp. *pallidum* strains. The biological basis for these differences is unclear at this time, in part due to the limitations of our understanding of *Tp* biology and syphilis pathogenesis. As the technical gap in the approaches to study this difficult pathogen narrows, and genomics, proteomics, and transcriptomics data populate public repositories, more light will be shed on the causes of differential reactivity. Overall, however, it is plausible to postulate that enhanced serological reactivity might be due to an overall increased expression of the target antigen in each strain. This hypothesis is supported by previous work where we showed the *tprC* mRNA level was higher in the Sea81-4 strain compared to other *Tp* subsp. *pallidum* strains (Nichols, Chicago, Bal73-1) used in this study (55).

Our studies further demonstrated that epitopes in TprC/D and $TprD_2$ are nearly-uniformly distributed across the length of the protein, even though the most reactive peptide epitopes are in the NH_2- and COOH-terminal regions. Previously published (36) experiments have shown that both of these regions in the Nichols TprC protein contain protective epitopes, as

immunization with these antigen fragments significantly attenuated lesion development upon infectious challenge (36), and polyclonal antisera elicited by immunization with these portions facilitated treponemal ingestion by macrophages in opsonophagocytosis assays compared to normal rabbit sera (36). Further work, however, will be necessary to identify which specific surface-exposed sequences provide targets for opsonic antibodies, which may not coincide with sero-dominant epitopes, as the pathogen gains an obvious advantage by exposing to the immune system epitopes with little or no protective value. Assuming that differentially recognized peptides are located on the antigens' protective epitopes, one could hypothesize that protective epitopes are immunologically sub-dominant during natural infection, which, *per se* could represent an additional strategy the pathogen uses to survive in the host despite the immune response that naturally develops to these antigens. If this was found to be the case, effort would need to be put into outflanking and overcoming immunodominance to target subdominant protective epitopes. Upon immunization, the immunodominance hierarchy is established at the germinal center, where B cells compete for the antigen through based on binding affinity, and subsequently undergo clonal expansion to become plasma cells or memory B cells (56). Controlling this process to drive antibody responses to increase recognition of subdominant protective epitopes would therefore be of primary importance. A possible strategy, referred to as germline targeting, relies upon the activation and expansion of rare but specific B cell lineages in naïve individuals (57, 58). For patients that are no longer naïve due to natural exposure to the pathogen, however, the same outcome would need to be achieved by manipulating established B cell immunodominance hierarchies and remodel antibody responses toward more desired targets (57, 58).

Protective B-cell epitopes (contrary to T-cell epitopes) are often conformational, and even when a significant portion of an epitope appears to be a short linear peptide, as in our study, it does not necessarily mean that the peptide represents the full epitope or, if it does, that the sequence will not require a certain conformation to elicit optimal bioactivity. For this reason, in the immunization studies performed in this study, we used CD-confirmed refolded recombinant antigens. The immunization-induced antibodies generally identified the same epitopes seen in infection, supporting the role of refolding in mimicking native structure, but immunization also resulted in recognition of a broader range of epitopes than seen during infection, including transmembrane scaffolding regions. This is likely because the scaffold regions are not shielded by the outer membrane in an immunization setting and are thus more easily processed for recognition. Thus, the design of vaccine immunogens is critical. Possible approaches vary from placing epitopes within chimeric antigens that could work as scaffold or, alternatively, using portions of the protein containing protective epitopes as structural elements of the antigen, or even using single β-hairpins instead of the full-length antigens. The work reported here represents an important step in evaluating TprC/D and $TprD_2$ epitopes as part of the process that will lead to an effective vaccine for syphilis.

MATERIALS AND METHODS

Strain Propagation and Experimental Infection

Outbred adult male New Zealand White rabbits ranging from 3.0-4.0 Kg were obtained from R&R Rabbitry (Stanwood, WA). Prior to entry into the study, serum from each animal was tested with both a treponemal (FTA-ABS) and a non-treponemal (VDRL; BD, Franklin Lakes, NJ) test to rule out infection with the rabbit syphilis agent *Treponema paraluiscuniculi*. Only rabbits seronegative in both tests were used for either propagation or experimental infection for sample collection. *Tp s*trains were propagated by intratesticular (IT) inoculation and harvested at peak orchitis as previously described (60). For experimental infections, groups of three rabbits were infected IT with a total of $5x10^7$ *Tp* cells per testis. In total, nine *Tp* isolates (one isolate per rabbit group) were used: seven *Tp* subsp. *pallidum* isolates (Nichols, Chicago, Bal73-1, Sea81-4, Bal3, MexicoA, and UW249), one *Tp* subsp. *endemicum* (IraqB) and one *Tp* subsp. *pertenue* (SamoaD) (**Table S9**). Briefly, on the day of infection bacteria were extracted from rabbit testes in sterile saline containing 10% normal rabbit serum (NRS), and testicular extract was collected in sterile 15-ml tubes. Extracts were centrifuged twice at 1,000 rpm (180 x g) for 10 minutes in an Eppendorf 5810R centrifuge (Eppendorf, Hauppauge, NY) to remove gross rabbit cellular debris. Treponemes were enumerated under a dark-field microscope (DFM) and percentage of motile organisms was recorded. Extracts were then diluted in serum-saline to the desired concentration ($5x10^7$/ml). Following IT injection, treponemal motility was assessed again to ensure that the time elapsed before injection into the new host did not affect pathogen viability. After IT inoculation, establishment of infection was assessed by monitoring development of orchitis during the following three weeks as well as by performing FTA-ABS and VDRL tests on sera collected at day 30 post-inoculation. Immune sera were collected from the animals at day 30, 60, and 90 post-infection. Animals were then euthanized. Extracted sera were heat-inactivated at 56°C for 30 min and stored at -20°C until use for ELISAs.

Amplification and Cloning of Full-Length *tprC* Gene Variants for Expression of Recombinant Antigens

Sequences for the *tprC* gene of *Tp* isolates (Nichols, Sea81-4, and MexicoA) were previously cloned (35). For expression, the *tprC* sequences were sub-cloned into the pET23b+ vector (Life Technologies) between BamHI and HindIII using the primers C-S (5'-cgggatccgatgg gcgtactcactccgca) and C-As (5'-gcaagcttccat gtcactttcattccac). For sub-cloning, the *tprC* ORF was amplified in a 100-µl final volume using 0.4 units of GoTaq polymerase (Promega) with approximately 10 ng of DNA template. $MgCl_2$ and dNTP final concentrations were 1.5 mM and 200 µM, respectively. Initial denaturation and final extension (72°C) were for 10 min each. Denaturation (94°C), annealing (60°C), and extension (72°C) were carried out for 1 min each for a total of 35 cycles. Amplicons were purified, digested, and ligated into the pET23b+ vector. As a result of cloning into pET23b+, 28 additional amino acids were added to the TprC ORFs (14 NH_2-terminal and

14 COOH-terminal amino acids), including the COOH-terminal 6×His tag for affinity purification. Ligation products were used to transform OneShot TOP10 chemically competent *E. coli* cells (Life Technologies) according to the provided protocol. Transformations were plated on LB-Ampicillin (100 μg/ml) agar plates for selection. For each cloning reaction, individual colonies were screened for the presence of insert-containing plasmids using primers annealing upstream and downstream of the pET23b+ vector poly-linker (T7 promoter and terminator primers). Positive plasmids were extracted from overnight liquid cultures obtained from replica colonies by using the Plasmid Mini kit (Qiagen, Germantown, MD), and two to five clones for each strain were sequenced to ensure sequence fidelity to the previously cloned templates (35). For expression of recombinant antigens, a suitable clone for each *tprC* gene variant was used to transform *E. coli* Rosetta (DE3) competent cells (Life Technologies).

Expression, Purification and Refolding of Recombinant Proteins

E. coli cells were grown overnight in LB media supplemented with ampicillin (100 μg/ml). The following day, multiple flasks containing 200 ml of auto-inducing media (61), were inoculated with 20 ml of overnight culture in a 2-liter baffled flask and grown at room temperature for 72 h at 175 rpm in a shaking incubator. Expression of recombinant antigens in induced and un-induced controls was assessed by immunoblot using a monoclonal anti-poly-histidine antibody (Millipore-Sigma, diluted 1:2000) after SDS-PAGE. Prior to purification, presence of the recombinant protein in the soluble and insoluble cellular fractions was evaluated by SDS-PAGE and immunoblot. Recombinant TprC purification was carried on under denaturing conditions. Briefly, *E. coli* cell pellets were resuspended in 5 ml/g of dry culture weight of binding buffer (50 mM NaH_2PO_4, 10 mM imidazole, pH 8.0) w/o denaturing agent, and the suspension was sonicated in ice with 100 pulses of 6 s each, with each pulse being separated by 10-s intervals. Insoluble components (containing the desired products) were precipitated by centrifugation and resuspended in 5 ml/g of culture weight of binding buffer (50 mM NaH_2PO_4, 5 mM imidazole, pH 8.0) containing 6M Guanidine-HCl denaturing agent and sonicated again as above. Insoluble components were precipitated again by centrifugation and the supernates were saved. For affinity chromatography, 5.0 ml of nickel-agarose (Ni-NTA agarose, Qiagen) was packaged into a 1.5x14 cm column (Bio-Rad, Carlsbad, CA) and washed with 3 column volumes of molecular-grade H_2O and 6 column volumes of binding buffer + denaturing agent. Cell lysate was then loaded, and the flow was adjusted to 1 ml/min. Unbound proteins were washed using 10 bed volumes of binding buffer, followed by 6 column volumes of wash buffer (50 mM NaH_2PO_4, 20 mM imidazole, pH 8.0) containing denaturing agent. Washing continued until the A_{280} of the flow through was <0.01 AU. Recombinant TprC was eluted with 15 ml of elution buffer (50 mM NaH_2PO_4, 300 mM imidazole, pH 8.0) containing denaturing agent. Eluted fractions devoid of visible contaminants by SDS-PAGE and Coomassie staining were pooled, and protein concentration was assessed by micro-bicinchoninic (BCA) assay (Thermo-Fisher). Pooled fractions were then dialyzed in PBS using a 10 kDa MWCO Slide-A-Lyzer dialysis cassette (Thermo-Fisher) over 12 hours, ensuring PBS change every ~4 hours. Precipitated protein, resulting from elimination of Guanidine-HCl during dialysis was transferred into microcentrifuge tubes and spun down at full speed. After removing the supernate, the pellet was resuspended in a volume of PBS containing 6M urea suitable to achieve a protein concentration of ~4 mg/ml, and protein concentration was then reassessed using the micro-BCA assay kit (Thermo-Fisher). Prior to immunizations, urea was eliminated using Profoldin (Hudson, MA) M7 renaturing columns for membrane proteins, which were used according to the manufacturer's protocol. M7 renaturing columns were found to provide the best yield when screened along with 19 other conditions offered by Profoldin. Lipid composition of the elute buffer included lysophosphatidylcholine (~5 mM), arginine (~150 mM), glycerol (~10%), dodecyl maltoside (0.7 mM), and Tris-HCl (0.1 mM), pH 7.5). Following buffer exchange, soluble protein concentration was evaluated using micro-BCA assay and analyzed by circular dichroism (CD) to evaluate percentage of β-sheet, alpha-helix, and random coil. CD spectra (190 to 260 nm) were acquired in triplicate at room temperature using 0.5 mg/ml of recombinant refolded TprC in a Jasco-1500 high-performance CD spectrometer. CD spectra were analyzed using the online platform Dichroweb (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml) (62) and the spectra from buffer alone for background subtraction.

Rabbit Immunization

Groups of three rabbits each were immunized with one of the purified, refolded recombinant TprC variants. Rabbits were injected with 125 μg of refolded protein every 3 weeks for a total of three immunizations. Prior to injection, antigen was mixed with an equal volume of in Titermax Gold Adjuvant (Millipore-Sigma), a water-in oil emulsion containing squalene, the block co-polymer CRL-8300, and a microparticle stabilizers to obtain a final volume of 1 ml. Immunogen-adjuvant preparation was performed according to the manufacturer`s instruction, and immunizations were performed *via* four 250 μl injections (each containing 31.25 μg of protein) into 4 intramuscular sites. Three weeks after the last boost, immunized animals were deeply anaesthetized, bled through cardiac puncture, and then euthanized.

ELISA Using Synthetic Peptides

Overlapping synthetic peptides (20-mers overlapping by 10 aa) were designed to represent the sequences of all TprC and TprD/D_2 loci present in each of the seven strains examined in this study starting after the predicted signal peptide (AA 1-22; **Figure 1** and **Figure 2**). Only the C56 peptide and its variants (**Table S6**), which represent the proteins' COOH-terminus, were synthesized as 26-mers. A total of 120 peptides (**Table S6**) were produced by Genscript (Piscataway, NJ). Upon receipt, lyophilized peptides were rehydrated in sterile PBS to a stock solution of 200 μg/ml. Solubility of hydrophobic peptides was increased by adding up to 4% (v/v) DMSO per manufacturer's instruction when needed (peptides C1, C4-7, C10, C15-16, C20, C25, C38-C39, C43-44, C53; **Table S6**). Reconstituted peptides were stored at -20°C until use. For ELISA, peptides were further diluted to 10 μg/ml in PBS, and 50 μl of working dilution (500 ng total) were used to coat the

wells of a 96-well Microwell Maxisorp flat-bottom plate (Thermo-Fisher, Waltham, MA) as previously described (44). Absorbance was measured at OD_{405} using a Molecular Devices SpectraMax Plus microplate reader (Molecular Devices, San Jose, CA). A micro-BCA protein assay (Thermo Fisher) was performed in plates coated with Ag and washed to demonstrate that all peptides bound to the well surfaces in the plates (data not shown). For each serum from each group, the value of each replicate experimental wells minus background reactivity (i.e., three times the mean of the wells tested with pooled uninfected rabbit serum) was calculated and plotted. If residual value for the No-antigen control wells was present after subtraction, statistical significance was calculated with one-way ANOVA with the Bonferroni correction of multiple comparisons or t-test, with significance set at $p<0.05$. Except for figures showing cumulative absorbance, graphs represent the mean ± SEM for triplicate wells tested with pooled sera from the 3 rabbits in each group after background subtraction.

TprC/D and D2 Structure Modeling

We used the ColabFold interface (63) to construct Multiple Sequence Alignments (MSA) for the TprC and $TprD_2$ query sequences by searching UniRef30 (64), Mgnify (65) and ColabFold sequence databases with MMSeq2 (66). The MSA was used as input for structure prediction with AlphaFold2 (38) using the default settings (template=False, amber_relax=False, 3 recycles). Visualization was performed using PyMol software (https://pymol.org) (67).

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**. Further inquiries can be directed to the corresponding author.

ETHICS STATEMENT

Animal care was provided in accordance with the procedures described in the Guide for the Care and Use of Laboratory Animals under protocols approved by the University of Washington Institutional Animal Care and Use Committee (IACUC, PI: Sheila Lukehart). The protocol number assigned by the IACUC committee that approved this study is 2090-08.

AUTHOR CONTRIBUTIONS

BM: Performed experiments, analyzed data, and reviewed manuscript. MF: performed experiments, analyzed data, and reviewed manuscript. CG: performed experiments, reviewed manuscript. AV: generated Tpr models, reviewed manuscript. SL: experiment conceptualization, analyzed data, reviewed manuscript. LG: analyzed data and wrote manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

Research reported in this publication was supported by National Institute of Allergy & Infectious Diseases of the National Institutes of Health under award number R01AI042143 grant (to SL). Tpr models using AlphaFold2 were generated thanks to support from Open Philanthropy (to LG). This work was also partially supported also by the National Institute for Allergy and Infectious Diseases of the National Institutes of Health grant number U19AI144133 (Project 2. Project 2 leader: LG; PI: Anna Wald, University of Washington). The funders had no role in study design, data collection and interpretation, or the decision to submit the work for publication.

ACKNOWLEDGMENTS

The authors are grateful to Janelle Deane for aiding with some of the experimental procedures.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fimmu.2022.862491/full#supplementary-material

REFERENCES

1. Giacani L, Lukehart SA. The Endemic Treponematoses. *Clin Microbiol Rev* (2014) 27(1):89–115. doi: 10.1128/CMR.00070-13
2. Rowley J, Vander Hoorn S, Korenromp E, Low N, Unemo M, Abu-Raddad LJ, et al. Chlamydia, Gonorrhoea, Trichomoniasis and Syphilis: Global Prevalence and Incidence Estimates, 2016. *Bull World Health Organ* (2019) 97(8):548–62p. doi: 10.2471/blt.18.228486
3. Savage EJ, Marsh K, Duffell S, Ison CA, Zaman A, Hughes G. Rapid Increase in Gonorrhoea and Syphilis Diagnoses in England in 2011. *Euro Surveill* (2012) 17(29). doi: 10.2807/ese.17.29.20224-en
4. Savage EJ, Hughes G, Ison C, Lowndes CM. Syphilis and Gonorrhoea in Men Who Have Sex With Men: A European Overview. *Euro Surveill* (2009) 14(47). doi: 10.2807/ese.14.47.19417-en
5. Simms I, Fenton KA, Ashton M, Turner KM, Crawley-Boevey EE, Gorton R, et al. The Re-Emergence of Syphilis in the United Kingdom: The New Epidemic Phases. *Sex Transm Dis* (2005) 32(4):220–6. doi: 10.1097/01.olq.0000149848.03733.c1
6. Tucker JD, Cohen MS. China's Syphilis Epidemic: Epidemiology, Proximate Determinants of Spread, and Control Responses. *Curr Opin Infect Dis* (2011) 24(1):50–5. doi: 10.1097/QCO.0b013e32834204bf
7. CDC. *2018 Sexually Transmitted Disease Surveillance*. Atlanta, GA: US Department of Health and Human Services: Centers for Disease Control and Prevention (2019).
8. LaFond RE, Lukehart SA. Biological Basis for Syphilis. *Clin Microbiol Rev* (2006) 19(1):29–49. doi: 10.1128/CMR.19.1.29-49.2006
9. Goldenberg RL, Thompson C. The Infectious Origins of Stillbirth. *Am J Obstet Gynecol* (2003) 189(3):861–73. doi: 10.1067/S0002-9378(03)00470-8

10. Moline HR, Smith JF Jr. The Continuing Threat of Syphilis in Pregnancy. *Curr Opin Obstet Gynecol* (2016) 28(2):101–4. doi: 10.1097/gco.0000000000000258
11. CDC. *The National Plan to Eliminate Syphilis in the United States*. Atlanta, GA: U.S. Department of Health and Human Services: Centers for Disease Control and Prevention (1999).
12. WHO. *The Global Elimination of Congenital Syphilis: Rationale and Strategy for Action*. Geneva: WHO Press (2007).
13. Mitjà O, Marks M, Konan DJ, Ayelo G, Gonzalez-Beiras C, Boua B, et al. Global Epidemiology of Yaws: A Systematic Review. *Lancet Glob Health* (2015) 3(6):e324–31. doi: 10.1016/s2214-109x(15)00011-x
14. Mitjà O, Houinei W, Moses P, Kapa A, Paru R, Hays R, et al. Mass Treatment With Single-Dose Azithromycin for Yaws. *N Engl J Med* (2015) 372(8):703–10. doi: 10.1056/NEJMoa1408586
15. Mitjà O, Godornes C, Houinei W, Kapa A, Paru R, Abel H, et al. Re-Emergence of Yaws After Single Mass Azithromycin Treatment Followed by Targeted Treatment: A Longitudinal Study. *Lancet* (2018) 391(10130):1599–607. doi: 10.1016/s0140-6736(18)30204-6
16. Pace JL, Csonka GW. Late Endemic Syphilis: Case Report of Bejel With Gummatous Laryngitis. *Genitourin Med* (1988) 64(3):202–4. doi: 10.1136/sti.64.3.202
17. Pace JL. Treponematoses in Arabia. *Saudi Med J* (1983) 4:211–20.
18. Julvez J, Michault A, Kerdelhue V. Serologic Studies of Non-Venereal Treponematoses in Infants in Niamey, Niger. *Med Trop (Mars)* (1998) 58 (1):38–40.
19. Galoo E, Schmoor P. Identification of a Focus of Bejel in Mauritania. *Med Trop (Mars)* (1998) 58(3):311–2.
20. Lieberman NAP, Lin MJ, Xie H, Shrestha L, Nguyen T, Huang ML, et al. Treponema Pallidum Genome Sequencing From Six Continents Reveals Variability in Vaccine Candidate Genes and Dominance of Nichols Clade Strains in Madagascar. *PloS Negl Trop Dis* (2021) 15(12):e0010063. doi: 10.1371/journal.pntd.0010063
21. Lithgow KV, Cameron CE. Vaccine Development for Syphilis. *Expert Rev Vaccines* (2017) 16(1):37–44. doi: 10.1080/14760584.2016.1203262
22. Cameron CE, Lukehart SA. Current Status of Syphilis Vaccine Development: Need, Challenges, Prospects. *Vaccine* (2014) 32(14):1602–9. doi: 10.1016/j.vaccine.2013.09.053
23. Turner TB, Hollander DH. *Biology of the Treponematoses*. Geneva: World Health Organization (1957).
24. Miller JN. Immunity in Experimental Syphilis. VI. Successful Vaccination of Rabbits With *Treponema Pallidum*, Nichols Strain, Attenuated by γ-Irradiation. *J Immunol* (1973) 110(5):1206–15.
25. Koebnik R, Locher KP, Van Gelder P. Structure and Function of Bacterial Outer Membrane Proteins: Barrels in a Nutshell. *Mol Microbiol* (2000) 37 (2):239–53. doi: 10.1046/j.1365-2958.2000.01983.x
26. Baker-Zander SA, Lukehart SA. Macrophage-Mediated Killing of Opsonized *Treponema Pallidum*. *J Infect Dis* (1992) 165(1):69–74. doi: 10.1093/infdis/165.1.69
27. Lukehart SA, Miller JN. Demonstration of the *In Vitro* Phagocytosis of *Treponema Pallidum* by Rabbit Peritoneal Macrophages. *J Immunol* (1978) 121(5):2014–24.
28. Edmondson DG, Norris SJ. *In Vitro* Cultivation of the Syphilis Spirochete Treponema Pallidum. *Curr Protoc* (2021) 1(2):e44. doi: 10.1002/cpz1.44
29. Walker EM, Borenstein LA, Blanco DR, Miller JN, Lovett MA. Analysis of Outer Membrane Ultrastructure of Pathogenic *Treponema* and *Borrelia* Species by Freeze-Fracture Electron Microscopy. *J Bacteriol* (1991) 173 (17):5585–8. doi: 10.1128/jb.173.17.5585-5588.1991
30. Radolf JD, Norgard MV, Schulz WW. Outer Membrane Ultrastructure Explains the Limited Antigenicity of Virulent *Treponema Pallidum*. *Proc Natl Acad Sci USA* (1989) 86(6):2051–5. doi: 10.1073/pnas.86.6.2051
31. Centurion-Lara A, Castro C, Barrett L, Cameron C, Mostowfi M, Van Voorhis WC, et al. *Treponema Pallidum* Major Sheath Protein Homologue Tpr K Is a Target of Opsonic Antibody and the Protective Immune Response. *J Exp Med* (1999) 189(4):647–56. doi: 10.1084/jem.189.4.647
32. Cox DL, Luthra A, Dunham-Ems S, Desrosiers DC, Salazar JC, Caimano MJ, et al. Surface Immunolabeling and Consensus Computational Framework to Identify Candidate Rare Outer Membrane Proteins of *Treponema Pallidum*. *Infect Immun* (2010) 78(12):5178–94. doi: 10.1128/IAI.00834-10
33. Fraser CM, Norris SJ, Weinstock GM, White O, Sutton GG, Dodson R, et al. Complete Genome Sequence of *Treponema Pallidum*, the Syphilis Spirochete. *Science* (1998) 281(5375):375–88. doi: 10.1126/science.281.5375.375
34. Anand A, Luthra A, Dunham-Ems S, Caimano MJ, Karanian C, LeDoyt M, et al. TprC/D (Tp0117/131), a Trimeric, Pore-Forming Rare Outer Membrane Protein of *Treponema Pallidum*, Has a Bipartite Domain Structure. *J Bacteriol* (2012) 194(9):2321–33. doi: 10.1128/JB.00101-12
35. Centurion-Lara A, Giacani L, Godornes C, Molini BJ, Brinck Reid T, Lukehart SA. Fine Analysis of Genetic Diversity of the *Tpr* Gene Family Among Treponemal Species, Subspecies and Strains. *PloS Negl Trop Dis* (2013) 16 (7):e2222. doi: 10.1371/journal.pntd.0002222
36. Sun ES, Molini BJ, Barrett LK, Centurion-Lara A, Lukehart SA, Van Voorhis WC. Subfamily I *Treponema Pallidum* Repeat Protein Family: Sequence Variation and Immunity. *Microbes Infect* (2004) 6(8):725–37. doi: 10.1016/j.micinf.2004.04.001
37. Centurion-Lara A, Sun ES, Barrett LK, Castro C, Lukehart SA, Van Voorhis WC. Multiple Alleles of *Treponema Pallidum* Repeat Gene D in *Treponema Pallidum* Isolates. *J Bacteriol* (2000) 182(8):2332–5. doi: 10.1128/JB.182.8.2332-2335.2000
38. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly Accurate Protein Structure Prediction With AlphaFold. *Nature* (2021) 596(7873):583–9. doi: 10.1038/s41586-021-03819-2
39. Holm L. Using Dali for Protein Structure Comparison. *Methods Mol Biol* (2020) 2112:29–42. doi: 10.1007/978-1-0716-0270-6_3
40. Geysen HM, Meloen RH, Barteling SJ. Use of Peptide Synthesis to Probe Viral Antigens for Epitopes to a Resolution of a Single Amino Acid. *Proc Natl Acad Sci USA* (1984) 81(13):3998–4002. doi: 10.1073/pnas.81.13.3998
41. Haynes AM, 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(8). doi: 10.1128/iai.00323-19
42. Giacani L, Lukehart S, Centurion Lara A. *Syphilis*. A,LS Barrett, editor. Cambridge, MA: Academic Press (2009).
43. Morgan CA, Lukehart SA, Van Voorhis WC. Protection Against Syphilis Correlates With Specificity of Antibodies to the Variable Regions of *Treponema Pallidum* Repeat Protein K. *Infect Immun* (2003) 71(10):5605–12. doi: 10.1128/IAI.71.10.5605-5612.2003
44. Morgan CA, Molini BJ, Lukehart SA, Van Voorhis WC. Segregation of B and T Cell Epitopes of *Treponema Pallidum* Repeat Protein K to Variable and Conserved Regions During Experimental Syphilis Infection. *J Immunol* (2002) 169(2):952–7. doi: 10.4049/jimmunol.169.2.952
45. Morgan CA, Lukehart SA, Van Voorhis WC. Immunization With the N-Terminal Portion of *Treponema Pallidum* Repeat Protein K Attenuates Syphilitic Lesion Development in the Rabbit Model. *Infect Immun* (2002) 70(12):6811–6. doi: 10.1128/IAI.70.12.6811-6816.2002
46. Cameron CE, Lukehart SA, Castro C, Molini B, Godornes C, Van Voorhis WC. Opsonic Potential, Protective Capacity, and Sequence Conservation of the *Treponema Pallidum* Subspecies *Pallidum* Tp92. *J Infect Dis* (2000) 181 (4):1401–13. doi: 10.1086/315399
47. Arroll TW, Centurion-Lara A, Lukehart SA, Van Voorhis WC. T-Cell Responses to *Treponema Pallidum* Subsp. *Pallidum* Antigens During the Course of Experimental Syphilis Infection. *Infect Immun* (1999) 67(9):4757–63. doi: 10.1128/IAI.67.9.4757-4763.1999
48. Cameron CE, Castro C, Lukehart SA, Van Voorhis WC. Function and Protective Capacity of *Treponema Pallidum* Subsp. *Pallidum* Glycerophosphodiester Phosphodiesterase. *Infect Immun* (1998) 66 (12):5763–70. doi: 10.1128/IAI.66.12.5763-5770.1998
49. LaFond RE, Molini BJ, Van Voorhis WC, Lukehart SA. Antigenic Variation of TprK V Regions Abrogates Specific Antibody Binding in Syphilis. *Infect Immun* (2006) 74(11):6244–51. doi: 10.1128/IAI.00827-06
50. Cretich M, Gori A, D'Annessa I, Chiari M, Colombo G. Peptides for Infectious Diseases: From Probe Design to Diagnostic Microarrays. *Antibodies* (2019) 8 (1):23. doi: 10.3390/antib8010023
51. Roy A, Kucukural A, Zhang Y. I-TASSER: A Unified Platform for Automated Protein Structure and Function Prediction. *Nat Protoc* (2010) 5(4):725–38. doi: 10.1038/nprot.2010.5
52. Hegedűs T, Geisler M, Lukács G, Farkas B. AlphaFold2 Transmembrane Protein Structure Prediction Shines. *bioRxiv* (2021). doi: 10.1101/2021.08.21.457196

53. Anand A, LeDoyt M, Karanian C, Luthra A, Koszelak-Rosenblum M, Malkowski MG, et al. Bipartite Topology of Treponema Pallidum Repeat Proteins C/D and I: OUTER MEMBRANE INSERTION, TRIMERIZATION, AND PORIN FUNCTION REQUIRE A C-TERMINAL β-BARREL DOMAIN. *J Biol Chem* (2015) 290(19):12313–31. doi: 10.1074/jbc.M114.629188
54. Leader BT, Hevner K, Molini BJ, Barrett LK, Van Voorhis WC, Lukehart SA. Antibody Responses Elicited Against the *Treponema Pallidum* Repeat Proteins Differ During Infection With Different Isolates of *Treponema Pallidum* Subsp. *Pallidum*. *Infect Immun* (2003) 71(10):6054–7. doi: 10.1128/IAI.71.10.6054-6057.2003
55. Giacani L, Molini B, Godornes C, Barrett L, Van Voorhis WC, Centurion-Lara A, et al. Quantitative Analysis of *Tpr* Gene Expression in *Treponema Pallidum* Isolates: Differences Among Isolates and Correlation With T-Cell Responsiveness in Experimental Syphilis. *Infect Immun* (2007) 75(1):104–12. doi: 10.1128/IAI.01124-06
56. Victora GD, Nussenzweig MC. Germinal Centers. *Annu Rev Immunol* (2012) 30(1):429–57. doi: 10.1146/annurev-immunol-020711-075032
57. Briney B, Sok D, Jardine JG, Kulp DW, Skog P, Menis S, et al. Tailored Immunogens Direct Affinity Maturation Toward HIV Neutralizing Antibodies. *Cell* (2016) 166(6):1459–70.e11. doi: 10.1016/j.cell.2016.08.005
58. Jardine JG, Ota T, Sok D, Pauthner M, Kulp DW, Kalyuzhniy O, et al. Priming a Broadly Neutralizing Antibody Response to HIV-1 Using a Germline-Targeting Immunogen. *Science* (2015) 349(6244):156–61. doi: 10.1126/science.aac5894
59. *National Research Council (US) Committee for the Update of the Guide for the Care and Use of Laboratory Animals*. Washington DC: The National Academic Press (2011).
60. Lukehart SA, Marra CM. Isolation and Laboratory Maintenance of Treponema Pallidum. *Curr Protoc Microbiol* (2007) 7:12A.1.1–A.1.8. doi: 10.1002/9780471729259.mc12a01s7
61. Studier FW. Protein Production by Auto-Induction in High Density Shaking Cultures. *Protein Expr Purif* (2005) 41(1):207–34. doi: 10.1016/j.pep.2005.01.016
62. Miles AJ, Ramalli SG, Wallace BA. DichroWeb, a Website for Calculating Protein Secondary Structure From Circular Dichroism Spectroscopic Data. *Protein Sci* (2021) 31(1):37–46. doi: 10.1002/pro.4153
63. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold - Making Protein Folding Accessible to All. *bioRxiv* (2021). doi: 10.1101/2021.08.15.456425
64. Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH. UniRef Clusters: A Comprehensive and Scalable Alternative for Improving Sequence Similarity Searches. *Bioinformatics* (2015) 31(6):926–32. doi: 10.1093/bioinformatics/btu739
65. Mitchell AL, Almeida A, Beracochea M, Boland M, Burgin J, Cochrane G, et al. MGnify: The Microbiome Analysis Resource in 2020. *Nucleic Acids Res* (2020) 48(D1):D570–d8. doi: 10.1093/nar/gkz1035
66. Steinegger M, Söding J. MMseqs2 Enables Sensitive Protein Sequence Searching for the Analysis of Massive Data Sets. *Nat Biotechnol* (2017) 35(11):1026–8. doi: 10.1038/nbt.3988
67. Janson G, Zhang C, Prado MG, Paiardini A. PyMod 2.0: Improvements in Protein Sequence-Structure Analysis and Homology Modeling Within PyMOL. *Bioinformatics* (2017) 33(3):444–6. doi: 10.1093/bioinformatics/btw638

Author Disclaimer: The content is solely the responsibility of the authors and does not necessarily represent the official views of the Funders.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.