984

Vaccine Optimization by Identification, Characterization, and Downselection of Human T Cell Epitopes from *Plasmodium falciparum* Circumsporozoite Protein.

Amy R. Noe[1], Frances E. Terry[2], Brian C. Schanen[3], Emily Sassano[3], Pooja Hindocha[2], Leonard Moise[2], Jayne M. Christen[1], Kenneth Tucker[1], Timothy W. Phares[4], William D. Martin[2], Anne S. De Groot[2], Donald R. Drake III[3], Lorraine Soisson[5], Carter Diggs[5], Robin Miller[5], Susan Youll[5], David R. Milich[6], and David C. Whitacre[6]

[1]Leidos Life Sciences, Leidos Inc., Frederick, MD, USA, [2]EpiVax Inc., Providence, RI, USA, [3]Sanofi Pasteur, VaxDesign Campus, Orlando, FL, USA, [4]Explorations in Global Health, Leidos Inc., Frederick, MD, USA, [5]Malaria Vaccine Development Program, United States Agency for International Development (USAID), Washington D.C., USA, [6]VLP Biotech, Inc., San Diego, CA, USA
Corresponding Author: Amy.R.Noe@leidos.com

ABSTRACT

An effective malaria vaccine must prevent disease in a wide range of populations living in regions with vastly different transmission rates and protect against genetically diverse *Plasmodium falciparum* strains. Helper CD4+ T cells, in particular those of the Th1 lineage, are central to development of protective immune responses to *P. falciparum*, playing roles in B cell activation and maturation processes as well as in cytokine production. Therefore, we took advantage of recent in silico modeling advances to predict and analyze the HLA-restricted CD4 (class II) epitopes (peptide sequences that bind HLA-DR) relevant to achieving broad human population coverage in a Pf circumsporozoite protein (CSP) vaccine, utilizing cognate T cell help. Based on in silico analysis, several class II epitope clusters are predicted in the N- and C-terminal regions of PfCSP. Further, one predicted class II cluster overlaps with the highly-variable region 2 (R2) in the C-terminus of PfCSP (3D7 strain). To identify additional predicted class II epitopes, 478 PfCSP R2 peptide sequence variants were analyzed for HLA-DR binding. Of these, nine predicted R2 variant class II clusters were identified through the silico analysis. The resulting peptide sequences were synthesized and assessed for HLA-DR binding in vitro as well as for the ability to prime/activate lymphocytes resulting in generation of cytokine responses using human PBMCs. The latter was conducted using MIMIC® technology (Modular IMmune In vitro Construct), a surrogate human immune system enabling researchers to test human immune responses to antigen. We found sufficient differential cellular activation and cytokine profiles among HLA-DR-matched PBMC donors to downselect class II epitope clusters for inclusion in a vaccine targeting PfCSP.

RESULTS

In Silico T Cell Epitope Analysis

In order to identify T cell epitopes capable of binding across a broad set of human leukocyte antigen (HLA)-DR alleles, *P. falciparum* (Pf) Circumsporozoite protein (CSP) sequences were parsed into overlapping 9-mer frames and each frame was evaluated for predicted binding affinity to a panel of nine class II HLA alleles using in silico analysis tools developed by EpiVax, Inc. These tools evaluate binding to the most common HLA molecules within each of the HLA supertypes (Lind 2004 and Reche 2005) to provide results representative of >95% of human populations worldwide without the need to test each individual haplotype (Southwood 1998 and Sette 1999). Data outputs from the EpiMatrix analysis are Z scores representing the probability that each specific 9-mer sequence will bind one or more of the nine class II HLA alleles. Significant Z scores are those in the top 5% and highly significant Z scores are those in the top 1%. Sequences with significant Z scores are predicted epitopes.

An in silico EpiMatrix analysis of PfCSP 3D7 was performed to determine the predicted class II T cell epitopes within this sequence. A composite representation of the predicted epitopes for PfCSP 3D7, across all nine class II HLA alleles, is shown in **Figure 1**. Within PfCSP 3D7, a total of four class II epitope clusters containing sequences predicted to bind a majority of the nine class II HLA alleles were identified (C1-C4). A pseudo-cluster overlapping the R1 domain of CSP was also identified (C'). Note that, as CSP is already on the surface of sporozoites within the mosquito salivary glands, predicted epitopes within the signal sequence were not further analyzed.

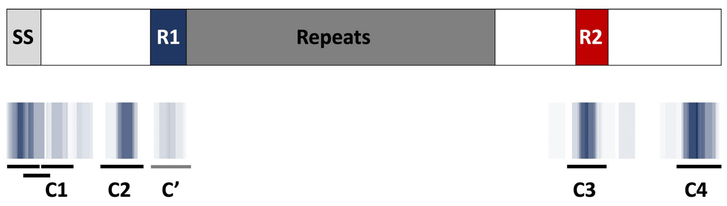

Figure 1. Predicted Class II T Cell Epitope Coverage for PfCSP 3D7. Areas of darker blue indicate higher numbers of predicted epitopes while white areas indicate a lack of predicted epitopes. C1-C4 represent the four identified epitope clusters, non-inclusive of predicted epitopes in the signal sequence (SS). C' represents a pseudo-cluster containing only a few predicted epitopes.

Further, as C3 overlaps the variable R2 domain of PfCSP, we also evaluated this region in PfCSP strains/isolates other than 3D7 to identify additional predicted class II epitopes. A total of 478 PfCSP isolates were analyzed in silico for HLA-DR binding and nine predicted R2 variant class II clusters were identified (ICS2-ICS10). These R2 sequence variants are termed Immunogenic Consensus Sequences (ICS). The number of significant Z scores for clusters 1-4, the pseudo-cluster (C'), and ICS2-ICS10 is shown in **Figure 2**. As the overall intent of this work is to identify epitopes with broad HLA allele coverage across PfCSP strains and given the relatively limited HLA allele coverage predicted for C' as well as the lack of highly significant Z scores, no further evaluations were performed for the pseudo-cluster.

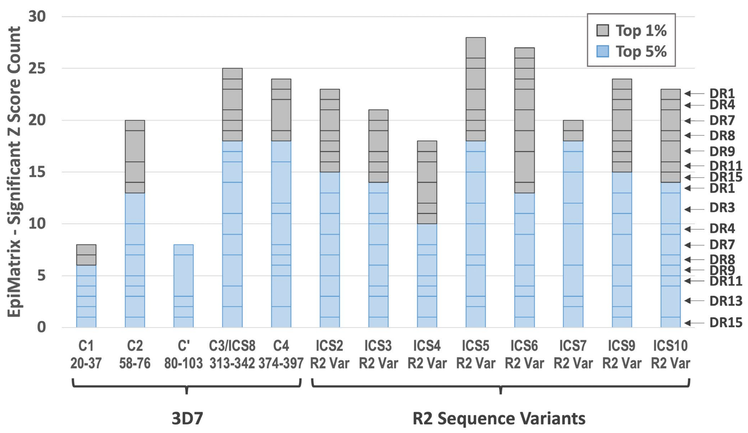

Figure 2. Significant Z Score Count for the Predicted PfCSP Class II T Cell Epitope Clusters. Highly significant Z scores in the top 1% are noted in grey, while significant Z scores in the top 5% are noted in blue. The segments represent the Z score count for each HLA allele, where binding is predicted. As an example, the predicted HLA coverage is shown for ICS10 where each segment is labeled by HLA-DR allele (on the right). The amino acid sequence address is shown at the bottom for each 3D7 cluster and the ICS are noted as R2 sequence variants. The sequences of C3 and ICS8 are identical.

In Vitro HLA-Peptide Binding Analysis

A subset of the predicted T cell epitope clusters/ICS (synthesized as peptides) were evaluated for HLA binding in vitro using a competition assay whereby test peptides compete for HLA binding with a fluorescently-labeled positive control peptide (Steere 2006). For each test peptide, the concentration that inhibits 50% of the specific HLA binding by the control peptide is calculated as IC_{50}. In turn, the IC_{50} values are used to gauge affinity of the test peptide (with lower IC_{50} values indicating greater affinity) and, when plotted against the significant Z score count for each epitope cluster by HLA, demonstrate that increased HLA affinity trends with an increase in significant Z score count by HLA allele **(Figure 3)**. Given the relatively low significant Z score counts for C1 and overall lack of high-affinity HLA binding with this sequence, the C1 cluster was dropped from further studies.

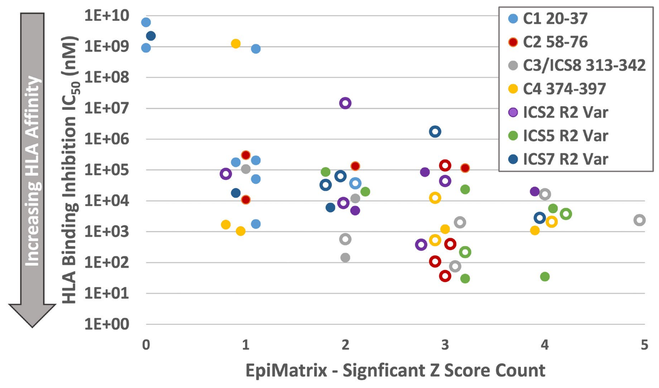

Figure 3. HLA Binding Affinity Trends with EpiMatrix Significant Z Score Count. In vitro HLA binding affinities to DR1, DR3, DR4, DR8, DR11, DR13, and DR15 were assessed for predicted T cell epitope clusters C1-C4, ICS2, ICS5, and ICS7, and plotted against the significant Z score count for each HLA. Clusters with higher Z score counts for a specific HLA allele and/or including at least one highly significant Z score trended toward a higher affinity for the relevant HLA allele. Closed circles represent data points with significant Z score counts for an HLA allele (i.e., all counts in the top 5%) and open circles represent data points with at least one highly significant Z score for an HLA allele (i.e., at least one count in the top 1%).

Ex Vivo Human PBMC Immunogenicity Assessments

The MIMIC platform simulates immune responses from a diverse human population using the circulating immune cells of individual donors to recapitulate each individual's human immune response (Dauner 2017). Within the system, autonomy of the donor is maintained resulting in an ex vivo test system that is functionally equivalent to the donor's own immune system and designed to respond in a similar manner as that seen in vivo. The MIMIC lymphoid tissue equivalent module (LTE) simulates adaptive immune responses through dendritic cell (DC) priming and production of activated T cells and cytokines, to mirror cellular immune response within a human lymph node. Ability of the predicted HLA class II T cell epitope clusters C2-C4 and ICS2-ICS10 (synthesized as peptides) to elicit cytokine responses as part of recall response assessment in primed DC/T cell co-cultures was evaluated. In **Figure 4**, the percentage of donors with IFN-γ recall responses to each T cell epitope cluster (C2-C4 and ICS2-ICS10) is shown for each HLA allele. The highest IFN-γ responses across a broad set of HLA alleles were seen with T cell epitope clusters overlapping the R2 domain (i.e., C3/ICS8 and several other ICS).

		HLA Allele								
	Cluster ID	DR1	DR3	DR4	DR7	DR8	DR9	DR11	DR13	DR15
3D7	C2	0%	17%	0%	0%	0%	0%	0%	0%	14%
	C3/ICS8	20%	33%	50%	50%	75%	50%	25%	0%	29%
	C4	0%	0%	0%	0%	25%	0%	0%	14%	0%
R2 Sequence Variants	ICS2	40%	33%	40%	50%	50%	0%	25%	57%	43%
	ICS3	80%	33%	30%	75%	0%	0%	25%	0%	14%
	ICS4	0%	0%	30%	0%	0%	0%	0%	0%	29%
	ICS5	80%	33%	60%	75%	25%	50%	100%	43%	43%
	ICS6	20%	50%	40%	50%	25%	50%	50%	43%	43%
	ICS7	60%	67%	70%	75%	75%	50%	25%	86%	43%
	ICS9	40%	50%	70%	75%	50%	50%	75%	43%	57%
	ICS10	20%	33%	40%	75%	0%	25%	50%	29%	57%
	N	5	6	10	4	4	4	4	7	7

Figure 4. Percentage of Donors by HLA Type Demonstrating IFN-γ Responses to Predicted Epitope Clusters/ICS. A Stimulation Index greater than or equal to 2-fold above baseline is used to specify a positive donor response. The percentage of responding donors, by HLA allele for each 3D7 cluster and ICS, is shown numerically along with color-coding to indicate higher percentages of responders with darker blues. The number of HLA-matching donors (N) for each allele is shown in the bottom row. A total of 30 donors were included in this study.

Multifunctional T cell profiling is the primary readout of immunogenicity assessments performed with the MIMIC LTE module. For each T cell epitope cluster (C2-C4 and ICS2-ICS10), the number and magnitude of CD4+ cytokine-producing T cells was evaluated for a panel of five cytokines. This panel included Th1 response markers (IFNγ, IL-2, and TNFα), a Th2 response marker (IL-4), and a marker for T regulatory response (IL-10). Multifunctional response profiling of the generated T cell sets was performed by multilayer Boolean data analysis and the data are represented as circular pie charts showing the number of functions (i.e., the number of cytokines secreted) in grey and type of function (i.e., cytokine secreted) in color **(Figure 5)**. Additionally, the overall magnitude of response is represented for each epitope cluster by size of the circular chart.

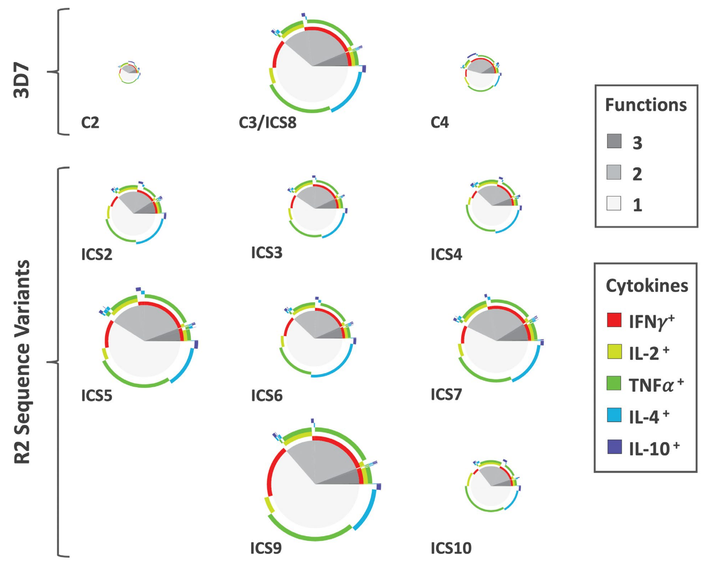

Figure 5. Multifunctional Cytokine-Producing CD4+ T Cell Response Profiles. Circular pie charts represent response across all 30 donors. The proportion of single-, double-, and triple-function T cells is shown in grey tones. The colored arcs highlight the total proportions of T cells secreting a particular cytokine or combination of cytokines. Relative magnitude of the overall median response (for all 30 donors to an epitope cluster) can be gauged by size of the pie chart, with greater magnitude response represented by larger pie charts.

Predicted Pf Strain Coverage for R2 Sequence Variants

An epitope sequence conservation analysis was performed to evaluate the number and geographical region of the 478 PfCSP isolates containing sequences matching the predicted epitopes within each of the ICS. The results are represented as a heat map showing the number of epitope matches (darker colors indicate a great number of matches) for each PfCSP isolate **(Figure 6)**.

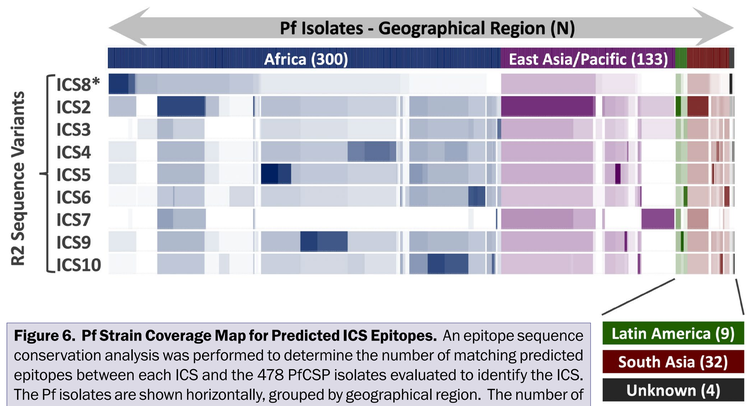

Figure 6. Pf Strain Coverage Map for Predicted ICS Epitopes. An epitope sequence conservation analysis was performed to determine the number of matching predicted epitopes between each ICS and the 478 PfCSP isolates evaluated to identify the ICS. The Pf isolates are shown horizontally, grouped by geographical region. The number of isolates for each geographical region (N) is indicated. Areas of darker colors indicate higher numbers of matching predicted epitopes while white areas indicate a lack of matching predicted epitopes. *ICS8 is the 3D7 sequence and a match for C3.

DISCUSSION AND ONGOING EFFORTS

Several aspects of this analysis strongly suggest that CSP is under immune pressure with regard to generation of cognate T cell help (CD4+) populations. Overall, based on size of this protein, the number of predicted class II T cell epitopes is fairly low. PfCSP 3D7 has an overall class II epitope EpiMatrix score of -20; however, the minimum threshold whereby a protein is considered to have strong immune potential is an overall EpiMatrix score of +20. Further, the most highly conserved class II T cell epitope clusters within PfCSP (C2 and C4) demonstrated the most limited ex vivo responses across the broad set of HLA alleles represented among the 30 donors, while the most variable epitope cluster (C3) and R2 region variants demonstrated the broadest responses across the 30 donors. In the human host, elicitation of poor/limited T cell responses to conserved regions of PfCSP is of advantage to the parasite.

Based on the ex vivo studies and predicted strain coverage, five ICS were downselected for CSP vaccine development (i.e., ICS2, ICS5, ICS7, ICS8, and ICS9) and incorporation into our leading CSP virus like particle (VLP) delivery platform construct: woodchuck Hepatitis core antigen VLP containing a surface-exposed B cell epitope (based on the CSP repeat region) and a large portion of the PfCSP 3D7 C-terminal region as inserts (Whitacre 2015).

REFERENCES

Dauner A, et al. The in vitro MIMIC(R) platform reflects age-associated changes in immunological responses after influenza vaccination. Vaccine. 2017;35:5487-94

Lindner SE, et al. Total and putative surface proteomics of malaria parasite salivary gland sporozoites. Mol Cell Proteomics. 2013 May;12(5):1127-43.

Lund O, et al. Definition of supertypes for HLA molecules using clustering of specificity matrices. Immunogenetics. 2004; 55(12):797–810.

Moise L, et al. iVAX: An integrated toolkit for the selection and optimization of antigens and the design of epitope-driven vaccines. Hum Vaccin Immunother. 2015;11(9):2312-21.

Reche PA and Reinherz EL. PEPVAC: a web server for multi-epitope vaccine development based on the prediction of supertypic MHC ligands. Nucleic Acids Res. 2005 Jul 1;33:W138-42.

Sette A and Sidney J. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. Immunogenetics. 1999 Nov;50(3-4):201-12.

Southwood S, et al. Several common HLA-DR types share largely overlapping peptide binding repertoires. J Immunol. 1998; 160(7):3363–73.

Steere AC, et al. Antibiotic-refractory Lyme arthritis is associated with HLA-DR molecules that bind a Borrelia burgdorferi peptide. J Exp Med. 2006; 203(4):961-71.

Whitacre DC, et al. *P. falc iparum* and P. vivax Epitope-Focused VLPs Elicit Sterile Immunity to Blood Stage Infections. PLoS One. 2015 May 1;10(5):e0124856.

ACKNOWLEDGEMENTS

This work was made possible through support provided by the Office of Infectious Diseases, Bureau for Global Health, U.S. Agency for International Development, under the terms of the Malaria Vaccine Development Program (MVDP) Contract AID-OAA-C-15-00071, for which Leidos, Inc. is the prime contractor. The opinions expressed herein are those of the authors and do not necessarily reflect the views of the U.S. Agency for International Development.