RESEARCH ARTICLE

WILEY

Selection-based design of in silico dengue epitope ensemble vaccines

David Murphy[1] | Pedro Reche[2] | Darren R. Flower[1]

[1]School of Life and Health Sciences, Aston University, Birmingham, UK

[2]Immunomedicine Group, Departamento de Inmunología, Facultad de Medicina, Universidad Complutense de Madrid, Madrid, Spain

Correspondence
Darren R. Flower, School of Life and Health Sciences, Aston University, Aston Triangle, Birmingham, B4 7ET, UK.
Email: d.r.flower@aston.ac.uk

Funding information
Aston University

Dengue virus affects approximately 130 countries. Twenty-five percentage of infections result in febrile, self-limiting illness; heterotypic infection results in potentially fatal dengue haemorrhagic fever or dengue shock syndrome. Only one vaccine is currently available. Its efficacy is very variable. Thus, to target dengue, we used an innovative immunoinformatics protocol to design a putative epitope ensemble vaccine by selecting an optimal set of highly conserved epitopes with experimentally verified immunogenicity. From 1597 CD4+ and MHC II epitopes, six MHC Class I epitopes (RAVHADMGYW, GPWHLGKLEM, GLYGNGVVTK, NMIIMDEAHF, KTWAYHGSY and WAYHGSYEV) and nine MHC Class II epitopes (LAKAIFKLTYQNKVV, GKIVGLYGNGVVTTS, AAIFMTATPPGSVEA, AAIFMTATPPGTADA, GKTVWFVPSIKAGND, KFWNTTIAVSMANIF, RAIWYMWLGARYLEF, VGTYGLNTFTNMEVQ and WTLMYFHRRDLRLAA) were selected; this candidate vaccine achieved a world population coverage of 92.49%.

KEYWORDS
dengue, epitope selection, immunoinformatics, vaccine

1 | INTRODUCTION

Drug Discovery and Chemical Biology encompasses many techniques and many perspectives: It is a set of disciplines that has been exploited but could be exploited more. Vaccine development in particular is an area in which Drug Discovery and Chemical Biology is yet to play its part as fully as it should. In the last decade, following the crisis in confidence and failure in performance that gripped the Pharmaceutical Industry, the search for new income streams has seen *inter alia* a rise in biologic drugs, medical devices and vaccines, as potential part-saviours of the Industry. We have recently exemplified an immunoinformatics-based approach to the selection-based design of optimal prevalidated epitope ensemble vaccines by demonstrating the reproducibility of this strategy by proposing a range of putative vaccines against

hepatitis C,[1] influenza[2] and malaria.[3] We focus here on the further exemplification of this methodology by designing potential epitope-based polyvalent putative vaccine candidates against the dengue virus.

Dengue virus (DENV) belongs to the genus *Flavivirus*, having four different serotypes (DENV-1, DENV-2, DENV-3 and DENV-4), with 65% sequence conservation across all serotypes.[4] DENV is transmitted to humans by female *Aedes albopictus* and *Aedes aegypti* peridomestic mosquitoes. *A. aegypti* is the more efficient vector. In the last decade, DENV has spread to areas between 30°N and 40°S of the equator, with cases of infection reported in over 128 countries.[5] Bhatt et al.[6] estimated about 390 million are infected with DENV annually; 96 million cases resulting in illness.

Only 25% of primary dengue infections are symptomatic. Serious complications are rare, due to the production

of neutralizing antibodies (nAb), which prevent viral entry into dendritic cells (DCs), and IFN-γ and TNF-α pro-inflammatory cytokines.[7] Postinfection, memory B cells provide lifelong homotypic immunity. Symptoms are more severe during secondary heterotypic infection, resulting in dengue haemorrhagic fever (DHF).

DHF is characterized by acute capillary leakage, clinically significant thrombocytopenia and varying degrees of liver injury.[8] Due to cross-reactivity and the low threshold needed for antigen stimulation, memory B cells from the primary infection bind epitopes on different DENV serotype, but with lower avidity, reducing the humoral response. This muted immune response, accompanied by IFN-γ, TNF-α, IL-2, IL-4 and IL-6 production, induce a so-called cytokine storm, resulting in widespread systemic inflammation and subsequent plasma leakage.[9] During secondary infection, production of low-affinity antibodies abrogates viral neutralization. Resulting antibody–antigen complexes bind to F_C receptors and are internalized faster than during primary infection.[4]

DENV is a 50-nm virion, spherical in shape, with a 10.7 kb genome comprising single-stranded, positive-sense RNA. Its genome is transcribed as a single polyprotein, which is cleaved into three structural proteins (capsid (C), membrane (M) and envelope (E)) and seven nonstructural proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5). The envelope protein is involved in cellular attachment, cell entry, membrane fusion and host cell virion assembly.[4]

Once transmission, dengue infects immature dendritic cells. The envelope protein binds to the nonspecific receptor DC-specific ICAM-3 grabbing nonintegrin, enabling viral entry into the cell.[10] DENV then becomes internalized into vesicles, where the viral envelope fuses to endosomal membranes, with single-stranded DNA then released into the DC's cytoplasm.[11]

Several dengue vaccines are in development. CYD-TDV is a live, attenuated, chimeric, tetravalent vaccine manufactured by *Sanofi Pasteur*, marketed as Dengvaxia®. Based on the YFV 7D vaccine, CYD-TVD consists of four recombinant, tetravalent, chimeric vaccines. It has completed two Phase III clinical trials: CD14 (South-East Asia) and CD15 (Latin America). The overall vaccine efficacy in CD14 was 56.5% and 60.8% in CD15.[12] For individual serotypes, CYD-TDV showed variable efficacy in CD14: 50% for DENV-1 and 35% for DENV-2, too low to ensure any immune protection, 78.4% against DENV-3 and 75.3% against DENV-4. In CD15, CYD-TDV also varied in efficacy across serotypes: 77.7% for DENV-4, 74% for DENV-3, 50.3% for DENV-1 and 42.3 for DENV-2. In a follow-up study, an unexplained increase in DHF/DSS was reported in children under nine in Asian and Latin American countries.[13] This may result from heterotypic infection by DENV-2. As CYD-TDV has significantly lower efficacy for DENV-2 compared to other serotypes, DENV-2 infection may have caused heterotypic Dengue infection.

"DENVax" from *Takeda Vaccines Inc*. is a chimeric vaccine comprising live, attenuated DENV-2, combined with preM and E genes from DENV-1, DENV-3 and DENV-4. It has completed a Phase II trial in Puerto Rico, Columbia, Singapore and Thailand, with no ADR reported, demonstrating the vaccine's safety. Seropositivity for DENV-1, DENV-2 and DENV-3 was reported to be >95% but for DENV4 ranged from 72.7% to 100%.[12]

LAV Delta 30 developed by *NAID/Butantan* is a serotype-specific live, attenuated vaccine.[12] Phase I studies have shown its safety and immunogenicity. DEN1-80E, a recombinant, envelope glycoprotein subunit vaccine, is being developed by Merck and *Hawaii Biotech Inc*.[12,14] The vaccine has undergone an initial Phase I clinical trial and was found to be well-tolerated, immunogenic, with no ADR reported. Another tetravalent vaccine containing monovalent DNA, encoding the prM and E genes of the four DENV serotypes, using Vaxfectin® (*Vical Inc.*) as an adjuvant, has undergone Phase I clinical trial. A subsidiary vaccine, D1ME100 (a DENV-1 vaccine construct) was also tested on healthy *Flavivirus* naïve adults over 5 months with either 1 or 5 mg doses.[12,15] In this context, additional vaccines for dengue are sorely needed. An ideal epitope-based vaccine would focus on highly conserved immunogenic epitopes with wide coverage. Here, we use our evolving approach to the selection-based design of prevalidated vaccine candidates[1–3] to create a putative antidengue epitope ensemble vaccine.

2 | METHODS AND MATERIALS

2.1 | Collection of MHC Class I and Class II epitopes

Experimentally verified CD8+ and CD4+ DENV epitopes were downloaded from IEDB (http://www.iedb.org).[16] CD8+ epitope search criteria were dengue virus (ID: 12637)-specific CD8+ linear epitopes positive in T-cell assays, known to infect humans, and involved in any disease. CD4+ epitope search criteria were dengue virus (ID: 12637)-specific CD4+ linear epitopes positive in T-cell assays, known to infect humans, and involved in any disease.

2.2 | Retrieval, processing and generation of MSA from the dengue polyprotein

Protein sequences corresponding to the dengue genomic sequence were initially retrieved from UniProt (http://www.uniprot.org)[17] and searched against the protein reference sequences (Refseq_protein) using BLASTp (https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins). A multiple sequence alignment (MSA) was generated from the top four related sequences.

2.3 | Variability analysis of MSA and identification of conserved sequences

The MSA was analysed using the protein variability server (PVS) (http://imed.med.ucm.es/PVS/index.html), which quantifies and masks sequence variability, returning conserved subsequences.[18] A threshold of 0.5 was used, returning conserved fragments of nine residues.

2.4 | Calculating MHC binding predictions

Conserved CD8+ epitope predictions were made using the IEDB MHC I binding tool (http://www.iedb.org/mhci) with default IEDB recommended method and the IEDB HLA allele reference set. Epitopes with predicted affinities less than 100 nM (ANN IC50) were chosen for further analysis. Conserved CD4+ epitope predictions were made using the IEDB MHC II binding tool (http://www.iedb.org/mhcii) with the "Human, HLA-DR" allele reference set selected. Epitopes with predicted affinities to HLA alleles of less than 100 nM (SMM IC50) were retained for further analysis.

2.5 | Calculating epitope-predicted population coverage

The percentage of the population that possess at least one allele able to present at least one epitope within a group is termed the population protection coverage (PPC).[2] The selected conserved CD8+ and CD4+ epitopes were entered into the IEDB population coverage tool (http://tools.iedb.org/tools/population/iedb_input). The tool enables the PPC for a set of alleles to be calculated in 78 different populations.[19] Different combinations of epitopes were tested until a world PPC of >90% was achieved.

3 | RESULTS

3.1 | Collection of MHC Class I and Class II epitopes and generation of a multiple sequence alignment and identification of conserved sequences in the dengue polyprotein

807 MHC Class I and 798 MHC Class II dengue virus epitopes were obtained from the IEDB. The dengue virus polyprotein genomic sequence [dengue virus type 2 (strain Thailand/NGS-C/1944) (DENV-2)] was retrieved from UniProt and searched against the reference protein database using BLASTp. The top four scoring sequences (dengue virus serotype 2; dengue virus serotype 1; dengue virus serotype 3; and dengue virus serotype 4) were used to create the MSA. Variability analysis of the MSA, using the protein variability server (PVS), identified 43 conserved fragments of nine or more residues within the dengue virus polyprotein.

3.2 | Identification of conserved CD8+ epitopes, predicting Class I binding and calculating population protection coverage

Epitopes overlapping conserved regions by at least 50% were selected, identifying 22 MHC Class I CD8+ epitopes. The IEDB MHC I binding prediction tool, using the human HLA allele reference set and a 2% cut-off, was used to predict epitope binding profiles. Epitopes which bound no MHC alleles were discarded. Where multiple epitopes had identical HLA binding profiles, a single representative epitope was chosen, selecting the affine epitope available: for example, GPWHLGKLEL and GPWHLGKLEM both bind HLA-B*07:02; however, GPWHLGKLEM has a lower IC50 value. Only epitopes with an IC50 < 100 nM were analysed further. Together, epitopes were reduced from 22 to 6: RAVHADMGYW, GPWHLGKLEM, GLYGNGVVTK, NMIIMDEAHF, KTWAYHGSY, and WAYHGSYEV (see Table 1).

The cumulative population protection coverage of the six selected MHC Class I epitopes was calculated using the IEDB PPC tool. The highest PPC achieved for an individual epitope was 30.92%, thus necessitating epitope combination. Using all six MHC Class I epitopes, a world population coverage of 67.84% was achieved (see Table 2). The same combination of epitopes has a population coverage of 31.76% in South America, 68.84% in South Asia, 57.60% in South-East Asia, 51.30% in Africa (50.32% in East Africa, 51.90% in West Africa, 43.72% in Central Africa, 53.03% in North Africa and 57.50% in South Africa) and 68.84% in South Asia.

3.3 | Identification of conserved MHC II epitopes, predicting MHC II binding and calculating population protection coverage

As mentioned above, epitopes overlapping conserved regions by at least 50% were selected, identifying 55 MHC Class II CD4+ epitopes. A high sequence redundancy was seen between epitopes. Binding profiles were calculated, using a threshold of IC50 < 100 nM to define binding. Epitopes exhibited redundant HLA binding profiles. Where multiple epitopes had identical HLA binding profiles, a single representative epitope was chosen. For example, SLMYFHRRDLRLASN, WTLMYFHRRDLRLAA, WSLMYFHRRDLRLAA, LMYFHRRDLRLAANA, LMYFHRRDLRLASNA and WQLMYFHRRDLRLAA have binding profiles covered by WQLMYFHRRDLRLAA. Using this approach, 11 conserved CD4+ epitopes were selected: GVFHTMWHVTRGSVI, GKIVGLYGNGVVTTS, EIVDLMCHATFTMRL, AAIFMTATPPGSVEA, AAIFMTATPPGTADA, GKTVWFVPSIKAGND, KFWNTTIAVSMANIF, RAIWYMWLGARYLEF, LAKAIF

KLTYQNKVV, VGTYGLNTFTNMEVQ and WTLMYFH RRDLRLAA (see Table S1).

The 11 MHC II epitopes had their PPC calculated. The highest world PPC for a single epitope was 23.19% for AAIFMTATPPGSVEA. Thus to reach a cumulative world PPC of >90%, a combination of epitopes is needed. A world PPC of 76.65% was achieved with a set of 11 MHC II epitopes (see Table S2). This combination PPC values are of: 72.40% in East Asia, 71.79% in South Asia, 52.73% in South-East Asia, 54.97% in South America and 49.81% in Africa (53.62% in East Africa, 55.68% in West Africa, 44.68% in Central Africa, 68% in North Africa and 27.07% in South Africa).

3.4 | Using combinations of MHC Class I and MHC Class II epitopes to generate potential vaccines with >90% population protection coverage

A combination of 15 epitopes (six CD8+ and nine CD4+) generated a putative universal vaccine with a world population coverage of 92.49% (see Table 3a). We also targeted Asia, South America and Africa, where dengue is endemic. The population coverage of epitopes from our potential universal vaccine was evaluated for East Asia. Selecting epitopes with a PPC value >10%, a maximum coverage of 85.83% was achieved. GPWHLGKLEM (PPC 9.44%) was added, increasing the cumulative PPC to 87.43% (see Table 3b). Likewise, a combination of five MHC Class I and nine MHC Class II epitopes achieved a combined PPC of 90.23% for South Asia (see Table 3c). Attempts were also made to design putative vaccines with a population coverage of >90% targeting South America, West Africa, East Africa and Central Africa; however, such efforts proved futile. Selecting 15 conserved epitopes, we identified potential vaccines with a population coverage of 73.84% for East Africa, 76.51 for West Africa and 66.5% for Central Africa. The

TABLE 1 Binding profile of the 22 conserved MHC I epitopes. N/A means no MHC I-restricted alleles were found to present the epitope

Epitope sequence	MHC I-restricted alleles/IC50 (nM)
NR**AVHADMGYW**IESA	N/A
K**AVHADMGYW**	**HLA-B*58:01/2.89**
	HLA-B*57:01/19.89
R**AVHADMGYW**	**HLA-B*58:01/2.86**
	HLA-B*57:01/21.14
	HLA-B*53:01/257.44
GPWHLGKLEL	**HLA-B*07:02/9.11**
GPWHLGKLEM	**HLA-B*07:02/4.7**
GLYGNGVVTK	**HLA-A*03:01/24.91**
	HLA-A*11:01/42.32
IMDEAHFTDPASIARRG	N/A
NMII**MDEAHF**	**HLA-B*15:01/41.74**
	HLA-A*23:01 101.08
LIV**MDEAHFTDP**SSVAA	N/A
NYNLII**MDEAHFTDP**A	N/A
NYNLII**MDEAHFTDP**ASI	N/A
NYNMII**MDEAHFTDP**A	N/A
MGEAAA**IFMTATPPG**SV	N/A
VEMGEAAA**IFMTATPPG**	N/A
LMRRGDLPVWLAYRV	N/A
MRRGDLPVWL	N/A
DLMRRGDLPV	HLA-A*02:01/200.17
	HLA-A*02:03/1162.7
WHYDEDNPYKTWAYHGSYEV	N/A
KTWAYHGSYETKQTG	N/A
KTWAYHGSY	**HLA-A*30:02/3.14**
	HLA-A*32:01/20.9
	HLA-B*57:01/27.08
	HLA-A*01:01/60.47
	HLA-B*15:01/153.39
	HLA-A*03:01/232.14
	HLA-B*58:01/255.91
	HLA-A*11:01/295.76
WAYHGSYET	HLA-B*35:01/3038.67
WAYHGSYEV	**HLA-A*02:06/42.32**
	HLA-B*51:01/77.05
	HLA-A*68:02/107.43
	HLA-B*35:01/9748.16

The HLA alleles and IC50 values in bold font were under the threshold of 100 nM required and were further analysed.

TABLE 2 Six selected conserved epitopes their binding profile and population protection coverage

Epitopes	MHC I-restricted allele(s)	Population coverage (world) (%)
RAVHADMGYW	HLA-B*58:01, HLA-B*57:01	7.26
GPWHLGKLEM	HLA-B*07:02	12.78
GLYGNGVVTK	HLA-A*03:01, HLA-A*11:01	30.92
NMIIMDEAHF	HLA-B*15:01	8.44
KTWAYHGSY	HLA-A*30:02, HLA-A*32:01, HLA-B*57:01, HLA-A*01:01	26.65
WAYHGSYEV	HLA-A*02:06, HLA-B*51:01	9.24
Epitope set	All of the above	67.84

TABLE 3 (A) The set of 15 epitopes (six MHC Class I and nine MHC Class II) and their HLA binding profiles used to generate a universal vaccine with a world population coverage of 92.49%. (B) The set of 11 epitopes (four MHC I and seven MHC II) and their MHC-restricted alleles that have a cumulative population coverage of 87.43% in East Asia. (C) The set of 14 epitopes (five MHC I and nine MHC II) and their HLA binding profiles, which have a cumulative population coverage of 90.23% in South Asia and 82.59% in Asia (88.97% in East Asia, 84.25% in Northeast Asia, 90.23% in South Asia, 76.67% in South-East Asia and 72.86% in Southwest Asia) (Bui et al., 2006)

(A)

Epitope sequence	MHC-restricted epitopes	Population coverage (world) (%)
RAVHADMGYW	HLA-A*30:02, HLA-A*32:01, HLA-B*57:01, HLA-A*01:01	7.26
GPWHLGKLEM	HLA-B*07:02	12.78
GLYGNGVVTK	HLA-A*03:01, HLA-A*11:01	30.92
NMIIMDEAHF	HLA-B*15:01	8.44
KTWAYHGSY	HLA-A*30:02, HLA-A*32:01, HLA-B*57:01, HLA-A*01:01	26.65
WAYHGSYEV	HLA-A*02:06, HLA-B*51:01	9.24
LAKAIFKLTYQNKVV	HLA-DRB1*07:01, HLA-DRB3*02:02, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*14:01	18.23
GKIVGLYGNGVVTTS	HLA-DRB1*15:01, HLA-DQA1*05:01/DQB1*03:01	18.41
AAIFMTATPPGSVEA	HLA-DRB1*04:01, HLA-DRB1*11:01, HLA-DRB3*02:02, HLA-DRB1*08:02	23.19
AAIFMTATPPGTADA	HLA-DRB1*04:01, HLA-DRB1*09:01	17.24
GKTVWFVPSIKAGND	HLA-DRB1*08:02, HLA-DRB1*11:01	12.74
KFWNTTIAVSMANIF	HLA-DRB1*07:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*08:02	20.33
RAIWYMWLGARYLEF	HLA-DPA1*01:03/DPB1*02:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*01/DPB1*04:01	11.53
VGTYGLNTFTNMEVQ	HLA-DRB1*04:05	3.02
WTLMYFHRRDLRLAA	HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB1*03:01	17.84
Epitope set	All of the above	92.49

(B)

Epitope Sequence	MHC allele(s)	Population coverage (East Asia) (%)
GLYGNGVVTK	HLA-A*03:01, HLA-A*11:01	17.76
GPWHLGKLEM	HLA-B*07:02	9.44
NMIIMDEAHF	HLA-B*15:01	17.64
WAYHGSYEV	HLA-A*02:06, HLA-B*51:01	28.09
GKIVGLYGNGVVTTS	HLA-DRB1*15:01, HLA-DQA1*05:01/DQB1*03:01	12.18
AAIFMTATPPGSVEA	HLA-DRB1*04:01, HLA-DRB1*11:01, HLA-DRB3*02:02, HLA-DRB1*08:02	16.18
AAIFMTATPPGTADA	HLA-DRB1*04:01, HLA-DRB1*09:01	23.96
GKTVWFVPSIKAGND	HLA-DRB1*08:02, HLA-DRB1*11:01	13.19
KFWNTTIAVSMANIF	HLA-DRB1*07:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*08:02	13.88
RAIWYMWLGARYLEF	HLA-DPA1*01:03/DPB1*02:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*01/DPB1*04:01	10.21
VGTYGLNTFTNMEVQ	HLA-DRB1*04:05	19.23
Epitope set	All of the above	87.43

(Continues)

TABLE 3 (Continued)

(C)

Epitope sequence	MHC binding allele(s)	Population coverage (South Asia) (%)
GLYGNGVVTK	HLA-A*03:01, HLA-A*11:01	33.60
GPWHLGKLEM	HLA-B*07:02	2.74
NMIIMDEAHF	HLA-B*15:01	4.29
WAYHGSYEV	HLA-A*02:06, HLA-B*51:01	15.19
GKIVGLYGNGVVTTS	HLA-DRB1*15:01, HLA-DQA1*05:01/DQB1*03:01	16.45
AAIFMTATPPGSVEA	HLA-DRB1*04:01, HLA-DRB1*11:01, HLA-DRB3*02:02, HLA-DRB1*08:02	15.28
AAIFMTATPPGTADA	HLA-DRB1*04:01, HLA-DRB1*09:01	4.93
GKTVWFVPSIKAGND	HLA-DRB1*08:02, HLA-DRB1*11:01	12.70
KFWNTTIAVSMANIF	HLA-DRB1*07:01, HLA-DQA1*01:02/DQB1*06:02, HLA-DRB1*08:02	29.09
RAIWYMWLGARYLEF	HLA-DPA1*01:03/DPB1*02:01, HLA-DRB1*01:01, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*14:01, HLA-DPA1*01/DPB1*04:01	9.25
VGTYGLNTFTNMEVQ	HLA-DRB1*04:05	1.01
RAVHADMGYW	HLA-A*30:02, HLA-A*32:01, HLA-B*57:01, HLA-A*01:01	26.42
LAKAIFKLTYQNKVV	HLA-DRB1*07:01, HLA-DRB3*02:02, HLA-DRB5*01:01, HLA-DPA1*02:01/DPB1*01:01, HLA-DPA1*02:01/DPB1*14:01	28.49
WTLMYFHRRDLRLAA	HLA-DRB3*01:01, HLA-DRB3*02:02, HLA-DRB1*03:01	16.37
Set of epitopes	All of the above	90.23

same 15 epitope combination had a population coverage of 68.81% for South America and 80.80% for Brazil.

4 | DISCUSSION

We designed several putative epitope ensemble vaccines using an optimized selection of conserved dengue epitopes of verified immunogenicity. From the 807 MHC Class I epitopes and 798 MHC Class II epitopes originally found in IEDB, a final combination of six conserved CD8+ and nine conserved CD4+ epitopes had a cumulative world PPC value of 92.49%. Compared to the universal influenza vaccine designed by Sheikh et al.,[2] our potential dengue vaccine had an extra epitope. Attempts to design putative vaccines targeting endemic regions only yielded a PPC value of 76.67%. However, a combination of four CD8+ and seven CD4+ epitopes had a population coverage of 87.43% across East Asia. For South Asia, 4 CD8+ and 11 CD4+ epitopes had a cumulative PPC value of 90.23%.

Currently CYD-TDV, manufactured under the name "Dengvaxia®" by *Sanofi* Pasteur, is the only vaccine to provide any prophylaxis against dengue virus infection. CYD-TDV is a live, attenuated tetravalent chimeric vaccine. Overall vaccine efficacy is low and varies significantly between geographical locations and between virus serotypes.[12] How this vaccine protects from the effect of secondary infection by a different genotype is unclear. A better approach might be to elicit an immune response against conserved T-cell epitopes. An effective vaccine against dengue should ideally have a much higher efficacy than CYD-TDV, especially in endemic regions. Our putative epitope-based vaccine comprises epitopes conserved between all four dengue virus serotypes and hopefully elicits a controlled immune response providing future homotypic and heterotypic immunity to vaccinated population. The estimated PPC values of several of our potential vaccines, such as those targeting South-East Asia and South America, exceed the equivalent measured efficacy of CYD-TDV.

Similar to CYD-TDV, "DENVax" is unlikely to provide heterotypic immunity, due to DENV-4 having a low seropositivity. A putative vaccine comprising epitopes conserved in all four dengue serotypes could be the best route to heterotypic immunity, without causing danger to the individual. Current vaccine trials and vaccination programmes are not giving the expected protection to dengue infection. In particular, the Philippines has stopped dengue vaccination and marketing of Sanofi Pasteur's Dengvaxia, the first licensed dengue vaccine. Postvaccination testing indicated that Dengvaxia increases the risk of severe dengue in those not

previously exposed to the virus. The molecular mechanism underlying this phenomenon remains unclear, but may act by overpriming the innate immune response. Both "DENVax" and "Dengvaxia®" are chimeric subunit vaccines, containing just the premembrane and membrane proteins from the different serotypes. The conserved epitopes we identified came from nonstructural proteins and the polyprotein. A combination of 15 epitopes, with a population coverage of 92.9%, if maintained in vivo, would likely provide effective protection against Dengue. The overpriming is highly unlikely to be an issue with our epitope-based approach, which targets recognition specifically rather than through a strategy based on subunit vaccines, with their much enhanced risk of inappropriate immune reactions.

Multiple epitope combinations within each designed potential vaccine increases the chances of cross-protection among all four dengue serotypes. Vaccines, as supramolecular entities, work primarily by potentiating the host immune system. Vaccines protect by inducing cellular or molecular effector mechanisms able rapidly to inactivate toxic components or control replicating pathogens. In history, vaccines have induced antibodies produced by B cells capable of binding specifically to a toxin or a pathogen. More recent, as diseases amenable to antibody-mediated vaccines have become rarer, attention has turned to cytotoxic CD8+ and CD4+ T cells as alternative effectors. CD8+ T cells limit the proliferation of infectious micro-organisms by recognizing and killing infected cells or producing specific antiviral cytokines, while CD4+ T-helper (Th) lymphocytes limit the protection by secreting cytokine and help support the generation of B and CD8+ T-cell responses. Effector CD4+ Th cells can be subdivided into T-helper (Th1) or T-helper 2 (Th2) subsets depending on their main cytokine production (interferon-γ or interleukin [IL]-4). On this basis, an epitope ensemble vaccine, derived solely from T-cell epitopes, benefits from inclusion of both CD8+ and CD4+ broad coverage epitopes. Class I peptide binding can be predicted, with high accuracy being achieved.[20,21] It is likely that the epitope-MHC I IC50 predictions in Results section are accurate. In the case of Class II binding predictions, the same accuracy is rarely achieved due to MHC II molecules having open binding grooves.[22] From the six MHC I epitopes selected in the potential universal vaccine, only two of the epitopes were 9-mers. However, our previous work only selected MHC I epitopes that were nine residues long.[1,2] Reche et al[23] state that most peptides presented in peptide-MHC-TCR interaction are 9-mers, although epitopes between 8 and 16 residues in length are also known to be presented by MHC I, but to a lesser extent.

Epitopes require delivery with the addition of adjuvants—which can be broadly defined as substance which when added to a molecule enhances its immunogenicity.[24] An example of this is the tetravalent vaccine containing monovalent DNA vaccine against DENV, which has undergone an initial Phase I clinical trial. It is combined with equal amounts of Vaxfectin® (Vical Inc.) to enhance its immunogenicity. Schubert and Kohlbacher assert that the use of string-of-bead polypeptides, employing spacers between epitopes, increases the prediction of correct in vivo cleavage epitope recovery rate by fivefold.[25] Using this method of delivery combined with adjuvant could enable the peptides in our potential vaccine to successfully provide immune protection against dengue infection.

The generation of successful peptide vaccines using in silico methods would provide a pathway to safer vaccines. Development of the vaccine itself would likely take less time compared to other vaccine types, and come at a lower cost, while being potentially more effective. Exploitation of peptide vaccines would especially benefit regions with low-quality health care, such as South-East Asia and Africa, as they should not necessitate such extensive cold chains. There remains the rare possibility that a live vaccine could revert to a disease-inducing form, as occurs with polio. An epitope-based vaccine would not have such disadvantages. Successful development of an epitope-based universal vaccine along with the other vaccines proposed—assuming that they retain efficacy in vivo—has the potential both to save lives and to reduce the economic burden of dengue virus. Currently, our putative vaccine contains no B-cell component. Even adequate prediction of B-cell epitopes lies well beyond the capacity of current immunoinformatics. However, immune responses against conserved epitopes will likely enhance antibody-mediated responses.

In this work, he has shown clearly and explicitly how the process of molecular design can be used to design vaccines in exactly the same way as it is used in other areas of synthetic chemistry, drug design, synthetic systems biology or translational biomedicine. Design-by-selection, as we use it here, is the hierarchical preprocessing of extant data, allowing the effective combination and exploitation of computational prediction and experimentally validated prior knowledge.[1–3] Other powerful methods and approaches are available: The main competitor to our approach uses a large number of different immunoinformatics and bioinformatics methods to predict from the virtual proteome a set of putative epitopes lacking independent experimental verification and then seeks to validate these using more speculative approaches, such as docking and molecular dynamics, or, more infrequently, using experimental post hoc validation.[26] Our strategy by contrast allows us to reduce a vast search space of potential vaccine ensembles to a handful of viable, pre-verified candidate solutions. As each is built from prevalidated epitopes, each of our vaccine ensemble candidates is likely to be widely immunogenic and should be prioritized for in vivo testing.

ACKNOWLEDGEMENTS

CONFLICT OF INTEREST

There are no conflict of interests to declare.

ORCID

Pedro Reche http://orcid.org/0000-0003-3966-5838

Darren R. Flower http://orcid.org/0000-0002-8542-7067

REFERENCES

[1] M. Molero-Abraham, E. M. Lafuente, D. R. Flower, P. A. Reche, *Clin. Dev. Immunol.* **2013**, *2013*, 601943.

[2] Q. M. Sheikh, D. Gatherer, P. A. Reche, D. R. Flower, *Bioinformatics* **2016**, *32*, 3233.

[3] S. A. Damfo, P. Reche, D. Gatherer, D. R. Flower, *J. Mol. Graph. Model.* **2017**, *78*, 195.

[4] M. S. Diamond, T. C. Pierson, *Cell* **2015**, *162*, 488.

[5] O. J. Brady, P. W. Gething, S. Bhatt, J. P. Messina, J. S. Brownstein, A. G. Hoen, C. L. Moyes, A. W. Farlow, T. W. Scott, S. I. Hay, *PLoS Negl. Trop Dis.* **2012**, *6*, e1760.

[6] S. Bhatt, P. W. Gething, O. J. Brady, J. P. Messina, A. W. Farlow, C. L. Moyes, J. M. Drake, J. S. Brownstein, A. G. Hoen, O. Sankoh, M. F. Myers, D. B. George, T. Jaenisch, G. R. W. Wint, C. P. Simmons, T. W. Scott, J. J. Farrar, S. I. Hay, *Nature* **2013**, *496*, 504.

[7] M. S. Diamond, E. Mehlhop, T. Oliphant, M. A. Samuel, *Front. Biosci. Landmrk* **2009**, *14*, 3024.

[8] S. B. Halstead, *Dengue. Lancet.* **2007**, *370*, 1644.

[9] A. Imrie, J. Meeks, A. Gurary, M. Sukhbaatah, T. T. Truong, C. B. Cropp, P. Effler, *Viral Immunol.* **2007**, *20*, 672.

[10] B. Tassaneetrithep, T. H. Burgess, A. Granelli-Piperno, C. Trumpfheller, J. Finke, W. Sun, M. A. Eller, K. Pattanapanyasat, S. Sarasombath, D. L. Birx, R. M. Steinman, S. Schlesinger, M. A. Marovich, *J. Exp. Med.* **2003**, *197*, 823.

[11] M. M. Alen, D. Schols, *J. Trop. Med.* **2012**, *2016*, 628475.

[12] J. Torresi, G. Ebert, M. Pellegrini, *Hum. Vaccin. Immunother.* **2017**, *13*, 1059.

[13] Y. Liu, J. Liu, G. Cheng, *Emerg, Microbes Infect.* **2016**, *5*, e77.

[14] S. B. Manoff, S. L. George, A. J. Bett, M. L. Yelmene, G. Dhanasekaran, L. Eggemeyer, M. L. Sausser, S. A. Dubey, D. R. Casimiro, D. E. Clements, T. Martyak, V. Pai, D. E. Parks, B. A. Coller, *Vaccine* **2015**, *33*, 7126.

[15] K. R. Porter, D. Ewing, L. Chen, S. J. Wu, C. G. Hayes, M. Ferrari, N. Teneza-Mora, K. Raviprakash, *Vaccine* **2012**, *30*, 336.

[16] R. Vita, J. A. Overton, J. A. Greenbaum, J. Ponomarenko, J. D. Clark, J. R. Cantrell, D. K. Wheeler, J. L. Gabbard, D. Hix, A. Sette, B. Peters, *Nucleic Acids Res.* **2015**, *43*, D405.

[17] The UniProt C, *Nucleic Acids Res.* **2017**, *45*, D158.

[18] M. Garcia-Boronat, C. M. Diez-Rivero, E. L. Reinherz, P. A. Reche, *Nucleic Acids Res.* **2008**, *36*, W35.

[19] H. H. Bui, J. Sidney, K. Dinh, S. Southwood, M. J. Newman, A. Sette, *BMC Bioinform.* **2006**, *7*, 153.

[20] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, *BMC Bioinform.* **2007**, *8*, 424.

[21] B. Peters, H. H. Bui, S. Frankild, M. Nielsen, C. Lundegaard, E. Kostem, D. Basch, K. Lamberth, M. Harndahl, W. Fleri, S. S. Wilson, J. Sidney, O. Lund, S. Buus, A. Sette, *PLoS Comput. Biol.* **2006**, *2*, 574.

[22] H. H. Lin, G. L. Zhang, S. Tongchusak, E. L. Reinherz, V. Brusic, *BMC Bioinform.* **2008**, *9*(Suppl. 12), S22.

[23] P. A. Reche, J. P. Glutting, H. Zhang, E. L. Reinherz, *Immunogenetics* **2004**, *56*, 405.

[24] J. Bayry, E. Tchilian, M. Davies, E. Forbes, S. Draper, S. Kaveri, A. V. Hill, M. D. Kazatchkine, P. C. Beverley, D. R. Flower, D. F. Tough, *Proc. Natl Acad. Sci. USA* **2008**, *105*, 10221.

[25] B. Schubert, O. Kohlbacher, *Genome Med.* **2016**, *8*, 9.

[26] M. M. Pourseif, G. Moghaddam, B. Naghili, N. Saeedi, S. Parvizpour, A. Nematollahi, Y. Omidi, *Comput. Biol. Chem.* **2018**, *72*, 150.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.