Published in final edited form as:

Vaccine. 2012 September 14; 30(42): 6081–6087. doi:10.1016/j.vaccine.2012.07.042.

Simplifying complex sequence information: a PCP-consensus protein binds antibodies against all four Dengue serotypes

David Bowen¹, Jessica A. Lewis², Wenzhe Lu¹, and Catherine H. Schein^{1,2,3,*}

¹Computational Biology, Sealy Center for Structural Biology and Molecular Biophysics, Department of Biochemistry and Molecular Biology, University of Texas Medical Branch Galveston, Texas 77555–0857

²Department of Microbiology and Immunology, University of Texas Medical Branch Galveston, Texas 77555–0857

³Sealy Center for Vaccine Development; Institute for Human Infections and Immunity WHO Collaborating Center for Tropical Diseases, University of Texas Medical Branch Galveston, Texas 77555–0857

Abstract

Designing proteins that reflect the natural variability of a pathogen is essential for developing novel vaccines and drugs. Flaviviruses, including Dengue (DENV) and West Nile (WNV), evolve rapidly and can “escape” neutralizing monoclonal antibodies by mutation. Designing antigens that represent many distinct strains is important for DENV, where infection with a strain from one of the four serotypes may lead to severe hemorrhagic disease on subsequent infection with a strain from another serotype. Here, a DENV physicochemical property (PCP)-consensus sequence was derived from 671 unique sequences from the Flavitrack database. PCP-consensus proteins for domain 3 of the envelope protein (EdomIII) were expressed from synthetic genes in *E. coli*. The ability of the purified consensus proteins to bind polyclonal antibodies generated in response to infection with strains from each of the four DENV serotypes was determined. The initial consensus protein bound antibodies from DENV-1-3 in ELISA and Western blot assays. This sequence was altered in 3 steps to incorporate regions of maximum variability, identified as significant changes in the PCPs, characteristic of DENV-4 strains. The final protein was recognized by antibodies against all four serotypes. Two amino acids essential for efficient binding to all DENV antibodies are part of a discontinuous epitope previously defined for a neutralizing monoclonal antibody. The PCP-consensus method can significantly reduce the number of experiments required to define a multivalent antigen, which is particularly important when dealing with pathogens that must be tested at higher biosafety levels.

© 2012 Elsevier Ltd. All rights reserved.

*To whom correspondence should be addressed: chschein@utmb.edu.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

INTRODUCTION

Most methods for analyzing aligned sequences of proteins are designed to identify the absolutely conserved residues, which are usually those that maintain the structure of the protein or are vital for function[1]. Here we deal with a different type of information contained in such data: identifying regions of maximum variability that may be correlated with function[2, 3] or antigenicity[4]. This second type of information is more difficult to define, as many studies have shown there is no absolute, linear relationship between changes in a protein's sequence and its immunogenicity. For Flaviviruses, even a single point mutations can prevent the binding of an antibody to its target[5].

We developed a new method that, given a sequence dataset with high superficial redundancy (i.e., many nearly identical sequences), containing unequal numbers of outliers, can be used to determine a sequence that is closest in its physicochemical properties (PCPs) to all others in the set. We previously used the PCP-consensus method to design reference sequences for genera of the Flaviviridae[6]. Here we tested whether PCP-consensus sequences derived from our DENV sequence set are also multivalent, and capable of binding antibodies generated against the most diverse strains. Such proteins could be useful for the design of multivalent vaccines, and for determining DENV prevalence using serum samples. Understanding variation in DENV strains is important, for diagnostic purposes[7, 8], for determining antibody cross reactivity after vaccination[9], and because mutations can greatly alter immunogenicity or infectivity[10–15]. There are four distinct serotypes of DENVs [16, 17], strains from several of which may be circulating at the same time in some populations[18]. There is a good deal of sequence variation between the different serotypes; for example, the sequences of the EdomIII of natural DENV isolates have >40% diversity. Due to increasingly severe epidemics in recent years [19, 20], antiviral agents and vaccines are being sought to protect against DENV in endemic areas[11, 12, 21–28]. Methods for defining vaccines require extensive experimental mutagenesis studies and testing, including those that rely on random methods for attenuation, such as serial passage[29, 30]. One of our goals is to use computational methods to limit the number of experimental studies that must be done to define suitable multivalent targets for vaccines and drug design[31], which is

especially important for the study of pathogens that require higher biosafety levels of containment.

Here we tested the application of a novel, physicochemical property (PCP) consensus method [6, 32] to the daunting task of generating one sequence that would adequately reflect the sequences of 671 unique DENV isolates. Our test for this was whether one consensus protein could bind to antibodies generated from infections with strains from all four serotypes of DENV, especially the more distantly related DENV-4 type. We show that by starting with PCP-consensus sequences, identifying regions of maximum variability in PCPs, and altering these residues as groups in an overall consensus, we could significantly simplify the initial stages of the search for a pan-DENV, multivalent antigen.

METHODS

Calculating an overall PCP-consensus sequence for DENV

The PCP consensus method, described in more detail in other publications[6, 32], was originally used for classifying Flaviviridae isolates and relating their sequences to observed phenotype. To document that the PCP-consensus proteins could also be multivalent antigens, we chose the 112 amino acid Domain 3 of the DENV envelope protein, which folds independently and contains serotype-specific epitopes[11, 12, 21, 22, 33–39]. The PCP-consensus was determined for DENV strains catalogued in Flavitrack (carnot.utmb.edu/flavitrack) according to our previously described method[6, 32]. Instead of simply determining the most common amino acid in each column, the PCP-consensus determines the average values of the PCPs of the 20 amino acids as defined by a set of numerical descriptors. These 5 Eigenvectors, obtained by multidimensional scaling of over 200 quantitative measures of physicochemical properties, reflect the intrinsic properties of the side chains, such as hydrophobicity, size, charge, helical propensity, etc. [40]. The average values of the PCPs of the amino acids in each column of a multiple alignment are

thus summarized by five discrete numerical values in the five PCP dimensions ($\overline{E}^p$ where $p=0-5$ in the equation shown). The program then determines which of the amino acids that naturally occur at that position is closest in its PCPs to this average value.

$$\overline{E}^p = \sum_{j=1}^N \frac{V_j^p}{N}$$

The maximum inter-sequence variability was 41% among all the 671 unique DENV sequences in Flavitrack at that time. Multiple sequence alignments for deriving the PCP-consensus sequences for each of the four DENV types were generated with MUSCLE[41, 42]. The four PCP-consensus sequences were then aligned using Clustalw 2.0.3[43] (at this level of identity (>59%), both alignment methods give equivalent results) to generate an overall PCP-consensus sequence for all the DENV (sequences **7**, **7B**, **8**, **7P8** in Figure 1B).

Comparison of the PCP-consensus with other consensus methods

Models of PCP-consensus sequences **7** and **7P8** were prepared with our MPACK modeling suite[44–47] using either the crystal structure of the DENV-2 protein (1OAN.pdb)[38] or that of the NMR structure of DENV-4 (2HOP.pdb)[48] as template. Figures were drawn with MolMol. The Jalview consensus was copied from the applet of the ClustalW program; the Ctree consensus was determined using the application within the protein tools section of the Biology Workbench (<http://seqtool.sdsc.edu>).

Protein expression and purification of PCP-consensus proteins

Synthetic gene sequences (from Bio Basic, Ontario, Canada) were obtained and cloned into the pET15 vector for expression as the untagged “antigen only” protein. Proteins were

purified as described elsewhere for earlier work with wild-type Flavivirus EdomIII antigens[48, 49]. For “7P8”, the gene sequences for 7 and 8 were cleaved at the PstI site within the gene (which occurs around amino acid 51 in both constructs) and two segments ligated to form a recombinant with the N-terminus of the 7 sequence and the C-terminus of 8. Proteins were pure according to PAGE; the consensus proteins were characterized by Maldi Mass spectroscopy before and after tryptic digest (which also confirmed the cystine bond between residues 14 and 45).

ELISA assay to detect binding of IgG from DENV infected mice

MaxiSorp plates (Nunc, Rochester, NY) were coated overnight at 4°C with 100ng/well of recombinant EIII antigens (prepared as described previously [48, 49] or the recombinant PCP-consensus proteins, diluted in borate saline. After coating, plates were washed 1x with PBS/0.5% Tween-20 (PBS-T), blocked with PBS-T/3% BSA for 1 hour at room temperature and washed 2x with PBS-T. The plates were incubated with DENV serotype specific murine immune ascites fluid (MIAF; 1:500 dilution of samples from the Tesh or WHO collection at the WRCEVA at the UTMB) from mice inoculated with DENV strains from each serotype (Dengue 1 (Hawaii), T30935; Dengue 2, T-34973; Dengue 3, MIAF-WHO; Dengue 4, T-30942) or with a commercially available monoclonal antibody reported to bind to all four serotypes (GeneTex GTX29202). After 1 hour at room temperature, plates were washed 3x with PBS-T. All four of the MIAF contained neutralizing antibodies against the serotype of the virus used for infection (data not shown). However, as they were generated against the whole virus, and not just the EdomIII protein, their content of antibodies that recognized the wild type proteins varied, being highest for DENV 1 and 2, and lower for DENV 3 and 4. Horseradish peroxidase-conjugated secondary antibody (anti-mouse IgG) (Sigma, St. Louis, MO) was allowed to bind to the primary antibody for 1 hour at room temperature; after which the plate was washed 4x with PBS-T. Plates were developed for 10 minutes with TMB solution before color development was stopped with 3M HCl. Absorbance values were read at 450nm, with absorbance at 650nm subtracted (450–650nm) and given here after subtracting the absorbance for a “no antigen” sample. The typical background values ranged from 0.02–0.1 units, with occasional background as high as 0.27–0.5 absorbance units. All values given are the average of two samples. The corrected numbers for the absorbance (450–650nm absorbance of sample - 450–650 nm of “no antigen” control) are given in the figures. For comparison, control proteins included the Domain III proteins of WNV and deer tick virus (DTV).

RESULTS

Determining a PCP-consensus sequence to represent multiple strains of DENV. Comparison of the PCP-consensus with other consensus methods

Conventional methods for calculating profiles and consensus sequences [50–52] calculate only the most common amino acid in a column[53]. We compared the consensus from two methods, one from the Jalview applet of Clustal W, and another from Ctree[54] with the PCP-consensus sequences. For easy comparison, we used a simplified alignment of four PCP consensus sequences, one for each of the four DENV types (DENV1-4c, figure 1A). The PCP-consensus sequences are based on individual alignments of 216 DENV-1,307 DENV-2,136 DENV-3, and 12 DENV4 sequences. DENV1 and 3 are most similar to one another. Although closest to DENV-2, DENV 4 is a clear outlier, with only 59% identity to any of the other serotypes. This is also consistent with the reactivity of the individual DENV EdomIII proteins with antibodies generated in mice (MIAFs) during infection with different serotypes of DENV (Figure 2,3). DENV-1,2,3 all have some degree of cross-reactivity with antibodies generated against the whole viruses, individually (see bars in Fig. 3), but the EdomIII protein of DENV-4 is not recognized at all by the MIAFs

generated against DENV 1 and 3, and only slightly by that against DENV-2. The reverse is also true: when an MIAF generated against a DENV-4 strain was used, these IgG bound to the EdomIII protein of a wild type DENV-4 strain, while the ELISA absorbance of the EdomIII proteins of DENV1-3 was only slightly above background. All three consensus methods chose the same residues for conserved areas of the sequences, which constitute 60% of the positions (Figure 1b). The differences are in the variable regions. Both the Jalview applet and Ctree calculate a “conventional” consensus sequence, and chose the amino acid that is most abundant in a given sequence alignment (Figure 1b, Sequences **5**, **6**). However, Jalview does not indicate an amino acid choice at variable positions. Ctree[54] selects a “best” amino acid according to an internal score function in positions where one amino acid is not in the majority of the sequences. The overall PCP-consensus (**7**) differs from the Ctree one (**6**) at 12 positions; the two methods chose the same amino acid at only 5 variable residues.

An additional problem with obtaining a consensus from large collections of virus sequences is data bias due to uneven sampling. This means that there may be many closely related sequences collected in one epidemic, and only a few representatives from other areas and times. To obtain a “balanced” alignment, those designing consensus sequences for virus antigens may thus select a few sequences[24, 55], leaving out much of the data on allowed substitutions that is contained in the whole sequence set [44, 56–58]. Although there have been methods suggested, such as selective sequence weighting[59], to deal with biased data sets, the problem may not be mathematically solvable for all examples. Here, we tested the effect of adding one additional sequence, of a single DENV-4 isolate, to our initial four sequence alignment (last line of Figure 1a). In this case, as Fig. 1b shows, the Ctree consensus changed considerably, choosing DENV-4 residues at 12 different positions (Sequence **6B**). The PCP-consensus for the alignment with two DENV-4 sequences (Sequence **7B**) changed at 8 positions. Thus the PCP-consensus method showed fewer changes in response to one outlier sequence. The sequences in Figure 1a indicated that DENV-4c differed significantly from the other 3 subtypes in two regions, 56–60 and 71–74, which were changed in **7B** to those of the DENV-4 consensus (to obtain sequence **8**).

Experimental tests of the PCP-consensus sequences

To determine whether the PCP- consensus sequences would be able to bind antibodies to all four DENV types, sequences **7**, **7B** and **8** were translated into synthetic gene sequences, cloned, the proteins expressed in *E. coli*, and purified. The proteins all bound with varying degrees of strength to antibodies in Western blotting to all four DENV types, although the binding to the DENV-4 specific antibodies (MIAF-4) was weaker than to those generated against DENV 1–3 (not shown). The proteins were then assayed quantitatively, using ELISA, for recognition by polyclonal sera that were specific for each of the four different DENV types, and by monoclonal antibodies with differing specificity (Figure 2). While the initial consensus sequence (**7** in Fig. 1B) bound as well as the wild type EdomIII regions to antibodies generated against DENV types 1–3, it bound only slightly to polyclonal antibodies induced by DENV-4. Our first attempt to correct for this bias, by adding an additional DENV-4 sequence to the alignment (**7B**), resulted in an unusable protein, that was not recognized by any of the antisera. Further changes to the PCP-consensus DENV-4 hypervariable regions (Sequence **8**), gave a protein that bound antibodies to all four types. However, its ability to bind to antibodies from any of the antisera, especially MIAF-3, was greatly reduced.

Reactivity with DENV1-3 antibodies is dependent on two amino acids

To determine which residue changes could account for the drop in reactivity, a recombinant, “**7P8**” (the P was short for the PstI site which occurred a few residues before the

hypervariable region in the synthetic genes for both **7** and **8**), was prepared. This restored 2 amino acids in the N-terminus to those of our original consensus sequence (**7**) while leaving the amino acids in the C-terminus that were added for recognition of DENV-4. The changes restored the ability of the protein to bind to antibodies specific for DENV1-3, and also enhanced binding to DENV-4 specific antibodies (Fig. 3). The ELISA values for binding of MIAF 1 and 3 to the **7** and **7P8** proteins were both at least 10x higher than those for the **7B** and **8** protein, when measured in duplicate in the same assay.

All the PCP-consensus proteins were also recognized by a DENV monoclonal antibody (GTx29202), which bound well to wild type DENV 1-3 proteins but did not recognize the wild type DENV-4 protein at all (the absorbance was less than the no-antigen control). The EdomIII of Deer Tick virus (DTV) was included as a negative control, as it is quite distant from the DENV sequences (only about 40% identity). This protein was not recognized by IgG from any of the MIAF (absorbance values were less than the “no antigen” control) but gave about 2x background (0.12 units, vs 0.072 for the control) with the monoclonal antibody (last set of bars in Fig. 3).

The two amino acids lie close to one another, in a variable region

We used structural analysis of the variable positions to understand why the MIAF-4 antibodies bound better to **7P8** than they did to either **7B** or **8** sequences, even though both of the latter contained the DENV-4 consensus residues. Mapping the variable residues on models of the initial (**7**) and final (**7P8**) DENV PCP-consensus antigens (Fig. 4) illustrates that the amino acids that enhance IgG recognition by all four of the MIAFs (K19 and D42) lie in a cluster of variable residues. This suggests the target for the neutralizing antibodies is a conformational epitope region that forms from residues that are separated by >30 amino acids in the linear sequence. Indeed, one of the amino acids needed for recognition by the DENV1 and 3 MIAF antibodies, K19, lies within regions that have been previously shown, by site directed mutagenesis and structural analysis[11, 12, 60] to bind neutralizing antibodies. The other, D42, (position 328 in the whole envelope numbering), a G in the **7B** and **8** sequences, is part of a BC loop important for folding in several flaviviruses[61]. This residue is next to one of the four amino acids implicated in immune response to a DENV-3 envelope fragment in another study[62]. The other 3 preferred amino acids positions in that study were selected as the PCP consensus from the beginning. Our results indicate that changing highly variable residues, as a group, gave the best results for multivalent binding.

DISCUSSION

This work shows that PCP-consensus sequences can be used to identify regions of maximum variability, and to design multivalent antigens. The regions of maximum variability in this case also correlated with residues involved in antibody recognition. This approach allowed us, in only three design cycles (Figure 1), to define a multivalent consensus protein that bound antibodies against all four types of DENV (Figs. 2,3). Altering a small number of amino acids, primarily in hypervariable regions, in the initial consensus greatly enhanced recognition by antibodies generated against the outlier, DENV-4. Both DENV-3 and DENV-4 are known to be difficult to neutralize with antisera from infection with heterologous DENV, which is a major problem in balancing vaccines based on combining viruses from the four types into one vaccine[28]. Thus, the fact that our final PCP-consensus was well recognized by antisera against all four types is quite significant.

When mapped onto the surface of a model protein for the consensus, all of the variable residues detected by our sequence analysis method mapped to two surface exposed areas of the protein (Figure 4). Two amino acids, in areas detected by previous mutagenesis experiments, were needed for optimal antibody recognition. These results suggest the

usefulness of the PCP based method for reducing the number of experiments required to define multivalent antigens for other emerging pathogens. The real value of our analysis was in limiting the number of residues that needed to be tested for optimizing our mean sequence.

Dealing with bias in sequence data sets with high superficial redundancy

Obtaining all the information from large datasets requires ways of dealing with inherent bias due to uneven sampling. Our initial consensus sequence (**7** in Figure 1B) reflected the common properties of DENV-1,2,3, and did not provide sufficient coverage of the DENV-4 outlier. Obtaining recognition by all four types required computationally guided, experimental optimization at variable positions. We tested two ways to better incorporate the changes specific for DENV-4, first by adjusting the alignment used to extract the PCP-consensus sequence by adding an additional DENV-4 sequence, and then, by specifically changing hypervariable regions to the DENV-4 consensus. The first strategy generated an essentially useless protein (Sequence **7B** in Figure 1B), with greatly reduced reactivity especially to antibodies specific for DENV 1 & 3. Converting the hypervariable region to that of DENV4 (Sequence **8** in Figure 1B) in this sequence indeed gave better recognition by the DENV-4 specific antibodies. Counterintuitively, changing the hypervariable region completely to that of DENV-4c in **7B** (to give sequence **8**) also gave better (although still unacceptably low) recognition by the polyclonal antibodies to DENV-1-3 (Figure 2,3).

We note that, as anticipated, the PCP-consensus approach was somewhat less affected by numerical bias than that from an algorithm in Ctree, where the choice was made based on relative abundance. However, our experimental results indicated that it is better to alter groups of residues identified as variable in the overall consensus than to try to “balance” the overall alignment to control for bias. The results for how well the different consensus sequences bound to MIAF-3 antibodies illustrate another point of caution in simply choosing a consensus with the highest % identity to all sequences. The identity of DENV-3c to **7B** was 70%, to **8**, 67%, yet only the latter bound the DENV-3 specific IgGs.

Uses of PCP-consensus sequences

We chose this example to validate our methodology as this region of the DENV envelope has been extensively studied, experimentally, and is the binding site of many neutralizing and type specific antibodies [11, 12, 21, 22, 33–39]. However, the method should be especially useful for analysis of viruses where little is known about the antigenic determinants, or for other biodefense targets that require experiments at higher biosafety containment levels [63, 64]. Preparing a PCP-consensus from many different viral strains gives reference values for PCPs at each position, a logical starting point for experimental analysis. The results can be coupled with structural data to reveal variance that is most likely to influence antigenicity. There are many other potential uses for PCP-consensus sequences, including determining type-specific or multivalent sequences for diagnostics, discriminating functional alterations in enzyme sequences, analysis of changes due to immune system pressure [65], and designing multivalent targets for drug design.

Acknowledgments

We thank David Beasley (UTMB, M&I) for directing the ELISA experiments and obtaining the antibodies used and Werner Braun (UTMB, BMB) for his thoughtful input on handling bias and alternate methods of consensus sequence determination. The computational resources of the Sealy Center for Structural Biology and Molecular Biophysics were also used in this project.

Funding: The DENV vaccine project was supported in part by grant 1UL1RR029876-01 from the National Center for Research Resources, NIH to the Institute for Translational Studies of the UTMB, (support for pilot protocols #763 and #809). JAL is supported by a pre-doctoral Fellowship from the Sealy Center for Vaccine Development at

the UTMB. Development of the PCP-consensus method was supported in part by NIH grant AI064913 (to Werner Braun and CHS), and EPA-STAR grant RE-83406601-0 to CHS.

References

1. Wilkins A, Erdin S, Lua R, Lichtarge O. Evolutionary trace for prediction and redesign of protein functional sites. *Methods Mol Biol.* 2012; 819:29–42. [PubMed: 22183528]
2. Rodriguez-Roche R, Sanchez L, Burgher Y, Rosario D, Alvarez M, Kouri G, et al. Virus role during intraepidemic increase in dengue disease severity. *Vector Borne Zoonotic Dis.* 2011 Jun; 11(6): 675–81. [PubMed: 21254883]
3. Wicker JA, Whiteman MC, Beasley DW, Davis CT, Zhang S, Schneider BS, et al. A single amino acid substitution in the central portion of the West Nile virus NS4B protein confers a highly attenuated phenotype in mice. *Virology.* 2006 Jun 5; 349(2):245–53. [PubMed: 16624366]
4. Soriani M, Petit P, Grifantini R, Petracca R, Gancitano G, Frigimelica E, et al. Exploiting Antigenic Diversity for Vaccine Design. *Journal of Biological Chemistry* 2010. Sep 24; 285(39):30126–38.
5. Zhang SL, Li L, Woodson SE, Huang CYH, Kinney RM, Barrett ADT, et al. A mutation in the envelope protein fusion loop attenuates mouse neuroinvasiveness of the NY99 strain of West Nile virus. *Virology.* 2006 Sep; 353(1):35–40. [PubMed: 16806383]
6. Danecek P, Lu W, Schein CH. PCP consensus sequences of flaviviruses: correlating variance with vector competence and disease phenotype. *J Mol Biol.* 2010 Feb 26; 396(3):550–63. [PubMed: 19969003]
7. Gijavanekar C, Anez-Lingerfelt M, Feng C, Putonti C, Fox GE, Sabo A, et al. PCR detection of nearly any dengue virus strain using a highly sensitive primer ‘cocktail’. *Febs J.* 2011 May; 278(10):1676–87. [PubMed: 21410643]
8. Munoz-Jordan JL, Collins CS, Vergne E, Santiago GA, Petersen L, Sun W, et al. Highly sensitive detection of dengue virus nucleic acid in samples from clinically ill patients. *J Clin Microbiol.* 2009 Apr; 47(4):927–31. [PubMed: 19225099]
9. Mansfield KL, Horton DL, Johnson N, Li L, Barrett ADT, Smith DJ, et al. Flavivirus-induced antibody cross-reactivity. *The Journal of general virology.* 2011; 92(Pt 12):2821–9. [PubMed: 21900425]
10. Barker WC, Mazumder R, Vasudevan S, Sagripanti JL, Wu CH. Sequence signatures in envelope protein may determine whether flaviviruses produce hemorrhagic or encephalitic syndromes. *Virus Genes.* 2009; 39(1):1–9. Note they say Asn 67 (glycosylated) should be characteristic of hemorrhagic viruses, with Asp in encephalitic ones. [PubMed: 19283462]
11. Gromowski GD, Barrett ND, Barrett AD. Characterization of dengue virus complex-specific neutralizing epitopes on envelope protein domain III of dengue 2 virus. *J Virol.* 2008 Sep; 82(17): 8828–37. [PubMed: 18562544]
12. Matsui K, Gromowski GD, Li L, Schuh AJ, Lee JC, Barrett AD. Characterization of dengue complex-reactive epitopes on dengue 3 virus envelope protein domain III. *Virology.* 2009 Dec 18; 384:16–20. [PubMed: 19101005]
13. Oliphant T, Nybakken GE, Engle M, Xu Q, Nelson CA, Sukupolvi-Petty S, et al. Antibody recognition and neutralization determinants on domains I and II of West Nile virus envelope protein. *Journal of Virology.* 2006 Dec; 80(24):12149–59. [PubMed: 17035317]
14. Sukupolvi-Petty S, Austin SK, Purtha WE, Oliphant T, Nybakken GE, Schlesinger JJ, et al. Type- and subcomplex-specific neutralizing antibodies against domain III of dengue virus type 2 envelope protein recognize adjacent epitopes. *Journal of Virology.* 2007 Dec; 81(23):12816–26. [PubMed: 17881453]
15. Rico-Hesse R. Dengue virus evolution and virulence models. *Clin Infect Dis.* 2007 Jun 1; 44(11): 1462–6. [PubMed: 17479944]
16. Sabin AB. Research on Dengue during World War II. *The American Journal of Tropical Medicine and Hygiene* 1952. Jan 1; 1952 1(1):30–50.
17. Rico-Hesse R. Microevolution and virulence of dengue viruses. *Adv Virus Res.* 2003; 59:315–41. [PubMed: 14696333]

18. Uzcategui NY, Comach G, Camacho D, Salcedo M, de Quintana MC, Jimenez M, et al. Molecular epidemiology of dengue virus type 3 in Venezuela. *Journal of General Virology*. 2003 Jun. 84:1569–75. [PubMed: 12771427]
19. Tomashek KM, Gregory CJ, Rivera Sanchez A, Bartek MA, Garcia Rivera EJ, Hunsperger E, et al. Dengue deaths in Puerto Rico: lessons learned from the 2007 epidemic. *PLoS Negl Trop Dis*. 2012 Apr. 6(4):e1614. [PubMed: 22530072]
20. Rodriguez-Barraquer I, Cordeiro MT, Braga C, de Souza WV, Marques ET, Cummings DAT. From Re-Emergence to Hyperendemicity: The Natural History of the Dengue Epidemic in Brazil. *PLoS Negl Trop Dis*. 2011; 5(1):e935. [PubMed: 21245922]
21. Babu JP, Pattnaik P, Gupta N, Shrivastava A, Khan M, Rao PVL. Immunogenicity of a recombinant envelope domain III protein of dengue virus type-4 with various adjuvants in mice. *Vaccine*. 2008 Aug; 26(36):4655–63. [PubMed: 18640168]
22. Chin JFL, Chu JJH, Ng ML. The envelope glycoprotein domain III of dengue virus serotypes 1 and 2 inhibit virus entry. *Microbes and Infection*. 2007 Jan; 9(1):1–6. [PubMed: 17196419]
23. Guy B, Guirakhoo F, Barban V, Higgs S, Monath TP, Lang J. Preclinical and clinical development of YFV 17D-based chimeric vaccines against dengue, West Nile and Japanese encephalitis viruses. *Vaccine*. 2010; 28(3):632–49. [PubMed: 19808029]
24. Ramanathan MP, Kuo YC, Selling BH, Li Q, Sardesai NY, Kim JJ, et al. Development of a novel DNA SynCon tetravalent dengue vaccine that elicits immune responses against four serotypes. *Vaccine*. 2009 Oct 30; 27(46):6444–53. [PubMed: 19580892]
25. Webster DP, Farrar J, Rowland-Jones S. Progress towards a dengue vaccine. *The Lancet Infectious Diseases*. 2009; 9(11):678–87. [PubMed: 19850226]
26. Murrell S, Wu S-C, Butler M. Review of dengue virus and the development of a vaccine. *Biotechnology Advances*. 2011; 29(2):239–47. [PubMed: 21146601]
27. Konishi E, Miyagawa Y. Balance of infection-enhancing and neutralizing antibodies induced by a dengue tetravalent DNA vaccine in a mouse model. *Microbes Infect*. 2011 Nov; 13(12–13):1091–8. [PubMed: 21763777]
28. Anderson KB, Gibbons RV, Edelman R, Eckels KH, Putnak RJ, Innis BL, et al. Interference and facilitation between dengue serotypes in a tetravalent live dengue virus vaccine candidate. *J Infect Dis*. 2011 Aug 1; 204(3):442–50. [PubMed: 21742844]
29. Kelly EP, Polo S, Sun W, Falgout B. Evolution of attenuating mutations in dengue-2 strain S16803 PDK50 vaccine and comparison of growth kinetics with parent virus. *Virus Genes*. 2011 Aug; 43(1):18–26. [PubMed: 21461924]
30. Kelly EP, Puri B, Sun W, Falgout B. Identification of mutations in a candidate dengue 4 vaccine strain 341750 PDK20 and construction of a full-length cDNA clone of the PDK20 vaccine candidate. *Vaccine*. 2010 Apr 9; 28(17):3030–7. [PubMed: 19874927]
31. Nwaka S, Besson D, Ramirez B, Maes L, Matheeußen A, Bickle Q, et al. Integrated dataset of screening hits against multiple neglected disease pathogens. *PLoS Negl Trop Dis*. 2011 Dec. 5(12):e1412. [PubMed: 22247786]
32. Danecsek P, Schein CH. Flavitrack analysis of the structure and function of West Nile non-structural proteins. *Int J Bioinform Res Appl*. 2010; 6(2):134–46. [PubMed: 20223736]
33. Bernardo L, Fleitas O, Pavon A, Hermida L, Guillen G, Guzman MG. Antibodies induced by dengue virus type 1 and 2 envelope domain III recombinant proteins in monkeys neutralize strains with different genotypes. *Clin Vaccine Immunol*. 2009 Dec; 16(12):1829–31. [PubMed: 19726617]
34. Bernardo L, Izquierdo A, Alvarez M, Rosario D, Prado I, Lopez C, et al. Immunogenicity and protective efficacy of a recombinant fusion protein containing the domain III of the dengue 1 envelope protein in non-human primates. *Antiviral Res*. 2008 Nov; 80(2):194–9. [PubMed: 18602424]
35. Guzman MG, Hermida L, Bernardo L, Ramirez R, Guillen G. Domain III of the envelope protein as a dengue vaccine target. *Expert Rev Vaccines*. 2010 Feb; 9(2):137–47. [PubMed: 20109025]
36. Hung JJ, Hsieh MT, Young MJ, Kao CL, King CC, Chang W. An external loop region of domain III of dengue virus type 2 envelope protein is involved in serotype-specific binding to mosquito but not mammalian cells. *Journal of Virology*. 2004 Jan; 78(1):378–88. [PubMed: 14671119]

37. Lazo L, Zulueta A, Hermida L, Blanco A, Sanchez J, Valdes I, et al. Dengue-4 envelope domain III fused twice within the meningococcal P64k protein carrier induces partial protection in mice. *Biotechnol Appl Biochem*. 2009 Apr; 52(Pt 4):265–71. [PubMed: 18636968]
38. Modis Y, Ogata S, Clements D, Harrison SC. Variable surface epitopes in the crystal structure of dengue virus type 3 envelope glycoprotein. *J Virol*. 2005 Jan; 79(2):1223–31. [PubMed: 15613349]
39. Sukupolvi-Petty S, Austin SK, Purtha WE, Oliphant T, Nybakken GE, Schlesinger JJ, et al. Type- and subcomplex-specific neutralizing antibodies against domain III of dengue virus type 2 envelope protein recognize adjacent epitopes. *J Virol*. 2007 Dec; 81(23):12816–26. [PubMed: 17881453]
40. Venkatarajan MS, Braun W. New quantitative descriptors of amino acids based on multidimensional scaling of a large number of physical-chemical properties. *Journal of Molecular Modeling*. 2001; 7(12):445–53.
41. Edgar RC. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *Bmc Bioinformatics*. 2004 Aug 19; 5:113. [PubMed: 15318951]
42. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004; 32(5):1792–7. [PubMed: 15034147]
43. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics*. 2002 Aug; Chapter 2(Unit 2):3. [PubMed: 18792934]
44. Ivanciuc O, Oezguen N, Mathura VS, Schein CH, Xu Y, Braun W. Using property based sequence motifs and 3D modeling to determine structure and functional regions of proteins. *Curr Med Chem*. 2004 Mar; 11(5):583–93. [PubMed: 15032606]
45. Oezguen N, Zhou B, Negi SS, Ivanciuc O, Schein CH, Labesse G, et al. Comprehensive 3D-modeling of allergenic proteins and amino acid composition of potential conformational IgE epitopes. *Mol Immunol*. 2008 Aug; 45(14):3740–7. [PubMed: 18621419]
46. Soman KV, Midoro-Horiuti T, Ferreon JC, Goldblum RM, Brooks EG, Kurosky A, et al. Homology modeling and characterization of IgE binding epitopes of mountain cedar allergen Jun a 3. *Biophys J*. 2000 Sep; 79(3):1601–9. [PubMed: 10969020]
47. Soman KV, Schein CH, Zhu H, Braun W. Homology modeling and simulations of nuclease structures. *Methods Mol Biol*. 2001; 160:263–86. [PubMed: 11265289]
48. Volk DE, Lee YC, Li X, Thiviyanathan V, Gromowski GD, Li L, et al. Solution structure of the envelope protein domain III of dengue-4 virus. *Virology*. 2007 Jul; 364(1):147–54. [PubMed: 17395234]
49. Volk DE, Anderson KM, Gandham SHA, May FJ, Li L, Beasley DWC, et al. NMR assignments of the sylvatic dengue 1 virus envelope protein domain III. *Biomolecular Nmr Assignments*. 2008 Dec; 2(2):155–7. [PubMed: 19636893]
50. Gribskov M. Profile analysis. *Methods Mol Biol*. 1994; 25:247–66. [PubMed: 8004170]
51. Gribskov M, McLachlan AD, Eisenberg D. Profile analysis: detection of distantly related proteins. *Proc Natl Acad Sci U S A*. 1987 Jul; 84(13):4355–8. [PubMed: 3474607]
52. Merkeev IV, Mironov AA. PHOG-BLAST--a new generation tool for fast similarity search of protein families. *BMC Evol Biol*. 2006; 6:51. [PubMed: 16792802]
53. Ramanathan MP, Kutzler MA, Kuo YC, Yan J, Liu H, Shah V, et al. Coimmunization with an optimized IL15 plasmid adjuvant enhances humoral immunity via stimulating B cells induced by genetically engineered DNA vaccines expressing consensus JEV and WNV E DIII. *Vaccine*. 2009 Jul 9; 27(32):4370–80. [PubMed: 19497647]
54. Archer J, Robertson DL. CTree: comparison of clusters between phylogenetic trees made easy. *Bioinformatics*. 2007. Nov 1; 23(21):2952–3.
55. Leng CH, Liu SJ, Tsai JP, Li YS, Chen MY, Liu HH, et al. A novel dengue vaccine candidate that induces cross-neutralizing antibodies and memory immunity. *Microbes Infect*. 2009 Feb; 11(2):288–95. [PubMed: 19114121]
56. Przybylski D, Rost B. Consensus sequences improve PSI-BLAST through mimicking profile-profile alignments. *Nucleic Acids Res*. 2007; 35(7):2238–46. [PubMed: 17369271]

57. Schein CH, Ozgun N, Izumi T, Braun W. Total sequence decomposition distinguishes functional modules, “molegos” in apurinic/apyrimidinic endonucleases. *BMC Bioinformatics*. 2002 Nov 25;3:37. [PubMed: 12445335]
58. Schein CH, Zhou B, Oezguen N, Mathura VS, Braun W. Molego-based definition of the architecture and specificity of metal-binding sites. *Proteins*. 2005 Jan 1; 58(1):200–10. [PubMed: 15505785]
59. May AC. Optimal classification of protein sequences and selection of representative sets from multiple alignments: application to homologous families and lessons for structural genomics. *Protein Eng*. 2001 Apr; 14(4):209–17. [PubMed: 11391012]
60. Lok SM, Kostyuchenko V, Nybakken GE, Holdaway HA, Battisti AJ, Sukupolvi-Petty S, et al. Binding of a neutralizing antibody to dengue virus alters the arrangement of surface glycoproteins. *Nature Structural & Molecular Biology*. 2008 Mar; 15(3):312–7.
61. Zhang S, Bovshik EI, Maillard R, Gromowski GD, Volk DE, Schein CH, et al. Role of BC loop residues in structure, function and antigenicity of the West Nile virus envelope protein receptor-binding domain III. *Virology*. 2010 Jul 20; 403(1):85–91. [PubMed: 20447672]
62. Zulueta A, Martin J, Hermida L, Alvarez M, Valdes I, Prado I, et al. Amino acid changes in the recombinant Dengue 3 Envelope domain III determine its antigenicity and immunogenicity in mice. *Virus Res*. 2006 Oct; 121(1):65–73. [PubMed: 16781791]
63. Mandell RB, Koukuntla R, Mogler LJ, Carzoli AK, Holbrook MR, Martin BK, et al. Novel suspension cell-based vaccine production systems for Rift Valley fever virus-like particles. *J Virol Methods*. 2010 Nov; 169(2):259–68. [PubMed: 20655330]
64. Reichert E, Clase A, Bacetty A, Larsen J. ALPHAVIRUS ANTIVIRAL DRUG DEVELOPMENT: SCIENTIFIC GAP ANALYSIS AND PROSPECTIVE RESEARCH AREAS. *Biosecur Bioterror*. 2009 Dec; 7(4):413–27. [PubMed: 20028250]
65. Binhua L, Ma L, Ball TB, Xiaojian Y, Van Domselaar G, Cuff WR, et al. Systematic Analysis of Host Immunological Pressure on the Envelope Gene of Human Immunodeficiency Virus Type 1 by an Immunobioinformatics Approach. *Current HIV Research*. 2008; 6(4):370–9. [PubMed: 18691035]

Highlights

a PCP-consensus antigen was designed based on the physicochemical properties of 671 unique DENV sequences

two regions of maximum variability were adjusted to obtain recognition by antibodies against DENV-4

the optimized protein was recognized by polyclonal antibodies against all four DENV serotypes

Two amino acids in the N-terminus were essential for good binding to all four serotype specific antibodies

Figure 1.

A) PCP-consensus sequences (DENV1c-4c) for the four individual DENV serotypes, based on 671 sequences from Flavitrack, were used to derive overall consensus sequences by three methods (B). The most variable regions of the DENV sequences are boxed; alignments are colored according to amino acid types. The sequence DENV-703-4 is a wild type sequence of the DENV-4 serotype that was added to correct for the bias against DENV-4 in the original alignment for consensus sequences 6B and 7B. B) Consensus sequences generated with three programs: the Jalview applet of Clustal W (**5**), Ctree of the original four sequence alignment (**6**), or with the added DENV-703-4 (**6B**), and PCP-consensus sequences of the original alignment (**7**), the alignment with the added DENV-703-4 (**7B**), the 7B sequence altered to reflect the hypervariable region (56–60) of DENV-4 (**8**). The last sequence, **7P8**, is a recombinant of the N-terminus of **7** and the C-terminus of **8**, cut and joined at a common Pst1 site around residue 51. The two amino acids in the N-terminus, essential for good recognition by all DENV antisera, are bold and indicated by arrows.

Figure 2.

ELISA assay showing the reactivity of three initial PCP-consensus DENV antigens (see Figure 1B for details) with murine immune ascites fluid (MIAF) from mice infected with DENV of each of the four types. While the initial consensus, **7** (first bar in each series) reacted as well as the individual wild type antigens (last bar in each series; WT means an individual wild type envelope protein domain 3 for each of the four types) for DENV-1-3, it lacked significant reactivity with antibodies generated against DENV-4. Adjusting the alignment for bias, and replacing the DENV-4 hypervariable region (underlined in Figure 1) (**8**, third box in each series) enhanced reactivity with the antibodies generated in response to DENV-4 inoculation, but greatly reduced reactivity with those induced by DENV1-3.

Figure 3.

Altering two amino acids (arrows in Figure 1B) enhances the reactivity of a multivalent consensus protein with polyclonal and a monoclonal antibodies against DENV types. The **7P8** protein (fourth bar in each series) is a recombinant with the first 51 amino acids of **7** (1st bar) and the 52–112 of **8** (3rd bar). The wt DENV-4 antigen was only recognized by MIAF-4 antibodies, and to a small extent by the MIAF-2. All of the PCP consensus proteins, but not the DENV-4 wild type protein, bound to the GTX29202 “pan-DENV” monoclonal antibody. The distantly related Deer Tick virus antigen (DTV), used as a control, was not significantly recognized by any of the antisera.

Figure 4.

Models of the PCP-consensus proteins **7** and **7P8** showing the residue changes between them (these side chains are shown space filling) are in two zones of variability. The hypervariable region (box in Fig. 1B) for both regions is shown as maroon or green, respectively; the two amino acids (K19, D42) essential for recognition by the DENV MIAF-antibodies are shown in khaki. Note these are both small amino acids (S, G) in DENV-4c, suggesting they may have arisen as “escape mutants”. The other amino acid differences between the initial and final consensus sequences are colored according to atom type (blue:N; red:O; black:C, gray:H).