Available online at http://www.journalcra.com

International Journal of Current Research
Vol. 6, Issue, 08, pp.7746-7753, August, 2014

INTERNATIONAL JOURNAL OF CURRENT RESEARCH

ISSN: 0975-833X

RESEARCH ARTICLE

INTEGRATION OF *IN-SILICO* APPROACHES FOR DESIGNING B CELL EPITOPES FOR DISEASE DIAGNOSIS

[1]Nehul Saxena, [2]Anubhav Pandey, [3]Aparna Jha Ahuja and [1*]Vijay Kumar

[1]International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110067, India
[2]Cryobanks International India Pvt. Ltd. 129-Pace City-1, Sector-37, Gurgaon-122001, Haryana, India
[3]Head of Medical Affairs - PAS EMA, BD Diagnostics, Dubai United Arab Emirates

ARTICLE INFO

Article History:
Received 17th May, 2014
Received in revised form
05th June, 2014
Accepted 20th July, 2014
Published online 06th August, 2014

Key words:
Calcium-binding fusion protein,
Dengue virus,
Diagnostics,
Serotypes.

ABSTRACT

In this technologically savvy world, we witness mushrooming of sophisticated line of softwares that are extensively used for predicting epitopes towards developing vaccines and therapeutics. However, most softwares also generate bulky junk data along with data of interest. Here we devised a combined usage of B cell epitope prediction softwares- ABCpred along with additional online tools such as CLUSTAL W and Sequence Manipulation Suite for stringent selection of B cell epitopes. Dengue was used as model disease since the present diagnostic kits give false positives owing to cross-reactivity with other flaviviruses. The *in silico* predicted 240 unique B cell epitopes of dengue viruses were reduced to 87 serotype-specific epitopes using combination of softwares. Three best ranking epitopes of each dengue serotype were functionally validated by cloning corresponding DNA sequences in thep CBP vector and expressing these epitopes as calcium-binding protein fusion proteins. Western blot analysis and indirect ELISAs of purified fusion proteins confirmed that nearly fifty percent of the epitopes were in close agreement with the predictions.Thus, stringent selection of epitopes using a combination of prediction softwares is likely to identify a compact library of B cell epitopes useful for developing more specific diagnostics in future.

INTRODUCTION

Prognosis of pathogenic diseases is a complicated multi-step process. Diagnosis usually begins with extrapolation of exhibited symptoms to a set of possible diseases. This is usually followed by serological tests to assess White blood cell, platelet, haemoglobin count, etc. accompanied by antigen/ antibody (IgG/ IgM) detection tests to confirm the nature of disease and the causative organism. Further, PCR tests for nucleic acid amplification (DNA/RNA) with pathogen specific primers can also be performed to detect the exact strain of infecting pathogen. Antigen detection can be carried out by performing Enzyme linked immunosorbent assays (ELISA) or Immuno histochemistry (IHC). Since, IHC is an expensive and invasive technique and requires skilled technician, ELISA tests are generally preferred over it for antigen detection. Disease specific epitope selection is of paramount importance for the development of a sensitive ELISA for disease diagnosis. Traditional epitope selection methods are usually cumbersome and require large resources. However, the advent of technologies related to B cell epitope prediction and databases such as ABCpred (Saha and Raghava, 2006), Immune epitope database (IEBD) (Vita *et al*., 2010), BepiPred 1.0 (Erik *et al*., 2006), IgPred (Gupta *et al*., 2013), BcePred (Saha and Raghava, 2004), etc. could aid to the development of new therapeutics and diagnostic kits. Despite the undeniable merits of these epitope prediction tools, the data generated by these softwares can be too bulky and this impractical to handle in small setups. The elaborate and carefully designed epitope prediction softwares over the years have incorporated several selection criterias like hydophilicity, homology, polarity, turns, mobility, flexibility, hidden Markov model, propensity scale methods, neural networks and many more. However these softwares donot provide users the flexibility to combine features of two or more softwares to further reduce the size of predicted data. Keeping these issues in mind, in the present study we devised a novel approach for selecting epitopes fordisease detection using a combination of freely available online tools. We chose Dengue as the disease model for this study.

Dengue is endemic to the tropical and subtropical areas of the world including India. It poses threat to 3.97 million people residing in these regions (Bhatt *et al*., 2013; Brady *et al*., 2012). Uptil 1970 dengue was epidemiologically restricted to 9 countries however now it is endemic to more than 128 countries**.** The World Health Organization (WHO) has estimated a 30% increase in dengue incidence in the last 50 years (WHO factsheet, 2012). The most recent dengue epidemic in the Indian subcontinent occurred in 2006 and caused an economic loss of approximately 27.4 million US

Corresponding author: Vijay Kumar,
International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi-110067, India.

dollars (Garg *et al*., 2008). Dengue Virus (DENV) is the causative agent of symptomatic infections ranging from Dengue Fever (DF) to Dengue Hemorrhagic Fever (DHF) in humans (Boonak *et al*., 2008; Chen *et al*., 2007; Guzman *et al*., 2010; Halstead, 2007; Huisman *et al*., 2009). The principal transmission vectors for this virus are the mosquitoes of *Aedes* genre, especially *Aedes aegypti and Aedes albopictus* (Halstead, 2008). Dengue shock syndrome (DSS) or DHF are the clinically more significant dengue infection and occurs on secondary encounter with heterologous dengue virus (Dejnirattisai *et al*., 2010; Rothman, 2010). There is nearly 75-80% homology among the four dengue virus serotypes, viz., DENV1 through DENV4 (Weaver and Vasilakis, 2009; Rodenhuis *et al*., 2010).

DENV is a positive-sense RNA virus that belongs to the *Flaviviridae* family (Harris *et al*., 2006). Its genome is positive sense, single stranded, 10.3 kDa RNA. The genome has a 5'-1,7 methyl guanosine cap, a 5' UTR, a single open reading frame and a 3' UTR. The open reading frame is translated onto 10 proteins, 3 structural proteins - Capsid, Pre-M and Envelope and 7 non-structural proteins - NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5 (Guzman *et al*., 2010). The human antibody response to dengue involves a polyclonal response to primary and secondary infections upon infection with 4 different DENV serotypes. The serum antibodies against dengue virus are detectable for structural proteins- Capsid, pr-M and Envelope proteins and non-structural proteins- NS1, NS3 and NS5 (though NS3 and NS5 are produced in low titers). Upon primary encounter with DENVs, IgMs are produced 4-5 days after the onset of fever while IgG1 and IgG3 production is initiated a week after the onset of fever but peaks only several weeks after infection (Halstead, 2007). The IgMs generated upon primary DENV infection may persist for 3 months however IgGs may persist for decades. Upon seconday infection, stimulation of B-cell memory leads to a rapid rise in DENV-specific IgGs detectable at very early stage in infection (Innis *et al*., 1989). Antibodies generated in response to dengue infection are mostly cross-reactive with heterologous dengue serotypes or other virus from flaviviridae family like Japanese encephalitis virus (JEV), West Nile fever virus (WNV) and Chikugunya virus (CGV) (due to high homology). Hence detection of the smaller fraction of dengue serotype specific and sensitive neutralizing antibodies is crucial for efficient dengue diagnosis (Dejnirattisai *et al*., 2010; Rothman, 2010; Ubol and Halstead, 2010).

Dengue specific IgG/IgM detection forms the basis of Dengue diagnostic for numerous kits. In addition, some other tests also use monoclonal antibodies for the detection of NS1 antigen in patient sera. (Bessoff *et al*., 2008; Sekaran *et al*., 2008; Wang and Sekaran, 2010). A major problem encountered with these kits is false positive case background with human sera samples (Sekaran *et al*., 2008; Hunsperger *et al*., 2009; Lima *et al*., 2010). Considering the accurate and efficient diagnosis of dengue crucial for pathogenesis surveillance and clinical care, we used a combination of bioinformatic tools-ABCpred for B-cell epitope prediction; Clustal W for multiple sequence alignmentand sequence manipulation suite for gravy index, to generate a compact database of unique dengue-specific B cell epitopes from the capsid, envelope, membrane and NS1 regions of the viral polyprotein. The rationale for using sequence homology and hydropathy as the selection parameters was to prevent cross-reactivity of the chosen epitopes with heterologous Dengue serotypes and other Flaviviridae (like West Nile Fever Virus, Japanese Encephalitis Virus and Chikungunya Virus) to reduce false positive case detection and for the selection of exposed epitopes (and not hidden ones) capable of eliciting immune response, respectively. In support, to establish the functional relevance of the database we expressed the three best ranking epitopes of each dengue serotype as recombinant fusion proteins and evaluated them for their efficacy in sero-diagnosis. We found that fifty percent of the three best ranking epitopes worked in close agreement with our final database. Thus, B-cell epitope prediction software-ABCpred when used along with other tools specific for adjunct selection parameters drastically reduced the size of data and also resulted in a more stringent selection of novel B cell epitopes.

MATERIALS AND METHODS

In silico study

We obtained the polyprotein sequence of the four pathogenic Dengue serotypes from NCBI in *FASTA* format. Further, we used the ABCpred software (Saha and Raghava, 2006) for the prediction of probable B-cell epitopes of 20 amino acid window (size) in the polyprotein sequence of all the four dengue serotypes. Multiple sequence alignment using Clustal W was performed for the four polyprotein sequence of DENV with the corresponding protein sequence of CGV, JEV and WNV (sequence extracted from NCBI) to differentiate between homologous peptides (similar or identical to peptide regions in other DENVs, CGV, JEV and WNV) capable of exhibiting cross-reactivity from non-homologous (dissimilar) regions. The dissimilar amino acids revealed by Clustal W were highlighted using Netgear software and threshold value of 12 unique amino acids or more out of the window of 20 amino acid (i.e., 60% or above dissimilarity) was set to differentiate the non-homologous epitopes from the homologous peptides. Thereafter, the conserved peptides were excluded from the study. On the assumption that most of the secreted proteins would be hydrophilic, the grand average of hydropathy (GRAVY using Sequence manipulation suite: Protein GRAVY (Stothard, 2002)) value for the protein sequences was used to determine the hydrophilicity of proteins. The hydrophobic proteins (with positive average hydropathy values) were also excluded from the study. ExPASY-reverse translate tool was used to convert the peptide sequence into nucleotide sequence. These nucleotide sequences and their reverse complement (generated by reverse complement tool) were used for the oligo designing using IDT–OligoAnalyzer 3.1 (http://www.idtdna.com/analyzer/applications/oligoanalyzer/). The oligos were cloned in The pCBP expression vector (Reddi *et al*., 2002) for further studies.

Chemical synthesis of deoxyoligonucleotides and their cloning

Sixty base pair long oligonucleotide sequences representing 20 amino acid individual epitope segments were designed for custom synthesis. Each pair of oligonucleotides carried EcoRI

and BamHI overhangs to facilitate their cloning in pCBP vector (Figure 1). The clones were identified by PCR by using T7 promoter primer 5'-GTAATACGACTCACTATA-3' (IMPERIAL LIFE SCIENCES, Gurgaon, India) as the forward primer and the antisense deoxyoligonucleotide of the individual epitope as the reverse primer.

recombinant proteins were purified by two step purification which involved boiling of the total cell lysate to separate calcium binding fusion protein fromthe thermolabile proteins, in the supernatant (a property unique to the calcium binding protein of present in the vector) followed by affinity

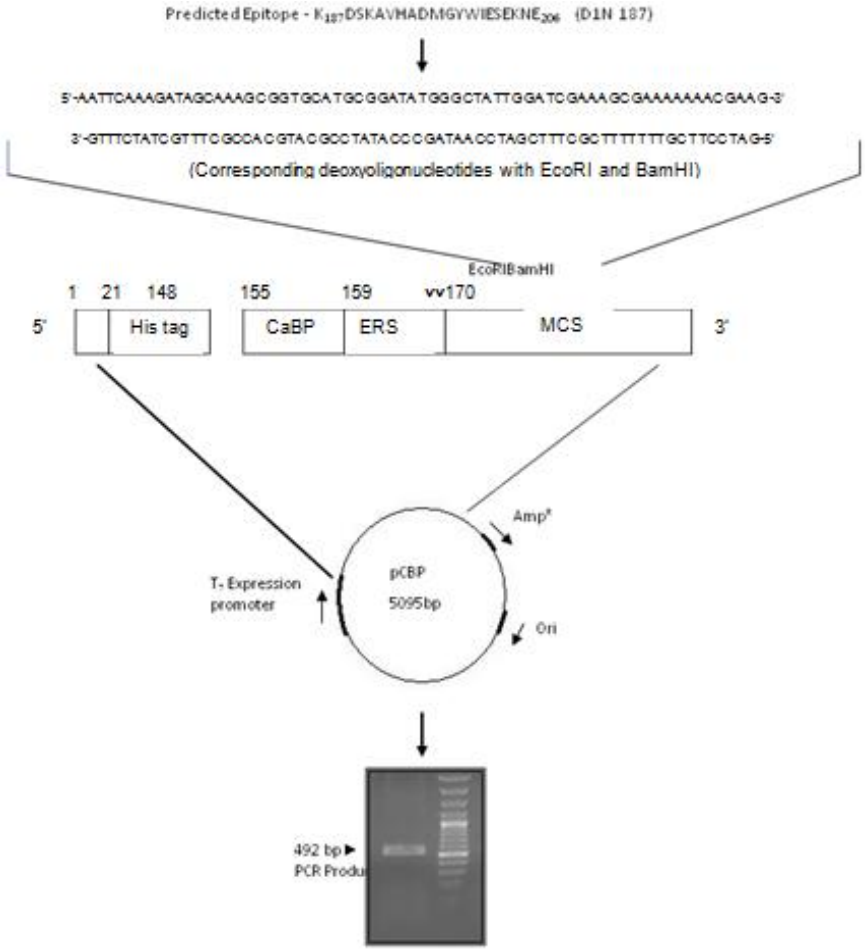

Figure 1. Cloning strategy for dengue epitopes in the pCBP prokaryotic expression vector. The pair of oligonucleotides corresponding to each dengue epitope were annealed and cloned between EcoR1 and BamH1 sites of pCBP vector

Identification of Dengue positive patient sera and naive patient sera

Sera samples obtained from Super Religare Laboratory Limited, Gurgaon, were confirmed as dengue positive by using Dengue Duo IgM/ IgG capture ELISA kit (PANBIO, Inverness Medical Innovations, Australia). We used Dengue-DUO IgG-IgM/NS1 detection kits for the re-confirmation of Dengue positive serum samples (STANDARD DIAGNOSTICS, Inc., South Korea). The samples that gave negative results in the test were considered uninfected.

Bacterial expression of Dengue-CBP fusion proteins

All clones representing individual epitopes were transformed into *Escherichia coli* [Strain: BL21(DE3)] and the fusion genes were expressed as per the protocol described earlier [30]. The purification by Ni-NTA beads (QUIAGEN, Germany), as described by Reddi *et al*., 2002. The CaBP fusion proteins were concentrated using Vivaspin columns (GE HEALTHCARE). The concentration of purified protein was determined by Bradford assay (BIO-RAD, USA). The fusion proteins were validated by western blotting with the control (uninfected human sera sample) and Dengue positive human sera sample (described previously). For this, each purified protein (20µg) was run on 15% SDS-PAGE gel and then transferred onto nitrocellulose membrane. The blot was blocked overnight at 4°C in blocking solution (5 % skimmed milk powder in 1X phosphate buffer saline containing 0.1% tween-20 (PBST)) and was then incubated with human sera sample (as primary antibody used at a dilution of 1:500 in blocking) at 37°C for 1 h. After washing the blots five times (for five minutes each with PBST) the blot was incubated with

anti-Human IgG-horseradish peroxidase (CALBIOCHEM, Switzerland) as the secondary antibody (at 1:10,000 dilution in blocking solution) for 40 min at 37°C in blocking solution. The blot was again washed five times and the protein bands were visualized using the Western blotting luminol reagent (SANTA CRUZ BIOTECHNOLOGY, USA).

Immunoglobulin G antibody detection ELISA (IgG-ELISA)

Medium Binding Nunc 96 well plates were coated with purified proteins at a serial dilution of 10, 5, 2.5 and 1.25µg/ml of Carbonate/bicarbonate coating buffer (pH 9.6) for 1 h at 37°C in a humidified chamber. After washing the plate 3 times with 0.5% tween in water, these coated proteins were incubated with blocking solution for 1hr followed by dengue positive human sera samples at 1:250 dilution in blocking solution (2% BSA in 1x PBST) at 37°C for 1hr in humidified chamber. Thereafter, plate was washed three times and bound antibodies were probed with anti-human IgG conjugated to horseradish peroxidase (at 1:10,000 dilution in PBST) and Ortho-phenylenediamine (OPD) was used as the substrate for the detection of signals. The samples absorbance was measured at 492 nm using PERKIN ELMER ELISA plate reader and soft Max Pro 4.8 software. The full length recombinant DV-2 envelope protein was used as positive control in these experiments (kindly provided by Dr. Navin Khanna, ICGEB, New Delhi). All the experiments were carried out in triplicates.

Ethics statement

The human sera samples (n = 60) used in this study were 'left-over' samples from the patients visiting Super Religare Diagnostic Laboratory, Gurgaon for check-ups and diagnosis. The samples were collected in accordance with the Declaration of Helsinki (2000) of the World Medical Association after informed written consent of the patients. All sera were pre-screened for anti-DENV antibodies and NS1 antigen using the PANBIO Dengue IgG/IgM capture ELISA test and Standard Diagnostics Dengue-DUO IgG-IgM/NS1 detection kits. The Institutional Review Committee and Institutional Review Board of the International Centre for Genetic Engineering and Biotechnology (ICGEB) approved this study (IEC/IRB No.ICGEB/IEC/2011/02, version 2).

RESULTS

Development of B-cell Epitope Library of DENVs

ABC pred algorithm was used to predict B-cell epitope from the capsid, envelope, pre-membrane and NS1 regions of DENV1, DENV2, DENV3 and DENV4. 240 unique sequences were retrieved as the primary data by using this algorithm. Of the 240 unique peptides, after exclusions based on the Multiple sequence alignment (using Clustal W) and GRAVY index calculation, a set of 17, 26, 19 and 25 peptide sequences (total 87) derived from of DENV1, DENV2, DENV3 and DENV4, respectively formed a library of dengue specific epitopes (Table 1). All the selected epitopes were hydrophilic in nature with a net negative GRAVY value. All the selected sequence were at least 60 percent dissimilar from corresponding sequences of other DENVs and other flavivirus (West Nile Fever Virus, Japanese Encephalitis Virus and Chickengunya Virus) thereby reducing their chances of cross-reactivity.

Identification of CBP-dengue epitope fusion recombinants

The deoxyoligonucleotides corresponding to the three best ranking epitopes of each Dengue serotype were cloned into EcoRI and BamHI restriction sites of pCBP vector [30]. The clones were selected on ampicillin plates and were confirmed by PCR. The positive clones were selected by performing PCR amplification of a ~492bp fragment (Figure 2) by using T7 Promoter forward primer and an intrinsic reverse primer specific for individual clones.

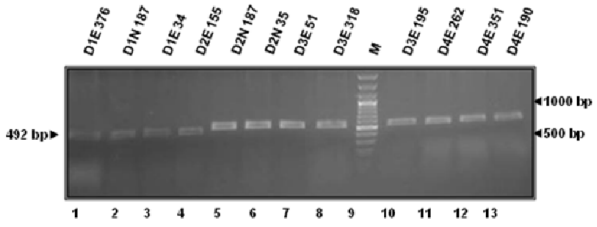

Figure 2. Selection of positive CBP- Dengue epitope fusion recombinant by PCR. The deoxyoligonucleotides corresponding the three best ranking epitopes from four dengue serotypes were cloned into the prokaryotic expression vector pCBP

Table 1. Dengue Serotype specific B cell epitope library

1. DENV-1 epitopes

S.No.	Rank	Epitopes	Hydropathy index
1	3	NS1 ^{187}KDSKAVHADMGYWIESEKNE^{206} (DIN 187)	-1.295
2	4	Envelope ^{376}ESYIVVGAGEKALKLSWFKK^{395} (D1E 376)	-0.015
3	6	Envelope ^{34}TMAKDKPTLDIELLKTEVTN^{53} (D1E 34)	-0.490
4	8	Envelope^{228}SQETWNRQDLLVTFKTAHAK^{247}(D1E228)	-0.930
5	10	Envelope^{339}SSQDEKGVTQNGRLITANPI^{358}(D1E 339)	-0.800
6	12	Envelope^{328}EGTDAPCKIPFSSQDEKGVT^{347}(D1E 34)	-0.825
7	13	NS1^{168}KLRDSYTQVCDHRLMSAAIK^{187}(D1N168)	-0.490
8	13	Capsid ^{83}KKEISNMLNIMNRRKRSVTM^{102}(D1C 83)	-0.940
9	13	Envelope^{147}GDQHQVGNETTEHGTTATIT^{166}(D1E 147)	-1.115
10	14	Envelope^{157}TEHGTTATITPQAPTSEIQL^{176}(D1E 157)	-0.470
11	15	Envelope^{75}CPTQGEATLVEEQDTNFVCR^{94}(D1E 75)	-0.565
12	18	NS1 ^{90}DVSGILAQGKKMIRPQPMEH^{109}(D1N90)	-0.585
13	21	Envelope^{315}ETQHGTVLVQVKYEGTDAPC^{334}(D1E 315)	-0.485
14	24	Envelope^{117}CAKFKCVTKLEGKIVQYENL^{136} (D1E 117)	-0.095
15	24	Envelope^{57}LRKLCIEAKISNTTTDSRCP^{76}(D1E 57)	-0.460
16	27	Envelope^{139}SVIVTVHTGDQHQVGNETTE^{158}(D1E 139)	-0.525
17	31	NS1 ^{160}IFTTNIWLKLRDSYTQVCDH^{179}(D1N160)	-0.230

2. DENV-2 epitopes			
1	3	Envelope [155]GKHGKEIKITPQSSITEAEL[174] (D2E 155)	-0.760
2	4	NS1 [35]ESPSKLASAIQKAHEEGICG[54] (D2N 35)	-0.455
3	5	NS1 [187]KDNRAVHADMGYWIESALND[206] (D2N 187)	-0.810
4	5	Envelope [210]QWFLDLPLPWLPGADTQGSN[229](D2E210)	-0.330
5	6	NS1 [14]GSGIFITDNVHTWTEQYKFQ[33](D2N14)	-0.585
6	6	Envelope[113]VTCAMFTCKKNMKGKVVQPE[132](D2E113)	-0.175
7	8	Envelope [313]ETQHGTIVIRVQYEGDGSPC[332](D2E313)	-0.580
8	9	Envelope[374]DSYIIIGVEPGQLKLNWFKK[393](D2E374)	-0.150
9	9	Envelope [330]SPCKIPFEIMDLEKRHVLGR[349](D2E330)	-0.345
10	11	Envelope [74]PTQGEPSLNEEQDKRFVCKH[93](D2E74)	-1.590
11	12	Envelope [164]TPQSSITEAELTGYGTVTME[183](D2E164)	-0.330
12	12	Envelope [51]QPATLRKYCIEAKLTNTTTD[70](D2E51)	-0.725
13	13	NS1 [118]KAKMLSTESHNQTFLIDGPE[137](D2N118)	-0.745
14	13	NS1 [24]HTWTEQYKFQPESPSKLASA[43](D2N24)	-1.200
15	13	Envelope [229]NWIQKETLVTFKNPHAKKQD[248] (D2E229)	-1.330
16	14	NS1 [126]SHNQTFLIDGPETAECPNTN[145](D2N126)	-0.940
17	16	Envelope [145]GEEHAVGNDTGKHGKEIKIT[164](D2E145)	-1.180
18	17	NS1 [88]TGDIKGIMQAGKRSLQPQPT[107](D2N88)	-0.820
19	18	Envelope [321]IRVQYEGDGSPCKIPFEIMD[340] (D2E321)	-0.355
20	19	NS1 [168]KLREKQDVFCDSKLMSAAIK[187] (D2N168)	-0.430
21	20	Envelope [126]GKVVQPENLEYTIVITPHSG[145](D2E126)	-0.160
22	22	Membrane[36]LGELCEDTITYKCPFLRQNE[55](D2M36)	-0.520
23	25	Envelope [62]AKLTNTTTDSRCPTQGEPSL[81](D2E62)	-0.960
24	28	Envelope [218]PWLPGADTQGSNWIQKETLV[237] (D2E218)	-0.565
25	29	Envelope [137]TIVITPHSGEEHAVGNDTGK[156] (D2E137)	-0.540
26	32	NS1 [160]VFTTNIWLKLREKQDVFCDS[179](D2N160)	-0.215
3. DENV-3 epitopes			
1	4	Envelope [51]QLATLRKLCIEGKITNITTD[70] (D3E 51)	-0.015
2	5	Envelope [318]LIKVEYKGEDAPCKIPFSTE[337] (D3E 318)	-0.400
3	6	Envelope [195]LLTMKNKAWMVHRQWFFDLP[214] (D3E 195)	-0.165
4	8	Envelope[150]VGNETQGVTAEITSQASTAE[169](D3E150)	-0.395
5	11	Envelope[70]DSRCPTQGEAILPEEQDQNY [89](D3E89)	-1.490
6	12	NS1 [92]TGVLEQGKRTLTPQPMELKY[111](D3N92)	-0.810
7	12	Envelope [296] YAMCLNTFVLKKEVSETQHG[215](D3E296)	-0.195
8	12	Membrane [42]DTVTYKCPHITEVEPEDIDC[61](D3M42)	-0.615
9	13	NS1 [111]YSWKTWGLAKIVTAETQNSS[130](D3N111)	-0.510
10	13	Envelope [337]EDGQGKAHNGRLITANPVVT[356](D3E337)	-0.650
11	13	Envelope[32]TMAKNKPTLDIELQKTEATQ[51](D3E32)	-0.975
12	17	Envelope[115]CAKFQCLESIEGKVVQHENL[134](D3E115)	-0.155
13	18	NS1 [63]NLLWKQIANELNYILWENDI[82](D3N161)	-0.225
14	18	Membrane[57]EDIDCWCNLTSTWVTYGTCN[76](D3M57)	-0.230
15	21	Envelope[188]LDFNEMILLTMKNKAWMVHR[207](D3E188)	-0.035
16	22	NS1 [124]AETQNSSFIIDGPSTPECPS[143](D3N124)	-0.560
17	23	Envelope [329] PCKIPFSTEDGQGKAHNGRL[348](D3E329)	-1.000
18	31	Envelope[77]GEAILPEEQDQNYVCKHTYV[96](D3E77)	-0.795
19	34	Envelope[304]VLKKEVSETQHGTILIKVEY[323](D3E 304)	-0.180
4. DENV-4 epitopes			
1	4	Envelope [262]ALAGATEVDSGDGNHMFAGH[281] (D4E 262)	-0.180
2	6	Envelope [351]ISSTPLAENTNSVTNIELEP[370] (D4E 351)	-0.305
3	6	Envelope [190]IDFNEMILMKMKKKTWLVHK[219] (D4E 190)	-0.275
4	8	Membrane [108]TWMSSEGAWKHAQRVESWIL[127](D4M 108)	-0.515
5	10	Envelope [270]DSGDGNHMFAGHLKCKVRME[289] (D4E 270)	-0.790
6	10	Envelope [145]GDTHAVGNDTSNHGVTAMIT[164](D4E 145)	-0.340
7	11	Envelope [218]PWTAGADTSEVHWNYKERMV[237](D4E 218)	-0.950
8	13	NS1 [158]WGKAKIFTPEARNSTFLIDG[177](D4N 158)	-0.315
9	13	Membrane [56]VNTEPEDIDCWCNLTSTWVM[75](D4M 56)	-0.185
10	14	Envelope [68]TTATRCPTQGEPYLKEEQDQ[87](D4E 68)	-1.625
11	15	NS1 [133]VKGVLTKGKRALTPPVSDLK[152] (D4N 133)	-0.200
12	16	NS1 [320]CPGTTVTIQEDCDHRGPSLR[339] (D4N 320)	-0.780
13	17	NS1 [113]NELNYVLWEGGHDLTVVAGD[132] (D4N 113)	-0.125
14	17	Envelope [163]ITPRSPSVEVKLPDYGELTL[182](D4E 163)	-0.205
15	18	Envelope [83]EEQDQQYICRRDVVDRGWGN[102](D4E 83)	-1.630
16	19	NS1 [73]KFQPESPARLASAILNAHKD[92](D4N 73)	-0.610
17	20	Membrane[43]MCEDTVTYKCPLLVNTEPED[62](D4M 43)	-0.430
18	21	NS1 [148]VSDLKYSWKTWGKAKIFTPE[167](D4N 148)	-0.680
19	21	Envelope [325]YEGAGAPCKVPIEIRDVNKE[344](D4E 325)	-0.580
20	25	Envelope [296]MSYTMCSGKFSIDKEMAETQ[315](D4E 296)	-0.500
21	28	NS1 [199]GFGMFTTNIWMKFREGSSEV[218] (D4N 199)	-0.155
22	29	Envelope[118]FSCSGKITGNLVQIENLEYT[137](D4E 118)	-0.020
23	30	Envelope [154]TSNHGVTAMITPRSPSVEVK[173](D4E 154)	-0.295
24	33	Envelope [75]TQGEPYLKEEQDQQYICRRD[94](D4E 75)	-1.945
25	34	Envelope [35]QGKPTLDFELTKTTAKEVAL[54] (D4E 35)	-0.425

The positive CBP-Dengue epitope fusion clones were isolated and inoculated from ampilcillin plates into Luria Broth supplemented with ampicillin and grown overnight at 37°C in shaker incubator. Their plasmid was isolated by plasmid isolation columns. Recombinant clones were identified by PCR using upstream T7 promoter as forward primer and an intrinsic reverse primer of individual clone.

Specificity of CBP-Dengue epitope fusion proteins for dengue patient sera

Purified recombinant fusion proteins corresponding to the twelve best ranking epitopes were separated on SDS page gel. Further, they were probed with Dengue positive and negative pooled serum samples followed by anti-human IgG conjugated to horseradish peroxidise. Out of the twelve purified proteins, only six viz. D1N 187 (^{187}KDSKAVHADMGYWIESE KNE ^{206}), D3E 51 (^{51}QLATLRKLCIEGKITNITTD^{70}), D3E 318 (^{318}LIKVEYKGEDAPCKIPFSTE^{337}), D3E 195 (^{195}LLTMKNKAWMVHRQWFFDLP^{214}), D4E 262 (^{262}ALAGATEVDSGDGNHMFAGH^{281}) and D4E 190 (^{190}IDFNEMILMKMKKKTWLVHKK^{219}) preferentially reacted with dengue-positive human sera samples (Figure 3). None of the twelve clones reacted with dengue negative human sera samples (Data not shown). Not surprisingly, the full length Dengue Virus 2 envelope protein (DV2 used as positive control) also reacted with the dengue positive pooled human sera samples.

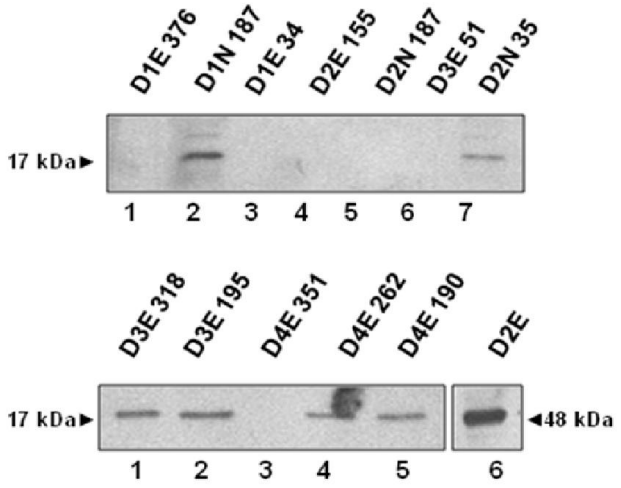

Figure 3. Expression analysis of recombinant dengue epitope clones by Western blotting. Selected Clones (D1N 187, D1E 34, D1E 376, D2E 155, D2N 187, D2N 35, D3E 51, D3E 318, D3E 195, D4E 351, D4E 262 and D4E 190) were transformed in BL21 (DE3) *Escherichiacoli* cells and induced with IPTG for 6h. The cells were lysed and protein samples were electrophoressed in 15% SDS-polyacrylamide gel, transferred and western blotted using pooled dengue patient sera.

M- pre-stained protein ladder, D2E- full length dengue envelope of dengue serotype 2.

Sensitivity of CBP-Dengue epitope fusion protein towards dengue febrile stage sera

Sample

The six reactive Dengue specific recombinant fusion proteins were further experimentally validated by an indirect ELISA. For this, the purified recombinant fusion proteins were coated in 96 well medium binding ELISA plates and were probed with Dengue positive pooled serum samples followed by anti-human IgG conjugated to horseradish peroxidise. OPD was used as sunstrate in ELISA. The six positive fusion protein lysates showed a marked increase (atleast 3 fold) in the absorbance at 492 nm when probed with pooled dengue patient sera in comparison to Non-reactive Dengue epitope D1E 34 which was used as the negative control (Figure 4A) and showed negligible sensitivity to dengue positive sera sample. The positive control, DV-2 showed the highest absorbance at 492nm. Upon serial dilution (from 10µg to 5 µg, 2.5 µg and 1.25 µg),all the six epitopes showed highest sensitivity to Dengue positive pooled sera samples at a dilution of 10ug/ml of coating solution (Figure 4B).

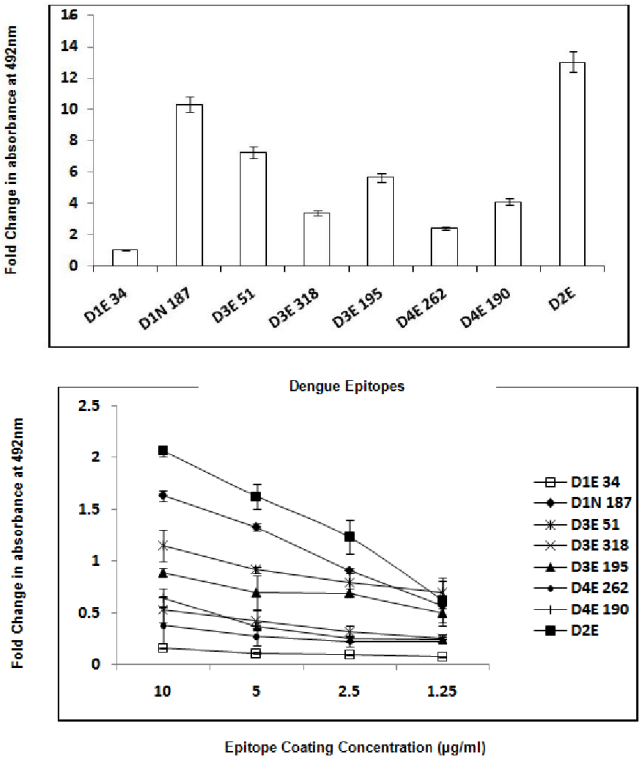

Figure 4. Sensitivity of recombinant CBP-dengue epitope fusion protein lysates towards naturally dengue infected pooled sera samples. A) Relative fold change in Absorbance of CBP-Dengue fusion proteins- D1N 187, D3E 51, D3E 318, D3E 195, D4E 262 and D4E 190 at 492 nm. The coated fusion protiens (Coating Concentration- 10µg/ml) were probed with pooled positive sera sample followed by peroxidase labelled secondary antibody and signals were detected using o-phenylenediamine as the substrate. (The t-test values of all the epitopes is ≤ 0.005) D1E 34 and D2E were used as the positive and the negative controls for the assay. B) Absorbance of

individual fusion protein at various coating dilutions (i.e. 10, 5, 2,5, 1.25 µg/µl). The values are average of three independent ELISA experiments and the error bars represent their respective standard deviation ±SD.

DISCUSSION

Several methods have been used to identify the antigenic epitopes and antibodies generated against structural epitopes, for the development of ELISA kits for disease diagnosis. These methods have helped in laying the foundation of diagnostic kits currently available in the market. The archaic methods used for epitope predictions were usually time consuming and utilized more resources. The recently available epitope prediction softwares on the contrary are highly efficient and cost-effective approach for epitope selection. Though these prediction softwares drastically the efforts needed by the researcher to select epitopes yet, a major drawback of using these software is dealing with the bulky junk data. Besides, it is cumbersome to experimentally validate such huge data sets. Now, we describe a systematic stepwise approach to overcome this limitation of the B cell epitope prediction software. We chose Dengue as our disease model as this disease is endemic to south-east Asia including India. We used ABCpred B-cell epitope prediction-software to fine map the epitopes on the genome sequence of the four dengue serotypes from the Capsid, Envelope, Pr-M and NS1 regions. ABCpred software uses recurrent neural network (Jordan network) for the prediction of B cell epitopes and claims to predict B cell epitopes with 65.93% accuracy. Our preliminary search using ABCpred predicted 240 novel B cell epitopes.

The Dengue diagnostic kits often suffer on account of specificity (due to cross-reactivity) and sensitivity (Bessoff *et al*., 2008; Sekaran *et al*., 2008; Wang and Sekaran, 2010; Hunsperger *et al*., 2009; Lima *et al*., 2010). This can be attributed to the fact that, anti-Dengue antibodies generated upon infection with DENVs are mostly capable of cross-reacting with heterologous DENV serotypes and other members of Flaviviridae family due to very high homology in their polypeptide sequences. In comparison, a very small pool of Dengue serotype specific antibodies are produced in response to Dengue infection. Hence, we reasoned that identification of non-homologous peptides (incapable of cross-reacting with other flavivirus and non-infecting Dengue serotype) could facilitate in more serotype specific diagnosis of dengue. Thus Multiple sequence alignment (using Clustal W) of polypeptide sequences of four Dengue serotypes along with Chickengunya, Japanese encephalitis virus (both endemic to India) and West nile fever (Non-endemic to India) was performed to discard the homologous sequences from the primary data generated using the ABCpred software. The combination of *in silico* approaches yielded 87 epitopes thereby reducing the size of data by approximately 2.8 folds.

As most of circulating antigens capable of eliciting host immune response have pre-dominantly hydrophilic peptide sequences, we further sieved the data based on the hydropathy indices using bioinformatic software- Sequence Manipulation Suite. To monitor the accuracy of our data we further extended our *in silico* study to determine the functional relevance of the database generated. For this, the oligonucleotides corresponding the three best ranking epitopes from the four dengue serotypes were cloned in pCBP vector and were expressed and purified as calcium Binding Fusion proteins. Further, the purified fusion proteins were functionally validated by western blotting and ELISA. In agreement with the results of our Dengue specific B-cell epitope database, nearly 50 percent epitopes tested in the pilot experiments (D1N 187, D3E 51, D3E 318, D3E 195, D4E 262 and D4E 190) reacted with dengue positive pooled serum but not with uninfected pooled serum samples thereby confirming the Dengue specific nature of these epitopes. Further, these six epitopes were used for the detection of anti-dengue antibodies in infected pooled serum samples using ELISA. All the six epitopes were found atleast three fold more sensitive than negative control D1E 34. Moreover the following three epitopes viz., D1N 187, D3E 51, D3E 195 were found atleast five fold more sensitive than D1E 34 to the dengue positive pooled serum. The full-length dengue serotype 2 envelope protein (D2E), which was used as the positive control for western blots and ELISA was shown to detect the dengue positive pooled serum sample with a much higher efficiency than individual epitopes. This could be attributed to the presence multiple reactive epitopes hidden in the 420 amino acid long recombinant protein. Thus, the serotype-specific epitopes could be tandemly stitched together to form a multi-epitope polypeptide (Kumar *et al*, 1992) to further increase their efficacy in detection of Dengue serotype specific immunoglobulins in patient sera sample. Thus, our combinatorial *in-silico* approach for the prediction of B-cell epitopes can act as a novel tool in selecting new epitopes for the development of more specific and sensitive diagnostic kit for Dengue and can have an far a reaching impact on clinical care, epidemiological studies, and therapeutic development.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

This work was supported by the core grant of ICGEB, New Delhi. We thank Dr. Navin Khanna (Group leader, RGP group, ICGEB, New Delhi) for the recombinant Dengue serotype 2 envelope purified protein provided to us.

REFERENCES

Bessoff K, Delorey M, Sun W and Hunsperger E. 2008. Comparison of two commercially available dengue virus (DENV) NS1 capture enzyme-linked immunosorbent assays using a single clinical sample for diagnosis of acute DENV infection. *Clin.Vaccine.Immunol*. 15:1513-1518.

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GR, Simmons CP, Scott TW, Farrar JJ and Hay SI. 2013. The global distribution and burden of dengue. Nature. 496:504-507.

Boonnak K, Slike BM, Burgess TH, Mason RM, Wu SJ, Sun P, Porter K, Rudiman IF, Yuwono D, Puthavathana P andMarovich MA. 2008. Role of dendritic cells in

antibody-dependent enhancement of dengue virus infection. *Journal of Virology*.82:3939-3951.

Brady OJ, Gething PW, Bhatt S, Messina JP, Brownstein JS, Hoen AG, Moyes CL, Farlow AW, Scott TW and Hay SI. 2012. Refining the global spatial limits of dengue virus transmission by evidence-based consensus. PLoS. Neglect. Tropl. Dis., 6:8, 1-15.

Chen HC, Hofman FM, Kung JT, Lin YD and Wu-Hsieh BA. 2007. Both virus and tumor necrosis factor alpha are critical for endothelium damage in a mouse model of dengue virus-induced hemorrhage. *Journal of Virology*. 81:5518-5526.

CLUSTAL W [http://www.genome.jp/tools/clustalw/]

Dejnirattisai.W, Jumnainsong A, Onsirisakul N, Fitton P, Vasanawathana S, Limpitikul W, Puttikhunt C, Edwards C, Duangchinda T, Supasa S, Chawansuntati K, Malasit P, Mongkolsapaya J andScreaton G. 2010. Cross-reacting antibodies enhance dengue virus infection in humans. Science. 328:745-748.

Garg P, Nagpal J, Khairnar P and Seneviratne SL. 2008. Economic burden ofdengueinfections in India. Trans. R. Soc. Trop. Med.Hyg.,102: 570-77.

Gupta S, Ansari HR and Gautam A. 2013. Open source Drug Discoveri Consortium, Raghava GP: Identification of B-cell epitopes in an antigen for inducing specific class of antibodies. *Biol. Direct*., 8:27.

Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, Hunsperger E, Kroeger A, Margolis HS, Martínez E, Nathan MB, Pelegrino JL, Simmons C, Yoksan S and Peeling RW. 2010. Dengue: a continuing global threat. Nat. Rev.Microbiol. , 8(12 Suppl):S7-16.

Halstead SB. 2007. Dengue.Lancet., 370(9599):1644-52.

Halstead SB. 2008. Dengue virus-mosquito interactions.Annu. Rev.Entomol. , 53:273-91.

Harris E, Holden KL, Edgil D, Polacek C and Clyde K. 2006. Molecular biology of flaviviruses. Novartis Foundation Symposium., 277:23-39.

Huisman W, MartinaBE, Rimmelzwaan GF, Gruters RA andOsterhaus AD. 2009. Vaccine-induced enhancement of viral infections.Vaccine., 27(4):505-12.

Hunsperger EA, Yoksan S, Buchy P, Nguyen VC, Sekaran SD, Enria DA, Pelegrino JL, Vázquez S, Artsob H, Drebot M, Gubler DJ, Halstead SB, Guzmán MG, Margolis HS, Nathanson CM, Rizzo Lic NR, Bessoff KE, Kliks S and Peeling RW. 2009. Evaluation of commercially available anti-dengue virus immunoglobulin M tests. Emerg. Infect. Dis., 15(3):436-40.

Innis BL, Nisalak A, Nimmannitya S, Kusalerdchariya S, Chongswasdi V, Suntayakorn S, Puttisri P and Hoke CH. 1989. An enzyme-linked immunosorbent assay to characterize dengue infections where dengue and Japanese encephalitis co-circulate. *Am. J. Trop. Med.Hyg*., 40(4):418-27.

Kumar V, Bansal VJ, Rao KV and Jameel S. 1992. Hepatitis B virus envelope epitopes: gene assembly and expression in Escherichia coli of an immunologically reactive novel multiple-epitope polypeptide 1 (MEP-1). Gene. , 110:137-144.

Lima MRQ, Nogueira RMR, Schatzmayr HG and Santos FBD. 2010. Comparison of three commercially available dengue NS1 antigen capture assays for acute diagnosis of Dengue in Brazil. PLoS. Neg. Trop. Dis., 4 (7): e738.

Reddi H, Bhattacharya A and Kumar V. 2002. The calcium-binding protein of Entamoebahistolytica as a fusion partner for expression of peptides in Escherichia coli. Biotechnol. Appl.Biochem., T 36(Pt 3):213-8.

Rodenhuis-Zybert IA, Wilschut J andSmit JM. 2010. Dengue virus life cycle: viral and host factors modulating infectivity. *Cell. Mol. Life. Sci*., 67(16):2773-86.

Rothman AL. 2010. Cellular immunology of sequential dengue virus infection and its role in disease pathogenesis.Curr.Top.Microbiol. , 338:83-98.

Saha S and Raghava GP. 2006. Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins., 65(1):40-8.

Saha S and Raghava GPS. 2004. BcePred:Prediction of Continuous B-Cell Epitopes in Antigenic Sequences Using Physico-chemical Properties. In G.Nicosia, V.Cutello, P.J. Bentley and J.Timis (Eds.) ICARIS., LNCS 3239: 197-204, Springer.

Sekaran SD, Lan EC and Subramaniam G. 2008 Comparison of five serological diagnostic assays for the detection of IgM and IgG antibodies to dengue virus. Afr J Microbiol. Res., 2(6): 141-147.

Stothard P. 2000. The Sequence Manipulation Suite: JavaScript programs for analyzing and formatting protein and DNA sequences. Biotechniques., 28:1102-1104.

Ubol S and Halstead SB. 2010. How innate immune mechanisms contribute to antibody-enhanced viral infections. Clin.Vaccine.Immunol. , 17(12):1829-35.

Vita R, Zarebski L, Greenbaum JA, Emami H, Hoof I, Salimi N, Damle R, Sette A and Peters B. 2010. The immune epitope database 2.0. Nuc.Acids. Res., 38(Database issue): D854-62.

Wang SM andSekaran SD. 2010. Early diagnosis of Dengue infection using a commercial Dengue Duo rapid test kit for the detection of NS1, IGM, and IGG. *Am. J. Trop. Med.Hyg*., 83(3):690-5.

Weaver SC and Vasilakis N. 2009. Molecular evolution of dengue viruses: contributions of phylogenetics to understanding the history and epidemiology of the preeminent arboviral disease. *Infect.Genet.Evol*., 9(4):523-40.

WHO Factsheet No 117. 2012. Dengue and dengue haemorrhagic fever. URL:www.who.int/mediacentre/ factsheets/fs117/en/

****