Insight Bioinformatics 1 (1): 6-10, 2011
ISSN: 2042-3233

Identifying the Potential Tetravalent Vaccine Candidate for Dengue Virus using Insilico Approach

[1]V. Baskar, [2]R. Madhan, [2]G. Srinivasan, [2]K. Selvakumar, [3]M. Radha and [3]Rajeswari
[1]Informatic Biology Division, Global BioLim, Chennai-600023, Tamilnadu, India
[2]Applied Biology Division, Global BioLim, Chennai-600023, Tamilnadu, India
[3]Department of Bioinformatics, Vels University, Chennai-138, India

Abstract: Background: Dengue fever has increased significantly in the past two decades and thus has been a major concern of public health globally. There are 4 serotypes of dengue virus (DEN 1, DEN 2, DEN 3 and DEN 4). Dengue fever caused by any one of these serotypes and when, it becomes severe it may cause fatal. Vaccines can be prophylactic (e.g., to prevent or ameliorate the effects of a future infection by any natural or wild pathogen), or therapeutic (e.g., vaccines against cancer). **Problem:** Identification of new vaccine is intended to produce immunity against any pathogenic infection by stimulating the production of antibodies. Vaccines available for dengue virus are specific for particular serotype. Four serotypes have different antigenic determinants thereby development of a vaccine against all these serotypes is still in lag phase and time consuming. **Result:** Genomic and proteomics information allows us to fast up the vaccine development by Insilco methods. This new approach is termed as reverse vaccinology; have identified antigenic epitopes present in both structural and non-structural proteins common for all 4 serotypes of dengue virus using well known algorithms. **Conclusion:** The computationally predicted epitopic regions in non-structural I protein component will improve our understanding of T cell immune response and helps in identifying the vaccine candidate for formulation of antigen based vaccine diagnostic kit against dengue virus by wet laboratory thereby reduces the time and cost of conventional methods.

Key words: Reverse vaccinology, tetravalent vaccines, trans-membrane proteins, serotypes, dengue virus

INTRODUCTION

Dengue Fever (DF), which is caused by Dengue virus infection, had become a major public health problem in the tropic and subtropical countries caused by any one of four related viruses transmitted by mosquitoes. Dengue fever is an acute febrile viral disease characterized by sudden onset, fever for 3-5 days by 4 different virus serotypes. Since, World War II, the 4 dengue viruses have progressively spread virtually to all geographical areas. This disease is now endemic in most tropical countries and found to be highly complex medical, economical and ecological problems primarily affecting human (Chang *et al*., 2004; Tambunan *et al*., 2009). The antigenic variants are highly seen in the proteins of all the 4 serotypes (Khan *et al*., 2006). There is no yet preventive vaccines are available for DF (Teixeira *et al*., 2005; Chatuverdi *et al*., 2005). With the advent of whole genome sequencing of pathogens has improved the vaccine identification mechanisms computationally. Moreover, it helps to identify the pathogen associated molecular patterns leading to activation of antigen presenting cells in the host system (Cassone and Rappuoli, 2010).

Dengue virus classified into 4 serotypes and each possesses unique structural and non-structural proteins that help the virus to infect the human. Identification of suitable antigen in each protein that is responsible for immune reaction is vital in characterizing the vaccines. Thus, there is a major role to identify the effective antigenic regions that can exhibit the immunologic reactions in humans using computational approach.

TYPES

Dengue Virus (DENV) has 4 serotypes based on their immunogenicity in the human body belongs to Flavivirus family. They are:

- DENV-1
- DENV-2
- DENV-3
- DENV-4

Corresponding Author: V. Baskar, Informatic Biology Division, Global BioLim, Chennai-600023, Tamilnadu, India

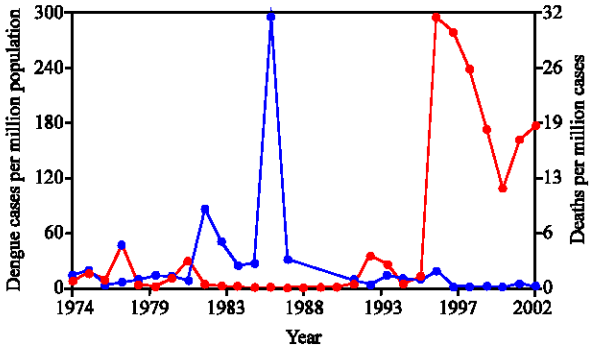

Fig. 1: Dengue Incidence and Death in India (1974-2002). Source: 1974-1995-Citizen's Health Report II statistical database (Edition: 1999), 1996-2002-CD alert

Annually, there are an estimated 50-100 million cases of dengue fever and 500,000 cases of the more severe and potentially lethal dengue hemorrhagic fever/shock syndrome (DHF/DSS) (Gubler, 2002). Figure 1 shows dengue incidence and death in India in 1974-2002.

FLAVIVIRUS PROTEINS

Virions contain only three virus-coded proteins called structural proteins. They are located at the amino terminus and incorporated into mature, infectious virions (Zhao *et al*., 1987). Several nonstructural proteins have also been identified in dengue virus infected cells that are located at the carboxy terminus are involved in the intracellular replication of the virus.

Structural proteins:

- Capsid (C)
- Premembrane (prM)
- Envelope (E) glycoprotein

Both C and M are internal proteins. The E protein is responsible for viral attachment to the putative cell surface receptors, fusion with the endosomal membranes upon entry and mediating protective immune responses in the infected host (Somvanshi and Seth, 2009; Crill and Roehrig, 2001; Kinney and Huang, 2001). The function of envelope protein is to control the receptor binding and fusion with the host cell. The envelope glycoprotein also exhibits hemagglutinating activity and is responsible for adsorption to the cell surface.

Nonstructural proteins:

- NS1
- NS2a
- NS2b
- NS3
- NS4a
- NS4b
- NS5

Nonstructural proteins play an important role in mediating immunity. Among the non-structural protein s, NS1 is considered an important vaccine component since it is expressed on the surface of infected cells making them targets for immune cytolysis. Recently, intracellular NS1 was shown to be involved in the early steps of viral replication in agreement with its retention in intracellular organelles of the infected cell and its ability to interact with membranes. The secretion of NS1 may be an important event in flavivirus infections in the human host.

RECOGNITION OF ANTIGEN BY B AND T LYMPHOCYTES

Antigens, which are generally very large and complex, are not recognized entirety by lymphocytes. Instead, both B and T-lymphocytes recognize discrete sites on the antigen called antigenic determinants, or epitopes. Epitopes are the immunologically active regions on a complex antigen, the regions that actually bind to B-cell or T-cell receptors.

MATERIALS AND METHODS

Bioinformatics tools were identified for analyzing all the Dengue virus pathogenic proteins for their antigenic properties. We have extracted all protein sequence of Dengue virus of 4 serotypes in FASTA format from Swissprot, a high quality curated protein sequence database. For Sequence similarity search of all the proteins present common in all serotypes, BLAST Algorithm is used. Each of the proteins is predicted for antigenicity properties by VaxiJen. VaxiJen is the first server for alignment-independent prediction of protective antigens present in the viral proteins (Doytchinova and Flower, 2007). Highly antigenicity proteins is analyzed for presence of MHC I and MHC II binding antigen peptides by ProPred I and ProPred. ProPred I allow to predict the 47 alleles MHC Class-I binding regions in antigenic protein sequence (Singh and Raghava, 2003). ProPred server allow to predict 57 MHC Class-II binding regions in an antigenic protein sequence, using quantitative matrices (Singh and Raghava, 2001). Each of the MHC I binding antigens is analyzed for binding affinity with TAP by TAPPred, which transport the antigen to MHC I binding receptor. TAPPred is an on-line tool for predicting binding affinity of peptides toward the TAP transporter (Bhasin and Raghava, 2004). HLAPRED predict HLA

binding regions from antigen sequence based on quantitative matrices and analyzed for homology with Human proteome.

MHC I and II binding Antigenic region identification: Both structural and non-structural proteins seen homology in all serotypes such as M, E Glycoprotein, NS1, NS3, NS4b and NS5 have analyzed for probable antigenic region using VaxiJen which have predicted protein which is having possibility of antigenic region above the threshold value of 0.5%. All the homology proteins which have passed above the threshold value of 0.5% are then predicted for potential antigenic region of length 9 mer peptides which promiscuously binds with MHC I for possible 47 different MHC I alleles by Propred1 tool. Similarly all the homology proteins predicted for potential antigenic region of length 9 mer peptides which promiscuously bind with MHC II for possible 51 different MHC II alleles by Propred tool. The predicted MHC I binding regions then analyzed for TAP binding affinity by TAPPred. Finally, All the predicted MHC I and II binding peptides regions where analyzed for homology with human proteome to verify those antigenic peptides do not trigger the auto-immunity.

RESULTS AND DISCUSSION

Antigenic proteins: Probable antigenic proteins that are suspected to be recognized by T_H and T_C cells have identified by VaxiJen using alignment-independent algorithm which are above the threshold value of 0.5% are prM, E, NS1, NS3 (Table 1).

MHC I binding epitopic regions: Human immune system as several factors to trigger immune response most such factors are epitopic regions that can effectively binds with MHC I. Epitopic regions of length 9 mer peptide present in each of the probable antigenic protein is predicted by ProPred I for binding with 47 different MHC I alleles to trigger cell mediated immunity in the human system (Table 2).

MHC II binding epitopic region: T_H lymphocyte recognizes antigens that are binding with MHC II alleles. Epitopic regions of length 9 mer peptide present in each of the probable antigenic protein is predicted by ProPred for binding with 51 MHC II alleles to trigger cell mediated and humoral immunity in human system (Table 3).

TAP binding affinity of MHC I binding epitopic region: Antigen presenting cells process the Antigens and by binding with TAP it has been transported to the site were MHC I protein is located. Thereafter, antigen and MHC I binds together and forms a complex and hence it is recognized by T_C cell receptors. The region of MHC I binding antigens that will be transported by the TAP is identified by TAPPred for calculating its binding affinity with the required antigens to be processed.

Table 1: Probable antigenic protein

Protein	Threshold value
prM	0.5795
E	0.7468
NS1	0.5976
NS3	0.5470
NS4b	0.4362
NS4	0.4490

Table 2: MHC I binding antigenic regions

Alleles	Regions
E-glycoprotein (29 alleles)	YIVIGAGEK-377-385
	VIGAGEKAL-379-387
	GAGEKALKL-381-389
NS 1(20 alleles)	ITPELNHIL-71-79
	ILSENEVKL-78-86
NS 3-(23 alleles)	FEPEREKSA-510-518
	AAIDGEYRL-518-526

Table 3: MHC II binding antigenic regions

Alleles	Regions
E-glycoprotein (50 alleles)	YGVLFSGVS-444-452
	FSGVSWTMK-448-456
	VSWTMKIGI-451-459
	WTMKIGIGV-453-461
	MKIGIGVLL-455-463
	IGIGVLLTW-457-465
	IGVLLTWLG-459-467
	LLTWLGLNS-462-470
	WLGLNSRST-465-473
	LGLNSRSTS-466-474
	LNSRSTSLS-468-476
NS 1(28 alleles)	WPKSHTLWS-225-233
	LWSNGVLES-231-239
NS 3(49 alleles)	MRLLSPVRV-268-276
prM (37 alleles)	FKTEDGVNM-25-33
	VNMCTLMAM-31-39
	MCTLMAMDL-33-41
	LMAMDLGEL-36-44
	MAMDLGELC-37-45

Antigenic regions non-homologous with human proteome: All the predicted MHC I and MHC II binding antigenic regions of corresponding proteins are analyzed for homologous with human proteome using HLAPred to determine any similar antigenic sequence present in human protein in order to avoid the auto immune response.

Tetravalent epitopic region for cell mediated and humoral immunity: In the present study all the pathogenic proteins were selected based on its homology from all the 4 serotypes by comparison using Blast algorithm. Within

Table 4: E-glycoprotein (YIVIGAGEKALKL)

Peptide rank	Start position	Sequence	Score	Predicted affinity
1	3	VIGAGEKAL	7.943	High
2	2	IVIGAGEKA	7.533	High
3	5	GAGEKALKL	3.938	Intermediate

Table 5: NS1 (ITPELNHILSENEVKL)

Peptide rank	Start position	Sequence	Score	Predicted affinity
1	6	NHILSENEV	7.285	High
2	8	ILSENEVKL	6.791	High
3	3	PELNHILSE	5.547	Intermediate
4	1	LITPELNHIL	3.536	Intermediate

Table 6: NS3 (FEPEREKSAAIDGEYRL)

Peptide rank	Start position	Sequence	Score	Predicted affinity
1	6	PVRVFEPER	7.241	High
2	4	LSPVRVFEP	5.215	Intermediate
3	12	PEREKSAAI	5.093	Intermediate
4	7	VRVFEPERE	4.265	Intermediate
5	17	SAAIDGEYR	3.841	Intermediate
6	10	FEPEREKSA	3.594	Intermediate

Table 7: Non-homologous with Human proteome

	Overlapped Antigenic Regions of MHC I and II
E-glycoprotein	YIVIGAGEKALKLYGVLFSGVSWTMKIGIGVLLTW
	LGLNSRSTSLS
NS1	IRSVTRLENLMWKQITPELNHILSENEVKL
NS3	LMCHATFTMRLLSPVRVFEPEREKSAAIDGEYRL
PrM protein	SRQEKGKSLFQRALIFILLTAVTPSM

all the homology proteins 4 proteins such as E Glycoprotein, prM, NS1, NS3 were predicted to have better antigenicity than other pathogenic proteins (Table 1). Antigenic sites of E Glycoprotein, prM, NS1, NS3 are having binding affinity with maximum number MHC I and MHC II alleles. On comparing each of these antigenic regions binding affinities with maximum number of MHC I and II alleles, antigenic peptides in NS1 and NS3 protein shows binding with more number MHC and II alleles (Table 2, 3). On more MHC I and MHC II binding antigenic peptides, more the immune response triggers the antibody production to fight against the pathogen. TAP binding affinity of NS1 antigenic peptides seems to be high compared to other proteins antigenic peptides (Table 4-6). None of these antigenic peptides were found homologous with human proteome (Table 7), thus confirmed there will no occurrence of auto immune reaction.

CONCLUSION

Immunoinformatics tools contributed major role to predict the vaccine candidate for dengue virus without using their live cultures. Experimental prediction of suitable tetravalent antigenic region present in Nonstructural protein 1 binds affectively with both MHC I and MHC II alleles with maximum number of alleles and also having high binding affinity to TAP; particularly the half-life of this antigen is higher than other membrane proteins, hence more number of repeated boosters is not required. So Non structural protein1 induce the active immunization and also generate the B-lymphocytes and T-lymphocytes to induce the humoral immunity and cell mediated immunity. This antigenic peptide from non-structural protein 1 is a preferable vaccine candidate against all the serotypes of dengue virus; hence, it may also be a tetravalent vaccine candidate for Dengue virus vaccine development.

ACKNOWLEDGMENT

The authors wish to thank Dr. G.P.S. Raghava's Group Imtech, Chandigarh for providing some of the immuno-informatics tools for my analysis.

REFERENCES

Bhasin, M. and G.P.S. Raghava, 2004. Analysis and prediction of affinity of TAP binding peptides using cascade SVM. Protein Sci., 13: 596-607.

Cassone, A. and R. Rappuoli, 2010. Universal vaccines: Shifting to one for many. American Society for Microbiology, Vol. 1. 10.1128/mBio.00042-10.

Chang, G.J., G. Kuno, D.E. Purdy and B.S. Davis, 2004. Recent advancement in Flavivirus vaccine development. Expert Rev. Vaccines, 3: 199-220.

Chatuverdi, U.C., R. Shrivastava and R. Nagar, 2005. Dengue vaccines: Problems and prospects. Rev. Art. Indian J. Med. Res., 121: 639-652.

Crill, W.D. and J.T. Roehrig, 2001. Monoclonal antibodies that bind to domain III of dengue virus E glycoprotein are the most efficient blockers of virus adsorption to vero cells. J. Virol., 75: 7769-7773.

Doytchinova, I.A. and D.R. Flower, 2007. VaxiJen: A server for prediction of protective antigens, tumourantigens and subunit vaccines. BMC Bioinform., 8: 4-4.

Gubler, D.J., 2002. Epidemic dengue/dengue hemorrhagic fever as a public health, social and economic problem in the 21st century. Trends Microbiol., 10: 100-103.

Khan, A.M., A.T. Heiny, K.X. Lee, K.N. Srinivasan, T.W. Tan, J.T. August and V. Brusic, 2006. Large-scale analysis of antigenicity diversity of T-cell epitopes in dengue virus. BMC Bioinform., 7: S4-S4.

Kinney, R.M. and C.Y. Huang, 2001. Development of new vaccines against dengue fever and Japanese encephalitis. Intervirology, 44: 176-196.

Singh, H. and G.P.S. Raghava, 2001. ProPred: Prediction of HLA-DR binding sites. Bioinformatics, 17: 1236-1237.

Somvanshi, P. and P.K. Seth, 2009. Prediction of T cell epitopes for the utility of vaccine development from structural proteins of dengue virus variants using insilico methods. Indian J. Biotechnol., 8: 193-198.

Tambunan, U.S.F., A.A. Parikesit, Hendra, R.I. Taufik and F. Amelia, 2009. In silico analysis of envelope dengue virus-2 and envelope dengue virus-3 protein as the backbone of dengue virus tetravalent vaccine by using homology modeling method. Online J. Biol. Sci., 9: 6-16.

Teixeira, M.G., M.C.N. Costa, M.L. Barreto and E.M.C.S. Publica, 2005. Dengue and dengue hemorrhagic fever epidemics in Brazil: What research is needed based on trends, surveillance and control experiences. Cad. Saude Publica, Vol. 21.

Zhao, B., G. Prince, R. Horswood, K. Eckels, P. Summers, R. Chanock and C.J. Lai, 1987. Expression of dengue virus structural proteins and nonstructural protein NS1 by a recombinant vaccinia virus. J. Virol., 61: 4019-4022.