



Contents lists available at ScienceDirect

Asian Pacific Journal of Tropical Disease

journal homepage: www.elsevier.com/locate/apjtd



Original article doi: 10.1016/S2222-1808(15)60852-0

©2015 by the Asian Pacific Journal of Tropical Disease. All rights reserved.

Analysis of E and NS proteins of dengue serotypes and identification of active binding sites for drug molecule

Vijayakumar Subramaniyan*, Ramesh Venkatachalam

P.G. and Research Department of Botany and Microbiology, A.V.V.M. Sri Pushpam College (Autonomous), Poondi-613 503, Thanjavur District, Tamil Nadu, India

ARTICLE INFO

Article history:

Received 8 Oct 2014

Received in revised form 20 Oct, 2nd

revised form 24 Nov 2014

Accepted 29 Nov 2014

Available online 1 Jun 2015

Keywords:

Dengue serotypes

E proteins

NS proteins

BLASTP analysis

ABSTRACT

Objective: To analyze the E and NS proteins of four dengue virus (DENV) serotypes which involve in the formation of dengue fever and to find out the active binding sites on the surfaces of the above proteins to access with the effective drug molecule to counter dengue.

Methods: ClustalW analysis for phylogram, BLASTP analysis for E and NS proteins of DENV serotypes and Swiss-Prot for 3D structure of E and NS proteins were used.

Results: E proteins were grouped into two clusters, DENV-1, DENV-2 and DENV-3 forming one cluster and DENV-4 alone forming a different serotype. On the surfaces of E proteins of four serotypes, five binding sites for ligands were located in different DENV which the fourth binding site was identified as the active ligand binding site in DENV-2, DENV-3 and DENV-4 serotypes and second binding site in the DENV-1. Among the seven NS proteins, NS-1, NS-3 and NS-5 were identified as significant molecules involved in dengue disease.

Conclusions: Binding sites, essential for ligand binding are found on all the E and NS proteins which active binding sites have been identified.

1. Introduction

Dengue hemorrhagic fever is one of the most important emerging tropical diseases in the 21st century[1]. In the later part of the 20th century, globalization and rapid urbanization of many developing tropical countries produced increased transmission and hyperendemicity of the disease[2]. World Health Organization estimates that there are as many as 50 million cases of dengue infection worldwide and global warming provide a significant selective advantage for dengue infection spreading into new areas[3]. Dengue infection is caused by dengue virus (DENV) and belongs to the genus *Flavivirus* in the family *Flaviviridae*.

Recent survey reveals that there are 50 million dengue cases

reported so far and approximately 2 billion people lives in the dengue endemic countries. The prevalence of dengue has dramatically increased in the recent years and it is now endemic over 100 countries like, Africa, America, Malaysia, Thailand, India, South East Asia and Western Pacific. The first dengue hemorrhagic fever was reported in Thailand and Philippines in 1950 where the first two DENV serotypes were identified, followed by third and fourth serotypes in 1954[4,5]. These four serotypes were recognized based on their antigenicity and immunogenic property in the human body.

The DENV has the inner nucleocapsid made of the viral genome and C protein. The nucleocapsid is surrounded by viral membrane in which the E protein is embedded. DENV genome is a single stranded RNA which is translated into a single polyprotein containing three structural proteins and seven non-structural proteins. The structural proteins are capsid protein C, membrane protein M and envelope protein E[6]. Seven NS proteins include NS-1, NS-2A, NS-2B, NS-3, NS-4A, NS-4B and NS-5 that are found only in the infected host cells where they are required for replication of the virus. Among the three structural proteins, the E protein is considered as the most obvious attractive target for the development of drugs and peptide

*Corresponding author: Vijayakumar Subramaniyan, P.G. and Research Department of Botany and Microbiology, A.V.V.M. Sri Pushpam College (Autonomous), Poondi-613 503, Thanjavur District, Tamil Nadu, India.

Tel: 04374-239523

Fax: 04374-239477

E-mail: svijaya_kumar2579@rediff.com

Foundation Project: Supported by the UGC through the Major Research project, [MRP R. No.: 41-472/2012(SR)].

vaccine candidates for the prevention of dengue. Among the seven NS proteins, NS-2A, NS-2B, NS-4A and NS-4B are involved in the anchoring of various particles in the membrane[7]. NS-1, NS-3 and NS-5 are involved in the synthesis of virus particles directly in the host cell. Therefore, these three NS proteins are the target of the study. NS-1 is a soluble protein detected very early during the infection, which regulates the replication of virus. NS-3 acts as the indicator protease molecule at the time of dengue infection and this molecule is also considered as the dengue protein target used to drug design for dengue. NS-5 is the most conserved dengue protein concerned with synthesis of RNA genome of DENV and plays an important role in the dengue pathogenesis[8]. Therefore, E proteins and NS-1, NS-3 and NS-5 have been selected for the present study.

2. Materials and methods

2.1. E and NS proteins

The search for E and NS protein sequences of four DENV serotypes (DENV 1-4) was done by accessing Viral Bioinformatics Resource Center (<http://www.athena.bioc.uvic.ca/>). The sequences format that was utilized for saving downloaded E and NS protein sequences in each DENV serotypes was FASTA format. This format was submitted for query of further analysis by Glide 6.3 module of Schrodinger suite[9].

2.2. Computational methods with Glide version 6.3

All computational studies were carried out using Glide version 6.3, installed in a single machine running on Intel Core i7 Duo processor with 1GB random access memory and 275 GB hard disk with Black Dell Inspiron version 7.0 as the operating system.

2.3. Phylogram analysis

The amino acid sequences obtained from individual sequencing reactions were combined for analysis and edited using the Auto Assembler 2.1 software (Perkin-Elmer Corp.-Applied Biosystems Inc.). Multiple alignments of sequences from the GenBank library were performed using the ClustalW 1.81 algorithm with default parameters[10]. A phylogram tree was constructed using the neighbor joining method[11], maximum likelihood distance parameters (Kimura two-parameter formula) and the Molecular Evolutionary Genetics Analysis software package[12].

2.4. Homology modeling

Homology analysis of E and NS proteins of dengue serotypes was accomplished by BLASTP to find out the similarity within the respective strains using Deep View/Swiss Pdb-Viewer3.7 (SP5) and Swiss Model server[13].

3. Results

3.1. E and NS proteins of DENV serotypes

E protein was basically formed by 495 amino acids residues. This number was not constant for all four serotypes. DENV-1 and DENV-4 had 495 amino acids in their sequences and DENV-3 showed 493 whereas only 480 amino acids were found in the case of DENV-2. This variation in the number was due to the deletion of some of the amino acids in their sequences in the middle or at the tail end. All the amino acids in the sequences of four serotypes were not similar and

showed variations. This variation was due to the mutations occurred in the position of some of the amino acids in the sequences. The replaced amino acids were qualitatively similar in their properties and characters of replaced amino acids. The NS proteins were formed by 295 to 352 amino acids residues. NS-1 showed 352 amino acids residues and NS-3 showed 322 whereas NS-5 had 295 amino acids residues.

3.2. Analysis of E proteins

The phylogram deduced from the E protein sequences of four dengue serotypes indicated a close relationship among DENV-1 and DENV-3 forming a single cluster whereas DENV-2 and DENV-4 formed another cluster. Thus, the DENV serotypes formed two clusters (Figure 1).

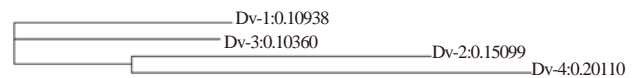


Figure 1. Phylogram for DENV 1-4.

A BLASTP analysis was carried out and it indicated the relationship among the dengue strains within the dengue serotypes[6]. In DENV-1, 100% identity was observed in 15 strains and other strains showed 99% identity. In DENV-3, 100% identity was observed in few strains and 99% in most of the cases. In DENV-2, all the strains showed only 99% identity. In DENV-4, very few strains showed 100% and 99% whereas most of the others showed 98% identity.

3.3. Analysis of NS proteins

Among the seven NS proteins, NS-1, NS-3 and NS-5 were considered to be more important in the drug development for the dengue disease treatment and management. NS-1 protein was formed by 352 amino acids residues in the proteome of DENV. BLASTP analysis of NS-1 protein molecules indicated the homology ranging from 95% to 100%. Out of 20 NS-1 sequences, only one showed 100% homology, while others were 99%, 98%, 97% respectively. NS-3 protein was formed by 322 amino acid residues. Out of 20 sequences studied, only one showed 100% similarity, and others were found to be 98%. NS-5 protein of the DENV proteome was formed by 295 amino acid residues (Tables 1-3 and Figures 2-4). Homology analysis of this protein showed 100% homology in all sequences. Thus among the NS proteins studied, NS-5 sequences showed more homology than that of NS-1 and NS-3.

Table 1

BLASTP analysis of protein sequence of NS-1 proteins.

Sample No.	Accession number	Identify percentage
1.	AAD115320.1	100
2.	AAD11531.1	99
3.	AAD11533.1	99
4.	CAA35218.1	99
5.	CAA48937.1	99
6.	AAB48937.1	99
7.	AAB48936.1	99
8.	AAB48938.1	99
9.	CAA35219.1	98
10.	CAA35220.1	98
11.	AEQ54950.1	98
12.	ACT68376.1	98
13.	ACT68377.1	98
14.	ACT68378.1	98
15.	ABS19903.1	98
16.	ABF83699.1	98
17.	AAW69297.1	97
18.	AGF00077.1	97
19.	AAL58460.1	97
20.	AAA42944.1	97

Table 2

BLASTP analysis of protein sequence of NS-3 strains.

Sample No.	Accession number	Identify percentage
1.	AAB22972.1	100
2.	NP739587.2	98
3.	AHB63926.1	98
4.	ACH61768.1	98
5.	AAF29146.1	98
6.	AAF29149.1	98
7.	ACH61768.1	98
8.	AAD32955.1	98
9.	ACA48809.1	98
10.	ACJO4237.1	98
11.	ACA49042.1	98
12.	ACY70823.1	98
13.	ACW82910.1	98
14.	ACO06134.1	98
15.	ACN42704.1	98
16.	ACL99224.1	98
17.	ACL99223.1	98
18.	ACL99161.1	98
19.	ACL99119.1	98
20.	ACQ44491.1	98

Table 3

BLASTP analysis of protein sequence of NS-5 strains.

Sample No.	Accession number	Identify percentage
1.	AFM29507.1	100
2.	AAM51540.1	100
3.	AAM51542.1	100
4.	AGH08163.1	100
5.	ADV76219.1	100
6.	AEA50923.1	100
7.	ABV03585.1	100
8.	ACL99076.1	100
9.	ACL99068.1	100
10.	ACL99046.1	100
11.	ACJ04219.1	100
12.	ADA60763.1	100
13.	AEF01546.1	100
14.	AEE99026.1	100
15.	ADK26427.1	100
16.	ACS31996.1	100
17.	ACQ44321.1	100
18.	AAM51544.1	99
19.	ADG22007.1	98
20.	YP001531176.2	98

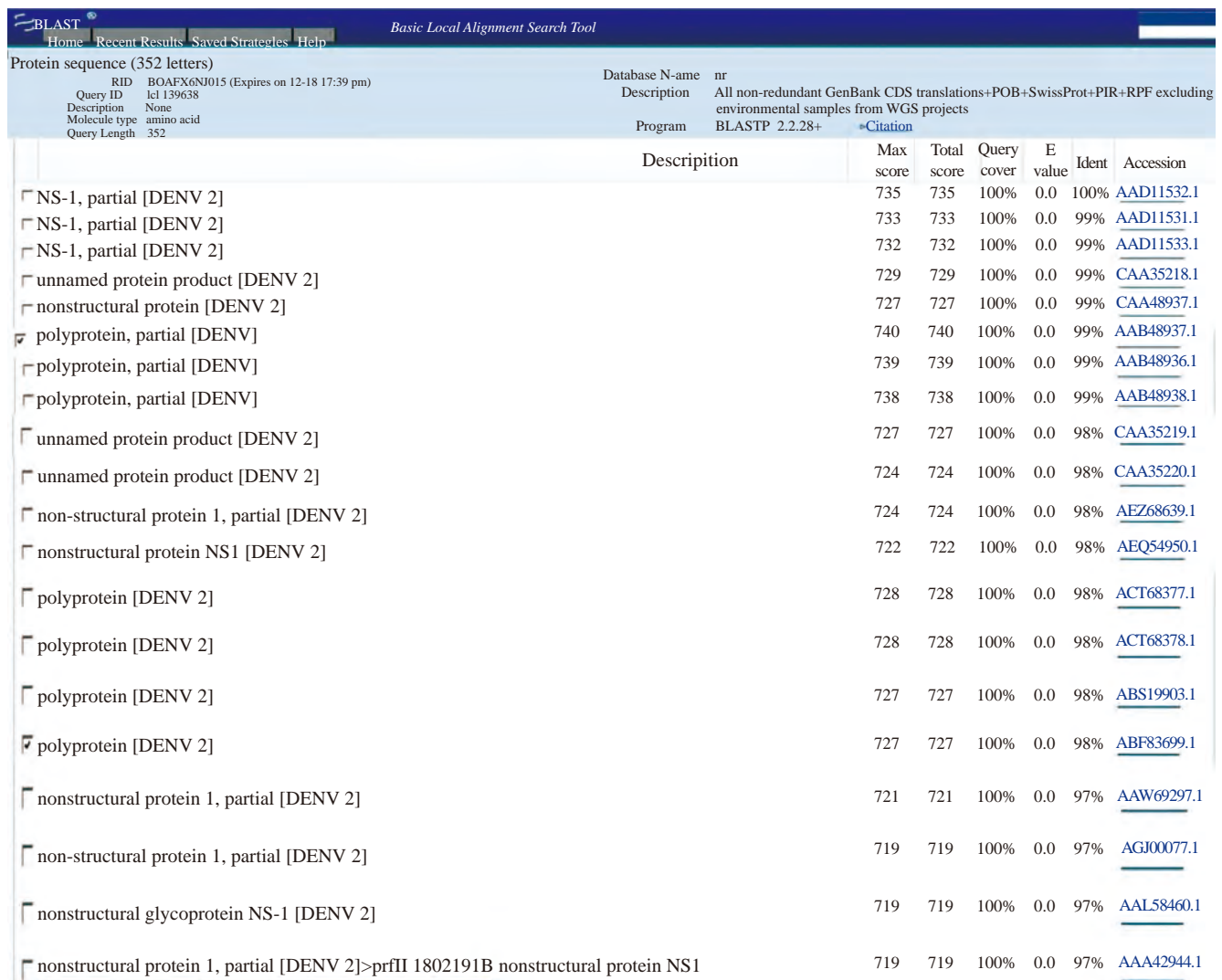


Figure 2. BLASTP analysis of protein sequence of NS-1 proteins.

Description	Max score	Total score	Query cover	E value	Ident	Accession
NS-3 [DENV 1]	663	663	100%	0.0	100%	AAB22972.1
Nonstructural protein NS3 [DENV 2]	647	647	100%	0.0	98%	NP_739587.2
nonstructural protein 3, partial [DENV]	647	647	100%	0.0	98%	AHB63926.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACH61768.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	AAF29146.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	AAF29149.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	AAD32955.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACI15880.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48809.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	AAD32957.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACJ04237.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACJ04266.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	CAR65147.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48925.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48919.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48912.1
polyprotein [DENV 2]>gb ACA48940.1 polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48911.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48731.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACA48730.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ABA61176.1
DEN polyprotein [DENV 2]	657	657	100%	0.0	98%	AAC40834.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACW82866.1
polyprotein [DENV 2]	657	657	100%	0.0	98%	ACN42706.1

Figure 3. BLASTP analysis of protein sequence of NS-3 strains.

3.4. Structure of E proteins

DENV-1 E protein was formed by 495 amino acid residues. The 3D structure of the E protein showed 36 strands with 9 helices and 57 turns. This structure was stabilized by 331 hydrogen bonds (Table 4). Five possible ligand binding sites were located on the surface of DENV-1 protein molecule. The score of the five binding sites was ranging from 0.997 to 1.097. The binding site-2, which secured the highest score of 1.097, was identified as the active binding site for the ligand (Table 5 and Figures 5a-5b).

Table 4
Characteristics features of E and NS proteins of DENV.

Characters	DENV-1	DENV-2	DENV-3	DENV-4	NS-1	NS-3	NS-5
Hydrogen bonds	331	280	329	327	52	205	161
Helices	9	12	9	9	1	12	15
Strands	36	41	36	36	11	16	10
Turns	57	52	56	57	11	40	26

DENV-2 E protein was formed by 480 amino acid residues, 41 strands, 12 helices and 52 turns (Table 5) in which the structure was stabilized by 280 hydrogen bonds. In DENV-2 E protein molecule, five possible ligand binding sites were located on the surface of E protein molecule (receptor). The binding site score of the five binding sites were calculated which was ranging from 0.957 to 1.068. The binding site-4, which secured the highest score of 1.068, was selected as the active binding site for the ligand (Table 5 and Figures 6a-6b).

DENV-3 E protein which was formed by 493 amino acid residues, showed 36 strands, 9 helices and 56 turns and the molecule was stabilized by 329 hydrogen bonds (Table 5). In DENV-3 E protein molecule, five possible ligand binding site were located on the surface of E protein molecule (receptor). The score of the five binding sites were ranging from 0.981 to 1.338. The binding site-4, which secured the highest score of 1.338, was the active binding site for the ligand (Table 5 and Figures 7a-7b).

Table 5
Prediction of ligand binding sites and identification of active sites of E and NS proteins.

DENV-1 E protein	Score	DENV-2 E protein	Score	DENV-3 E protein	Score	DENV-4 E protein	Score	NS-1	Score	NS-3	Score	NS-5	Score
Site-2	1.097	Site-4	1.068	Site-4	1.338	Site-4	1.296	Site-1	0.574	Site-1	1.067	Site-2	1.091
Site-5	1.080	Site-5	1.054	Site-2	1.117	Site-2	1.051	Site-2	0.548	Site-2	1.030	Site-1	0.940
Site-1	1.029	Site-3	1.035	Site-1	1.030	Site-1	0.997			Site-3	1.002	Site-4	0.777
Site-4	1.004	Site-2	0.979	Site-5	1.006	Site-3	0.994			Site-4	0.891	Site-3	0.760
Site-3	0.997	Site-1	0.957	Site-3	0.981	Site-5	0.782			Site-5	0.716	Site-5	0.706

Home		Recent Results		Saved Strategies		Help	
Protein Sequence (295 letters)							
RID	BOBYEWMXO15 (Expires on 12-18 18:20 pm)			Database N-ame	nr		
Query ID	lcl13899			Description	All non-redundant GenBank CDS translations+POB+SwissProt+PIR+RPF excluding environmental samples from WGS projects		
Description	None			Program	BLASTP 2.2.28+ Citation		
Molecule type	amino acid						
Query Length	322						

Description	Max score	Total score	Query cover	E value	Ident	Accession
NS-5 protein, partial [DENV 3] > gb AFM29508.1 NS5 protein, partial [DENV 3]	615	615	100%	0.0	100%	AFM29507.1
polyprotein [DENV 3]	615	615	100%	0.0	100%	AAM51540.1
polyprotein [DENV 3]>gb AAM51546.1 polyprotein [DENV 3]>gb AAM51548.1 polyprotein [DENV 3]	615	615	100%	0.0	100%	AAM51542.1
polyprotein, partial [DENV 3]	620	620	100%	0.0	100%	AGH08163.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ADU76219.1
polyprotein [DENV 3]>gb AEA50924.1 polyprotein [DENV 3]	622	622	100%	0.0	100%	AEA50923.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ABV03585.1
polyprotein [DENV 3]>gb ACQ44511.1 polyprotein [DENV 3]	622	622	100%	0.0	100%	ACL99076.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ACL99068.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ACL99058.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ACL99046.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ACL04218.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ADA60766.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	AEF01546.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ABE99026.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ADK26427.1
polyprotein [DENV 3]	622	622	100%	0.0	100%	ACS31996.1
polyprotein [DENV 3]	612	612	100%	0.0	99%	AAM51544.1
non-structural protein 4B [DENV 3]	580	580	95%	0.0	98%	ADG22007.1
Nonstructural protein NS5 [DENV 3]	580	580	95%	0.0	98%	YP001531176.2

Figure 4. BLASTP analysis of protein sequence of NS-5 strains.

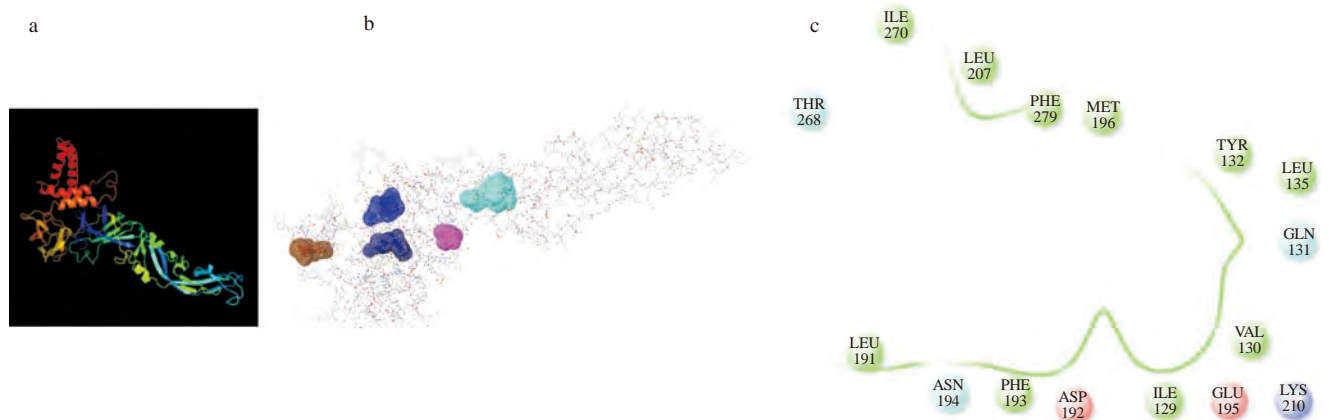


Figure 5. 3D structure of E protein from DENV-1 (a), DENV-1 E protein molecule showing possible ligand binding sites (b), DENV-1 E protein showing the amino acid residues lined around the binding site (c).

DENV-4 E protein was formed by 495 amino acid residues. The 3D structure of the E protein showed 36 strands with 9 helices and 57 turns and the molecular structure was stabilized by 327 hydrogen bonds (Table 5). In DENV-4 E protein molecule, five possible ligand binding sites were located on the surface of E protein molecule (receptor). The score of the five binding sites were ranging from 0.782 to 1.296. The binding site-4, which secured the highest score of 1.296, was the active binding site for the ligand (Table 5 and Figures 8a-8b).

3.5. Structure of NS proteins

NS-1 protein was formed by 352 amino acid residues. The 3D

structure of the NS-1 protein showed 11 strands with 12 helices and 46 turns and the molecular structure was stabilized by 52 hydrogen bonds (Table 5). In NS-1 protein molecule, two possible ligand binding sites were located on the surface of NS-1 protein molecule (receptor). The binding site-1, which secured the highest score of 0.574, was the active binding site for the ligand (Table 5 and Figures 9a-9b).

NS-3 protein formed by 322 amino acid residues, showed 16 strands, 12 helices and 40 turns. This structure was stabilized by 205 hydrogen bonds (Table 4). In NS-3 protein molecule, five possible ligand binding sites were located on the surface of NS-3 protein molecule (receptor). The binding site score of the five binding sites was ranging from 0.716 to 1.067. The binding site-1, which secured

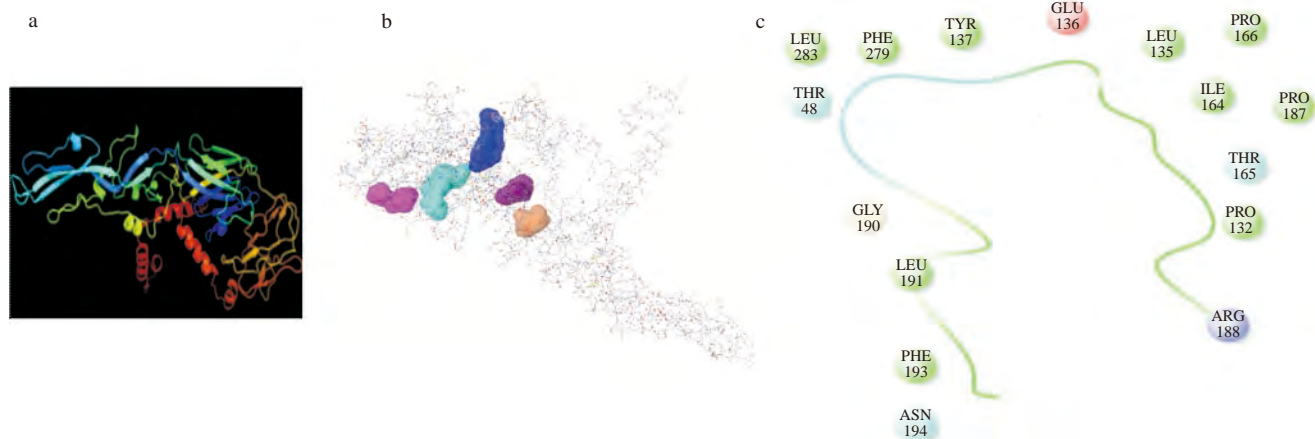


Figure 6. 3D structure of E protein from DENV-2 (a), DENV-2 E protein molecule showing possible ligand binding sites (b), DENV-2 E protein showing the amino acid residues lined around the binding site (c).

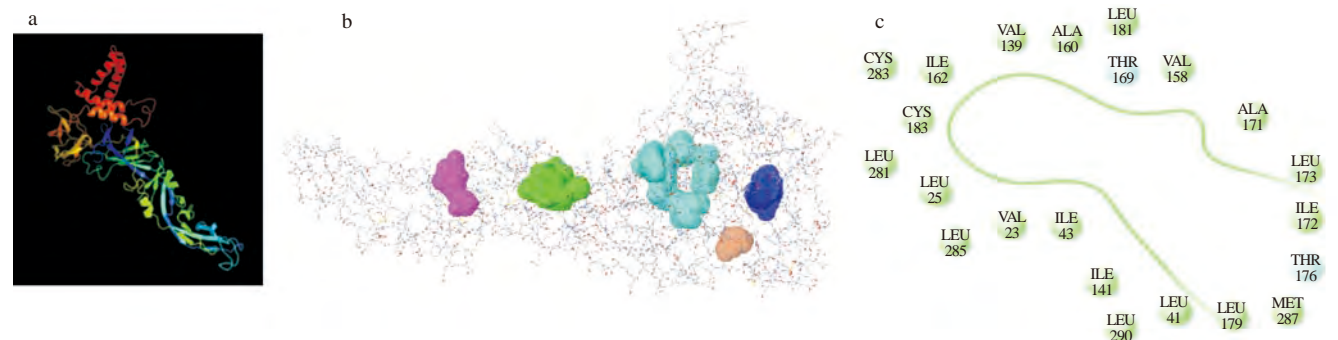


Figure 7. 3D structure of E protein from DENV-3 (a), DENV-3 E protein molecule showing possible ligand binding sites (b), DENV-3 E protein showing the amino acid residues lined around the binding site (c).

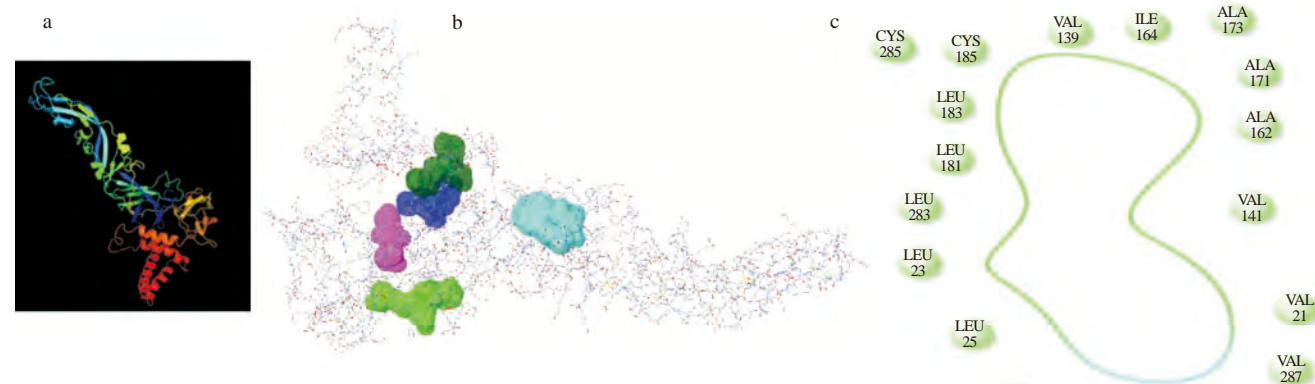


Figure 8. 3D structure of E protein from DENV-4 (a); DENV-4 E protein molecule showing possible ligand binding sites (b), DENV-4 E protein showing the amino acid residues lined around the binding site (c).

the highest score of 1.067, was the active binding site for the ligand (Table 5 and Figures 10a-10b).

NS-5 protein was formed by 295 amino acid residues, 10 strands with 15 helices and 26 turns. The 3D structure is stabilized by 161 hydrogen bonds (Table 4). In NS-5 protein molecule, five possible ligand binding sites were located on the surface of NS-5 protein molecule (receptor). The score of the five binding sites were ranging from 0.706 to 1.091. The binding site-2, which secured the highest score of 1.091, was the active binding site for the ligand (Table 5 and Figures 11a-11b).

3.6. Binding site

The active binding site of DENV-1 E protein molecule was

surrounded by 17 amino acids in which 13 of them were hydrophobic (Figure 5c). In the active binding site of DENV-2 E protein, 17 amino acids were surrounded the pocket; 13 of them were hydrophobic nature (Figure 6c). The active binding site of DENV-3 showed 22 amino acids from which 20 of them were hydrophobic indicating the high affinity between the receptor and ligand molecules (Figure 7c). In the E protein of DENV-4, the active binding site was lined with 19 amino acids from which 18 of them were hydrophobic (Figure 8c). The pocket of the NS 1 active binding site was surrounded by 13 amino acids. Only three were hydrophobic (Figure 9c). This NS-1 protein may not show effective binding during docking. In NS-3 molecule, the active binding site was lined with 22 amino acids from which only 7 of them were hydrophobic indicating loose binding

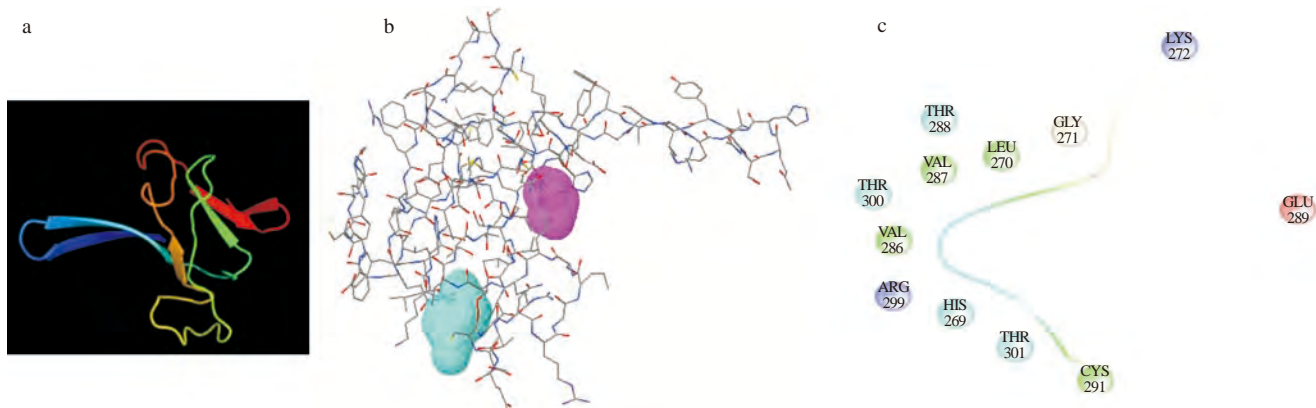


Figure 9. 3D structure of NS-1 protein from DENV (a), DENV NS-1 protein molecule showing possible ligand binding sites (b), DENV NS-1 protein showing the amino acid residues lined around the binding site (c).

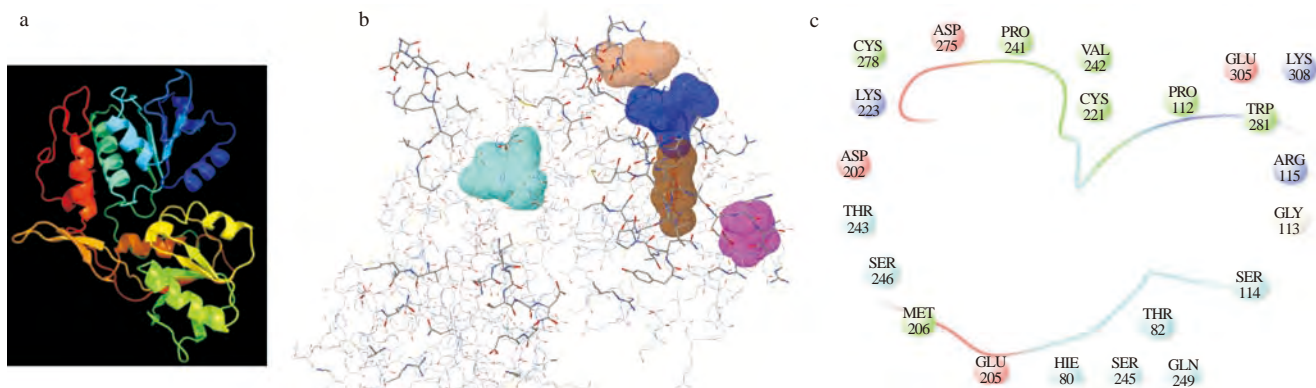


Figure 10. 3D structure of NS-3 protein from DENV (a), DENV NS-3 protein molecule showing possible ligand binding sites (b), DENV NS-3 protein showing the amino acid residues lined around the binding site (c).

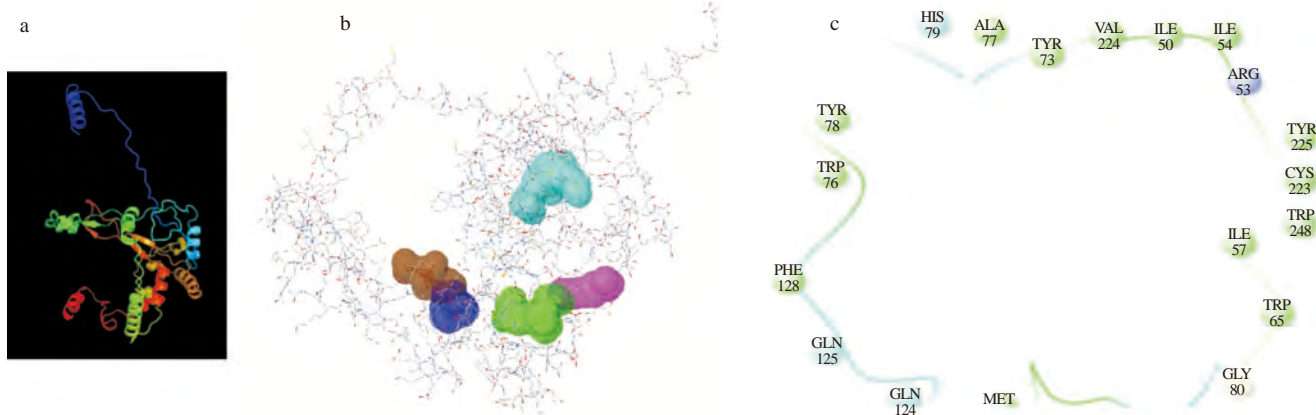


Figure 11. 3D structure of NS-5 protein from DENV (a), DENV NS-5 protein molecule showing possible ligand binding sites (b), DENV NS-5 protein showing the amino acid residues lined around the binding site (c).

during docking (Figure 10c). In NS-5 molecule, the active binding site was surrounded by 18 amino acids from which 10 of them were hydrophobic (Figure 11c).

4. Discussion

The study of homology is very important from the dendrographic point of view. Homology indicates the evolution of polymorphic genotypes. Different serotypes evolve from a single basic genotype through the mutation occurred in a single or a group of amino acids in the sequences. This may occur in the middle or the end of the sequence. This phenotypic change of the amino acids may be due to the need based and environmentally related process. This development of diversity is found in various kinds of living organisms in the world. Similar evolution also has occurred in the genome and proteome of *Flavivirus* causing dengue fever and resulted in the evolution of four serotypes like DENV-1, DENV-2, DENV-3 and DENV-4. The E protein of the DENV is basically formed by 495 amino acid residues. However, the number of amino acids in all the four types is not similar. All 495 amino acids are found in DENV-1 and DENV-4. DENV-3 showed 493 after the deletion of two amino acids in the sequence and 480 amino acids are found in the case of DENV-2 due to the deletion of 15 amino acids either in the middle or the end of the sequence. Such a kind of analysis and in depth study in this area has not been carried out in the recent past. Based on the dendrogram, in the present study it has been concluded that dengue serotypes 1, 3 and 2 formed a single group and the fourth one is distant from the above three serotypes, forming the second group. Based on the BLASTP analysis of various strains of four serotypes of DENV, the strains of DENV-1, DENV-2 and DENV-3 showed 99% homology. On the other hand DENV-4 showed only 98% homology. We have studied the homology of E proteins dengue serotypes and reported that the four serotypes of DENV (DENV 1-4) share approximately 65%-75% homology at the amino acid level[6]. In the present study the homology among the dengue serotypes is found to be 63% to 78.7%.

BLASTP analysis indicates the homology of this NS-1 protein ranging from 95% to 100%. NS-3 protein is not homologous and shows 96% to 100% similarity. Most of the sequences of NS-5 show 100% homology. Thus, among the NS proteins studied, NS-5 sequences show more homology than that of NS-1 and NS-3. Similarly, Mairiang *et al.* also reported that, the homology of NS-5 proteins had shown more homology than NS-3[14].

In biology, the active site is a small portion of the receptor molecules (enzyme) where ligand (drug) molecule binds and undergoes a chemical reaction. This chemical reaction occurs when a ligand collides with the active site of the enzyme. The active site is usually found in a 3D groove or pocket of the enzyme, lined with amino acid residues. These residues are involved in recognition of the ligand molecules. Residues that directly participate in the catalytic reaction mechanism are called active site residues. A tighter fit between an active site and the ligand molecule is believed to increase efficiency of a reaction[15].

The simulated ligand circles the target protein extensively before finding the active binding site. The ligand correctly identifies its

target binding site and forms a complex. Ligand molecules bind to the active site of the enzyme through hydrogen bonds, hydrophobic interactions, temporary covalent interactions (van der Waals) or a combination of all of these to form the receptor-ligand complex. Residues of the active site will act as donors or acceptors of protons or other groups on the drug (ligand) to facilitate the reaction. In other words, the active site modifies the reaction mechanism in order to change the activation energy of the reaction[16].

An emerging challenge in drug discovery concerns the identification of allosteric ligand binding sites, which drugs can modulate the effects of ligands that bind at the primary site. In general, an important limitation of traditional virtual drug screening is that it must start with a well-defined binding site[17,18]. By allowing the identification of previously unknown binding sites, molecular dynamics simulations of protein-ligand binding, such as those presented here, may substantially broaden the applicability of computational techniques to the drug development. Thus, the protein-ligand interaction plays an important role in determining the suitable drug for the treatment of dengue disease.

Identification of active sites was crucial in the process of target based drug design. The 3D structure of the enzyme was analyzed to identify the active sites and design drugs which can fit into them. The 3D structure of the E and NS proteins were provided with five ligand binding sites except in NS-1 in which only two ligand binding sites were present. The ligands of the bioactive compounds and antiviral drug were not recognizing the same binding site as an active site for the molecular docking. The active binding sites were four in the case of DENV-2, DENV-3 and DENV-4 and two sites were in DENV-1 and NS-5. For NS-1 and NS-3, the active binding site was found to be site-1 on the protein molecule.

The active binding site is surrounded by amino acid residues. Some of them are hydrophobic and others are hydrophilic. For the effective binding between the receptor molecule and the ligand molecule, hydrophobic amino acids are very essential. The active binding site of DENV-1 E protein molecule was surrounded by 17 amino acids in which 13 of them were hydrophobic. This shows the high affinity of the E protein of DENV-1 with the appropriate ligand molecule during docking. The distance of the hydrogen bonds between the above amino acids and the ligand molecules will be very less.

In the active binding site of DENV-2 E protein, 17 amino acids are surrounding the pocket and 13 of them are hydrophobic nature. This E protein also will show the high affinity towards the ligand molecules since most of the amino acids are hydrophobic in nature. The active binding site of DENV-3 showed 22 amino acids from which 20 of them were hydrophobic indicating the high affinity between the receptor and ligand molecules. In the E protein of DENV-4, the active binding site was lined with 19 amino acids from which 18 of them were hydrophobic.

The pocket of the NS-1 active binding site was surrounded by 13 amino acids. Only three were hydrophobic. This NS-1 protein may not show effective binding during docking. In NS-3 molecule, the active binding site was lined with 22 amino acids from which only 7 of them were hydrophobic indicating loose binding during docking. In NS-5 molecule, the active binding site was surrounded by 18 amino acids from which 10 of them were hydrophobic.

Hydrophobic interactions are the most important non-covalent forces that are responsible for different phenomena such as structural stabilization of proteins, binding of receptor with the ligand and folding of proteins[19]. Sun *et al.* pointed out the importance of more hydrophobic residues in the antigen-antibody interaction[20]. Fifty percent of the attractive force between the ligand and the receptor is contributed by the hydrophobic amino acids lined on the cavity of the binding site[21].

Chandra *et al.* also reported that hydrophobic amino acids are more frequent in binding group than in the non-binding group[22]. In the present investigation, all the active binding sites of E proteins showed the occurrence of more than 60% of the hydrophobic amino acid. Therefore active binding sites in the E proteins of the four DENV serotypes are eligible for tight fit molecular docking with the corresponding ligand molecules. On the other hand, the number of hydrophobic amino acids in the NS proteins were less in number than that of E proteins. In NS-1 and NS-3 and NS-5, they were less than 60%. From the above study, it is concluded that the active binding site with more number of hydrophobic amino acid residues (E proteins) will be more effective than that of binding sites with less number of hydrophobic amino acid residues (NS proteins).

Four kinds of structural E proteins are 2 clusters, out of seven kinds of NS proteins-3 are more important (NS-1, NS-3 and NS-5). All the four E proteins show the five binding sites on the surfaces of the protein molecules; the fourth binding site is identified as an active binding site in DENV-2, DENV-3 and DENV-4 serotypes. In DENV-1, E protein first binding site is an active binding site. Among the NS proteins, NS-1 and NS-3 have shown the first binding site as the active binding site whereas in NS-5, the second binding site is identified as the active binding sites.

Conflict of interest statement

We declare that we have no conflict of interest.

Acknowledgments

The authors are grateful to the UGC New Delhi, India, assistance for the financial given through the Major Research project, [MRP R. No.: 41-472/2012(SR)]. The authors specially express their thanks to the management of A.V.V.M. Sri Pushpam College (Autonomous), Poondi, for providing necessary facilities and support to carry out this work.

References

- [1] Banu S, Hu W, Guo Y, Naish S, Tong S. Dynamic spatiotemporal trends of dengue transmission in the Asia-Pacific region, 1955-2004. *PLoS One* 2014; **9**(2): e89440.
- [2] Murray NEA, Quam MB, Wilder-Smith A. Epidemiology of dengue: past, present and future prospects. *Clin Epidemiol* 2013; **5**: 299-309.
- [3] Rodriguez-Roche R, Gould EA. Understanding the dengue viruses and progress towards their control. *Biomed Res Int* 2013; doi: 10.1155/2013/690835.
- [4] Pigili RK, Runja C. Medicinal plants used in dengue treatment: an overview. *Int J Chem Nat Sci* 2014; **2**(1): 70-6.
- [5] Abd Kadir SL, Yaakob H, Mohamed Zulkifli R. Potential anti-dengue medicinal plants: a review. *J Nat Med* 2013; **67**(4): 677-89.
- [6] Venkatachalam R, Subramaniyan V. Homology and conservation of amino acids in E-protein sequences of dengue serotypes. *Asian Pac J Trop Dis* 2014; **4**(Suppl 2): S573-7.
- [7] Alcaraz-Estrada SL, Yocupicio-Monroy M, Del Angel RM. Insights into dengue virus genome replication. *Future Virol* 2010; **5**(5): 575-92.
- [8] de Oliveira AS, da Silva ML, Oliveira AFCS, da Silva CC, Teixeira RR, De Paula SO. NS3 and NS5 proteins: important targets for anti-dengue drug design. *J Braz Chem Soc* 2014; **25**: 1-11.
- [9] LigPrep. Version 2.3. New York: Schrodinger, LLC; 2009. [Online] Available from: http://isp.ncifcrf.gov/files/isp/uploads/2010/07/lp23_user_manual.pdf [Accessed on 21st September, 2014]
- [10] Jiang C, Wu LL, Zhao GC, Shen PH, Jin K, Hao ZY, et al. Identification and characterization of a novel fumarase gene by metagenome expression cloning from marine microorganisms. *Microb Cell Fact* 2010; **9**: 91.
- [11] DeGiorgio M, Syring J, Eckert AJ, Liston A, Cronn R, Neale DB, et al. An empirical evaluation of two-stage species tree inference strategies using a multilocus dataset from North American pines. *BMC Evol Biol* 2014; **14**: 67.
- [12] Tamura K, Stecher G, Peterson D, Filipinski A, Kumar S. MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol Biol Evol* 2013; **30**(12): 2725-9.
- [13] Tambunan USF, Parikesit AA, Hendra, Taufik RI, Amelia F, Syamsudin. 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* 2009; **9**(1): 6-16.
- [14] Mairiang D, Zhang H, Sodja A, Murali T, Suriyaphol P, Malasit P, et al. Identification of new protein interactions between dengue fever virus and its hosts, human and mosquito. *PLoS One* 2013; **8**(1): e53535.
- [15] Praba D, Menakha M, Jeyanthi KA. Ideal drug for blood pressure. *Int J Pharm Sci Rev Res* 2014; **25**(2): 7-12.
- [16] Alberts B, Bray D, Hopkin K, Johnson A, Lewis J, Raff M, et al. *Essential cell biology*. 3rd ed. Abingdon: Taylor & Francis; 2009.
- [17] Wootten D, Christopoulos A, Sexton PM. Emerging paradigms in GPCR allostery: implications for drug discovery. *Nat Rev Drug Discov* 2013; **12**: 630-44 .
- [18] Shan YB, Kim ET, Eastwood MP, Dror RO, Seeliger MA, Shaw DE. How does a drug molecule find its target binding site? *J Am Chem Soc* 2011; **133**(24): 9181-3.
- [19] Snyder PW, Lockett MR, Moustakas DT, Whitesides GM. Is it the shape of the cavity, or the shape of the water in the cavity? *Eur Phys J Spec Top* 2014; **223**: 853-91.
- [20] Sun P, Ju H, Liu Z, Ning Q, Zhang J, Zhao X, et al. Bioinformatics resources and tools for conformational B-cell epitope prediction. *Comput Math Methods Med* 2013; doi: 10.1155/2013/943636.
- [21] Parulek J, Turkay C, Reuter N, Viola I. Visual cavity analysis in molecular simulations. *BMC Bioinformatics* 2013; **14**(Suppl 19): S4.
- [22] Chandra S, Singh D, Singh TR. Prediction and characterization of T-cell epitopes for epitope vaccine design from outer membrane protein of *Neisseria meningitidis* serogroup B. *Bioinformation* 2010; **5**(4): 155-61.