Available online at www.worldscientificnews.com

World Scientific News

An International Scientific Journal

WSN 127(3) (2019) 177-190 EISSN 2392-2192

Dengue virus (NS2B/NS3 protease) protein engineering. Part I: An automated design for hotspots stability and site-specific mutations by using HotSpot Wizard 3.0 tool

Madhumita Lahiri[1], **Ipsita Ghosh**[1], **Partha Talukdar**[1,*] and **Soumendra Nath Talapatra**[2]

[1]Department of Botany, Serampore College, University of Calcutta, William Carey Road, Hooghly, West Bengal, India

[2]Department of Biological Science, Seacom Skills University, Kendradangal, Shantiniketan, Birbhum – 731236, West Bengal, India

*E-mail address: parthatalukdar4@gmail.com

*Phone: +919231531769

ABSTRACT

The non-structural dengue virus (DNV) protein, DNV-2 NS2B/NS3 protease is a combination of two proteins as 2B and 3 and these two proteins in complex replicate faster during dengue fever. The objective of the present study was to detect hot spots and design of smart libraries for engineering protein stability, substrate specificity, tunnels and cavities as well as suitable mutability position of studied protein by using an online tool, HotSpot Wizard (version 3.0). The prediction results were obtained in output interface for functional hot spots, stability hot spots (structural flexibility), correlated hot spots and stability hot spots (sequence consensus) from the sequence string. It is concluded that the prediction of pocket and mutability of this protein can easily be identified the structural alternation especially in disease diagnosis and space for ligand binding site in pocket for the potential of new drug design. Moreover, this computational prediction is suggested to compare with experimental hotspots for studied protein in relation to therapeutic efficacies, which are lacking to prevent viral infection.

Keywords: Non-structural protein, NS2B/NS3 protease, protein engineering, HotSpot Wizard, computational tool

(Received 30 March 2019; Accepted 18 April 2019; Date of Publication 19 April 2019)

1. INTRODUCTION

The dengue fever (DF) is caused by dengue virus transmitted through mosquito *Culex* sp. Several researchers have reported that dengue virus (DNV) has four serotypes DENV-1, DENV-2, DENV-3 and DENV-4 having structural and sequence similarities (Drumond et al. 2013; Elahi et al., 2014; Mirza et al., 2016). Drumond et al. (2013) stated that the RNA genome contains a single open reading frame that encodes a precursor polyprotein, which is co- and post-translationally cleaved into three structural proteins such as capsid (C), pre-membrane (prM), envelope (E) and seven non-structural proteins such as NS1, NS2A, NS2B, NS3, NS4A, NS4B, NS5. Other researchers have also been observed the above-mentioned protein structures (Pereira et al., 2008; Rothwell et al., 2009; Suganya and Mahendran, 2016). Among four serotypes, any one can cause DF in human and sometimes characterized by haemorrhage and capillary leakage, dengue haemorrhagic fever and dengue shock syndrome (DHF/DSS). Sometimes, the DF occurred with or without any above-mentioned symptoms (WHO, 1997; 2009).

Generally, protein-protein interaction, which indicated residues of $\Delta\Delta G \geq 2$kcal/mol, is termed as hot spot. In other words, certain residues in protein-protein interactions, also termed as hot spots. These residues have unique and variety of energetic properties, can be designed an important target of protein-protein complex (Morrow and Zhang, 2012). In different experiments it was observed that only a small subset of contact residues obtained significance binding free energy. These residues have been termed 'hot spots' and if mutated then they can disrupt the interaction (Lise et al., 2011). The hot spot residues are contained mostly conserved amino acids. Generally, NS2B/NS3 protease is a combination of two dengue viral proteins as 2B and 3 and these two proteins in complex replicate faster during dengue fever. It was reported that NS2B cofactor is essential for the protease catalytic activity of NS3 (de Almeida et al., 2013). This protein complex is very important for the recent research on anti-dengue drug development. Till date several small molecules against DENV-2 NS2B/NS3 protease showed less success in case of *in silico* study and confusing results have obtained in drug development (Schüller et al., 2011; de Almeida et al., 2013).

Bloom et al. (2005) have emphasized that stability of proteins is of great concern, those who are working on protein engineering research for the implementation of enzymes in industrial sector. The recent study of protein stability is concerned about future utilization of biomolecules in different sectors viz. biocatalyst, disease diagnosis and therapy, nanoscience etc. (Bednar et al., 2015). In general, stability means protein gets unfold and refold during unfavourable environmental conditions as temperature or solvent, etc.

In general, all proteins are simplest form and suitable example of evolvable biological systems as per their potent biochemical functions in which alterations can be noted due to few mutations (Aharoni et al., 2005). Wagner (2005) has revealed that evolvability is robustness to mutations, and proteins are often quite mutationally robust. From past to recent experimental study revealed that several proteins are retaining their native functions due to more than half number of single mutant (Bloom et al., 2005; Rennell et al., 1991; Markiewicz et al., 1994; Guo et al., 2009). From decades, it could be known the function and properties with the molecular mechanisms through computational approach of any protein by researchers (Romero and Arnold, 2009; Currin et al., 2015) and the exact function was unclear by the sequences of protein encode but Sumbalova et al. (2018) predicted protein encode from the sequence in Hotspots Wizard (version 3.0). Generally, enzyme is known as biocatalyst, which has specific substrate

binding ability as lock and key strategy for maintaining biochemical reactions in an organism. In recent trend of research, several computational tools for protein engineering have been developed by researchers mainly detection for tunnel and cavity, smart libraries, mutation positions, functions etc. (Bednar et al., 2015; Pavelka et al., 2009; Zhang et al., 2011; Chovancova et al., 2012; Brezovsky et al., 2013; Damborsky and Brezovsky, 2014; Kozlikova et al., 2014; Sebestova et al., 2014; Bendl et al., 2014; 2016) as well as design mutation and library from protein sequence (Sumbalova et al., 2018). The present study was to predict dengue virus protein (NS2B/NS3 protease) hot spots and design of smart libraries for engineered protein stability, substrate specificity, tunnels and cavities as well as suitable mutability position through the design of library and mutation by using HotSpot Wizards, version 3.0.

2. MATERIALS AND METHODS

2. 1. Selection of protein

The Dengue virus protein, NS2B/NS3 protease, the crystal structure of protein, .pdb files as PDB ID: 2FOM was selected for the present prediction. The crystal structure has been deposited by Erbel et al. (2006) in the protein data bank (www.rcsb.org).

2. 2. Retrieval of protein sequence

The protein sequence as Chain A: GSHMLEADLELERAADVRWEEQAEISGSSP ILSITISEDGSMSIKNEEEEQTLGGGGSGGGG and Chain B: AGVLWDVPSPPPVGK AELEDGAYRIKQKGILGYSQIGAGVYKEGTFHTMWHVTRGAVLMHKGKRIEPSWA DVKKDLISYGGGWKLEGEWKEGEEVQVLALEPGKNPRAVQTKPGLFKTNTGTIGAV SLDFSPGTSGSPIVDKKGKVVGLYGNGVVTRSGAYVSAIANTEKSIEDNPEIEDDIFRK for NS2B/NS3 protease was taken from FASTA file developed by Pearson (1990).

2. 3. Study of automated protein engineering

The protein sequence downloaded from FASTA file and incorporated separately in the input interface of HotSpot Wizard (version 3.0) online software. In this automated prediction study, chains were not specified manually. The HotSpot Wizard 3.0 is free online software for academic purpose to detect automatically hot spots and design of smart libraries for engineered proteins' stability, cavity and tunnels, catalytic activity, substrate specificity and enantioselectivity (Pavelka et al., 2009; Sebestova et al., 2014; Bendl et al., 2014; 2016; Sumbalova et al., 2018). This tool is an upgradation version of of HotSpot Wizard 2.0. On the other hand, this present tool can be utilized for the annotation of protein structures. The older version from 2009 is modified version 2.0 and the version 3.0 is modified from version 2.0 (Pavelka et al., 2009; Sebestova et al., 2014).

This present online server comprises sequence, structural and evolutionary information obtained from 3 databases and 20 computational tools. In previous version (2.0), the online tool integrates annotated residues, in which it can be known easily mutagenesis and designed structure for suitable codons for each implemented strategy. Ultimately, this software helps in comprehensive annotations of protein structures and engineering with the stable design of site-specific mutations and targeted libraries (Bendl et al., 2014; Sebestova et al., 2014). Sumbalova et al. (2018) have developed the workflow steps in HotSpot Wizard 3.0, the calculation is based on four different phases such as particular protein sequence input and homology modelling

(phase-1), quality assessment of the model (phase-2), functional hotspots, stability hotspots (consensus), correlated hotspots and stability hotspots as structural flexibility (phase-3) and smart library and design mutations as single point and multipoint (phase-4) respectively.

For statistical analysis, Z scoring values were obtained for each computational tools such as DCA (Direct Coupling analysis), ELSC (Explicit Likelihood of Subset Variation), McBASC (McLachlan Based Substitution correlation), MI (Mutual Information), aMIc (All Microarray Clustering), OMES (Observed Minus Expected Squared) and SCA (Statistical Coupling Analysis).

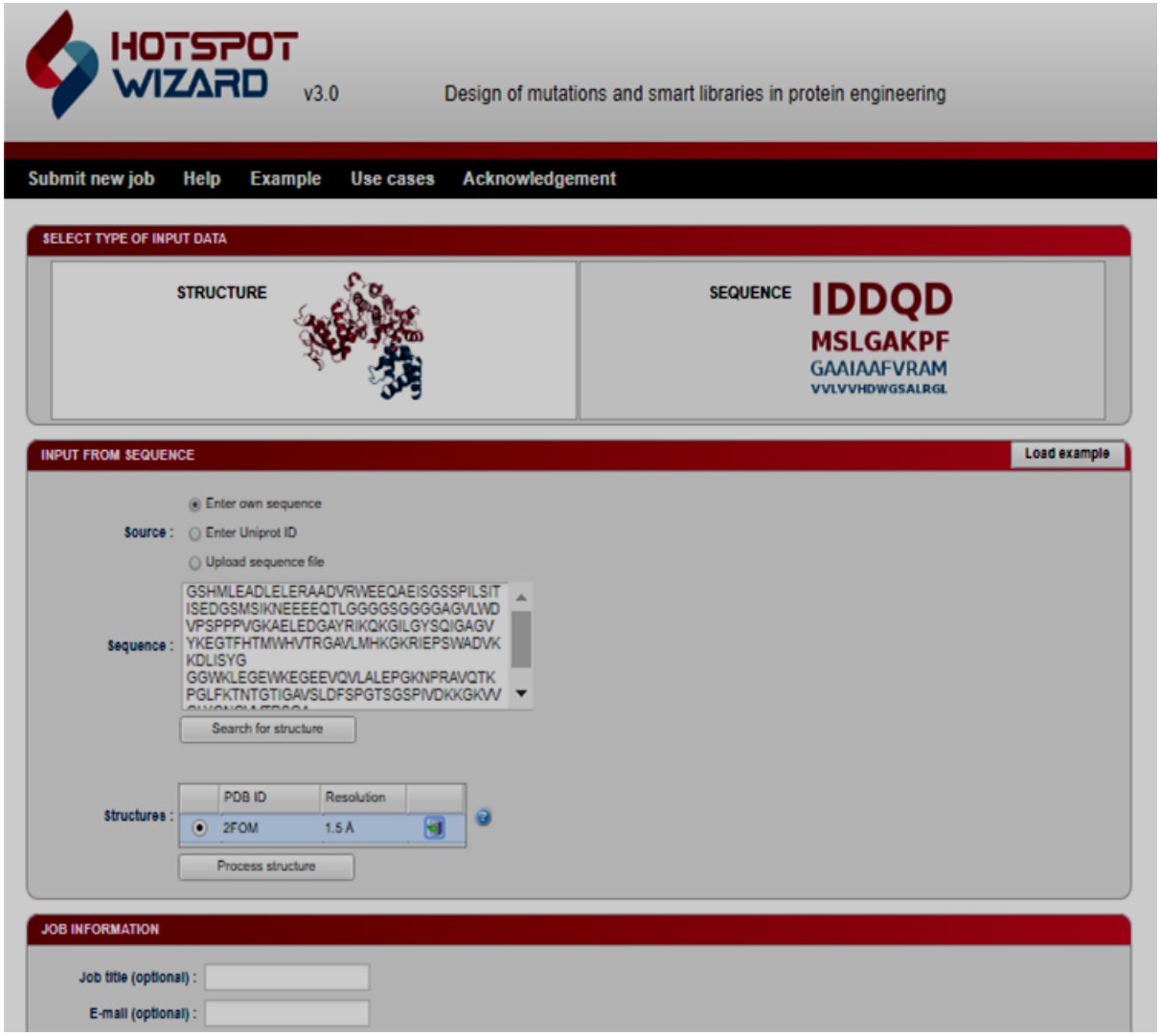


Figure 1. Hotspot wizard (3.0) input interface for NS2B/NS3 protease (PB ID: 2FOM)

3. RESULTS AND DISCUSSION

In the present predictive results, the studied protein engineering strategies through automated computational prediction were observed. Four separate prediction data such as

functional hot spots, stability hot spots (structural flexibility), correlated hot spots and stability hot spots (sequence consensus) for the studied protein through output interface of Hotspots wizard 3.0 is depicted in Figure 2.

In functional hot spots, the data were obtained for activity, substrate specificity and selectivity. Moreover, this step identified residues, which were forming catalytic pocket or accessible tunnel that were not directly participated in the catalysis or located at the evolutionary-conserved position. For stability hot spots (structural flexibility), the prediction was done to identify the residues of flexible structure, which is observed mainly residues with highest B-factors (Table 1). In case of the study of correlated hot spots (Table 2), the data were obtained same as functional hot spots along with the identification of correlated position through consensus approach resulted data from other computational tools viz. aMIc (all Microarray Clustering), DCA (Direct Coupling Analysis), ELSC (Explicit Likelihood of Subset Variation), McBASC (McLachlan Based Substitution correlation), MI (Mutual Information), OMES (Observed Minus Expected Squared) and SCA (Statistical Coupling Analysis). For stability hot spots (sequence consensus), consensus design is an important strategy for the stabilization of proteins. It helps amino acid conservation in sets of homologous protein to identify likely beneficial as well as deleterious mutations of the target protein.

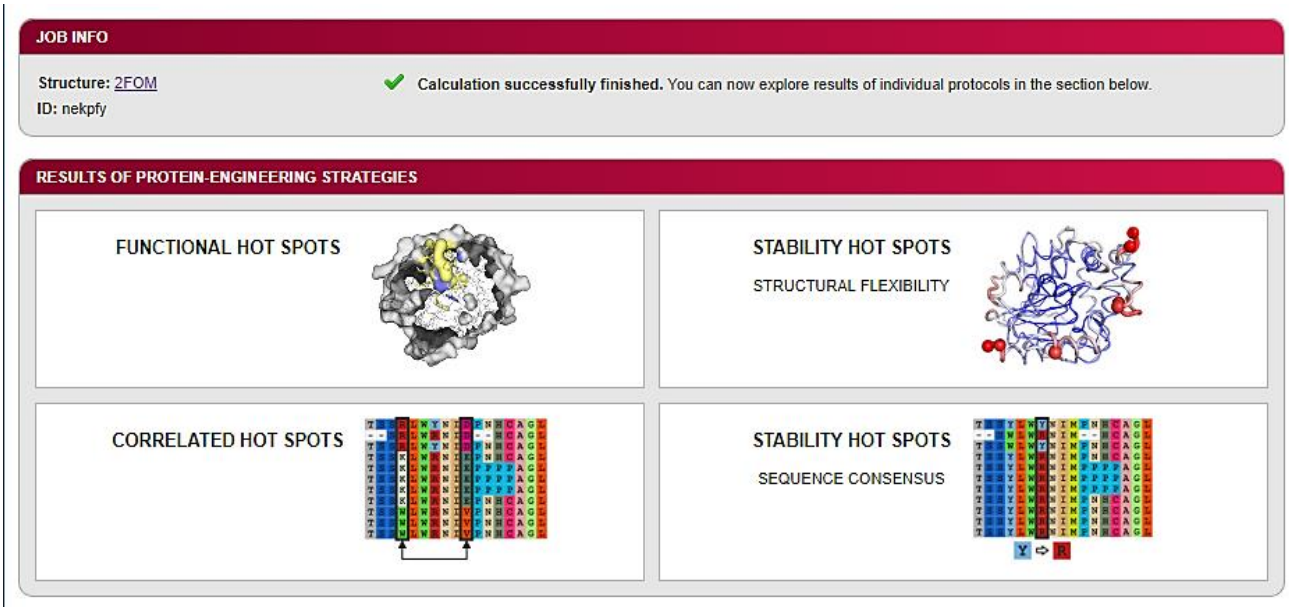


Figure 2. Protein engineering strategies of NS2B/NS3 protease

Figure 3 (A, B and D), revealed that NS2B/NS3 protein showed different hotspots through HotSpot Wizard tool (version 3.0). In general, hot spot determines the energy distribution along with the interface region without homogenous in nature, where certain residues do not contribute majorly for free energy binding (Wells, 1991; Clackson and Wells, 1995; Bogan and Thorn, 1998; Gune et al., 2008; Tuncbag et al., 2009; 2010). The hot spot prediction detects the exact protein binding sites, which helps for designing specific therapeutic agents in protein interactions (Tuncbag et al., 2009). In Figure 3 (C), sequence consensus was obtained for target protein. In this observation, wild-type and mutated consensus sequences were obtained based

on hot spots (Richter et al., 2007). The molecular mechanism as structure and function of any protein can easily be known (Romero and Arnold, 2009; Currin et al., 2015) from the exact function was unclear by the sequences of protein encode but Sumbalova et al. (2018) predicted protein encode from the sequence in Hotspots Wizard (version 3.0). It was documented that experimental evolution work suffered major problems when occurred by several irregular study of mutagenesis and detecting of large sequence libraries to evaluate the mutational landscape and proteins showed important structural and functional properties (Currin et al., 2015; Bendl et al., 2016).

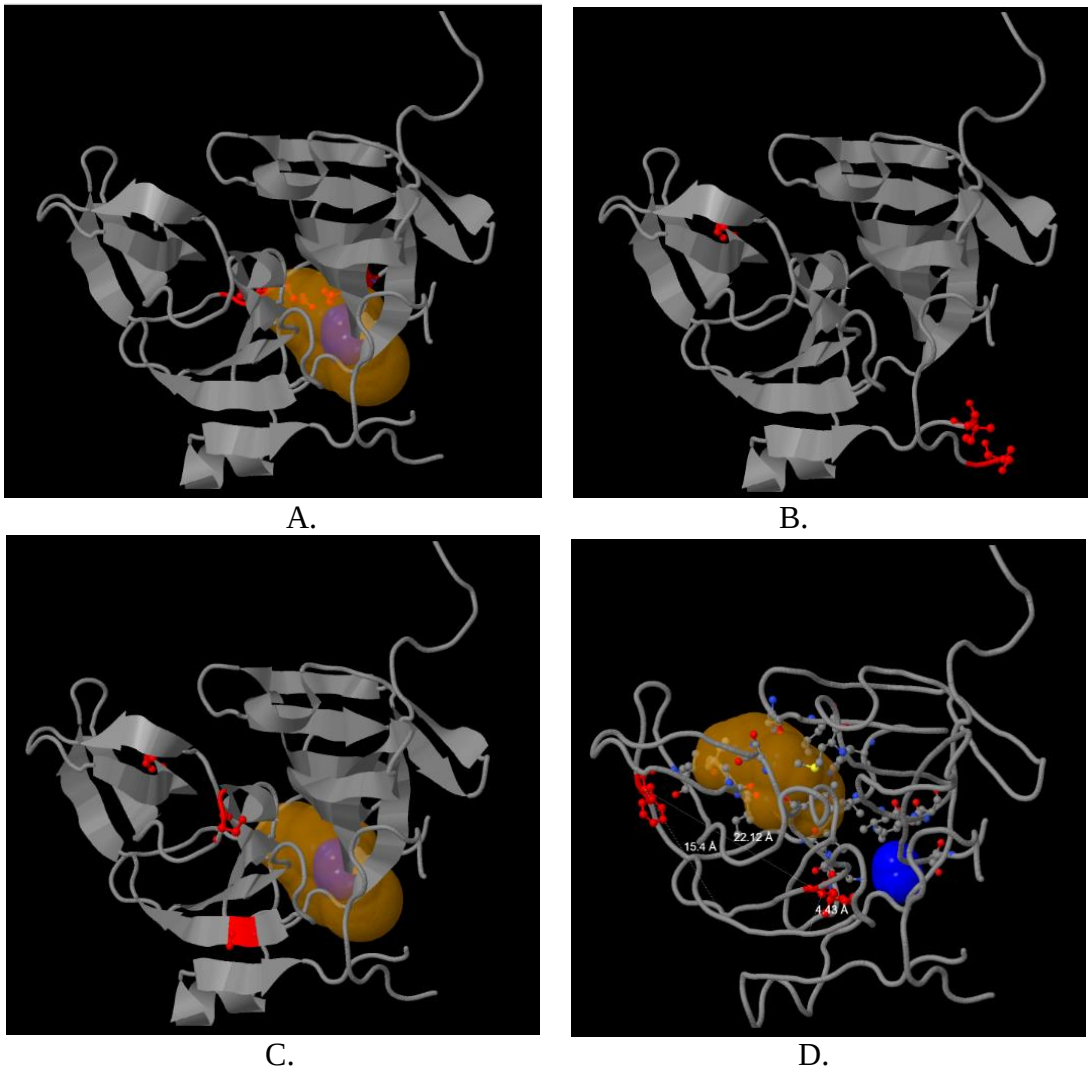


Figure 3. NS2B/NS3 protease protein (2FOM): A = functional hotspot; B = stability hotspot; C = sequence consensus and D = correlated hotspot (yellow ball = pocket; blue ball = tunnel)

Table 1 describes the functional hotspot of NS2B/NS3 where only chain B attached to residues like Glu at 86 position, Gly at 87 position and Lys at 42 position while correlated residues were not observed. The tunnels were not observed while pockets were showed catalytic

as 1 no. in each case and 2, 7 for second one and 1 (from pocket 1) for third one in which B-factor values 18.34, 16.46 and 12.88 Å2 respectively. The B-factor values mainly influenced by crystal contacts and solvent conditions, various theoretical methods have used to predict flexible regions, which help to determine the targets for stabilization (Reetz et al., 2006; Wijma et al., 2013). The mutability rate was observed high and score values were 8, 8 and 7 respectively.

Table 1. Study of functional hotspots of studied protein.

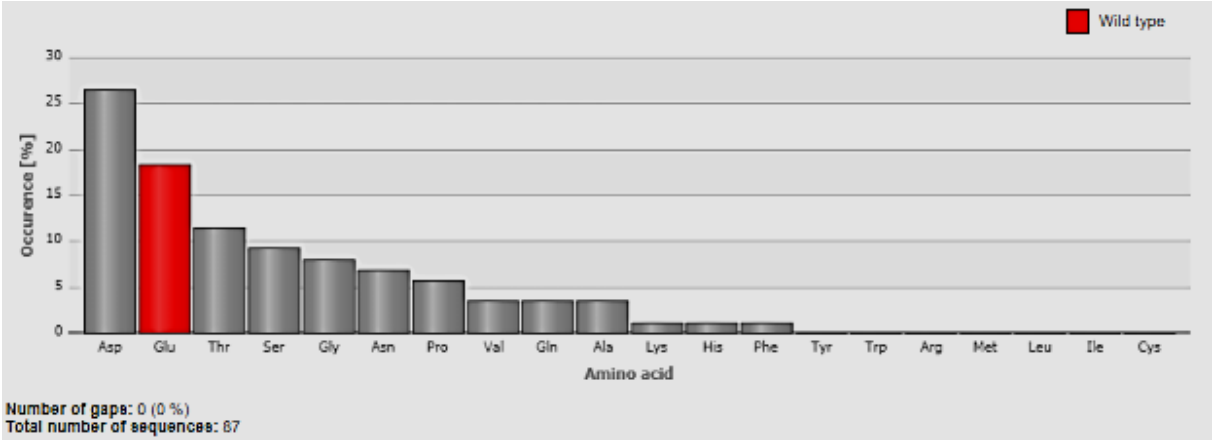
Studied Protein	Chains	Residues & position	Secondary structures	Pockets & tunnels	Average B-factor (in $Å^2$)	Mutability rate & score	Correlated residues & position
2FOM	B	Glu & 86	Extended strand (E)	1 (catalytic)	18.34	High & 8	---
	B	Gly & 87	Extended strand (E)	1 (catalytic) 2,7	16.46	High & 8	---
	B	Lys & 42	Extended strand (E)	1 (catalytic) & 1 (from pocket 1)	12.88	High & 7	---

Zuo et al. (2009) stated that the residues Glu92 and Asp50 in the NS2B cofactor and residues Gln27, Gln35, and Arg54 in NS3pro have been directly involved in interaction between the ligand and the binding pockets on the enzyme. The kinetic study of the proteolytic activity of wild-type and mutated NS2B-NS3pro enzymes supported that the residues identified by modelling are involved in substrate binding. The study on computational and biofunctional features it was found an evidence for potential dual role for NS2B cofactor-mediated activation of NS3pro and they have explained that the folding of the enzyme complex to a catalytically active conformational state and provided direct but weak secondary interaction sites for substrate binding.

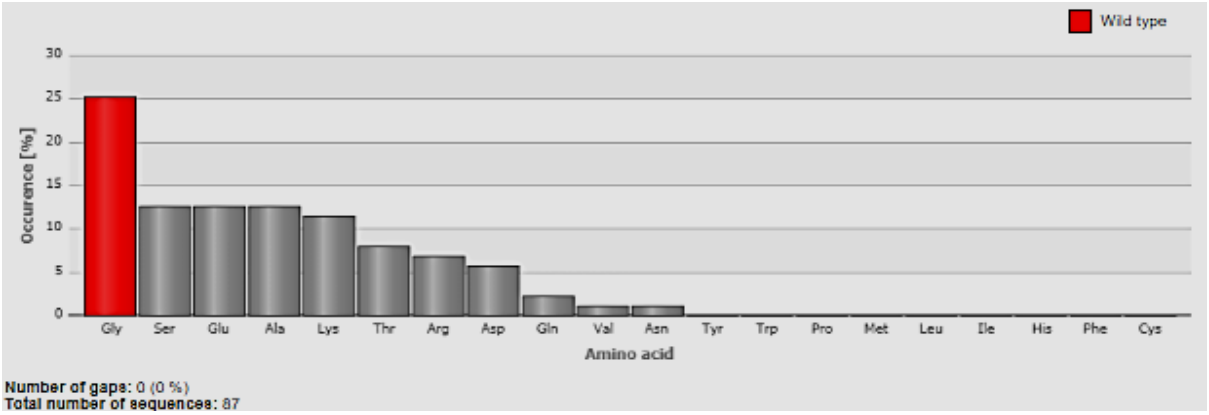
Table 2. Values obtained from different tools for functional hot spots of studied protein

Studied Protein	Consensus z-scoring values							
	Chains	MI	aMIc	DCA	ELSC	McBASC	OMES	SCA
2FOM	B	1.66	2.60	1.46	6.65	1.24	7.84	3.81
	B	1.88	2.58	1.57	4.79	1.19	9.58	5.75
	B	2.91	2.95	2.47	4.42	1.18	7.69	4.41

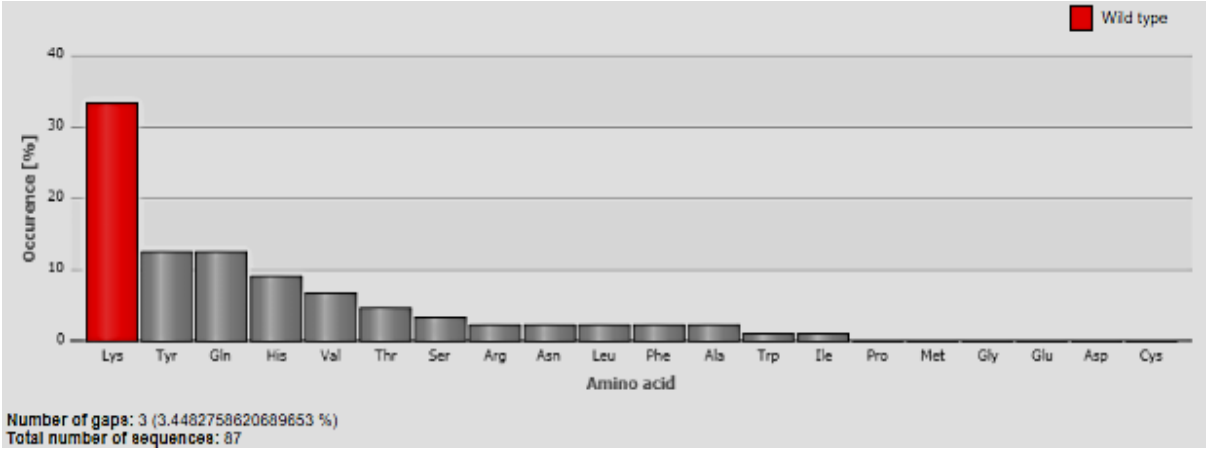
In Table 2, consensus z-scoring value was obtained for different tools such as MI 1.66, 1.88 and 2.91, aMIc 2.60, 2.58 and 2.95; DCA 1.46, 1.57 and 2.47; ELSC 6.65, 4.79, 4.42; McBASC 1.24, 1.19 and 1.18; OMES 7.84, 9.58 and 7.69 and SCA 3.81, 5.75 and 4.41 were obtained by using this tool for NS2B/NS3.



A.



B.



C.

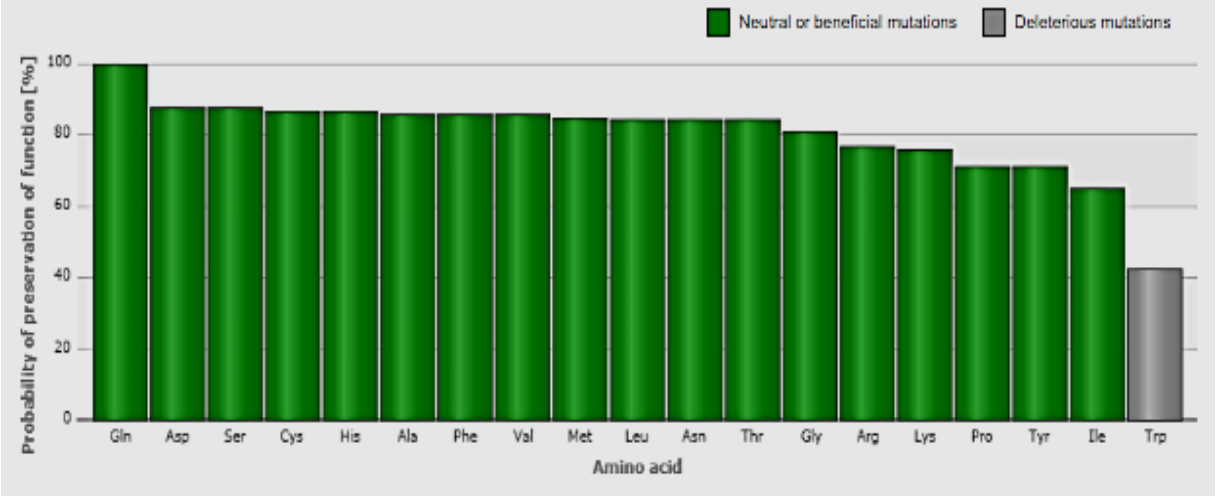
Figure 4. Amino acids frequencies as per positions.

In Fig. 4, it was obtained that the amino acid residues fulfilling the criterion of minimal frequency in the multiple sequence alignment. The wild type variety was observed Glu (19%), Gly (25%) and Lys (33%) as per positions of different amino acids frequencies of Ns2B/NS3 (Fig. 4A, B and C). Fig. 5 states that mutational landscape, which mainly showed the estimation of the probability in relation to preservation of protein function for individual substitution at the particular site of NS2B/NS3. It was obtained that deleterious mutation was only found in Trp and Pro (Fig. 5 A and C) but no deleterious mutation was observed (Fig 5 B). In the present computational study, the results were obtained for B chain, which indicated a strong linking with the transitions than the A chain for NS2B/NS3 enzyme.

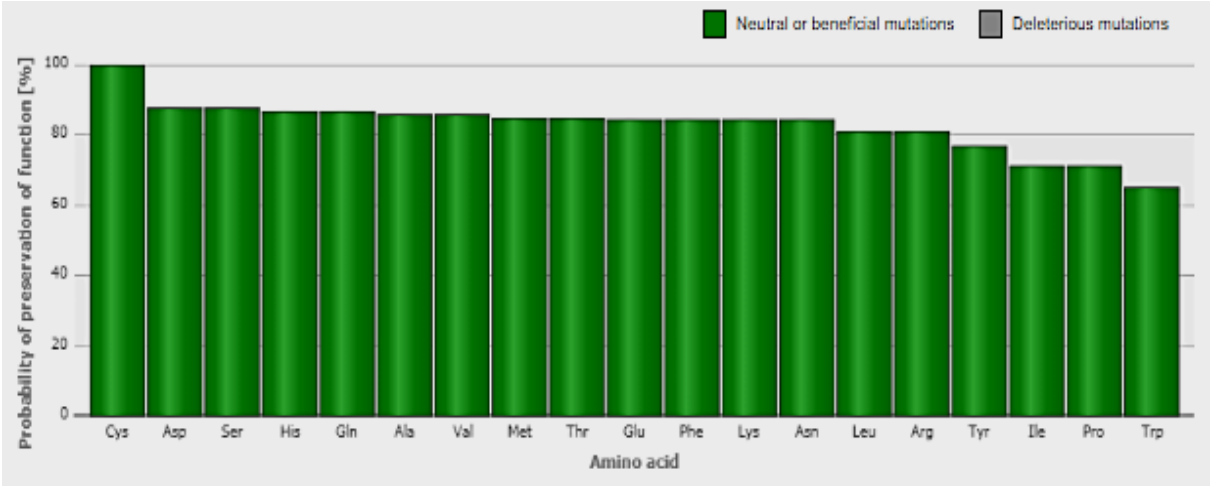
It was reported that mutability scale is ranged between 1 to 9 i.e. lower to higher rate. In the present study high mutability rate was observed 8, 8 and 7 respectively (Table 1). In Fig. 5A and C, amino acids viz. Trp (52%) and Pro (24%) were observed deleterious mutation of amino acids but in Fig. 5B, no deleterious mutation. According to Zuo et al. (2009), the kinetic fitness of the three enzyme variants compared with the wild type, which provided some indication that the residues of Asp50 and Glu92 on NS2B and Gln27, Gln35, and Arg54 on NS3pro interacted directly with the substrate during the proteolytic process. On the other hand, the experiment on yellow fever virus NS2B-NS3pro indicated that charged-to-alanine mutations in the NS2B region have less deleterious effects on cleavage activity compared to the protease domain (Chambers et al., 2005). Moreover, the effects of the mutations on substrate binding and catalytic efficiency of the NS2B/NS3 proteases are weak (Zuo et al., 2009).

However, the prediction of different hotspots can be facilitated drug designing and development. It was suggested that the starting point of a binding site of a receptor in the hotspots may be granted to analyse docking of ligands (Gonzalez-Ruiz and Gohlke, 2006). On the other hand, rigid docking lead to an achievement the comparatively least flexible hotspots, which lead to an upgradation in protein docking has been performed by creating dominant conformation of the hotspot side chains resulted through molecular dynamics probing rather than the unbound X-ray conformation (Rajamani et al., 2004; Ozbabacan et al., 2010).

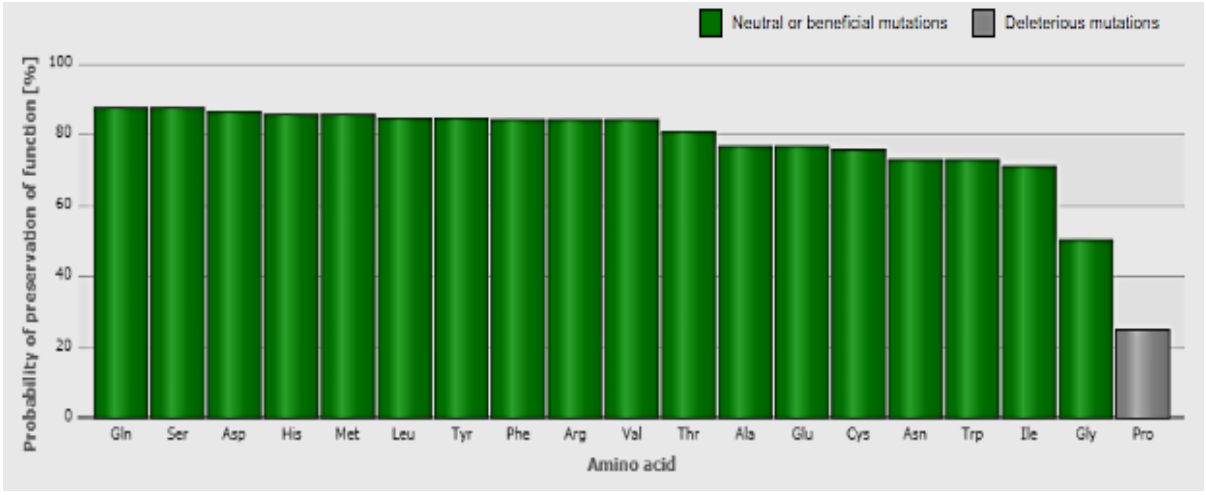
Thus, the prediction of hotspots is a suitable tool to identify exact functional mechanisms of particular protein to identify mutant residue(s) in relation to cause of disease and new drug development.



A.



B.



C.

Figure 5. Amino acids mutability landscape

4. CONCLUSIONS

In conclusion, HotSpot Wizard (version 3.0) is an online computational tool, which helped easily to obtain results for oxy-haemoglobin through protein engineering protocol by the integration of several inbuilt databases derived from other bioinformatics tools and all the data generated within short duration to prevent laborious jobs of experiment (Kozlikova et al., 2014). This software also helped to incorporate only FASTA sequence file as an input of studied protein without prior knowledge of computational biology to set up input interface. The parameters like pocket identification and mutability prediction of NS2B/NS3 protease can lead to know structural alternation in the particular disease diagnosis as well as space for ligand binding pocket in new anti-dengue drug design. It is suggested in future to compare with experimental hotspots for the validation of the present prediction work for therapeutic efficacies and druggability assessment.

Acknowledgement

Authors are thankful to the developers of online software HotSpot Wizards (Version 3.0), European Protein Databank and FASTA sequence file for NS2B/NS3 protease enzyme (2FOM), used in the present study.

References

[1] Drumond BP, Mondini A, Schmidt DJ, Bronzoni RVdM, Bosch I, Nogueira ML. Circulation of different lineages of dengue virus 2, genotype American/ Asian in Brazil: Dynamics and molecular and phylogenetic characterization. *Plos One* 8(3) (2013) e59422. doi:10.1371/journal.pone.0059422

[2] Mirza SB, Salmas RE, Fatmi MQ, Durdagi S. Virtual screening of eighteen million compounds against dengue virus: Combined molecular docking and molecular dynamics simulations study. *Journal of Molecular Graphics and Modelling* 66 (2016) 99-107

[3] Elahi M, Islam MM., Noguchi K, Yohda M, Toh H, Kuroda Y. Computational prediction and experimental characterization of a "size switch type repacking" during the evolution of dengue envelope protein domain III (ED3). *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics* 1884(3) (2014) 585-592

[4] Pereira R, Kunh JR. Structural proteomics of dengue virus. *Current Opinion in Microbiology* 11 (2008) 369

[5] Rothwell C, Lebreton A, Young NgC, Lim JYH, Liu W, Vasudevan S, Labow M, Gu F, Gaither LA. Cholesterol biosynthesis modulation regulates dengue viral replication. *Virology* 389 (2009) 8-19

[6] Suganya. J, Mahendran R. Insilico QSAR and molecular docking studies of selected medicinal plant compounds against NS5 & NS3 protein of dengue virus: a comparative approach. *International Journal of Pharma and Bio Sciences* 7(3) (2016) B1135-1144

[7] World Health Organization (2009). Dengue, guidelines for diagnosis, treatment, prevention and control. Accessed 26 February 2019 Available: http://whqlibdoc.who.int/publications/ 2009/9789241547871_eng.pdf

[8] World Health Organization (1997). Dengue Hemorrhagic Fever: Diagnosis, Treatment, Prevention and Control. Accessed 26 February 2019. Available: http://www.who.int/csr/ resources/publications/dengue/Denguepublication/en/

[9] Morrow JK, Zhang S. Computational prediction of hot spot residues. *Current Pharmaceutical Design* 18(9) (2012) 1255-1265

[10] Lise S, Buchan D, Pontil M, Jones DT. Predictions of hot spot residues at protein-protein interfaces using support vector machines. *Plos One* 6(2) (2011) e16774. doi:10.1371/journal.pone.0016774

[11] de Almeida H, Bastos IMD, Ribeiro BM, Maigret B, Santana JM New binding site conformations of the Dengue Virus NS3 protease accessed by molecular dynamics simulation. *Plos One* 8(8) (2013) e72402. doi:10.1371/journal.pone.0072402

[12] Schüller A, Yin Z, Brian Chia CS, Doan DNP, Kim H-K et al. Tripeptide inhibitors of dengue and West Nile virus NS2B-NS3 protease. *Antiviral Research* 92 (2011) 96-101. doi:10.1016/j.antiviral.2011.07.002

[13] Bloom JD, Silberg JJ, Wilke CO, Drummond DA, Adami C Arnold FH. Thermodynamic prediction of protein neutrality. *Proceedings of the National Academy of Sciences USA* 102 (2005) 606-611

[14] Bednar D, Beerens K, Sebestova E, Bendl J, Khare S, Chaloupkova R, Prokop Z, Brezovsky J, Baker D, Damborsky J. FireProt: Energy- and evolution based computational design of thermostable multiple-point mutants. *Plos Computational Biology* 11 (2015) e1004556

[15] Aharoni A, Gaidukov L, Khersonsky O, Gould SM, Roodveldt C, Tawfik DS. The 'evolvability' of promiscuous protein functions. *Nature Genetics* 37 (2005) 73-76.

[16] Wagner A. Robustness and evolvability in living systems. Princeton University Press, Princeton (2005).

[17] Rennell D, Bouvier SE, Hardy LW and Poteete AR. Systematic mutation of bacteriophage T4 lysozyme. *Journal of Molecular Biology* 222 (1991) 67-87

[18] Markiewicz P, Kleina LG, Cruz C, Ehret S, Miller JH. Genetic studies of the lac repressor. XIV. Analysis of 4000 altered Escherichia coli lac repressors reveals essential and non-essential residues, as well as "spacers" which do not require a specific sequence. *Journal of Molecular Biology* 240 (1994) 421-433

[19] Guo HH, Choe J, Loeb LA. Protein tolerance to random amino acid change. *Proceedings of the National Academy of Sciences USA* 101 (2004) 9205-9210

[20] Romero RA, Arnold FH. Exploring protein fitness landscapes by directed evolution. *Nature Reviews Molecular Cell Biology* 10 (2009) 866-876

[21] Currin A, Swainston N, Day PJ, Kell DB. Synthetic biology for the directed evolution of protein biocatalysts: navigating sequence space intelligently. *Chemical Society Reviews* 44 (2015) 1172-1239

[22] Pavelka A, Chovancova E, Damborsky J. HotSpot Wizard: A web server for identification of hot spots in protein engineering. *Nucleic Acids Research* 37 (2009) W376-W383

[23] Zhang Z, Li Y, Lin B, Schroeder M, Huang B. Identification of cavities on protein surface using multiple computational approaches for drug binding site prediction. *Bioinformatics* 27 (2011) 2083-2088

[24] Chovancova E, Pavelka A, Benes P, Strnad O, Brezovsky J, Kozlikova B, Gora A, Sustr V, Klvana M, Medek P, Biedermannova L, Sochor J, Damborsky J. CAVER 3.0: A tool for analysis of transport pathways in dynamic protein structures. *Plos Computational Biology* 8 (2012) e1002708

[25] Brezovsky J, Chovancova E, Gora A, Pavelka A, Biedermannova L, Damborsky J. Software tools for identification, visualization and analysis of protein tunnels and channels. *Biotechnology Advances* 31 (2013) 38-49

[26] Damborsky J, Brezovsky J. Computational tools for designing and engineering enzymes. *Current Opinion in Chemical Biology* 19 (2014) 8-16

[27] Kozlikova B, Sebestova E, Sustr V, Brezovsky J, Strnad O, Daniel L, Bednar D, Pavelka A, Manak M, Bezdeka M, Benes P, Kotry M, Gora A, Damborsky J, Sochor J. CAVER Analyst 1.0: Graphic tool for interactive visualization and analysis of tunnels and channels in protein structures. *Bioinformatics* 30 (2014) 2684-2685

[28] Sebestova E, Bendl J, Brezovsky J, Damborsky J. Computational tools for designing smart libraries. *Methods in Molecular Biology* 1179 (2014) 291-314

[29] Bendl J, Stourac J, Sebestova E, Vavra O, Musil1 M, Brezovsky J, Damborsky J. HotSpot Wizard 2.0: automated design of site-specific mutations and smart libraries in protein engineering. *Nucleic Acids Research* 44(W1) (2016) W479-W487

[30] Sumbalova L, Stourac J, Martinek T, David Bednar D, Jiri Damborsky J. HotSpot Wizard 3.0: web server for automated design of mutations and smart libraries based on sequence input information. *Nucleic Acids Research* 46 (2018) W356-W362

[31] Pearson WR. Rapid and sensitive sequence comparison with FASTP and FASTA. *Methods in Enzymology* 183 (1990) 63-98

[32] Prokop Z, Sato Y, Brezovsky J, Mozga T, Chaloupkova R, Koudelakova T, Jerabek P, Stepankova V, Natsume R, Leeuwen JGE, Janssen DB, Florian J, Nagata Y, Senda T, Damborsky J. Enantioselectivity of haloalkane dehalogenases and its modulation by surface loop engineering. *Angewandte Chemie International Edition* 49 (2010) 6111-6115

[33] Sebestova E, Bendl J, Brezovsky J, Damborsky J. Computational tools for designing smart libraries, in: Directed evolution library creation: Methods and protocols, methods in molecular biology, Gillam EMJ, Copp JN, Ackerley DF, Eds. 2nd edition, Humana Press, New York (2014) 291-314.

[34] Wells JA. Systematic mutational analyses of protein-protein interfaces. *Methods in Enzymology* 202 (1991) 390-411

[35] Clackson T, Wells JA. A hot spot of binding energy in a hormone-receptor interface. *Science* 267 (1995) 383-386

[36] Bogan AA, Thorn KS. Anatomy of hot spots in protein interfaces. *Journal of Molecular Biology* 280(1) (1998) 1-9

[37] Guney E, Tuncbag N, Keskin O, Gursoy A. HotSprint: database of computational hot spots in protein interfaces. *Nucleic Acids Research* 36 (2008) D662-D666.

[38] Tuncbag N, Gursoy A, Keskin O. Identification of computational hot spots in protein interfaces: combining solvent accessibility and inter-residue potentials improves the accuracy. *Bioinformatics* 25(12) (2009) 1513-1520

[39] Tuncbag N, Keskin O, Gursoy A. HotPoint: hot spot prediction server for protein interfaces. *Nucleic Acids Research* 38 (2010) W402-W406

[40] Richter SN, Giaretta G, Comuzzi V, Leo E, Mitchenall LA, Fisher LM, Maxwell A, Palumbo M. Hot-spot consensus of fluoroquinolone-mediated DNA cleavage by Gram-

negative and Gram-positive type II DNA topoisomerases. *Nucleic Acids Research* 35(18) (2007) 6075-6085

[41] Lutz S. Beyond directed evolution–semi-rational protein engineering and design. *Current Opinion in Biotechnology* 21 (2010) 734-743

[42] Cheng F, Zhu L, Schwaneberg U. Directed evolution 2.0: improving and deciphering enzyme properties. *Chemical Communications. (Camb.)* 51 (2015) 9760-9772

[43] Acevedo-Rocha CG, Reetz MT, Nov Y. Economical analysis of saturation mutagenesis experiments. *Scientific Reports* 5 (2015) 10654

[44] Reetz MT, Carballeira JD, Vogel A. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. *Angewandte Chemie International Edition England* 45 (2006) 7745-7751

[45] Wijma HJ, Floor RJ, Janssen DB. Structure- and sequence-analysis inspired engineering of proteins for enhanced thermostability. *Current Opinion in Structural Biology* 23 (2013) 588-594

[46] Zuo Z, Liew OW, Chen G, Chong PC, Lee SH, Chen K, Jiang H, Puah CM, Zhu W. Mechanism of NS2B-mediated activation of NS3pro in dengue virus: molecular dynamics simulations and bioassays. *Journal of Virol*ogy 83(2) (2008) 1060-1070

[47] Chambers TJ, Droll DA, Tang Y, Liang Y, Ganesh VK, Murthy KH, Nickells M. Yellow fever virus NS2B-NS3 protease: characterization of charged-to-alanine mutant and revertant viruses and analysis of polyprotein-cleavage activities. *Journal of General Virology* 86 (2005) 1403-1413

[48] Gonzalez-Ruiz D, Gohlke H. Targeting protein-protein interactions with small molecules: challenges and perspectives for computational binding epitope detection and ligand finding. *Current Medicinal Chemistry* 13 (2006) 2607-2625