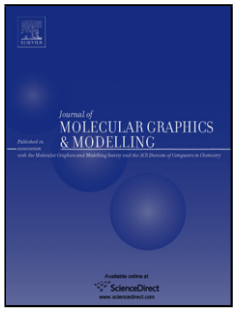


Title: Binding Specificity of Polypeptide Substrates in NS2B/NS3pro Serine Protease of Dengue Virus Type 2: A Molecular Dynamics Study

Author: Pathumwadee Yotmanee Thanyada Rungrotmongkol Kanin Wichapong Sy Bing Choi Habibah A. Wahab Nawee Kungwan Supot Hannongbua

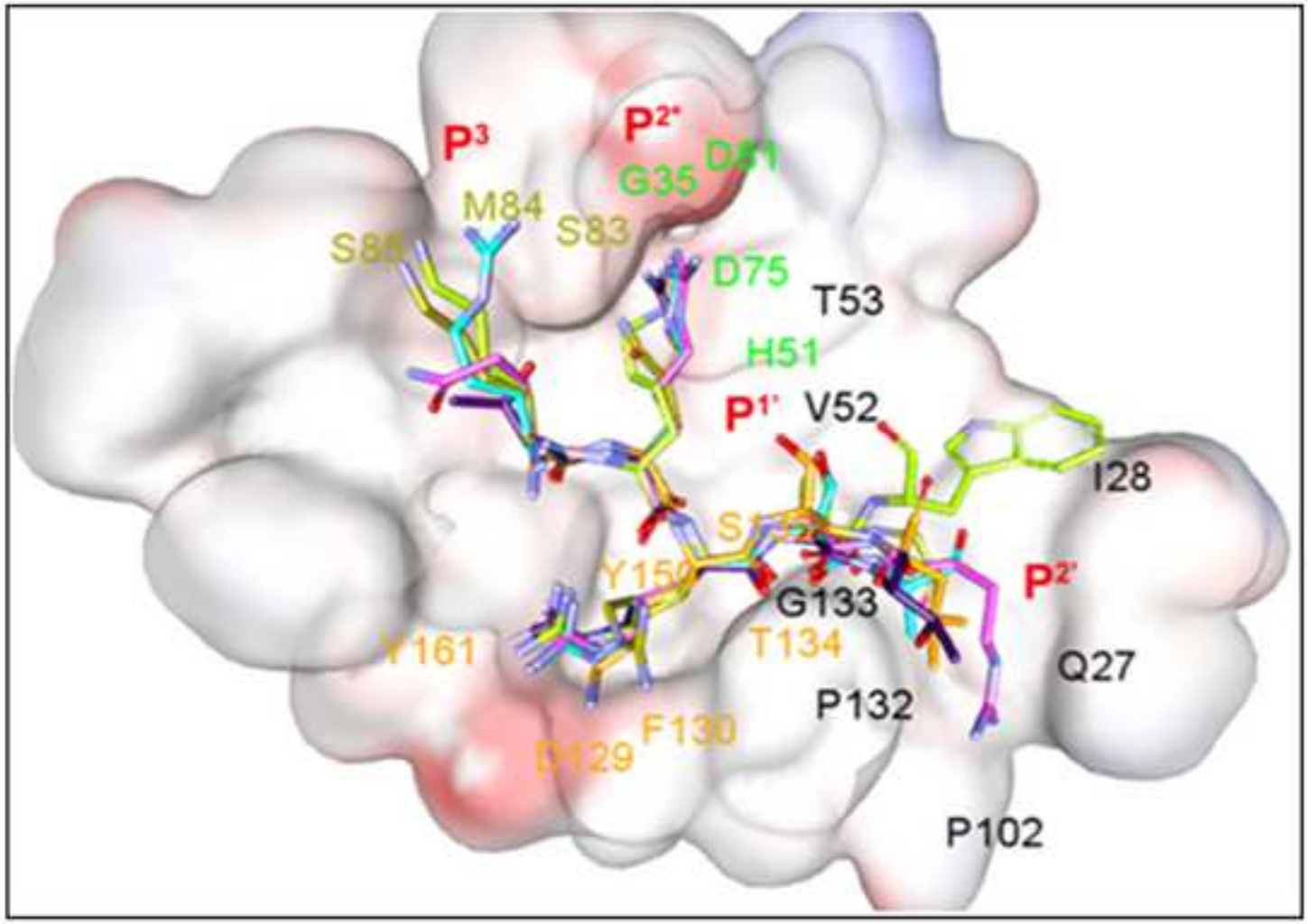
PII: S1093-3263(15)00086-8
DOI: http://dx.doi.org/doi:10.1016/j.jmgm.2015.05.008
Reference: JMG 6548

To appear in: *Journal of Molecular Graphics and Modelling*

Received date: 26-1-2015
Revised date: 14-5-2015
Accepted date: 15-5-2015

Please cite this article as: P. Yotmanee, T. Rungrotmongkol, K. Wichapong, S.B. Choi, H.A. Wahab, N. Kungwan, S. Hannongbua, Binding Specificity of Polypeptide Substrates in NS2B/NS3pro Serine Protease of Dengue Virus Type 2: A Molecular Dynamics Study, *Journal of Molecular Graphics and Modelling* (2015), http://dx.doi.org/10.1016/j.jmgm.2015.05.008

Molecular dynamics study revealed that the P1 and P2 subsite of the substrate play a major role in the binding with NS2B/NS3pro of dengue virus type 2.



Highlights

- Binding specificity of polypeptide substrates in DV2 NS2B/NS3 protease is ranked as capsid > intNS3 > 2A/2B > 4B/5 > 3/4A > 2B/3.

- Basic residues at P^1 and P^2 subsites of substrate are a vital role in the binding with NS2B/NS3pro.

- The polar and charged NS3 residues within S^1 (D^{129}, G^{133} and S^{135}), S^2 (D^{75} and N^{152}) and S^3 (G^{153} and Y^{161}) pockets are preserved for important interaction with the substrate, while the binding of prime site with $S^{1'}$, $S^{2'}$ and $S^{4'}$ pockets is possessed by hydrophobic interaction.

Binding Specificity of Polypeptide Substrates in NS2B/NS3pro Serine Protease of Dengue Virus Type 2: A Molecular Dynamics Study

Pathumwadee Yotmanee[1,2], Thanyada Rungrotmongkol[3],[4] Kanin Wichapong[2], Sy Bing Choi[5], Habibah A. Wahab[5,*], Nawee Kungwan[6], Supot Hannongbua[2,*]

[1]*Department of Chemistry, Faculty of Science, Ramkhamhaeng University, Bangkok 10240, Thailand.*

[2]*Computational Chemistry Unit Cell, Department of Chemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.*

[3]*Ph.D. Program in Bioinformatics and Computational Biology, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.*

[4]*Department of Biochemistry, Faculty of Science, Chulalongkorn University, Bangkok 10330, Thailand.*

[5]*School of Pharmaceutical Sciences, Universiti Sains Malaysia, Pulau Pinang 11800 Malaysia.*

[6]*Department of Chemistry, Faculty of Science, Chiang Mai University, Chiang Mai 50200, Thailand.*

*Corresponding authors. S. Hannongbua: Tel.: +66 22 187602; Fax: +66 22 187603;E-mail: supot.h@chula.ac.th. H. A. Wahab: Tel.: +60 46532236; Fax: +60 46570017; E-mail: habibahw@usm.my.

Abstract

The pathogenic dengue virus (DV) is a growing global threat, particularly in South East Asia, for which there is no specific treatment available. The virus possesses a two-component (NS2B/NS3) serine protease that cleaves the viral precursor proteins. Here, we performed molecular dynamics simulations of the NS2B/NS3 protease complexes with six peptide substrates (capsid, intNS3, 2A/2B, 4B/5, 3/4A and 2B/3 containing the proteolytic site between P^1 and $P^{1'}$ subsites) of DV type 2 to compare the specificity of the protein-substrate binding recognition. Although all substrates were in the active conformation for cleavage reaction by NS2B/NS3 protease, their binding strength was somewhat different. The simulated results of intermolecular hydrogen bonds and decomposition energies suggested that among the ten substrate residues (P^5-$P^{5'}$) the P^1 and P^2 subsites play a major role in the binding with the focused protease. The arginine residue at these two subsites was found to be specific preferential binding at the active site with a stabilization energy of < -10 kcal/mol. Besides, the P^3, $P^{1'}$, $P^{2'}$ and $P^{4'}$ subsites showed a less contribution in binding interaction (< -2 kcal/mol). The catalytic water was detected nearby the carbonyl oxygen of the P^1 reacting center of the capsid, intNS3, 2A/2B and 4B/5 peptides. These results led to the order of absolute binding free energy (ΔG_{bind}) between these substrates and the NS2B/NS3 protease ranked as capsid > intNS3 > 2A/2B > 4B/5 > 3/4A > 2B/3 in a relative correspondence with previous experimentally derived values.

K**eywords**: *Dengue virus, NS2B/NS3 protease, polypeptide, pathogenic, molecular dynamic*s

1. Introduction

Dengue virus (DV), transmitted by *Aedes aegypti* and *Ae. albopictus* mosquitoes, has become a major international public health concern in recent decades since it causes dengue fever, dengue hemorrhagic fever and dengue shock syndrome in the tropical and sub-tropical regions around the world, predominantly in Asian countries. The World Health Organization currently estimates that there may be 50 million dengue infections worldwide every year (http://www.who.int/mediacentre/factsheets/fs117/en/index.html). To date, there is no effective vaccine or antiviral drug available against these dengue diseases [1, 2]. Thus, to overcome this or any other future DV endemics, a detailed understanding at the molecular level of DV replication and inhibition is crucial, particularly for the design of novel dengue antivirus inhibitors.

DV belongs to the flavivirus genus, which also contains several other important human-pathogens, such as yellow fever, tick-borne encephalitis, west nile virus (WNV) and Japanese encephalitis viruses [3]. Currently DV is classified into four genetically distinct serotypes (DV1, DV2, DV3 and DV4), but the whole genome of all DV serotypes is made up of a single stranded (ss) RNA inside a 50-nm diameter virion that consists of an icosahedral nucleocapsid enclosed in a lipoprotein envelope [4]. The ssRNA nucleotides encode for a single polyprotein precursor that after processing is comprised of three structural (capsid (C), membrane (prM) and envelope (E)) and seven non-structural (NS) proteins (NS1, NS2A, NS2B, NS3, NS4A, NS4B and NS5) arranged as C-prM-E-NS1-NS2A-NS2B-NS3-NS4A-NS4B-NS5 [5]. This polyprotein precursor is simply referred to hereafter as C-prM-E-1-2A-2B-3-4A-4B-5. The replication and maturation of DV is very much dependent upon the correct cleavage of the precursor, and requires both host cell proteases (furin and secretase) and a virus-encoded protease [6] that is composed of two components, namely NS2B and NS3, respectively. NS3 is a multifunctional protein of approximately 69 kDa that consists of

a serine protease, helicase, nucleoside 5'-triphosphatase and 5'-terminal RNA triphosphatase domain and plays an important role in viral polyprocessing and genome replication.[7] The 180 amino acid residues trypsin-like serine protease domain at the N-terminal of NS3 suggests its role as a putative viral protease. For NS2B, it is a 14 kDa protein that acts as a cofactor containing the hydrophilic 40 amino acids important for NS3 activation. It has previously been reported that the optimal catalytic activity of the NS3 protease (NS3pro) can only be achieved when NS2B is bound [8]. Therefore, the existence of the NS2B/NS3pro complex is essential to cleave the viral polyprotein at the internal regions within the capsid and NS3 (intNS3) protein as well as at the 2A/2B, 2B/3, 3/4A and 4B/5 boundaries (junctions).

The crystal structures of DV2 NS3pro (1BEF.pdb [9]) and DV2 NS3pro complexed with the Mung-Bean Bowman-Birk inhibitor (1DF9.pdb [10]) are devoid of the cofactor NS2B yet it is crucial for the NS2B/NS3 protease functionality. Although the crystal structure of the DV2 NS2B/NS3pro complex (2FOM.pdb) was available, the information on substrate/inhibitor binding was not inferable. Alternatively, the two X-ray structures of the related WNV NS2B/NS3pro complexed with inhibitor are available (3E90.pdb and 2FP7.pdb [7, 11]). The lower resolution structure is WNV NS2B/NS3pro complexed with the peptidic inhibitor, benzoyl-norleucine(P^4)-lysine(P^3)-arginine(P^2)-arginine(P^1)-aldehyde (Bz-Nle-K-R-R-H) (2FP7.pdb [11]). The NS2B/NS3pro homologs from WNV and DV2 (2FP7 and 2FOM) share a 56% amino acid sequence identity that indicates their relative structural similarity. Interestingly, the conformation of the cofactor NS2B in the inhibitor WNV protease complex (2FP7) significantly differs from the apo DV2 protease (2FOM), while the conformation of the NS3 domain is more conserved. It was recently reported that the DV2 NS2B domain directly interacts with the substrate-binding site of NS3pro to contribute in the formation of the S^2 and S^3 sub-pockets, but this is not observed in the DV2 crystal structure

in the apo-form (2FOM [12]). Thus, the apo DV2 NS2B/NS3pro structure does not appear to represent a suitable starting structure for studying the substrate/inhibitor binding mode. Based on the available experimental data, it is suggested that the active inhibitor complex form of DV2 NS2B/NS3pro should be close to that of the WNV protease. Therefore, a homology model of the DV2 NS2B/NS3pro with inhibitor bound was constructed by using the WNV protease crystal structure as template [13]. Up to date, some potential inhibitors have been reported [14-23], but unfortunately no compound has been shown to be appropriate for further drug development. This is probably due to the fact that the active site is flat and negatively charged, which might lead to a weak binding of the inhibitor to the enzyme [24, 25]. One accepted NS2B/NS3pro inhibitor is aprotinin, which envelops the enzyme and prevents the substrate from accessing the protease active site [12]. From various published studies on the activity of NS2B/NS3pro [26-28], the preferential substrate contains basic residues (R or K) at the P^1 and P^2 sites and a short side-chain amino acid (G, S or A) at the $P^{1'}$ site [29, 30], where the protease cleaves at the $P^2P^1 \downarrow P^{1'}$ sites.

To search for potent antiviral drugs against DV, detailed information obtained at the atomic level would most likely provide essential information towards a better understanding and identification of the key residues that are important to formulate a more efficient inhibitor (stronger binding) of NS2B/NS3pro. Here, we examine the influence of the substrate binding specificity and recognition of NS2B/NS3pro on different substrates. Molecular dynamics simulations (MDSs) of the NS2B/NS3pro complex bound to the six different substrates (capsid, intNS3, 2A/2B, 2B/3, 3/4A and 4B/5) were performed. The NS2B/NS3pro-substrate binding, as well as the structural and dynamical properties potentially involved in the cleavage mechanism, were analyzed and extensively discussed. To our knowledge, this is the first study focusing on the interaction between substrates of DV2 NS2B/NS3pro by means of MDS.

2. Materials and Methods

2.1. Initial structure of the NS2B/NS3pro-substrate complexes

The homology model of the DV2 NS2B/NS3pro complexed with the peptidic inhibitor Bz-Nle-KRR-H based on WNV as obtained from our previous work [13] was used as the template in construction of the enzyme-substrate complexes. The amino acid sequence of DV2 was taken from the UniProtKB reference Q91H74, which has an identical amino acid sequence to that in the crystal structure of apo DV2 NS2B/NS3pro (2FOM).

In this study, the six different polypeptide substrates used were the capsid, intNS3, 2A/2B, 2B/3, 3/4A and 4B/5. The polypeptide cleavage region of the capsid ($N^5RRRR^1S^{1'}AGMI^{5'}$) complexed with DV2 NS2B/NS3pro (Fig. 1) was generated as follows. The three dimensional-structure of the capsid's cleavage region was built based on the X-ray structure of WNV residues 11–20 (TGPCLARIIR) as the template (2IJO.pdb [12]), using homology modeling in the module implemented in Discovery Studio 2.5. (DS2.5, Accelrys Software Inc., San Diego, CA, USA). The sequence alignment revealed a 50% similarity relative to its template. The capsid heavy atoms of -$R^3R^2R^1$- were superimposed with the -KRR-coordinate of the Bz-Nle-KRR-H inhibitor bound to DV2 NS2B/NS3pro [13], while the substrate backbone atoms of N^5R^4 and $S^{1'}$–$I^{5'}$ were superimposed with the X-ray structures of the WNV NS2B/NS3pro complexed with the inhibitor aprotinin (2IJO). Note that the aprotinin inhibitor, also known as bovine pancreatic trypsin inhibitor, is a long peptide of 56 amino acids [12]. Thus, the coordinates of the -$R^3R^2R^1$- residues were changed according to the –KRR– coordinates of the peptidic inhibitor, while the N^5R^4_ _ _ $S^{1'}AGMI^{5'}$ residues were changed according to the backbone of the aprotinin inhibitor (-$T^{11}G$_ _ _ $ARIIR^{20}$-). Then, the side chains of the capsid sequence were added using DS2.5's tool with the newly formed conformation of the NS2B/NS3pro-capsid complex. The complex was subsequently minimized by keeping the $P^3P^2P^1$- residues and the NS2B/NS3pro fixed. Then,

the geometry and stereochemistry of the loops were validated using Ramachandran plots in the Discovery Studio 2.5 program. Ten residues of the capsid cleavage region were located in the favored (92%) and the allowed (8%) regions of the plot, indicating the suitability of this model for further simulation. For the remaining systems, the initial structure of their cleavage regions were constructed using the atomic coordinates of the capsid built as above and then the mutated amino acid residue along each substrate was generated using the build and edit protein tool in DS2.5 program. The sequences and superimposition of the six modeled polypeptide cleavage regions in the binding site of DV2 NS2B/NS3pro are presented in Fig. 1.

2.2. Molecular Dynamics Simulations (MDSs)

For each NS2B/NS3pro-substrate complex, the protonation state of all possible charged residues (Arg, Lys, His, Asp and Glu) was assigned at pH 7. Afterwards, the complexes were randomly neutralized by Cl^- ions (3, 3, 2, 1, 1 and 2 ions for the capsid, intNS3, 2A/2B, 4/B5, 3/4A and 2B/3 systems, respectively), and the TIP3P [31] water box was applied with the minimum distance of 10 Å from the protein surface to the edge of the simulation box that the dimensions were set to 86 x 90 x 88 $Å^3$. In total, the NS2B/NS3pro-capsid, intNS3, 2A/2B, 4/B5, 3/4A and 2B/3 complexes contains 27378, 27373, 27396, 27395, 27391 and 27380 atoms, respectively. The periodic boundary condition with the NPT ensemble was applied. Energy minimizations and MDSs were performed using the SANDER module of AMBER 10 [32] with the ff03 force field [33, 34]. A Berendsen coupling time of 0.2 ps was used to maintain the temperature and standard pressure of the system [35]. The SHAKE algorithm [36] was applied to constrain all bonds involving hydrogen. A simulation time step of 2.0 fs was used. All MDSs were run with a 12 Å residue-based cutoff for non-bonded interactions and the particle mesh Ewald method was applied to get an adequate treatment of the long-range electrostatic interactions [37]. Every system was subjected to

four stages of restrained MDSs at 298 K with force constants of 10, 7.5, 5.0 and 2.5 kcal.mol^{-1}.Å^2 for 500 ps each stage to relax the systems. These subsequent steps could allow the substrate to adapt its geometry and orientation (from the initial model) to fit better into the NS2B/NS3pro cavity. Afterwards, the fully unrestrained MDS was performed for 20 ns. The convergences of energy, temperature and global root mean-square displacement (RMSD) were used to verify the stability of the systems. The MDS trajectories were collected every 0.2 ps from the production phase (12–20 ns) for analysis in terms of the structural dynamics, protein substrate interactions and binding free energies, using the Ptraj and MMPB-SA modules of AMBER 10.

3. Results and discussion

3.1. Stability of the NS2B/NS3pro-substrates

The known cleavage sites of the DV polyprotein by NS2B/NS3pro are well established and the substrate models in this study were built based on the sequence alignment reported by Niyomrattanakit et al. [27]. Functional substrate profiling of the P^1–P^5 and $P^{1'}$ – $P^{5'}$ regions of the proteases has been described previously for all four DV serotypes and revealed that the four enzymes appear to share very similar substrate specificities [27]. The sequence of each substrate model was generally composed of 10 amino acid residues containing the proteolytic site as well as important residues that were reported to interact with the protease active site, as shown in Fig. 1. The protease cleavage sites for the substrates are known to occur at $P^2P^1 \downarrow P^{1'}$, which the P^1 and P^2 sites are two basic residues (R or K) and a short side-chain amino acid (G, S or A) is at the $P^{1'}$ site (Fig. 1).

[Figure 1]

To monitor the stability of the six simulated systems, the individual capsid, intNS3, 2A/2B, 4/B5, 3/4A and 2B/3 binding to DV2 NS2B/NS3pro, the RMSD of the structures obtained during the 20-ns MDS relative to the initial structure for all heavy atoms in the complex and the 10 substrate residues (P^5-$P^{5'}$) were evaluated and plotted in Fig. 2. The RMSD *vs*. simulation time plots indicated that all six systems reached equilibrium at 12 ns. Thus, the obtained coordinates from 12–20 ns onwards were taken as the production phase for further analyses.

[Figure 2]

3.2. Hydrogen bond patterns of substrate binding

The intermolecular interaction between the substrate and the surrounding residues at the binding pocket of NS2B/NS3pro is an important determinant in the NS2B/NS3pro-mediated cleavage of the capsid, intNS3, 2A/2B, 4B/5, 3/4A and 2B/3 substrates. To access the likely strength of the substrate-enzyme interaction, the number and percentage of hydrogen bond formation between the ten individual residues of the substrate and those of the NS2B and NS3pro proteins were probed using the two following criteria: *(i)* a proton donor-acceptor distance of ≤ 3.5 Å, and *(ii)* a donor-H-acceptor bond angle of ≥ 120°. The number of strong, medium and weak hydrogen bond interactions was determined using hydrogen bond occupations of ≥ 80%, 50–79% and < 50%, respectively, and the results are summarized in Table 1, whilst more details of the hydrogen bond strength are given in Fig. S1 in the supplementary information (SI). For each studied system, the averaged MD structure derived during the 12–20 ns simulation period as the representative substrate binding to the DV2 NS2B/NS3pro is depicted in Fig. 3.

[Table 1]

From the results in Table 1, not all of the substrate residues formed hydrogen bonds with the NS2B and NS3 residues. No hydrogen bond (grey box) was detected at the four residues of the capsid ($P^4/P^{2'}/P^{3'}/P^{4'}$), intNS3 ($P^5/P^4/P^{4'}/P^{5'}$) as well as 2A/2B and 4B/5 ($P^5/P^4/P^{3'}/P^{4'}$), and the six residues of 3/4A ($P^5/P^4/P^{1'}/P^{3'}/P^{4'}/P^{5'}$) and 2B/3 ($P^5/P^4/P^{2'}/P^{3'}/P^{4'}/P^{5'}$). In almost all complexes, the highest number of hydrogen bond formations was likely to be observed at the P^1, P^2 and P^3 positions, with the exception of the 2A/2B and 2B/3 systems which showed only two weak hydrogen bonds at the P^2 position.

In the case of the reaction at the center of cleavage on the P^1 position, the positively charged substrate residue (R or K), showed the highest number of hydrogen bonds among the ten substrate residues. It was noted that the interactions formed at this site were only contributed from the NS3 residues with different interaction strengths. This center residue was found to interact with the targeted protein through two strong hydrogen bond formations with $_{NS3}S^{135}$ and $_{NS3}G^{133}$ (or $_{NS3}F^{130}$ in the 2B/3 system). An additional strong interaction with $_{NS3}D^{129}$ was observed in the 2A/2B and 4B/5 complexes, while the other remaining systems showed medium and /or weak interactions with this residue. Moreover, the P^1 residue was moderately/weakly stabilized by $_{NS3}F^{130}$ and $_{NS3}T^{134}$ (Fig. 3).

The second highest number of hydrogen bond formations between NS2B/NS3pro and the respective substrate was found at the P^2 position. The P^2(R) residue of the capsid, intNS3, 4B/5 and 3/4A substrates was intensively stabilized by the two negatively charged residues $_{NS2B}D^{81}$ and $_{NS3}D^{75}$. In contrast, the interactions with the $_{NS3}D^{75}$ and $_{NS2B}D^{81}$ residues were dramatically decreased and totally disappeared, respectively, in the 2A/2B system due to the introduction of the positively charged short side chain of lysine (K) at this site. In the 2B/3 system, the change of the residue at this position to the polar glutamine (Q) residue led to the loss of two hydrogen bonds with these two residues, but formed rather weak interactions with $_{NS2B}S^{83}$ and $_{NS3}N^{152}$ instead.

At the P^3 site, two strong ($_{NS2B}M^{84}$ and $_{NS3}Y^{161}$ residues) and one medium ($_{NS3}G^{153}$ residue) hydrogen bonds were found in the capsid and intNS3 complexes, while the interaction with the $_{NS2B}M^{84}$ residue was significantly reduced or completely lost in the other systems. The P^3(R) residue of the capsid substrate was weakly stabilized by the $_{NS2B}S^{83}$ residue. For the other substrate sites, a lower number of medium and/or weak hydrogen interactions was observed, excluding the strong interaction with $_{NS3}I^{30}$ at the $P^{5'}$ residue of the capsid substrate.

By considering the total strong/medium/weak hydrogen bond formations in Table 1, the order of the interaction strength in the substrate binding to the DV2 NS2B/NS3pro was: capsid (8/3/11) > intNS3 (7/8/10) > 4B/5 (7/6/9) > 2A/2B (5/5/9) ~ 3/4A (5/3/12) > 2B/3 (3/4/5). A high number of hydrogen bond interactions were likely formed at the P^1, P^2 and P^3 substrate sites, whereas no interaction was detected at the P^4 and $P^{4'}$ sites in all complexes. While the NS3 residues stabilized the substrate at almost all the sites, the NS2B residues only interacted with the P^2 and P^3 sites. The strong interactions at the P^1 residue of substrates with the $_{NS3}G^{133}$ and $_{NS3}S^{135}$ residues are likely conserved in almost all complexes. Changing the residue at the P^2 site from R to the others caused a significant reduction in the hydrogen bond interactions at this site.

[Figure 3]

3.3. Distances at the catalytic triad and oxyanion hole

The distances at the catalytic regions, namely the catalytic triad of $_{NS3}D^{75}$, $_{NS3}H^{51}$ and $_{NS3}S^{135}$ *(d1–d4)*, nucleophilic attack *(d5)* and oxyanion hole constructed by the $_{NS3}S^{135}$, $_{NS3}G^{133}$ and $_{NS3}T^{134}$ *(d6–d8)*, are schematically represented in Fig. 4A. The distribution of these *d1–d8* distances is plotted in Fig. 4B where the corresponding distances obtained from the two co-crystal structures of WNV NS2B/NS3pro with the inhibitor bound in the

tetrahedral conformation [7, 11] are also shown as dashed lines in the same figure for comparison.

At the catalytic triad, the simulated distances of *d1* and *d2* mostly appeared at ~ 2.9 Å (Fig. 4B), as determined from the carboxyl oxygens of $_{NS3}D^{75}$ to the N1-imidazole nitrogen of $_{NS3}H^{51}$, for all six DV2 NS2B/NS3pro systems suggesting the presence of two hydrogen bonds. In contrast to the experimentally derived results from the WNV NS2B/NS3pro crystal structures, the O1 atom was located closer than the other one (*d1* = 2.6 and 2.8 Å, *d2* = 3.4 Å). The nucleophilic attack distance involved in the first step of acylation is determined as being between the O3-hydroxyl oxygen of $_{NS3}S^{135}$ and the carbonyl carbon of the P^1(R) [38], shown as *d4* in Fig. 4A. The *d4* bond-making distance was found to be ~3.0 Å in all systems, which indicates that the three catalytic residues are in a suitable configuration for initiating the cleavage reaction.

The oxyanion hole of the NS2B/NS3pro active site is generally formed by the three back bone nitrogens of $_{NS3}S^{135}$, $_{NS3}T^{134}$ and $_{NS3}G^{133}$ (*d5*, *d6* and *d7*, respectively), which significantly stabilize the carbonyl oxygen (Oγ) of the P^1 reacting residue in the cleavage reaction. These distances were mostly at ~3 Å in all complexes (Fig. 4B), suggesting the well-formed oxyanion hole plays an important role in supporting the negatively charged Oγ after the nucleophilic attack. Additionally, it is of note that from the six NS2B/NS3pro-substrate simulations, the shortening *d7* distance (by 1.2 Å) from 4.2 Å in the starting structure modeled from Wichapong et al. [13] found here is in good agreement with that obtained from the 3E90 crystal structure of the WNV NS2B/NS3pro (2.8 Å). Besides the $_{NS3}T^{134}$ contribution through its backbone nitrogen, the hydroxyl group also stabilizes the Oγ atom of P^1 via attractive electrostatic interaction, as implied from the *d8* value (range of 3.7–4.1 Å depending on the system). The overall substrate conformation and intermolecular interaction inside the DV2 binding site are likely to be similar to those obtained from

catalytic activity of the WNV NS2B/NS3pro X-ray structures [7,11], especially at the conserved distances within catalytic triad (*d1*), nucleophilic attack (*d5*) and oxyanion hole (*d7*) regions.

[Figure 4]

3.4. Decomposition energy (DC)

By correlating the hydrogen bond interaction at each subsite of the substrates with the corresponding NS2B/NS3pro pocket (Table 1), the decomposition energies (DCs; the interaction energy of individual NS2B/NS3pro residues with the substrate and vice versa) were plotted (Fig. 5). At subsite P^1, residue R was able to bind very well, as was residue K, in the DV2 NS2B/NS3pro, where the interaction between the P^1(R) subsite with the S^1 pocket of all systems was around -18 kcal.mol^{-1} (except for 2B/3) and around -12 kcal.mol^{-1}(3/4A) in the case of P^1(K). This result also implies that the S^1 pocket of NS2B/NS3pro is rather specific for charged residues, such as R or K. Comparison of the pair interaction at the other substrate subsites with the NS2B/NS3pro protein pocket concluded that the preference of the interaction between the P^2 subsite with the S^2 pocket is R > K > Q; between the P^3 subsite with the S^3 pocket is R/Q/T > K/G, and no attractive interaction occurs between the P^4 and P^5 subsites with the S^4 and S^5 pockets. Note that only these P^1–P^3subsites were considered for the relative interaction because the non-prime site are more important than the prime site for interacting with the NS2B/NS3pro, as discussed before. These theoretically derived results broadly agree with the reported experimental data [39] that the preference of each subsite for interacting with NS2B/NS3pro are R > K, R > T > Q/N/K, K > R > N and Nle > L > K for P^1, P^2, P^3 and P^4, respectively.

Interestingly for 2A/2B, the P^2(K) subsite of this substrate can only form rather weak H-bond with $_{NS2B}S^{83}$ and $_{NS3}D^{75}$ (Table 1). However, from the DC analysis, the $P^{2'}$(W) and $P^{4'}$(L) subsites of 2A/2B were found to have van der Waals interactions with $_{NS3}I^{36}$ and

$_{NS3}V^{52}$-R^{54}, respectively, as represented in Fig. 5B (the substrate lies in the pocket as shown in Fig. 3). This effect might lead to the better binding of the 2A/2B substrate than 4B/5 with NS2B/NS3pro. In addition, the interaction of P^1(R) with $_{NS3}D^{129}$, $_{NS3}G^{133}$ and $_{NS3}S^{135}$ located at the S^1 pocket of 2A/2B is stronger than that with the P^1(K) of 3/4A, while the interaction between $_{NS3}D^{75}$, located at the S^2 pocket, with P^2 of the 2A/2B substrate is strongly evident yet it was absent in the 2B/3 system (Fig. 5B). Therefore, the 2A/2B substrate can bind better with NS2B/NS3pro than with 4B/5, 3/4A and 2B/3. These MDS derived results agreed well with the experimental data that the *Km* values (Table 2) of 2A/2B are lower than those for 4B/5, 3/4A and 2B/3 [40], which suggests a better binding between that substrate and the enzyme. Moreover, these results reflect that van der Waals interactions between the $P^{2'}$ and $P^{4'}$ subsites with non-polar residues in the pocket as well as the H-bond between the key subsites (P^1, P^2 and P^3) of the substrate with residues in NS2B/NS3pro play an important role. Altogether, the order of the interaction between these substrates with NS2B/NS3pro is in the order of capsid > intNS3 > 2A/2B > 4B/5 > 3/4A > 2B/3, and these results agree well with the MDS derived binding free energy calculation (section 3.7). The MDS derived results not only substantiate the experimental data, as explained before, but also predict that the capsid or intNS3 likely bind with DV2 NS2B/NS3pro better than 2A/2B does.

From the perspective of the designed inhibitory drugs, the results suggested that the most important parts of the substrate for interacting with the DV2 NS2B/NS3pro are the P^1 and P^2 subsites through strong electrostatic interactions with S^1 ($_{NS3}D^{129}$ and $_{NS3}S^{135}$) and S^2 ($_{NS2B}D^{81}$and $_{NS3}D^{75}$) pockets, respectively (Fig. 5B). Thus H-bonds at these subsites with these residues were quite stable during the simulation period, especially for $_{NS3}D^{75}$ and $_{NS3}S^{135}$ (Table 1). Therefore, competitive inhibitors can be designed based on the non-prime sites (P^1–P^2), and substrate-based inhibitors should contain a positively charged group, such as a guanido group to mimic the side chain of R or K, for optimal interaction with $_{NS3}D^{129}$ and

$_{NS3}S^{135}$ at the S^1 pocket and also with $_{NS2B}D^{81}$ and $_{NS3}D^{75}$ at the S^2 pocket. In addition, the P^3 subsite of all substrates interacted with $_{NS2B}M^{84}$ and $_{NS3}Y^{161}$ (Fig. 5B). The stable H-bonds between the $_{NS2B}M^{84}$, $_{NS3}G^{153}$ and $_{NS3}Y^{161}$ subsites with residues at the S^3 pocket are shown in Table 1. By considering the P^3(G) subsite of 3/4A, it was found that the P^3 subsite can also form a stable H-bond with $_{NS3}G^{153}$ and $_{NS3}Y^{161}$. Hence, these results highlighted that the main H-bond interaction between the P^3 subsite with the S^3 pocket of NS2B/NS3pro is the interaction between the backbone atom, as reported before [13]. In addition, a structure-activity relationship revealed that the modification of the P^3 subsite by a smaller side chain did not significantly reduce the activity of the inhibitors [41]. Thus, the side chain of the P^3(K) subsite can be replaced by another short side chain, such as G or A. In conclusion, the potent inhibitors for NS2B/NS3pro should contain a positive side chain, such as a guanidine group, R or K or another positive group, at the P^1 and P^2 subsites. Also, the backbone atom of the P^3subsite is important for forming H-bonds with NS2B/NS3pro at $_{NS3}G^{153}$ and $_{NS3}Y^{161}$, and so this part of the inhibitor must be conserved but the side chain of P^3 can be replaced by a short side chain.

[Figure 5]

3.6. Solvation effect at the substrate binding site

The occurrence of water in the cavity region of enzyme-substrate complex leads to a hydrophilic effect that is known to play a significant role in the catalytic mechanism of biological systems. To elucidate the water accessibility, the radial distribution function (RDF; the probability of finding the type *y* particle in a spherical radii, *r*, around the particle of type *x* (*gxy(r)*)), together with its integration coordination number (CN) were evaluated. Here, the RDFs from the hydrogen acceptor/donor atoms of residues at the catalytic site ($_{NS3}H^{51}$, $_{NS3}D^{75}$, $_{NS3}T^{134}$ and $_{NS3}S^{135}$) and the substrate reaction center residue at the P^1 subsite (see

label in Fig. 4A), to the oxygen atom of waters were calculated. The results for the six simulated systems are shown in Fig. 6 together with the corresponding running integration number up to the first minimum of the RDF plot (marked by an arrow).

The RDFs for both carboxyl oxygen atoms of $_{NS3}D^{75}$ (O1 and O2) of the 2A/B and 2B/3 show the first sharp and narrow peak at ~2.9 Å with CNs of 0.7/0.8 and 0.8/1.3, respectively, while in the other systems a water molecule was only found surrounding atom O1 (CN range of 0.5–1.0) except for intNS3 that had no water surrounding these two atoms within ~4 Å. For the O3(S^{135}), N1(H^{51}), N2(H^{51}) and N4(T^{134}) of NS3 residues, there were no water molecules within a distance of less than 3.5 Å in all systems.

Focusing on the N3 atom of the essential nucleophillic $_{NS3}S^{135}$ residue, where the catalytic water was observed and is proposed to contribute to the peptide cleavage reaction mechanism by the serine protease [42], these MDS results were in good agreement with a water molecule clearly having access to and coordinate with the N3 atom with the first peak centered at ~3.2 Å and CNs of 1.2, 0.5, 1.0 and 1.2 for the capsid, intNS3, 2A/B and 4B/5, respectively. In contrast, for the 3/4A and 2B/3 systems the first peak maximum was shifted to > 4 Å. Apart from the catalytic water surrounding the N3 atom of the $_{NS3}S^{135}$ nucleophile residue, more water molecules appeared at the NS2B/NS3pro pocket and would interfere with the substrate-protein binding. In other words, the substrate does not fit well with the binding site of NS2B/NS3pro. The solvation data based on the catalytic and other detected waters within the binding pocket suggesting a substrate binding strength is in the order of: 2B/3 < 3/4A < 4B/5 < 2A/2B < intNS3 < capsid.

[Figure 6]

3.7. Binding free energy

The binding free energies between the six different polypeptide substrates and NS2B/NS3pro were estimated using the MM/GBSA approach, implemented in AMBER. The calculated binding free energies (ΔG_{bind}) and the corresponding energy components averaged over 100 snapshots during the production period (12–20 ns) are given in Table 2. In this study, the entropic contribution ($T\Delta S$) was neglected based on the assumption that the difference was assumed to be small amongst the six simulated systems due to their similarities, as previously mentioned [43, 44]. In addition, this study focused on the relative values of the absolute binding free energy.

The major contribution to the protein-substrate binding energy in all complexes comes from the electrostatic term (Table 2). The van der Waal interactions also contributed significantly in the binding of the substrates, although the effect was not as pronounced as the electrostatic interaction. The predicted ΔG_{bind} of NS2B/NS3pro-capsid, -intNS3, -2A/2B, -4B/5, -3/4A and -2B/3 were -83.90, -74.44, -70.30, -64.69, -56.66, -49.83 kcal.mol^{-1}, respectively. These values indicate a strong binding between NS2B/NS3pro and the substrates in all complexes, but especially to the NS2B/NS3pro-capsid complex that had the most favorable binding.

In line with the DC analysis, it was reported that the key residues at the prime and non-prime sites of NS2B/NS3pro's substrate were annotated into the contribution groups of the (i) primary (R > K and R > K > Q for P^1 and P^2, respectively), (ii) secondary (R/Q/T > K/G, A > S/G, W > L > R > A/T/G and M/L > N/G for P^3, $P^{1'}$, $P^{2'}$ and $P^{4'}$, respectively), and the (iii) tertiary (P^4, P^5, $P^{3'}$ and $P^{5'}$) that have no interaction within the binding pocket (Fig. 5A). Consideration of the six substrates at the primary and secondary sites (P^1, P^2, P^3, $P^{1'}$, $P^{2'}$ and $P^{4'}$) revealed the key residues as $\underline{R}^3\underline{R}^2\underline{R}^1S^{1'}A^{2'}\underline{M}^{4'}$, $Q^3\underline{R}^2\underline{R}^1G^{1'}\underline{R}^{2'}G^{4'}$, $\underline{K}^3\underline{K}^2\underline{R}^1S^{1'}\underline{W}^{2'}\underline{L}^{4'}$, $T^3\underline{R}^2\underline{R}^1G^{1'}T^{2'}N^{4'}$, $G^3\underline{R}^2\underline{K}^1S^{1'}\underline{L}^{2'}\underline{L}^{4'}$ and $\underline{K}^3Q^2\underline{R}^1\underline{A}^{1'}G^{2'}L^{4'}$ for the capsid, intNS3, 2A/2B, 4B/5, 3/4A and 2B/3, respectively, where the underlined residues represent the amino acid types at

any P position that have the most effective contribution to the binding energy. The MM/PBSA results clearly show that the capsid has the highest absolute ΔG_{bind} of ~ 83.9 kcal/mol^{-1} due to the existence of R and M at position P^1–P^3 and $P^{4'}$, respectively. In comparison with the capsid, the absolute ΔG_{bind} decreases when the amino acid at position $P^{4'}$ was substituted by glycine in intNS3 (~10 kcal.mol^{-1}) and asparagine in 4B/5 (~20 kcal.mol^{-1}). The reduction in the absolute ΔG_{bind} value of intNS3 is less than that for 4B/5, which may be due to the strong contribution from the interaction of $P^{2'}$(R). For 2A/2B, the changing of P^2 and P^3 to a shorter base ($R^3R^2 \rightarrow K^3K^2$) resulted in a decreased absolute ΔG_{bind} value of 13 kcal.mol^{-1}, but the contribution from the $P^{2'}$(W) and $P^{4'}$(L) helped the binding. In the cases of 3/4A and 2B/3, a high reduction in the absolute ΔG_{bind} values of about 28 and 34 kcal.mol^{-1}, respectively, was observed, which could be due to the change of the amino acids in key positions, such as $R^1 \rightarrow K^1$ in 3/4A and $R^2 \rightarrow Q^2$ in 2B/3. The absolute ΔG_{bind} order between these six substrates with NS2B/NS3pro (capsid > intNS3 > 2A/2B > 4B/5 > 3/4A > 2B/3) was in good agreement with the previously reported experimentally derived *Km* and %binding values [26, 40, 45, 46].

[Table 2]

4. Conclusion

The elucidation of the key residues at the prime and non-prime sites for the polypeptide substrate of DV2 NS2B/NS3pro was studied by the classical MD simulations. The results of NS2B/NS3pro in complex with six peptide substrates (capsid, intNS3, 2A/2B, 4B/5, 3/4A and 2B/3) revealed that the basic residues at P^1 and P^2 subsites of the substrate play a major role in the binding with NS2B/NS3pro, whereas the P^3, $P^{1'}$, $P^{2'}$ and $P^{4'}$subsites have a minor contribution. In addition, the preferred substrate residues in each subsite for

interacting with NS2B/NS3pro were: R > K, R > K > Q, R/Q/T > K/G, A > S/G, W > L > R > A/T/G and M/L > N/G for P^1, P^2, P^3, $P^{1'}$, $P^{2'}$ and $P^{4'}$, respectively. The substrate binding affinity may result to the different cleavage ability by NS2B/NS3pro. The interaction of the substrate with the S^1 pocket involves only those residues from the NS3 domain. On the other hand, both residues from the C-terminal domain of NS2B as well as $_{NS3}D^{75}$ and $_{NS3}N^{152}$ are important for maintaining the interaction with the P^2 residue of the substrate at the S^2 pocket. At the S^3 pocket, the main interactions were observed between the backbone of the P^3 residue and $_{NS3}G^{153}$ as well as $_{NS3}Y^{161}$. The interaction of the substrate at the $S^{1'}$, $S^{2'}$ and $S^{4'}$ pocket of NS3 is maintained only by hydrophobic interaction. In short, the detailed interaction analysis from all six decapeptidic substrates complexed with NS2B/NS3pro provided significant and useful information for the structure-based design and identification of potent inhibitors.

5. Acknowledgements

This is a collaborative project under the South-East Asia's Collaborations on Research and Education in Computer-Aided Drug Discovery. This study was supported by the Research Chair Grant, the National Science and Technology Development Agency (NSTDA), Thailand. P.Y. and K.W. thank the postdoctoral fellowship from the Ratchadaphiseksomphot Endowment Fund from Chulalongkorn University. The authors would like to thank the Computational Chemistry Unit Cell, Faculty of Science, Chulalongkorn University, the School of Pharmaceutical Sciences, Universiti Sains Malaysia, and Research and Development Institute, Ramkhamhaeng University for providing research facilities, software packages and computing times. HAW and CSB gratefully acknowledged the funding from Malaysian Ministry of Science, Technology and Innovation through the dengue R&D initiative grant no: 09-05-IFN-MEB-007. We thank Dr. Panita Decha for helpful suggestions.

Reference

[1] Kautner, I., Robinson, M.J., Kuhnle, U. Dengue virus infection: Epidemiology, pathogenesis, clinical presentation, diagnosis, and prevention. Journal of Pediatrics. 1997, 131, 516-24.

[2] Monath, T.P. Dengue - The risk to developed and developing-countries. Proceedings of the National Academy of Sciences of the United States of America. 1994, 91, 2395-400.

[3] Burke, D.S., Monath, T.P. Flaviviruses. Philadelphia., In Fields Virology (eds. D.M. Knipe and P.M. Howley), Lippincott Williams & Wikins, , 2001.

[4] Kuhn, R.J., Zhang, W., Rossmann, M.G., Pletnev, S.V., Corver, J., Lenches, E., et al. Structure of dengue virus: implications for flavivirus organization, maturation, and fusion. Cell. 2002, 108, 717-25.

[5] Irie, K., Mohan, P.M., Sasaguri, Y., Putnak, R., Padmanabhan, R. Sequence analysis of cloned dengue virus type 2 genome (New Guinea-C strain). Gene. 1989, 75, 197-211.

[6] Falgout, B., Pethel, M., Zhang, Y.M., Lai, C.J. Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of Dengue virus nonstructural proteins. J. Virol. . 1991, 65, 2467–75.

[7] Robin, G., Chappell, K., Stoermer, M.J., Hu, S.H., Young, P.R., Fairlie, D.P., et al. Structure of West Nile Virus NS3 Protease: Ligand Stabilization of the Catalytic Conformation. Journal of Molecular Biology. 2009, 385, 1568-77.

[8] Bianchi, E., Pessi, A. Inhibiting viral proteases: challenges and opportunities. Biopolymers. 2002, 66, 101-14.

[9] Murthy, H.M.K., Clum, S., Padmanabhan, R. Dengue virus NS3 serine protease - Crystal structure and insights into interaction of the active site with substrates by molecular modeling and structural analysis of mutational effects. Journal of Biological Chemistry. 1999, 274, 5573-80.

[10] Murthy, H.M.K., Judge, K., DeLucas, L., Padmanabhan, R. Crystal structure of dengue virus NS3 protease in complex with a Bowman-Birk inhibitor: Implications for flaviviral polyprotein processing and drug design. Journal of Molecular Biology. 2000, 301, 759-67.

[11] Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S.P., et al. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nature Structural & Molecular Biology. 2006, 13, 372-3.

[12] Aleshin, A.E., Shiryaev, S.A., Strongin, A.Y., Liddington, R.C. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. Protein Science. 2007, 16, 795-806.

[13] Wichapong, K., Pianwanit, S., Sippl, W., Kokpol, S. Homology modeling and molecular dynamics simulations of Dengue virus NS2B/NS3 protease: insight into molecular interaction. Journal of Molecular Recognition. 2009, 23, 283-300.

[14] Chang, J.H., Schul, W., Butters, T.D., Yip, A., Liu, B.P., Goh, A., et al. Combination of alpha-glucosidase inhibitor and ribavirin for the treatment of dengue virus infection in vitro and in vivo. Antiviral Research. 2011, 89, 26-34.

[15] Frecher, V., Miertus, S. Design, structure-based focusing and in silico screening of combinatorial library of peptidomimetic inhibitors of Dengue virus NS2B-NS3 protease. Journal of Computer-Aided Molecular Design. 2010, 24, 195-212.

[16] Knehans, T., Schuller, A., Doan, D.N., Nacro, K., Hill, J., Guntert, P., et al. Structure-guided fragment-based in silico drug design of dengue protease inhibitors. Journal of Computer-Aided Molecular Design. 2011, 25, 263-74.

[17] Lim, P.Y., Keating, J.A., Hoover, S., Striker, R., Bernard, K.A. A Thiopurine Drug Inhibits West Nile Virus Production in Cell Culture, but Not in Mice. *Plos One*. 2011, 6.
[18] Nitsche, C., Steuer, C., Klein, C.D. Arylcyanoacrylamides as inhibitors of the Dengue and West Nile virus proteases. *Bioorganic & Medicinal Chemistry*. 2011, 19, 7318-37.
[19] Pierra, C., Amador, A., Benzaria, S., Cretton-Scott, E., D'Amours, M., Mao, J., et al. Synthesis and pharmacokinetics of valopicitabine (NM283), an efficient prodrug of the potent anti-HCV agent 2 '-C-methylcytidine. *Journal of Medicinal Chemistry*. 2006, 49, 6614-20.
[20] Tomlinson, S.M., Malmstrom, R.D., Russo, A., Mueller, N., Pang, Y.P., Watowich, S.J. Structure-based discovery of dengue virus protease inhibitors. Antiviral Research. 2009, 82, 110-4.
[21] Tomlinson, S.M., Watowich, S.J. Anthracene-based inhibitors of dengue virus NS2B-NS3 protease. *Antiviral Research*. 2011, 89, 127-35.
[22] Wahab, H.A., Yusof, R., Rahman, N.A. A search for vaccines and therapeutic for dengue: A review. Current Computer-Aided Drug Design. 2007, 3, 101-12.
[23] Yang, C.C., Hsieh, Y.C., Lee, S.J., H., W.S., Liao, C.L., Tsao, C.H., et al. Novel Dengue Virus-Specific NS2B/NS3 Protease Inhibitor, BP2109, Discovered by a High-Throughput Screening Assay. Antimicrobial Agents and Chemotherapy. 2011, 55, 229-38.
[24] Ekonomiuk, D., Su, X.C., Ozawa, K., Bodenreider, C., Lim, S.P., Yin, Z., et al. Discovery of a Non-Peptidic Inhibitor of West Nile Virus NS3 Protease by High-Throughput Docking. Plos Neglected Tropical Diseases. 2009, 3.
[25] Mueller, N.H., Yon, C., Ganesh, V.K., Padmanabhan, R. Characterization of the West Nile virus protease substrate specificity and inhibitors. International Journal of Biochemistry & Cell Biology. 2007, 39, 606-14.
[26] Gouvea, I.E., Izidoro, M.A., Judice, W.A.S., Cezari, M.H.S., Caliendo, G., Santagada, V., et al. Substrate specificity of recombinant dengue 2 virus NS2B-NS3 protease: Influence of natural and unnatural basic amino acids on hydrolysis of synthetic fluorescent substrates. Archives of Biochemistry and Biophysics. 2007, 457, 187-96.
[27] Niyomrattanakit, P., Yahorava, S., Mutule, I., Mutulis, F., Petrovska, R., Prusis, P., et al. Probing the substrate specificity of the dengue virus type 2 NS3 serine protease by using internally quenched fluorescent peptides. Biochemical Journal. 2006, 397, 203-11.
[28] Yusof, R., Clum, S., Wetzel, M., Murthy, H.M., Padmanabhan, R. Purified NS2B/NS3 serine protease of Dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. J. Biol. Chem. . 2000, 275, 9963–9.
[29] Chambers, T.J., Weir, R.C., Grakoui, A., McCourt, D.W., Bazan, J.F., Fletterick, R.J., et al. Evidence that the N-terminal domain of nonstructural protein NS3 from yellow fever virus is a serine protease responsible for site-specific cleavages in the viral polyprotein. . Proc. Nat. Acad. Sci. U.S.A. . 1990 87, 8898–902.
[30] Preugschat, F., Yao, C.W., Strauss, J.H. In vitro processing of Dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. J. Virol. . 1990, 64, 4364–74.
[31] Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. . 1983, 79, 926-35.
[32] Case, D.A., Darden, T.A., Cheathamlll, T.E., Simmerling, C.L., Wang, J., Duke, R.E., et al. AMBER 10. University of California, San Francisco., 2008.
[33] Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G.M., Zhang, W., et al. A point-charge force field for molecular mechanics simulations of proteins based on condensed-

phase quantum mechanical calculations. Journal of Computational Chemistry. 2003, 24, 1999-2012.

[34] Lee, M.C., Duan, Y. Distinguish protein decoys by using a scoring function based on a new AMBER force field, short molecular dynamics simulations, and the generalized born solvent model. Proteins-Structure Function and Bioinformatics. 2004, 55, 620-34.

[35] Berendsen, H.J.C., Postma, J.P.M., Gunsteren, W.F.V., DiNola, A. Molecular dynamics with coupling to an external bath. J. Chem. Phys. . 1984, 81 3684-90.

[36] Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. . J. Comput. Phys. . 1997, 23 327-41.

[37] York, D.M., Darden, T.A., Pedersen, L.G. The effect of long-range electrostatic interactions in simulations of macromolecular crystals: a comparison of the Ewald and truncated list methods J. Chem. Phys. . 1993, 99 8345-8.

[38] Steuer, C., Gege, C., Fischl, W., Heinonen, K.H., Bartenschlager, R., Klein, C.D. Synthesis and biological evaluation of alpha-ketoamides as inhibitors of the Dengue virus protease with antiviral activity in cell-culture. *Bioorg Med Chem*. 2011, 19, 4067-74.

[39] Li, J., Lim, S.P., Beer, D., Patel, V., Wen, D.Y., Tumanut, C., et al. Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. Journal of Biological Chemistry. 2005, 280, 28766-74.

[40] Khumthong, R., Niyomrattanakit, P., Chanprapaph, S., Angsuthanasombat, C., Panyim, S., Katzenmeier, G. Steady-state cleavage kinetics for dengue virus type 2 NS2B-NS3(pro) serine protease with synthetic peptides. Protein and Peptide Letters. 2003, 10, 19-26.

[41] Yin, Z., Patel, S.J., Wang, W.L., Chan, W.L., Rao, K.R.R., Wang, G., et al. Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. Bioorganic & Medicinal Chemistry Letters. 2006, 16, 40-3.

[42] Molloy, S.S., Bresnahan, P.A., Leppla, S.H., Klimpel, K.R., Thomas, G. Human furin is a calcium-dependent serine endopro-tease that recognizes the sequence Arg-X-X-Arg and efficiently cleaves anthrax toxin protective antigen. J. Biol. Chem. . 1992, 267, 16396-402.

[43] Kuhn, B., Kollman, P.A. Binding of a diverse set of ligands to avidin and streptavidin: An accurate quantitative prediction of their relative affinities by a combination of molecular mechanics and continuum solvent models. Journal of Medicinal Chemistry. 2000, 43, 3786-91.

[44] Villa, J., Strajbl, M., Glennon, T.M., Sham, Y.Y., Chu, Z.T., Warshel, A. How important are entropic contributions to enzyme catalysis? Proceedings of the National Academy of Sciences of the United States of America. 2000, 97, 11899-904.

[45] Leung, D., Schroder, K., White, H., Fang, N.X., Stoermer, M.J., Abbenante, G., et al. Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. Journal of Biological Chemistry. 2001, 276, 45762-71.

[46] Shiryaev, S.A., Aleshin, A.E., Ratnikov, B.I., Smith, J.W., Liddington, R.C., Strongin, A.Y. Expression and purification of a two-component flaviviral proteinase resistant to autocleavage at the NS2B-NS3 junction region. Protein Expression and Purification. 2007, 52, 334-9.

Figure 1. (A) Superimposition of the six modeled polypeptide substrates (stick model) in the electrostatic surface of DV2 NS2B/NS3pro, where the blue and red colors represent the positively and negatively charged amino acid residues, respectively. (B) Comparison of the six substrate sequences.

Figure 2. RMSDs relative to the initial structure for all heavy atoms of the substrate-NS2B/NS3pro complexes (black) and the individual substrate (orange) for the six studied systems: capsid, intNS3, 2A/2B, 3/4A, 4/B5 and 2B/3.

Figure 3. Substrate conformation in the binding site of DV2 NS2B/NS3pro, averaged from the 12–20 ns MDS for the six systems (capsid-, intNS3-, 2A/2B-, 4B/5-, 3/4A- and 2B/3-NS2B/NS3pro). The line model represents the binding residues of NS2B and NS3pro proteins. The key substrate residues contributing in the P1, P2 and P3 domains are shown as ball and stick models in orange, green and brown, respectively.

Figure 4. (A) Definition of the intermolecular distances involved in the cleavage reaction at the catalytic triad (d1–d4), nucleophilic attack (d5) and oxyanion hole (d6–d8). (B) Distribution plot of the d1–d8 distances for all six simulated systems, sampling from 12 to 20 ns in MDSs, and the corresponding distance from the two X-ray structures of WNV NS2B/NS3pro inhibitor complexes.

Figure 5. Decomposition of total binding free energy of (A) each residue of the NS2B/NS3pro and (B) each residue of the respective substrate.

Figure 6. Radial distribution functions, *g(r)*, and their integrated coordination numbers, *n(r)*, from hydrogen donor/acceptor atoms of NS2B/NS3pro catalytic residues and its substrate (see Fig. 4A for definition) to water oxygen atoms for the six studied systems.

Figure S1. Percentage hydrogen bond occupation between the 10 residues, P5-P5', of the substrates and the binding residues of NS2B/NS3pro

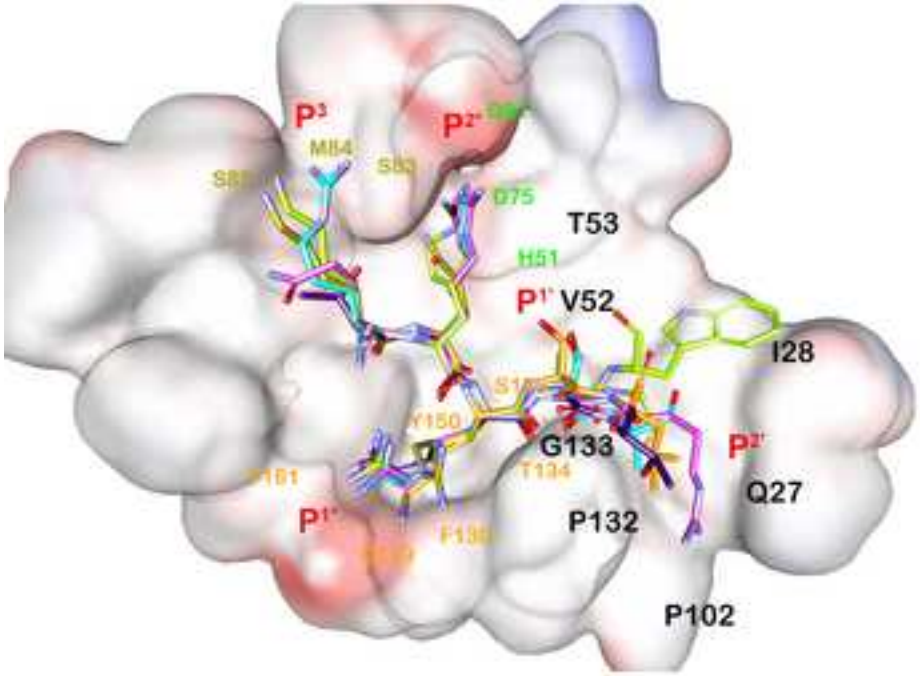
Table 1. Number of strong/medium/weak hydrogen bonds, defined as the percentage of occupancy ≥ 80/50–79/< 50, respectively, formed between the 10 residues, P^5–$P^{5'}$, of the substrates and the binding residues of NS2B/NS3pro.

NS2B/NS3pro-substrate	number of strong/medium/weak hydrogen bonds					
	capsid	**intNS3**	**2A/2B**	**4B/5**	**3/4A**	**2B/3**
***P*^{5}**	***N***	***A***	***T***	***T***	***A***	***V***
$_{NS3}D^{120}$	- / - / 1	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
$_{NS3}G^{159}$	- / - / 1	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
***P*^{4}**	***R***	***A***	***S***	***N***	***A***	***K***
	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
***P*^{3}**	***R***	***Q***	***K***	***T***	***G***	***K***
$_{NS2B}S^{83}$	- / - / 1	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
$_{NS2B}M^{84}$	1 / - / -	1 / - / -	- / - / 1	- / - / -	- / - / -	- / - / 1
$_{NS3}G^{153}$	- / 1 / -	- / 1 / 1	- / 1 / -	- / 1 / -	- / 1 / -	- / 1 / -
$_{NS3}Y^{161}$	1 / - / -	1 / - / -	1 / - / -	1 / - / -	1 / - / -	1 / - / -
***P*^{2}**	***R***	***R***	***K***	***R***	***R***	***Q***
$_{NS2B}D^{81}$	1 / - / 3	1 / 1 / 2	- / - / -	1 / - / 2	1 / - / 3	- / - / -
$_{NS2B}S^{83}$	- / - / -	- / - / -	- / - / 1	- / - / 1	- / - / 1	- / - / 1
$_{NS3}D^{75}$	2 / - / -	2 / - / -	1 / - / -	2 / - / -	1 / 1 / -	- / - /-
$_{NS3}N^{152}$	- / - / 1	- / - / 1	- / - / -	- / - / 1	- / - / -	- / - / 1
***P*^{1}**	***R***	***R***	***R***	***R***	***K***	***R***
$_{NS3}D^{129}$	- / 1 / 2	- / 2 / 2	1 / 1 / 2	1 / 1 / -	- / - / 3	- / - / 1
$_{NS3}F^{130}$	- / 1 / -	- / - / 1	- / 1 / -	- / - / 1	- / - / 3	1 / 1 / -
$_{NS3}G^{133}$	1 / - / -	1 / - / -	1 / - / -	1 / - / -	1 / - / -	- / 1 / -
$_{NS3}T^{134}$	- / - / 1	- / 1 / -	- / 1 / -	- / 1 / -	- / 1 / -	- / 1 / -
$_{NS3}S^{135}$	1 / - / -	1 / - / -	1 / - / -	1 / - / -	1 / - / -	1 / - / -
***P*$^{1'}$**	***S***	***G***	***S***	***G***	***S***	***A***
$_{NS3}G^{133}$	- / - / 1	- / - / 1	- / - / 1	- / - / 1	- / - / -	- / - / 1
***P*$^{2'}$**	***A***	***R***	***W***	***T***	***L***	***G***
$_{NS3}Q^{35}$	- / - / -	- / 1 / -	- / - / -	- / 1 / -	- / - / -	- / - /-
$_{NS3}I^{36}$	- / - / -	- / 1 / 1	- / - / -	- / - / 1	- / - / -	- / - / -
$_{NS3}V^{52}$	- / - / -	- / - / -	1 / - / -	- / - / -	- / - / 2	- / - / -
***P*$^{3'}$**	***G***	***I***	***P***	***G***	***T***	***V***
$_{NS3}H^{51}$	- / - / -	- / 1 / 1	- / - / -	- / - / -	- / - / -	- / - / -
***P*$^{4'}$**	***M***	***G***	***L***	***N***	***L***	***L***
	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
***P*$^{5'}$**	***I***	***R***	***Q***	***I***	***N***	***W***
$_{NS3}I^{30}$	1 / - / -	- / - / -	- / - / -	- / - / -	- / - / -	- / - / -
$_{NS3}R^{54}$	- / - / -	- / - / -	- / 1 / 3	- / 2 / 2	- / - / -	- / - / -
Total	**8 / 3 / 11**	**7 / 8 / 10**	**5 / 5 / 9**	**7 / 6 / 9**	**5 / 3 / 12**	**3 / 4 / 5**

Table 2. Calculated binding free energy and its components (kcal.mol^{-1}) as well as the experimental *Km* (in nM) of the six substrate-NS2B/NS3pro complexes (capsid-, intNS3-, 2A/2B-, NS4B/5-, NS3/4A- and NS2B/3-NS2B/NS3pro).

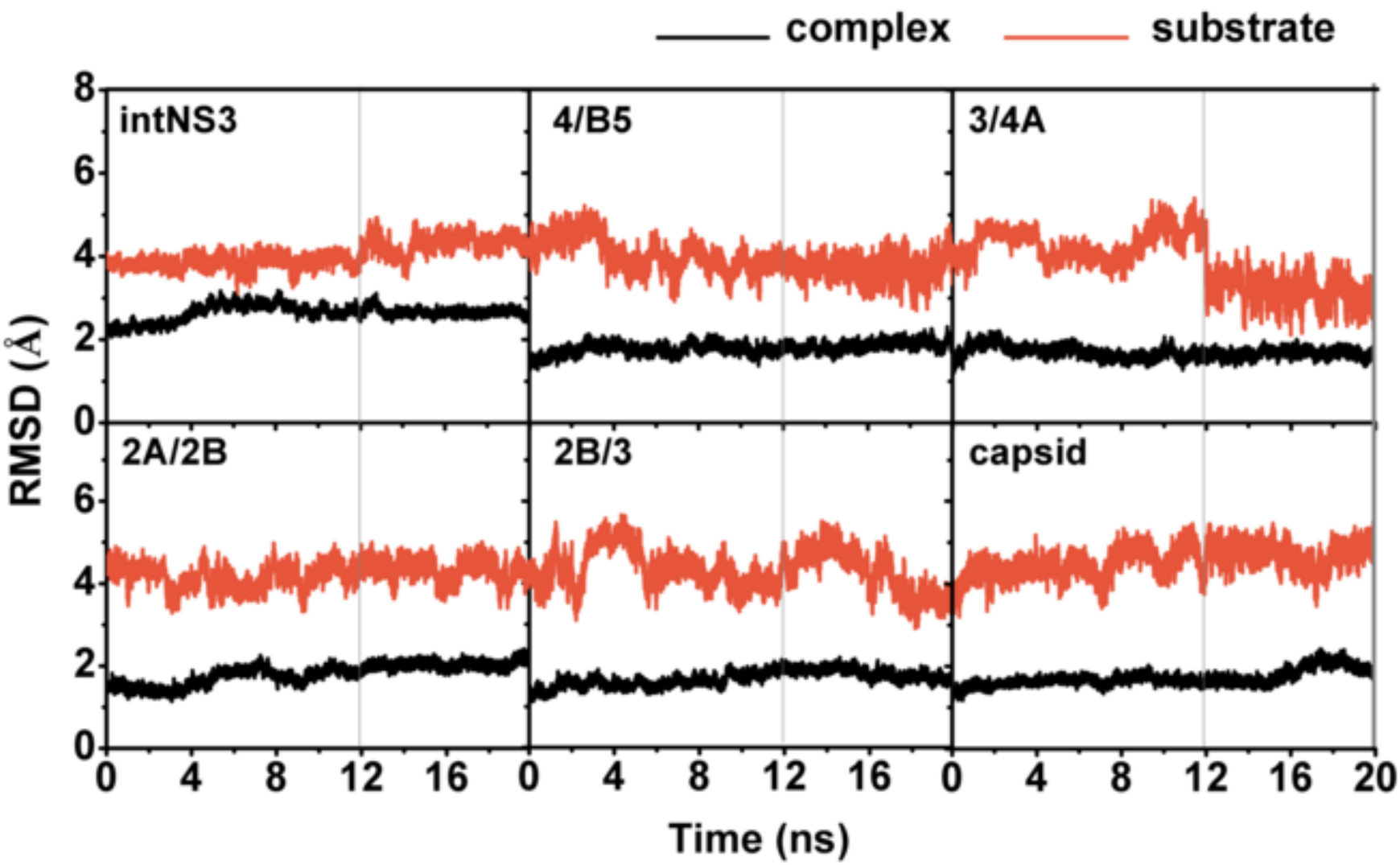
	capsid	**intNS3**	**2A/2B**	**NS4B/5**	**NS3/4A**	**NS2B/3**
ΔE_{ele}	-472.9 ± 33.2	-327.4 ± 28.1	-408.9 ± 37.1	-264.7 ± 21.7	-301.5 ± 20.9	-265.1 ± 22.0
ΔE_{vdw}	-86.6± 5.2	-72.2 ± 5.4	-73.0 ± 5.0	-60.9 ±4.7	-58.6 ± 6.2	-72.1 ± 4.1
ΔE_{MM}	-559.4 ± 32.7	-399.5 ± 30.9	-481.9 ± 37.6	-325.6 ± 23.0	-360.2 ± 21.4	-337.1 ± 22.5
ΔG_{sur}	-9.6 ± 0.4	-9.7 ± 0.4	-10.3 ± 0.6	-8.3 ± 0.5	-8.6 ± 0.6	-9.5 ± 0.5
ΔG_{cal}	479.0 ± 31.1	334.3 ± 28.8	423.4 ± 31.5	270.9 ± 20.1	312.2 ± 18.3	296.4 ± 22.6
ΔG_{sol}	469.4 ± 31.1	324.6 ± 28.5	413.1 ± 31.3	262.5 ± 19.9	303.7 ± 18.0	286.9 ± 22.5
ΔG_{elec}	6.2 ± 7.7	6.9 ± 6.4	14.4 ± 8.2	6.1 ± 7.5	10.7 ± 7.3	31.3 ± 7.4
$\Delta G_{bind(GB)}$	-83.9 ± 4.8	-74.4 ± 5.7	-70.3 ± 6.2	-64.7 ± 5.9	-56.7 ± 6.4	-49.8 ± 4.5
k_m[40]	-	-	184	416	432	3890
k_m[45]	-	-	96	326	100	984
%binding[46]	64	45	5	1	1	11

Figure1



	P5	P4	P3	P2*	P1*	P1'	P2'	P3'	P4'	P5'
capsid	N	R	R	R	R	S	A	G	M	I
intNS3	A	A	Q	R	R	G	R	I	G	R
3/4A	A	A	G	R	K	S	L	T	L	N
4B/5	T	N	T	R	R	G	T	G	N	I
2A/2B	T	S	K	K	R	S	W	P	L	Q
2B/3	V	K	K	Q	R	A	G	V	L	W

Figure2


complex
substrate
intNS3
4/B5
3/4A
2A/2B
2B/3
capsid
RMSD (Å)
Time (ns)
0
2
4
6
8
0
4
8
12
16
20

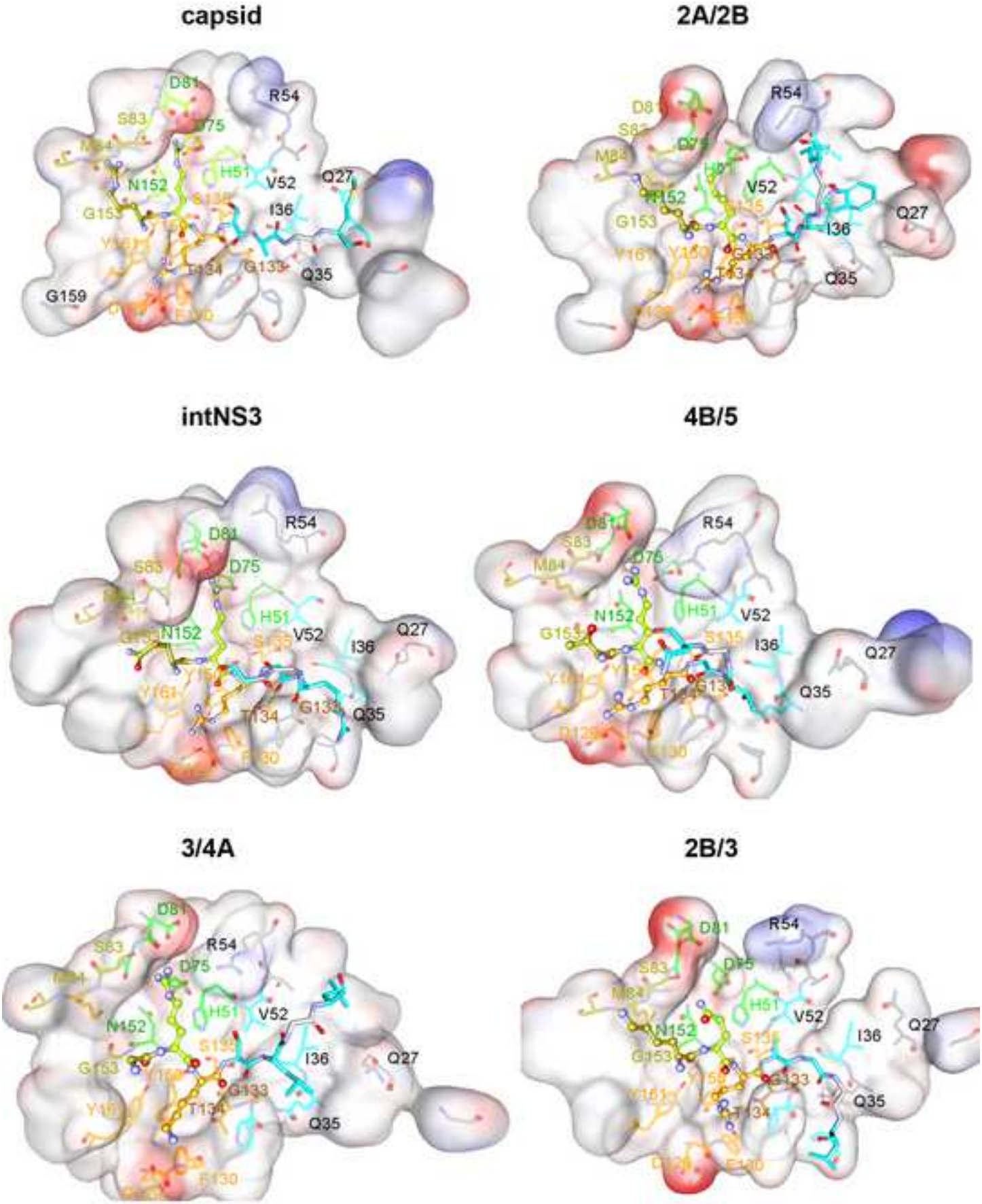
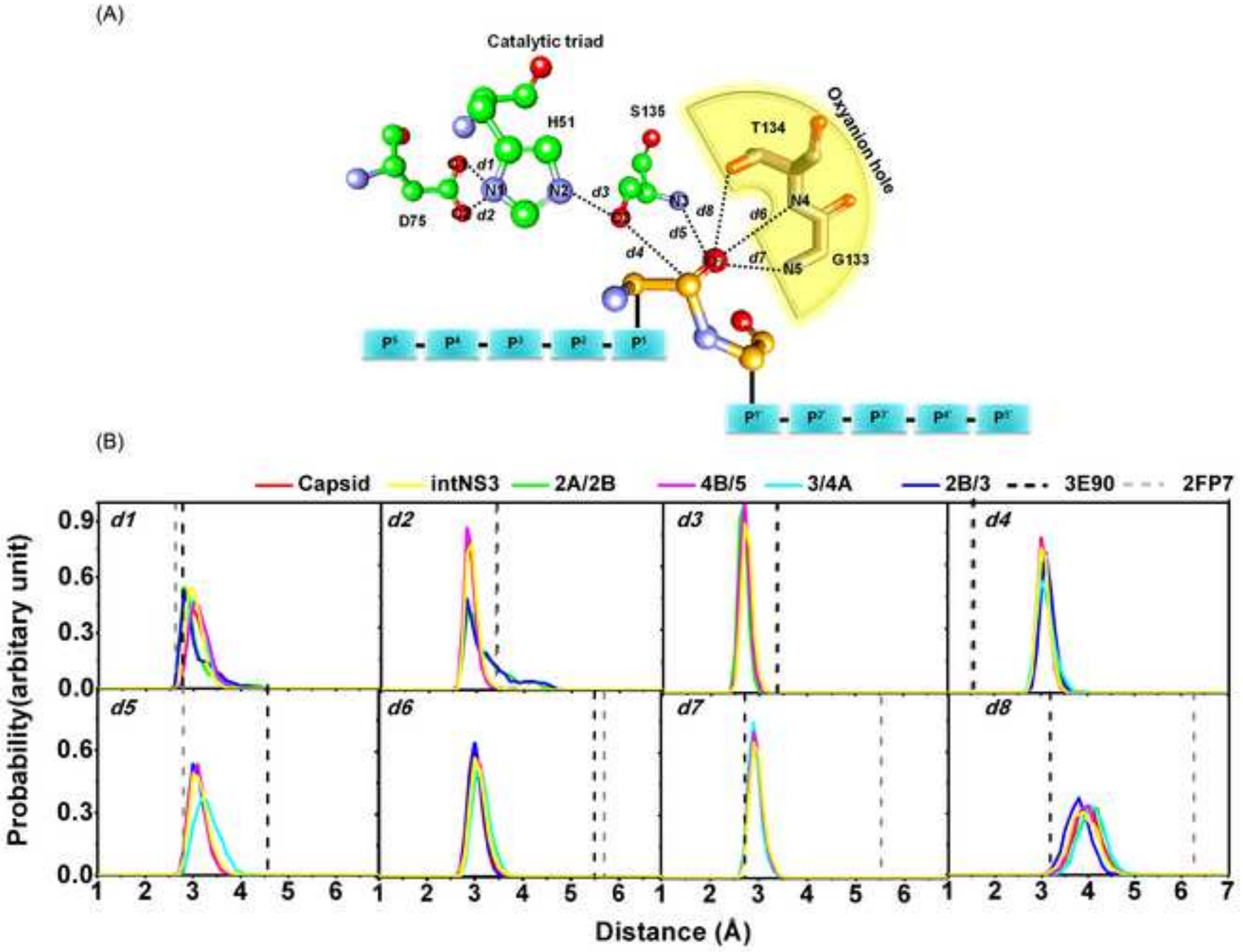


Figure4


(A)
Catalytic triad
H51
S135
T134
Oxyanion hole
D75
N1
N2
N3
N4
N5
G133
d1
d2
d3
d4
d5
d6
d7
d8
(B)
Capsid
intNS3
2A/2B
4B/5
3/4A
2B/3
3E90
2FP7
d1
d2
d3
d4
d5
d6
d7
d8
Probability(arbitary unit)
Distance (Å)

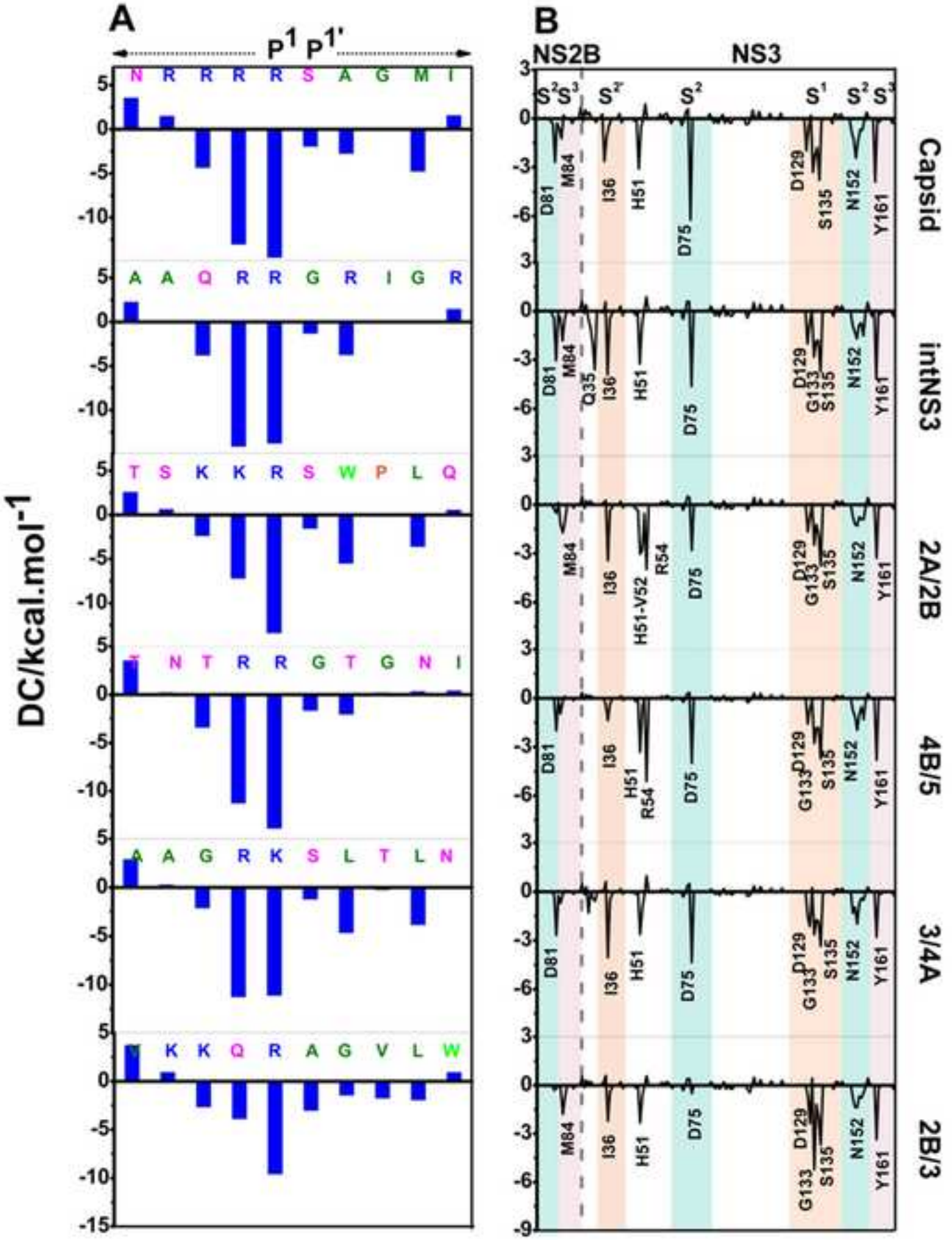
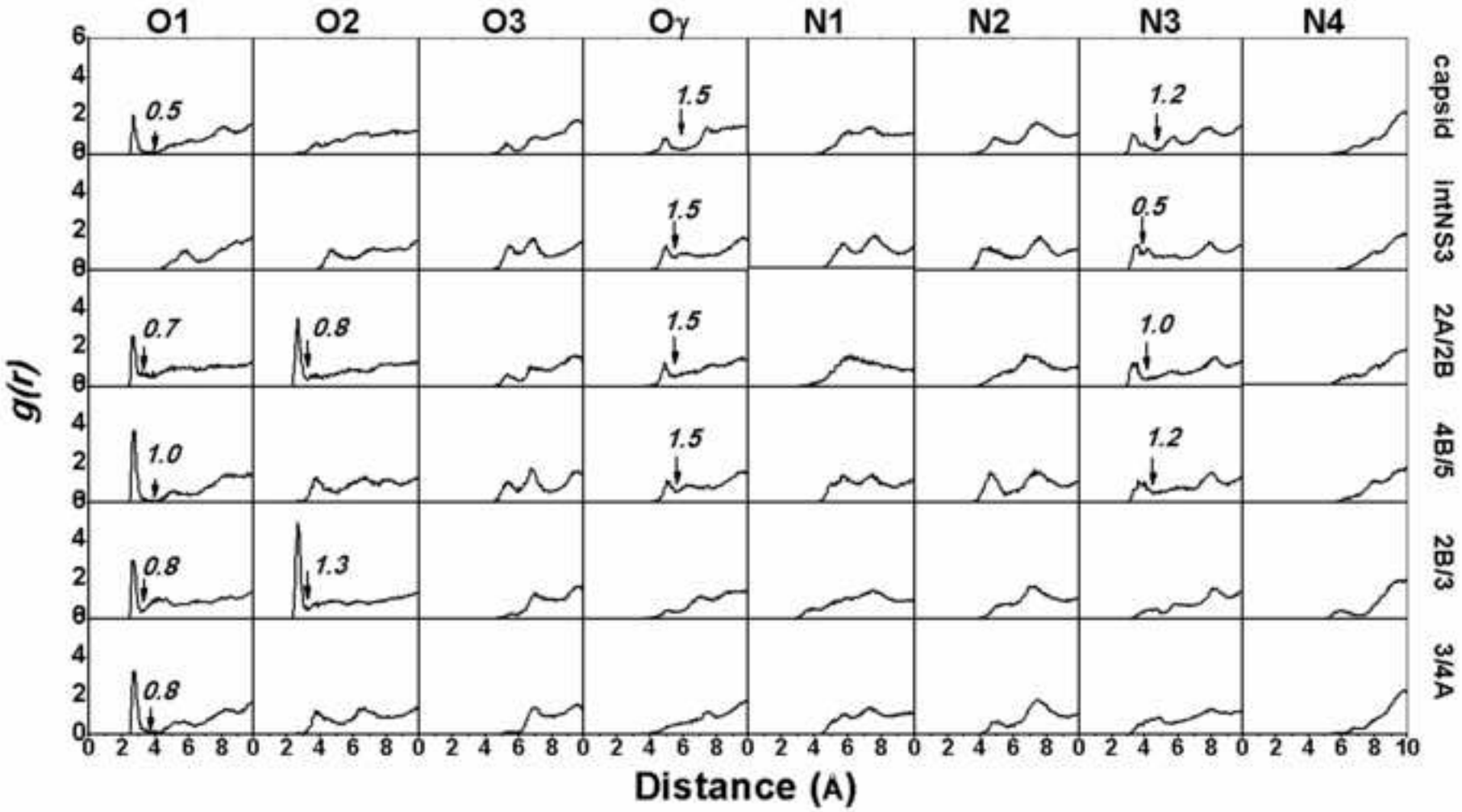


Figure6


O1
O2
O3
Oγ
N1
N2
N3
N4
capsid
intNS3
2A/2B
4B/5
2B/3
3/4A
g(r)
0.5
1.5
1.2
1.5
0.5
0.7
0.8
1.5
1.0
1.0
1.5
1.2
0.8
1.3
0.8
Distance (Å)