Antiviral Research 93 (2012) 126–137

Contents lists available at SciVerse ScienceDirect

Antiviral Research

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

Molecular modeling study on the resistance mechanism of HCV NS3/4A serine protease mutants R155K, A156V and D168A to TMC435

Weiwei Xue [a], Dabo Pan [b], Ying Yang [a], Huanxiang Liu [b,*], Xiaojun Yao [a,c,*]

[a] State Key Laboratory of Applied Organic Chemistry, Department of Chemistry, Lanzhou University, Lanzhou 730000, China
[b] School of Pharmacy, Lanzhou University, Lanzhou 730000, China
[c] Key Lab of Preclinical Study for New Drugs of Gansu Province, Lanzhou University, Lanzhou 730000, China

ARTICLE INFO

Article history:
Received 13 October 2011
Revised 5 November 2011
Accepted 14 November 2011
Available online 22 November 2011

Keywords:
HCV NS3/4A protease
TMC435
Drug resistance
Substrate envelope
Molecular dynamics simulations

ABSTRACT

Hepatitis C virus (HCV) NS3/4A protease represents an attractive drug target for antiviral therapy. However, drug resistance often occurs, making many protease inhibitors ineffective and allowing viral replication to occur. Herein, based on the recently determined structure of NS3/4A–TMC435 complex, atomic-level models of the key residue mutated (R155K, A156V and D168A) NS3/4A–TMC435 complexes were constructed. Subsequently, by using molecular dynamics simulations, binding free energy calculation and substrate envelope analysis, the structural and energetic changes responsible for drug resistance were investigated. The values of the calculated binding free energy follow consistently the order of the experimental activities. More importantly, the computational results demonstrate that R155K and D168A mutations break the intermolecular salt bridges network at the extended S2 subsite and affect the TMC435 binding, while A156V mutation leads to a significant steric clash with TMC435 and further disrupts the two canonical substrate-like intermolecular hydrogen bond interactions (TMC435(N1–H46)···Arg155(O) and Ala157(N–H)···TMC435(O2)). In addition, by structural analysis, all the three key residue mutations occur outside the substrate envelope and selectively weaken TMC435's binding affinity without effect on its natural substrate peptide (4B5A). These findings could provide some insights into the resistance mechanism of NS3/4A protease mutants to TMC435 and would be critical for the development of novel inhibitors that are less susceptible to drug resistance.

1. Introduction

Hepatitis C virus (HCV) infection is now recognized as a serious and growing threat to human health. It is estimated that approximately 180 million people (a minimum of 3% of the world's population) are chronically infected. There are about 3–4 million new cases of HCV infection each year. HCV-infected patients are at a risk of developing cirrhosis and (or) hepatocellular carcinoma (World Health Organization, 2010; Manns et al., 2007). Until now, six major HCV genotypes (genotype 1–6) with numerous subtypes (such as genotype 1a, 1b, etc.) have been identified based on the degree of genetic similarity (World Health Organization, 2010; Simmonds et al., 1993, 2005), varying in their worldwide prevalence and response to treatments (Smith, 2006).

HCV NS3/4A protease is a trypsin-like serine protease which is essential for the generation of components of the viral RNA replication complex. It is a prime and most extensively studied anti-HCV therapeutic target (Bartenschlager, 1999; De Francesco and Migliaccio, 2005; Johnson and Gale, 2006; Kwong et al., 2008). *In vitro* and *in vivo* studies revealed that the 181-residue N-terminal protease domain of the NS3 protein formed a heterodimer with the small 54-residue NS4A peptide cofactor, leading to the subsequent downstream cleavage of the HCV polyprotein at the junctions between NS3/4A, NS4A/4B, NS4B/5A and NS5A/B (Bartenschlager, 1999; Halfon and Locarnini, 2011; Kolykhalov et al., 2000; Kwong et al., 2008). Therefore, drug design targeting the NS3/4A to block viral replication and restore hepatocyte innate immune control of HCV replication has been significantly investigated. For example, TMC435 (Fig.1), as a competitive macrocycle inhibitor of NS3/4A protease with subnanomolar K_i values for genotype 1a and 1b NS3/4A proteases (Lenz et al., 2010; Raboisson et al., 2008), is in phase 3 of clinical trials development by Tibotec (2011). Although NS3/4A is a promising drug target for special antiviral treatment, point mutations in the NS3/4A active site pose a major challenge for development of antivirals as several mutations are known to cause a serious loss of sensitivity to reported chemotherapeutics (Halfon and Locarnini, 2011; Lenz et al., 2010).

Recently, the crystal structure of TMC435 bound to its genotype 1b HCV NS3/4A protease (Cummings et al., 2010) was reported. This is the first reported non-covalent NS3/4A protease–inhibitor complex (Fig. 2a), which can be a starting point to understand the binding and resistance mechanism due to point mutations. In this structure,

* Corresponding authors at: State Key Laboratory of Applied Organic Chemistry, Department of Chemistry, Lanzhou University, Lanzhou 730000, China (X. Yao). Tel.: +86 931 891 2578; fax: +86 931 891 2582.
E-mail addresses: hxliu@lzu.edu.cn (H. Liu), xjyao@lzu.edu.cn (X. Yao).

0166-3542/$ - see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.antiviral.2011.11.007

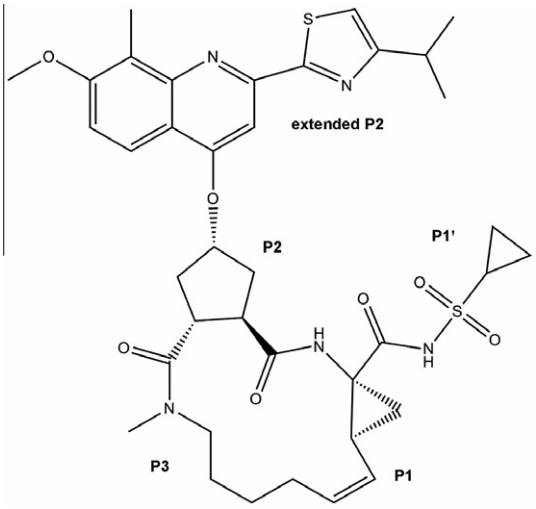

Fig. 1. Chemical structure of TMC435. The indicated P1, P1′, P2 and P3 are substrate positions (Schechter and Berger, 1967) from NS3/4A protease complex structures.

TMC435 bound at the active site of NS3/4A, spanning the S3–S1′ subsites (Figs. 2b and 3). It is noteworthy that the large quinoline–thiazole P2 substituent of TMC435 involves an induced-fit mechanism with Arg155 in the extended S2 subsite (Fig. 3), and the induced conformation of Arg155 is stabilized by interacting with Asp168 and Gln80 (Fig. 2b) (Cummings et al., 2010). Additionally, the P3 substituent of TMC435 (conserved in most inhibitors) makes several close contacts with Ala156 positioned at the edge of the proximal region of S2 (Fig. 3) (Cummings et al., 2010). Importantly, it is demonstrated that these critical interaction residues outside the defined substrate envelope often can lead to drug resistance according to the previous studies (Romano et al., 2010). Currently, the detailed molecular mechanism of the dynamic conformational changes resulting from the key resistant mutations (R155K, A156V and D168A) of NS3/4A protease during the process of recognition with TMC435 is still unknown. Therefore, finding out why several substitutions, including those at position 155, 156 and 168 of the protease, have such a large effect on the susceptibility of NS3/4A to TMC435, is of high relevance and will provide a valuable guidance for the design of next generation inhibitors less susceptible to drug resistance.

Several computational studies have been performed to investigate the molecular mechanism leading to drug resistance of NS3/4A mutants that are in direct or indirect contact with protease inhibitors (PIs). Molecular docking as well as modeling analysis of VX-950 and BILN 2061 at the active site of NS3/4A suggests that there are different mechanisms of resistance for mutations (R155K, A156V/T and D168A/V) induced by VX-950 and BILN 2061 (Lin et al., 2004, 2005). In addition, based on the binding mode of VX-950 by protein–ligand docking, a network of non-covalent interactions between amino acid and the ligand is analyzed to explain the development of viral breakthrough variants (V36 and T54) (Welsch et al., 2008). Moreover, a free energy perturbation (FEP) approach is used to evaluate the effects of the resistance mutations A156T and D168V/Q in HCV protease on the binding of PIs (SCH, SCH 503034, VX-950 and BILN 2061) (Guo et al., 2006). These studies have shown the successful use of molecular modeling in the drug resistance mechanism study of NS3/4A mutants. However, none of the theoretical studies to data have been carried out to investigate the effects of the three dominant mutations R155K, A156V and D168A on the binding of TMC435.

In the present work, we performed all-atom molecular dynamics (MD) simulations followed by Molecular Mechanics Generalized Born Surface Area (MM-GBSA) (Kollman et al., 2000; Massova and Kollman, 2000; Onufriev et al., 2000; Tsui and Case, 2000) calculations and substrate envelope analyzes on the complexes of TMC435 that bind to the wild-type NS3/4A protease and three variants, R155K, A156V and D168A. According to the analyzes of the structural conformational change caused by mutations occurring outside the substrate envelop, as well as the calculations of the relative binding energy and the inhibitor residue interaction spectrums, a detailed picture of the mechanism of drug resistance was depicted.

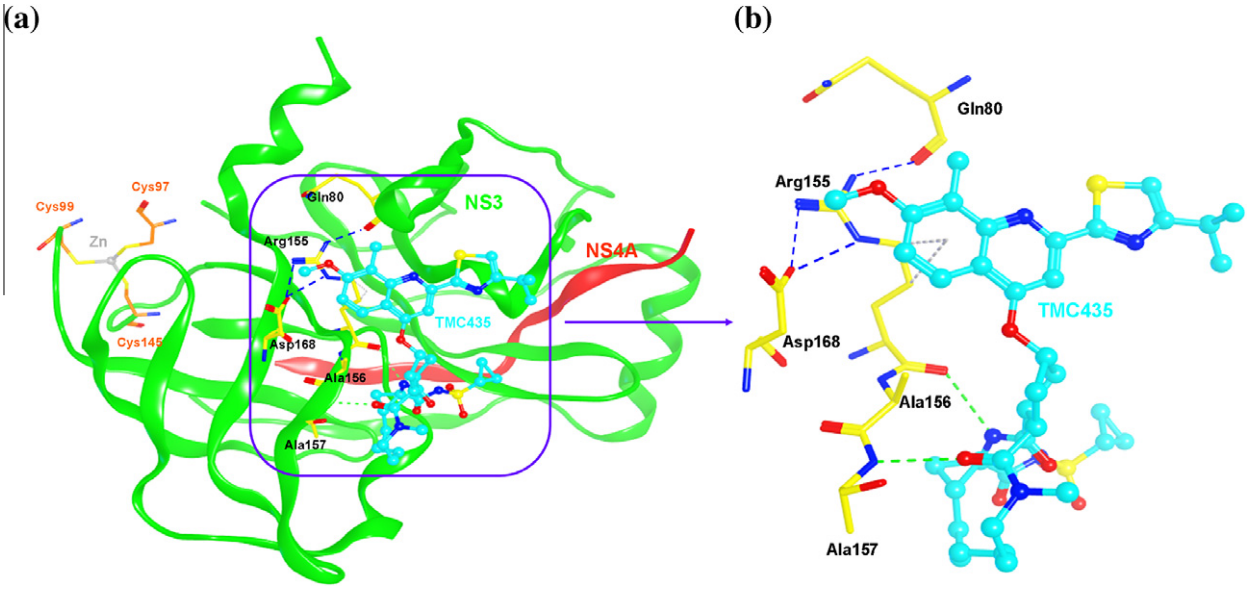

Fig. 2. (a) Crystal structure of TMC435 bound to wild-type NS3/4A protease (PDB ID: 3KEE). The NS3 protease (in green) and NS4A cofactor peptide (in red) are shown as ribbon. The zinc atom (in gray) is shown in ball-stick presentation, which is coordinated by Cys97, Cys99 and Cys145 (in orange sticks). (b) Detailed view of the interactions between Gln80, Arg155, Ala157, Asp168 (in yellow sticks) and TMC435 (in cyan ball-stick). The salt bridge and hydrogen bond interactions are labeled in blue and green dotted lines, respectively. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

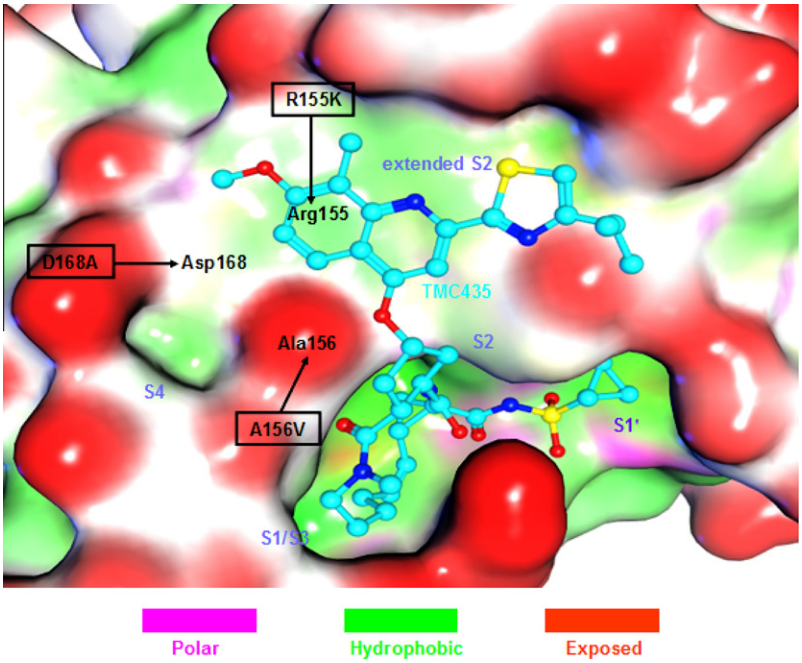

Fig. 3. Molecular surface representation of the active site of the NS3/4A protease with TMC435 and the selected subsites are indicated. The mutated NS3/4A protease residues which confer drug resistance are particularly labeled.

2. Materials and methods

2.1. Preparation of the initial structures

The atom coordinates of the X-ray crystal structure of wild-type HCV NS3A/NS4A protease-TMC435 ternary complex was obtained from the Protein Data Bank (PDB ID: 3KEE (Cummings et al., 2010), 2.4 Å resolution). As the X-ray structures of the mutants are unknown, Discovery Studio 2.5 software (Accelrys, 2009) was applied to generate the 3D structures of the studied mutants (R155K, A156V and D168A) in complex with TMC435 by substituting specific residues using the wild-type model as the template.

Before the MD simulations were started, geometric optimization was performed on TMC435 and the electrostatic potential was calculated at the HF/6-31G* level using the Gaussian 09 program (Frisch et al., 2009). Then, atomic charges of the TMC435 were generated by the restrained electrostatic potential (RESP) fitting method (Bayly et al., 1993; Cieplak et al., 1995; Fox and Kollman, 1998) with the Antechamber module of AMBER. The general AMBER force field (GAFF) (Wang et al., 2004) was used as the parameters for TMC435, and the standard AMBER force field (ff03) (Duan et al., 2003) was used to describe protein parameters. A nonbonded approach was used for the catalytic zinc atoms. Zinc was assigned a formal charge of +2.0 and a van der Waals radius 1.10 Å according to previously derived parameters (Stote and Karplus, 1995).

All missing hydrogen atoms of the proteins were added using the LEaP module in the AMBER 10 software package. Then, the appropriate number of chloride counterions were added to maintain the electro-neutrality of all the systems and then each system was immersed in a rectangular periodic box of pre-equilibrated TIP3P (Jorgensen et al., 1983) water molecules with at least 10 Å distance around the complex, with all of the crystal structural water molecules remaining.

2.2. Molecular dynamics simulations

The energy minimizations and equilibration protocols of the four complexes were carried out using the Sander module of the AMBER 10 software package (Case et al., 2005). For each system, a nine step equilibration protocol was adopted before production phase, including several minimization steps of 1000 cycles followed by additional short MD simulations to relax the structures. First, energy minimization for 1000 cycles followed by 50 ps of MD simulation at 298.15 K was performed on each complex by restraining all heavy atoms of the NS3/4A–TMC435 complex with a harmonic restraint weight of 5.0 kcal/mol Å^2. This was followed by three additional steps of energy minimization for 1000 cycles each with a decreased restraint weight on the heavy atoms from 2.0, to 0.1, to 0.05 kcal/mol Å^2. Additional three MD equilibrations of 50 ps each were performed using decreasing restraints weights from 1.0, to 0.5, to 0.1 kcal/mol Å^2, respectively. These were followed by the last MD equilibration step of 50 ps by releasing all the restraints. Subsequently, 10 ns production MD simulations were performed without any restraint on these four systems in the NPT ensemble at a temperature of 298.15 K and a pressure of 1 atm. A time step of 1 fs was used for the equilibration stages and a time step of 2 fs was used for the production simulations. During the simulations, periodic boundary conditions were employed and all electrostatic interactions were calculated using the particle-mesh Ewald (PME) method with a dielectric constant of unity (Darden et al., 1993). A 10.0 Å cutoff was used to calculate the direct space sum of PME. The SHAKE algorithm was used to restrain bond lengths involving hydrogen atoms (Ryckaert et al., 1977).

2.3. Binding free energy calculations (MM-GBSA)

The binding free energy of TMC435 to the NS3/4A protease was analyzed by the MM-GBSA methods, integrated in the AMBER 10 software package. Despite the approximation made in MM-GBSA, the method has been used successfully to study a wide variety of problems (Chachra and Rizzo, 2008; Gohlke et al., 2003; Hou et al., 2008; Kormos et al., 2007; Liu et al., 2010; Rizzo et al., 2004). The first step of MM-GBSA is to generate multiple snapshots from the stable MD production trajectory of the complex. Here, 500 snapshots were collected, equally spaced at 10 ps intervals. For each snapshot, a free energy is

calculated for each molecular species (complex, receptor and ligand), and the ligand binding free energy is estimated as follows:

$$\Delta G_{binding} = G_{complex} - G_{receptor} - G_{ligand} \quad (1)$$

where $G_{complex}$, $G_{receptor}$ and G_{ligand} are the free energy of complex, receptor and ligand molecules, respectively. The free energy (G) was calculated based on an average over the extracted snapshots from the MD trajectories. Each state is estimated from the molecular mechanics energy E_{gas}, the solvation free energy G_{sol} and the solute entropy S, respectively (Gouda et al., 2003).

$$G = E_{gas} + G_{sol} - TS \quad (2)$$

$$E_{gas} = E_{int} + E_{vdw} + E_{else} \quad (3)$$

$$G_{sol} = G_{GB} + G_{nonpolar} \quad (4)$$

where E_{gas} is the gas-phase energy; E_{int} is the internal energy; E_{ele} and E_{vdw} are the Coulomb and van der Waals energies, respectively. E_{gas} was calculated using the AMBER03 molecular mechanics force field. G_{sol} is the solvation free energy and can be decomposed into polar and nonpolar contributions. G_{GB} is the polar solvation contribution calculated by solving the GB equation. Dielectric constants for solute and solvent were set to 1 and 80, respectively. $G_{nonpolar}$ is the nonpolar solvation contribution and was estimated by the solvent accessible surface area (SAS) determined using a water probe radius of 1.4 Å. The surface tension constant γ was set to 0.0072 kcal/mol/Å^2 (Sitkoff et al., 1994).

2.4. Energy decomposition

To obtain a detailed view of the NS3/4A–TMC435 interactions, MM-GBSA calculations were used to decompose the interaction energies to each residue by considering molecular mechanics energies and solvation energies without considering the contribution of entropies.

2.5. Normal mode calculations

Normal mode calculations were carried out in the AMBER 10 NMODE module to find the entropic contributions. Because of the computational expense of the NMODE calculations, 50 snapshots were used from the last equilibrated 5 ns of the molecular dynamics simulations with 100 ps time intervals.

3. Results and discussion

3.1. Stability of the system simulations

Ten nanoseconds molecular dynamics simulations of four systems were performed to obtain the equilibrated and stable systems. Here, to evaluate the dynamic stability of each simulation, root-mean-square-deviation (RMSD) of all protein backbone atoms (Cα) and the ligand heavy atoms relative to the initial structure coordinates was monitored, as shown in Fig. 4a and b, respectively. In addition, we also monitored the time evolution of the RMSD of the Cα for the residues within 5 Å around the ligand for each system (Fig. 4c). From Fig. 4a and c, the RMSD values of the protein backbone atoms and binding pocket atoms were always stable

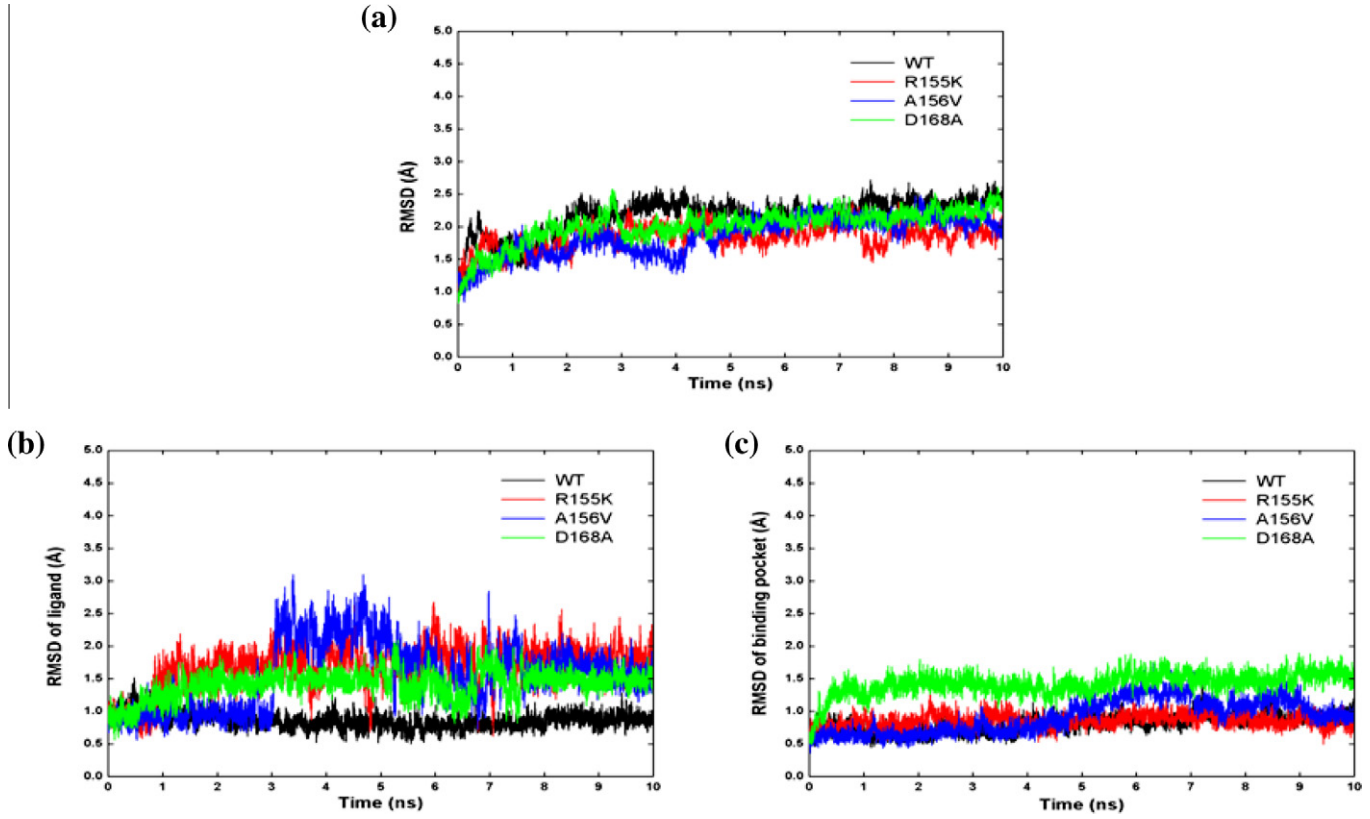

Fig. 4. Plots of the RMSD values relative to the initial structure of wild-type (in black), R155K mutant (in red), A156V mutant (in blue) and D168A mutant (in green) NS3/4A–TMC435 complex during the MD simulations. (a) Time evolution of the RMSD of all protein backbone atoms. (b) Time evolution of the RMSD of heavy atoms for the ligand. (c) Time evolution of the RMSD of Cα atoms for the residues around 5 Å of ligand. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

along the simulation. But from Fig. 4b, it can be seen that the ligand heavy atoms of the three mutant systems have large fluctuations during the first 5 ns, especially for the A156V mutant complex. Therefore, the last 5 ns of the four NS3/4A–TMC435 complex simulations were taken for the following structural and energetic discussion.

3.2. Structural and energetic basis of NS3/4A–TMC435 interaction

To obtain the effect of specific mutations on the mode of binding, we performed classical MD simulations on the wild-type and three mutated complexes. By analyzing the binding site of the protease, four structures with TMC435 and the key residues of the

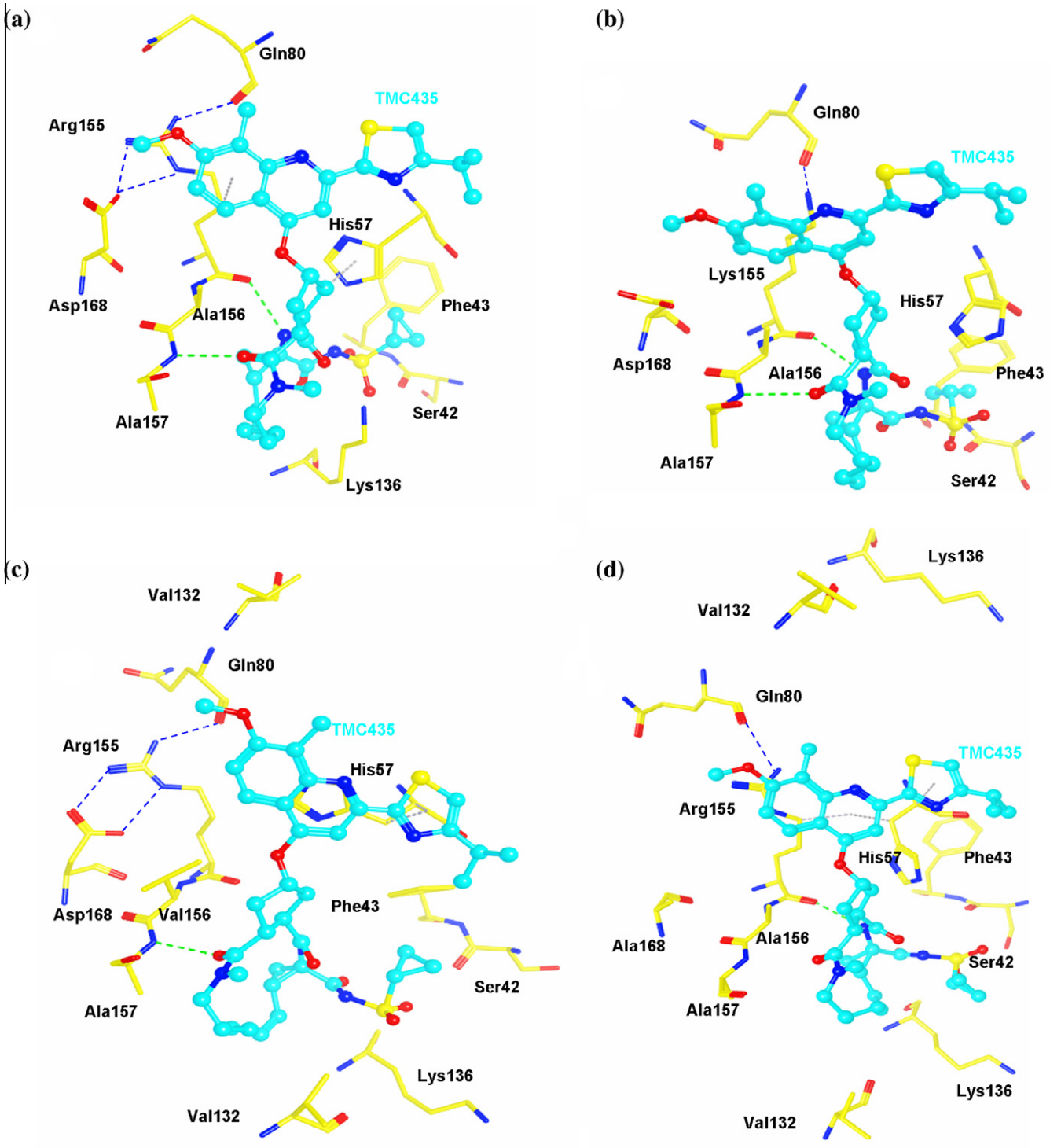

Fig. 5. Average structures taken from the last 5 ns of the MD simulations with the key residues of the binding pocket of the complexes. The side chains of the active site residues are shown as a stick drawing (in yellow) and TMC435 is represented as a ball-stick model (in cyan) for (a) wild-type NS3/4A–TMC435 complex, (b) R155K mutant NS3/4A–TMC435 complex, (c) A156V mutant NS3/4A–TMC435 complex, (d) D168A mutant NS3/4A–TMC435 complex. The salt bridge and hydrogen bond interactions are labeled in blue and green dotted lines, respectively. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

Table 1
The calculated binding free energies based on MM-GBSA method ($\Delta G_{MM\text{-}GBSA}$).

Complex	Contribution (kcal/mol)							Resistance level [g] (Lenz et al., 2010)
	ΔE_{ele} [a]	ΔE_{vdw} [b]	ΔE_{MM}	ΔG_{solv} [c]	ΔE_{bind} [d]	$-T\Delta S$ [e]	ΔG_{bind} [f]	
WT	−32.86	−57.46	−90.32	40.74	−49.58	29.04	−20.54	1
R155K	−23.48	−48.31	−71.79	30.14	−41.65	28.06	−13.59	24
A156V	−22.39	−52.86	−75.25	34.10	−41.15	28.89	−12.26	195
D168A	−28.31	−54.63	−82.94	45.15	−47.78	28.77	−19.01	578

[a] Electrostatic energy.
[b] van der Waals energy.
[c] Solvation free energy.
[d] Total binding free energy.
[e] Entropy value.
[f] Relative binding free energy.
[g] Resistance level is defined as EC_{50} (mutant)/EC_{50} (wild-type) in the references.

binding pocket taken from the average MD-simulated structure are shown in Fig. 5. Moreover, the energetic components and the total binding free energies for TMC435 in the wild-type and mutant NS3/4A binding pocket were calculated during the last 5 ns of simulation by applying the MM-GBSA approach. Table 1 lists the calculated binding free energies ($\Delta G_{MM\text{-}GBSA}$), for all four simulations together with the experimentally derived resistance fold level. As shown in Table 1, the calculated binding free energies (ΔG_{bind}) for TMC435 with WT NS3/4A and R155K, A156V and D168A mutants were −20.54 kcal/mol, −13.59 kcal/mol, −12.26 kcal/mol and −19.01 kcal/mol, respectively. Here, the relative values of the calculated binding free energy exhibited a consistent order with the experimental results except for D168A (Lenz et al., 2010). According to the energy components of the binding free energies (Table 1), both van der Waals and electrostatic terms play important roles for TMC435 binding. In addition, in order to gain a detailed picture of the binding free energy, the total binding free energy was decomposed to each residue, and the corresponding

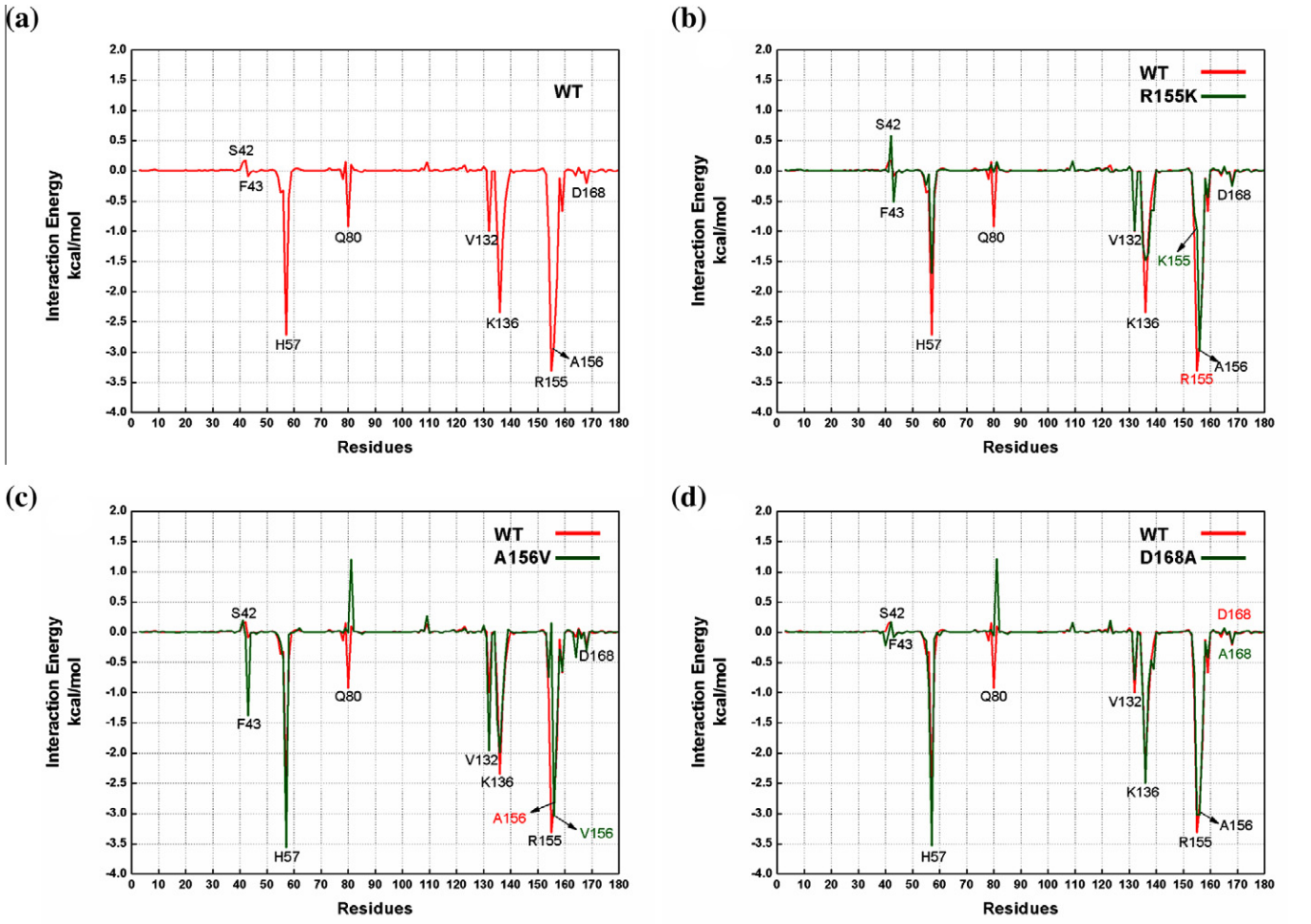

Fig. 6. Intermolecular ligand–protein (per-residue) interaction spectrum of the wild-type and mutant NS3/4A–TMC435 complex according to the MM-GBSA method. Red lines represent the wild-type and green lines represent the mutant. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

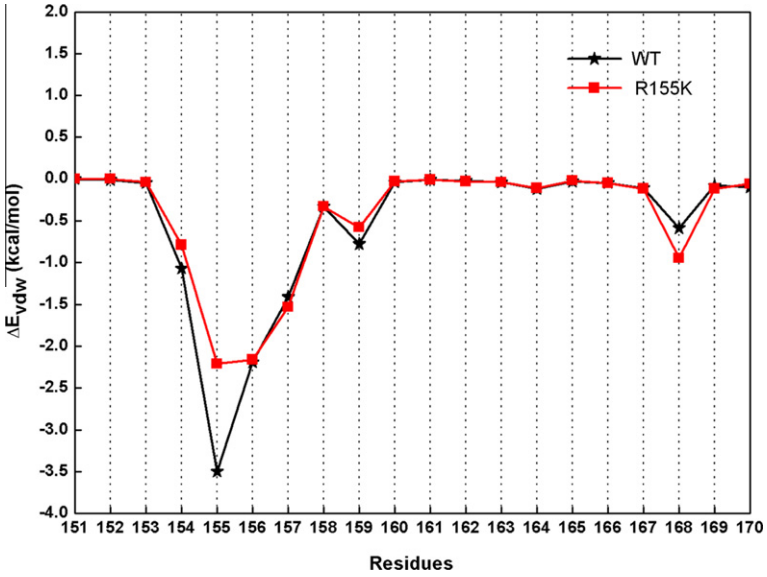

Fig. 7. The nonbonded intermolecular van der Waals (ΔE_{vdw}) interactions between NS3/4A protease residues (number from 151 to 170) with TMC435.

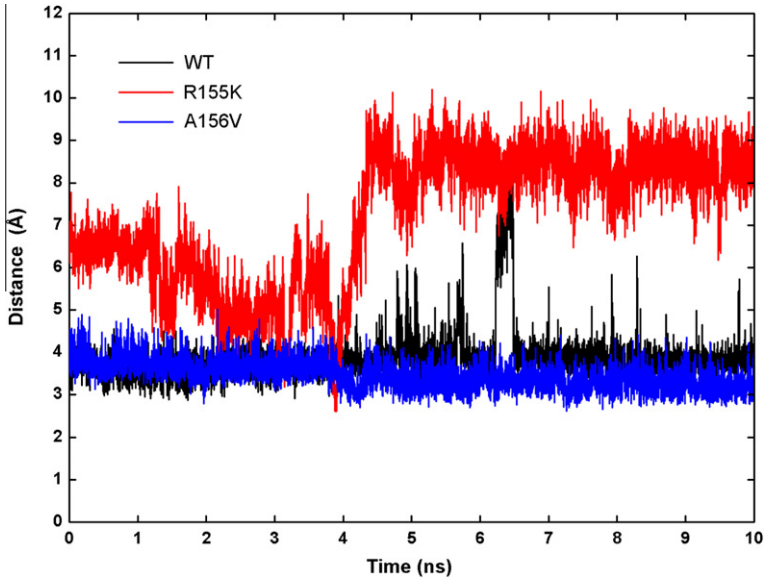

Fig. 8. Plot of the salt bridge distance formed by Arg155–Asp168 of wild-type NS3/4A–TMC435 complex (in black), Lys155–Asp168 of R155K mutant NS3/4A–TMC435 complex (in red) and Arg155–Asp168 of A156V mutant NS3/4A–TMC435 complex (in blue) during the MD simulations. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

results are depicted in Fig. 6. It can be seen that we successfully predicted the key interaction residues (Phe43, His57, Gln80, Val132, Lys136, Arg155, Ala156 and Asp168) verified by experiment (Cummings et al., 2010). Interestingly, TMC435 has strong interactions with Arg155 and Ala156 (Fig. 6a), whose mutations will significantly lower the binding affinity of TMC435 (Lenz et al., 2010). However, to give a more deep understanding of TMC435 binding mechanism, we will analyze the structural features in details in the following part.

3.2.1. Salt bridges network in the extended S2 subsite

From the MD-averaged structure of the NS3/4A–TMC435 complex (Fig.5), we see that the side chain conformation of Arg155 is stabilized by salt bridge interactions with Gln80 and Asp168, resulting in a perfect contact of the extended P2 quinoline moiety of TMC435 with Arg155. Actually, Figs. 5a and 7 (in black) indicated that the alkyl side chain conformation of the arginine allows the maximal number of van der Waals interactions with the quin-

Table 2
The occupancy (%) of two canonical substrate-like intermolecular hydrogen bonds.

Complex	Donor	Acceptor	Occupancy (%)
WT	TMC435(N1–H46)	Arg155(O)	72.77
	Ala157(N–H)	TMC435(O2)	46.22
R155K	TMC435(N1–H46)	Lys155(O)	61.76
	Ala157(N–H)	TMC435(O2)	52.78
A156V	TMC435(N1–H46)	Arg155(O)	27.38
	Ala157(N–H)	TMC435(O2)	29.68
D168A	TMC435(N1–H46)	Arg155(O)	69.42
	Ala157(N–H)	TMC435(O2)	57.93

The hydrogen bonds interactions were defined by acceptor···donor atom distances of <3.5 Å and acceptor···H-donor angles of <120°.

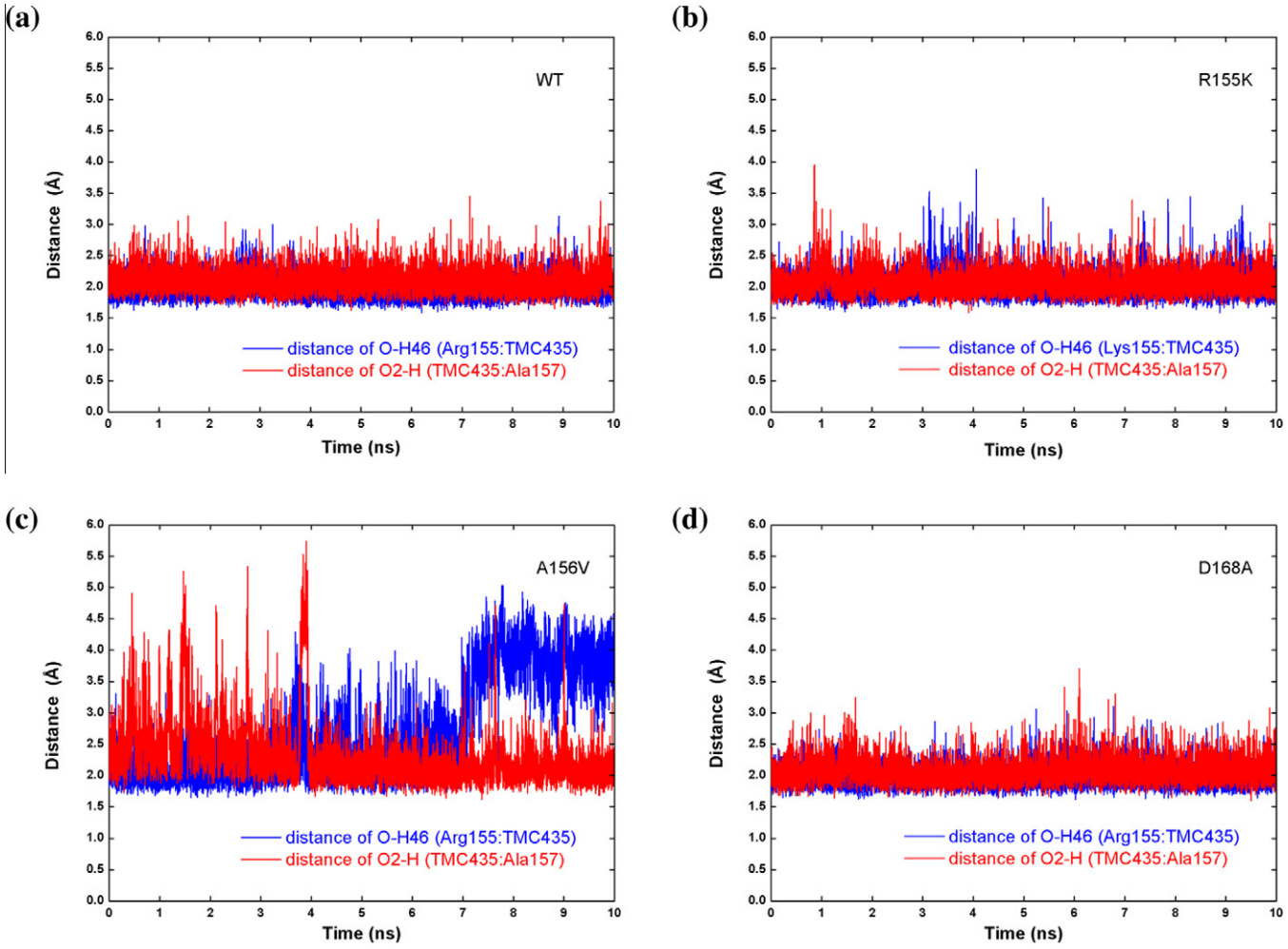

Fig. 9. The monitoring for the two canonical substrate-like intermolecular hydrogen bonds during the all time simulation. Plots show the time evolution of hydrogen bond distance of O–H46 (Arg155:TMC435) and O2–H (TMC435:Ala157) for (a) wild-type NS3/4A–TMC435 complex, (b) R155K mutant NS3/4A–TMC435 complex, (c) A156V mutant NS3/4A–TMC435 complex, (d) D168A mutant NS3/4A–TMC435 complex.

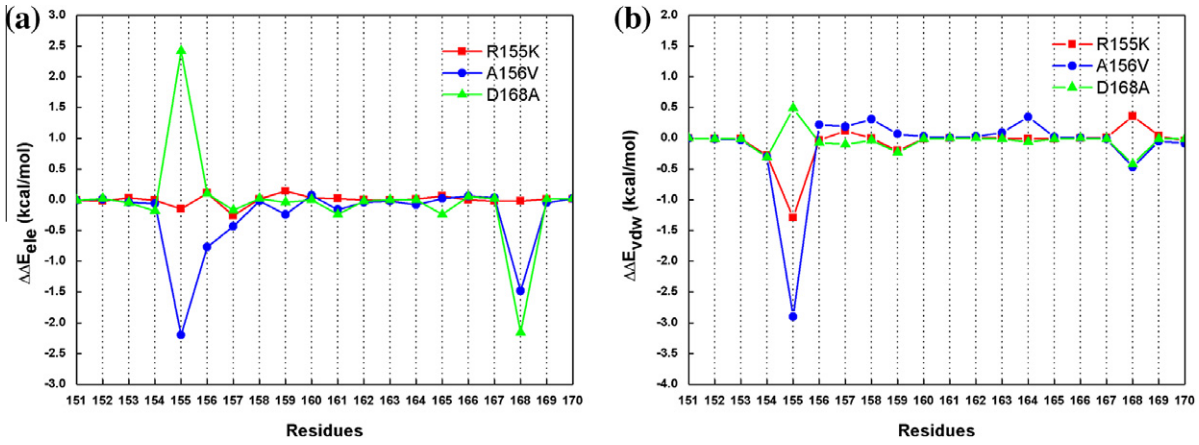

Fig. 10. Pre-residue differential (wild type minus mutant) for NS3/4A protease (residue number from 151 to 170) with TMC435 representing electrostatic (a) and van der Waals (b) energy changes due to mutations.

oline. In addition, the decomposition calculation of the interaction energies to each residue shown in Fig. 6a illustrates that Arg155 makes the biggest contribution to TMC435 binding. These computational results are in good agreement with the experimental results (Cummings et al., 2010).

By contrast, as shown in Fig. 5b and d, the substitutions by lysine at 155 and alanine at 168 could severely disrupt the salt bridge network formed by Arg155 and Asp168. Here, to investigate profoundly this interesting phenomenon, the salt bridge distances (defined by the distance of the nitrogen atoms of residue 155 and

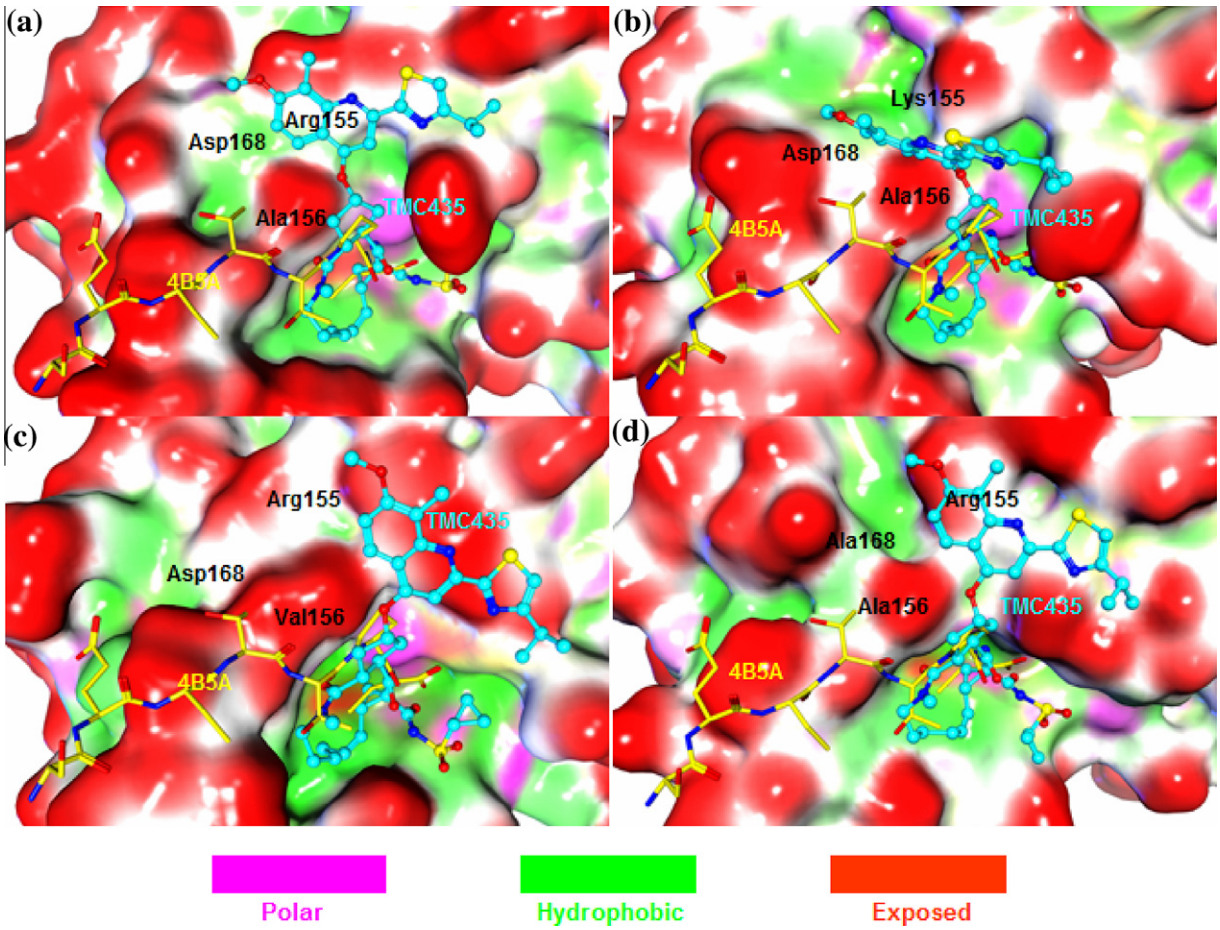

Fig. 11. Molecular surface representation of the average active site structure of the wild-type and mutant NS3/4A–TMC435 complex as well as the superimposed 4B5A substrate. TMC435 is represented as a ball-stick model (in cyan), substrate 4B5A is shown as a stick drawing (in yellow). (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

the oxygen atoms of residue 168) were calculated over the whole MD simulation and the plot of distance vs. the simulation time is shown in Fig. 8. It can be seen that the salt bridge interactions between Lys155 and Asp168 disappears completely after 4 ns for the R155K mutant. As for the D168A mutant, the salt bridge between residue 168 and Arg155 is not stable throughout the whole simulation. The disruption of this salt bridge will change the conformation of Arg155 and further affect seriously the interaction between the residue155 and the P2 quinoline group of TMC435, which can be further verified by the calculated energetical loss shown in Fig. 6b and d. In detail, the R155K mutation generates a much larger effect on TMC435 binding to the extended S2 subsite than D168A mutation which can be observed both from the structure (Fig. 5b and d) and the calculated relative binding free energy (−13.59 kcal/mol for R155K and −19.01 kcal/mol for D168A, respectively, as shown in Table 1).

3.2.2. Two canonical substrate-like intermolecular hydrogen bonds

In the reported crystal complexes, eight hydrogen bonds are conserved between protease residues (Ser159, Ala157, Arg155, Ser139, Ser138 and Gly137) and the viral substrate residues (Romano et al., 2010). It is well known that TMC435 is an inhibitor mimicking the N-terminal side of the viral substrates and can form two canonical substrate-like intermolecular hydrogen bonds with NS3/4A (TMC435(N1–H46)···Arg155(O) and Ala157(N–H)···TMC435(O2)), shown in Fig. 2b) in the crystallographic structure of the TMC435–NS3/4A complex (Cummings et al., 2010), which are very important for TMC435 binding at the active site. Therefore, we monitored the occupancies of two hydrogen bonds in four complexes shown in Table 2. From Table 2, it can be seen that for the wild-type complex that TMC435 can form hydrogen bond with Arg155 (TMC435(N1–H46)···Arg155(O), with 72.77% occupancy) and with Ala157 (Ala157(N–H)···TMC435(O2), with 46.22% occupancy). In addition, the distance between the donor and acceptor for each system was also computed and shown in Fig. 9. For example, as shown in Fig. 9a, the distance between the donor and acceptor atoms of two hydrogen bonds in the wild-type complex is very stable. Hence, the two canonical substrate-like intermolecular hydrogen bonds maintained well for wild-type NS3/4A–TMC435 complex during the 10 ns MD simulation. Compared with the wild type complex, the canonical substrate-like intermolecular hydrogen bonds are well maintained for R155K and D168A (Fig. 9b and d and Table 2). But for A156V, the mutation does have an obvious effect on the two canonical substrate-like intermolecular hydrogen bonds to some degree, where the hydrogen bond occupancies reduced to 27.38% and 29.68% (Table 2) for TMC435(N1–H46)···Arg155(O) and Ala157(N–H)···TMC435(O2), respectively. Corresponding with the reduction of hydrogen bond occupancies, there are significant

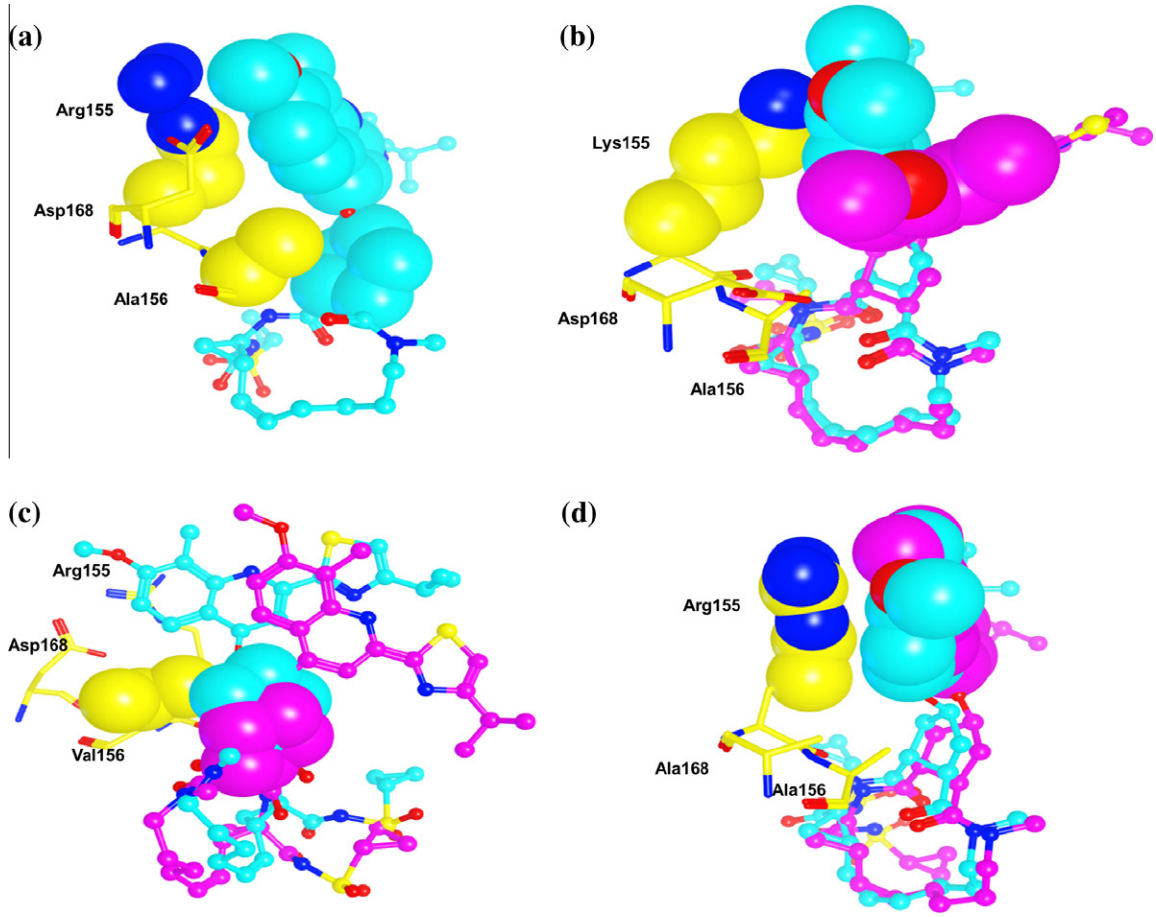

Fig. 12. The superposition of the average structure of TMC435 taken from the wild-type and mutant NS3/4A–TMC435 complex simulation. TMC435 is represented as a ball-stick model (in cyan for wild-type and in pink for mutant complex, respectively). The three major resistance related residues are taken from the wild-type NS3/4A–TMC435 complex simulation and shown as a stick drawing (in yellow). The van der Waals surface of some particular amino acid side chains and TMC435 groups are also displayed. (For interpretation of the references to colors in this figure legend, the reader is referred to the web version of this article.)

fluctuations in hydrogen bond distance (Fig. 9c) and a decreased nonbonded electrostatic interaction energy ($\Delta\Delta E_{ele}$ shown in Fig. 10a) between residue 156 and TMC435.

3.3. How NS3/4A mutants cause drug resistance

3.3.1. Mutations destroy the induced-fit binding of NS3/4A–TMC435 complex

It has been reported that binding of TMC435 involves an induced-fit mechanism in an extended S2 subsite of NS3/4A (Fig. 3) (Cummings et al., 2010). In particular, as discussed above, Arg155 adopts a conformation favorable to interact with the quinoline of TMC435 and it should be emphasized that the salt bridge network between Gln80, Arg155 and Asp168 in the complex plays a key role to stabilize this meaningful conformation (Fig. 2b). Our molecular modeling study of the wild-type NS3/4A–TMC435 complex perfectly reproduced the induced-fit binding mode (Figs. 5, 11 and 12a) as in the crystallographic structure. Furthermore, the large contribution of the van der Waals energetic component of Arg155 in the extended S2 subsite (black line in Fig. 7) is reproduced play a major role in antiviral activity. Therefore, both experimental and simulation results suggest that an extended S2 subsite is produced for the binding of TMC435 in the active site of the NS3/4A protease.

For the R155K mutation, one of the major class-specific mutations has been found to affect HCV NS3/4A protease susceptibility to its inhibitors (Halfon and Locarnini, 2011). In the present study, the R155K mutation results in the loss of the salt bridge of residue 155 with Asp168 (Fig. 5b, red line in Fig. 8) and furthermore causes the large quinoline–thiazole group of TMC435 to move away (Figs. 5 and 12b). Accordingly, the van der Waals interaction energy contribution of the Lys155 to TMC435 was largely decreased in the mutant complex (red line in Figs. 7 and 10b).

Similar to the R155K mutation, substitution of Asp168 with a non-acidic residue, such as Ala, also results in the loss of the salt-bridge interaction with the Arg155 side-chain on the neighboring β-strand (Fig. 5d). Therefore, the conformation of the Arg155 is no longer energetically favored in the NS3/4A D168A mutant TMC435 complex because the loss of the salt bridge makes Arg155 more flexible and its interactions with the large extended P2 group of TMC435 decrease (Figs. 6 and 11d).

As for the A156V mutation, our computational modeling analysis suggests that the possible conformation of substituted valine side chain has an unfavorable steric contact with TMC435 at the P3 carbonyl group. Accordingly, in the A156V mutant, the P3 group of TMC435 is forced to leave the favorable position by this additional methyl group of Val156, which further makes the quinoline–thiazole P2 substituent of TMC435 to move away from the extend S2 subsite (Fig. 5c), and the induced-fit binding mode of

the quinoline–thiazole P2 substituent with extended S2 subsite is obviously influenced (Fig. 6c, blue line in Figs. 10 and 11c).

3.3.2. Mutations help NS3/4A escape from TMC435 without influencing the substrate binding

Previous studies on the NS3/4A protease domain in complex with the N-terminal products of viral substrates proposed that the protease substrates occupied a consensus volume defined as the substrate envelope (Romano et al., 2010, 2011).

Herein, to survey the locations where TMC435 protrude from the substrate envelope, the NS3/4A molecular surface around the PIs and substrate binding site was generated for the simulated average wild-type and mutant complexes (Fig. 11). Then, we superimposed the structure of product complex 4B5A (PDB ID: 3M5N (Romano et al., 2010)) in our four systems based on the coordinates of the enzyme backbone atoms (Fig. 11).

The superposition of active sites for the wild-type TMC435–NS3/4A and product 4B5A complexes shown in Fig. 11a reveals that TMC435 interact with several protease residues of the substrate envelope. Moreover, a specific interaction between the Arg155 side chain of the extended S2 subsite and P2 quinoline of TMC435 (Figs. 2b, Fig. 5a, 11a and 12a) was proven both by experimental data (Cummings et al., 2010; Romano et al., 2010) and our theoretical computation. Thus, mutation R155K and D168A out of the substrate envelop domain can seriously affect the perfect extended S2 subsite (Figs. 11b, 11d, 12b and 12d, respectively), which is very important for binding of the TMC435 extended P2 group.

In addition, the P2 group of TMC435 is found to form a more close interaction with Ala156 relative to the side chain of the 4B5A peptide (Fig. 11a). Therefore, the mutation occurring at the Ala156 position (A156V)forces the P2 group of TMC435 to move away due to a steric clash (Fig 5c) and further leads to a decrease of the binding affinity (Table 1) because the extended P2 quinoline of TMC435 loses the induced binding mode at the extended S2 subsite (Figs. 11c and 12c).The above results prove that the three mutations R155K, A156V and D168A outside the substrate envelope selectively affect TMC435 activity but not the catalytic activity of protease and consequently can automatically help NS3/4A escape from TMC435 without influencing the substrate binding.

4. Conclusions

In summary, the molecular mechanism of drug resistance due to the key residue mutations R155K, A156V and D168A has been explored by using all-atom molecular dynamics simulations followed by MM-GBSA calculations and substrate envelope analysis. Drug resistant mutations R155K and D168A could seriously affect the favorable induced-fit binding with TMC435 through breaking the salt bridges network that stabilizes the extended S2 pocket. The A156V mutation disturbs hydrogen bond interaction by introduction of a steric clash in the complex. Interestingly, we found that mutations in the present study occurring outside the substrate envelop and could easily help the NS3/4A protease escape from TMC435 without influencing the natural substrate peptide (4B5A) binding. These detailed insights into the mechanism of drug resistance will be very useful to structure-based design of novel inhibitors that fit within the substrate envelope that will be less susceptible to resistance.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant Nos. 20905033, 21175063) and the Fundamental Research Funds for the Central Universities (Grant No. lzujbky-2011-19). The authors also would like to thank the Super-computing Center of Gansu Province for providing the computing resources.

References

Accelrys, 2009. Discovery Studio, 2.5; Accelrys: San Diego, CA.
Bartenschlager, R., 1999. The NS3/4A proteinase of the Hepatitis C virus: unravelling structure and function of an unusual enzyme and a prime target for antiviral therapy. J. Viral Hepat. 6, 165–181.
Bayly, C.I., Cieplak, P., Cornell, W., Kollman, P.A., 1993. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J. Phys. Chem. 97, 10269–10280.
Case, D.A., Cheatham, T.E., Darden, T., Gohlke, H., Luo, R., Merz, K.M., Onufriev, A., Simmerling, C., Wang, B., Woods, R.J., 2005. The amber biomolecular simulation programs. J. Comput. Chem. 26, 1668–1688.
Chachra, R., Rizzo, R.C., 2008. Origins of resistance conferred by the R292K neuraminidase mutation via molecular dynamics and free energy calculations. J. Chem.Theory Comput. 4, 1526–1540.
Cieplak, P., Cornell, W.D., Bayly, C., Kollman, P.A., 1995. Application of the multimolecule and multiconformational RESP methodology to biopolymers: charge derivation for DNA, RNA and proteins. J. Comput. Chem. 16, 1357–1377.
Cummings, M.D., Lindberg, J., Lin, T.-I., de Kock, H., Lenz, O., Lilja, E., Felländer, S., Baraznenok, V., Nyström, S., Nilsson, M., Vrang, L., Edlund, M., Rosenquist, Å., Samuelsson, B., Raboisson, P., Simmen, K., 2010. Induced-fit binding of the macrocyclic noncovalent inhibitor TMC435 to its HCV NS3/NS4A protease target. Angew. Chem. Int. Ed. 49, 1652–1655.
Darden, T., York, D., Pedersen, L., 1993. Particle mesh Ewald: an N [center-dot] log(N) method for Ewald sums in large systems. J. Chem. Phys. 98, 10089–10092.
De Francesco, R., Migliaccio, G., 2005. Challenges and successes in developing new therapies for Hepatitis C. Nature 436, 953–960.
Fox, T., Kollman, P.A., 1998. Application of the RESP methodology in the parametrization of organic solvents. J. Phys. Chem. B 102, 8070–8079.
Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E.; Robb, M. A., Cheeseman, J. R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G. A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H. P., Izmaylov, A. F., Bloino, J., Zheng, G., Sonnenberg, J. L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, Jr., J. A., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J. J., Brothers, E., Kudin, K. N., Staroverov, V. N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A., Burant, J. C., Iyengar, S. S., Tomasi, J., Cossi, M., Rega, N., Millam, N. J., Klene, M., Knox, J. E., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R. E., Yazyev, O., Austin, A. J., Cammi, R., Pomelli, C., Ochterski, J. W., Martin, R. L., Morokuma, K., Zakrzewski, V. G., Voth, G. A., Salvador, P., Dannenberg, J. J., Dapprich, S., Daniels, A. D., Farkas, Ö., Foresman, J. B., Ortiz, J. V., Cioslowski, J., Fox, D. J., 2009. In Gaussian 09; Gaussian, Inc.: Wallingford, CT
Gohlke, H., Kiel, C., Case, D.A., 2003. Insights into protein-protein binding by binding free energy calculation and free energy decomposition for the Ras-Raf and Ras–RalGDS complexes. J. Mol. Biol. 330, 891–913.
Gouda, H., Kuntz, I.D., Case, D.A., Kollman, P.A., 2003. Free energy calculations for theophylline binding to an RNA aptamer: comparison of MM-PBSA and thermodynamic integration methods. Biopolymers 68, 16–34.
Guo, Z., Prongay, A., Tong, X., Fischmann, T., Bogen, S., Velazquez, F., Venkatraman, S., Njoroge, F.G., Madison, V., 2006. Computational study of the effects of mutations A156T, D168V, and D168Q on the binding of HCV protease inhibitors. J. Chem. Theory Comput. 2, 1657–1663.
Halfon, P., Locarnini, S., 2011. Hepatitis C virus resistance to protease inhibitors. J. Hepatol. 55, 192–206.
Hou, T., Zhang, W., Case, D.A., Wang, W., 2008. Characterization of domain-peptide interaction interface. A case study on the amphiphysin-1 SH3 domain. J. Mol. Biol. 376, 1201–1214.
Johnson, C.L., Gale, J.M., 2006. CARD games between virus and host get a new player. Trends Immunol. 27, 1–4.
Jorgensen, W.L., Chandrasekhar, J., Madura, J.D., Impey, R.W., Klein, M.L., 1983. Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. 79, 926–935.
Kollman, P.A., Massova, I., Reyes, C., Kuhn, B., Huo, S., Chong, L., Lee, M., Lee, T., Duan, Y., Wang, W., Donini, O., Cieplak, P., Srinivasan, J., Case, D.A., Cheatham, T.E., 2000. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc. Chem. Res. 33, 889–897.
Kolykhalov, A.A., Mihalik, K., Feinstone, S.M., Rice, C.M., 2000. Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3′ nontranslated region are essential for virus replication *in vivo*. J. Virol. 74, 2046–2051.
Kormos, B.L., Benitex, Y., Baranger, A.M., Beveridge, D.L., 2007. Affinity and specificity of protein U1A–RNA complex formation based on an additive component free energy model. J. Mol. Biol. 371, 1405–1419.
Kwong, A.D., McNair, L., Jacobson, I., George, S., 2008. Recent progress in the development of selected Hepatitis C virus NS3 4A protease and NS5B polymerase inhibitors. Curr. Opin. Pharm. 8, 522–531.
Lenz, O., Verbinnen, T., Lin, T.I., Vijgen, L., Cummings, M.D., Lindberg, J., Berke, J.M., Dehertogh, P., Fransen, E., Scholliers, A., Vermeiren, K., Ivens, T., Raboisson, P., Edlund, M., Storm, S., Vrang, L., de Kock, H., Fanning, G.C., Simmen, K.A., 2010. *In*

vitro resistance profile of the Hepatitis C virus NS3/4A protease inhibitor TMC435. Antimicrob. Agents Chemother. 54, 1878–1887.
Lin, C., Lin, K., Luong, Y.-P., Rao, B.G., Wei, Y.-Y., Brennan, D.L., Fulghum, J.R., Hsiao, H.-M., Ma, S., Maxwell, J.P., Cottrell, K.M., Perni, R.B., Gates, C.A., Kwong, A.D., 2004. *In vitro* resistance studies of Hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061: structural analysis indicates different resistance mechanisms. J. Biol. Chem. 279, 17508–17514.
Lin, C., Gates, C.A., Rao, B.G., Brennan, D.L., Fulghum, J.R., Luong, Y.-P., Frantz, J.D., Lin, K., Ma, S., Wei, Y.-Y., Perni, R.B., Kwong, A.D., 2005. *In vitro* studies of cross-resistance mutations against two Hepatitis C virus serine protease inhibitors, VX-950 and BILN 2061. J. Biol. Chem. 280, 36784–36791.
Liu, H., Yao, X., Wang, C., Han, J., 2010. In silico identification of the potential drug resistance sites over 2009 influenza A (H1N1) virus neuraminidase. Mol. Pharm. 7, 894–904.
Manns, M.P., Foster, G.R., Rockstroh, J.K., Zeuzem, S., Zoulim, F., Houghton, M., 2007. The way forward in HCV treatment-finding the right path. Nat. Rev. Drug Discov. 6, 991–1000.
Massova, I., Kollman, P., 2000. Combined molecular mechanical and continuum solvent approach (MM-PBSA/GBSA) to predict ligand binding. Perspect. Drug Discovery Des. 18, 113–135.
Onufriev, A., Bashford, D., Case, D.A., 2000. Modification of the generalized born model suitable for macromolecules. J. Phys. Chem. B 104, 3712–3720.
Raboisson, P., de Kock, H., Rosenquist, A., Nilsson, M., Salvador-Oden, L., Lin, T.-I., Roue, N., Ivanov, V., Wähling, H., Wickström, K., Hamelink, E., Edlund, M., Vrang, L., Vendeville, S., Van de Vreken, W., McGowan, D., Tahri, A., Hu, L., Boutton, C., Lenz, O., Delouvroy, F., Pille, G., Surleraux, D., Wigerinck, P., Samuelsson, B., Simmen, K., 2008. Structure-activity relationship study on a novel series of cyclopentane-containing macrocyclic inhibitors of the Hepatitis C virus NS3/4A protease leading to the discovery of TMC435350. Bioorg. Med. Chem. Lett. 18, 4853–4858.
Rizzo, R.C., Toba, S., Kuntz, I.D., 2004. A molecular basis for the selectivity of thiadiazole urea inhibitors with stromelysin-1 and gelatinase-A from generalized born molecular dynamics simulations. J. Med. Chem. 47, 3065–3074.
Romano, K.P., Ali, A., Royer, W.E., Schiffer, C.A., 2010. Drug resistance against HCV NS3/4A inhibitors is defined by the balance of substrate recognition versus inhibitor binding. Proc. Natl. Acad. Sci. USA 107, 20986–20991.
Romano, K.P., Laine, J.M., Deveau, L.M., Cao, H., Massi, F., Schiffer, C.A., 2011. Molecular mechanisms of viral and host-cell substrate recognition by HCV NS3/4A protease. J. Virol. 85, 6106–6116.
Ryckaert, J.-P., Ciccotti, G., Berendsen, H.J.C., 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23, 327–341.
Schechter, I., Berger, A., 1967. On the size of the active site in proteases. I. Papain. Biochem. Biophys. Res. Commun. 27, 157–162.
Simmonds, P., Holmes, E.C., Cha, T.-A., Chan, S.-W., McOmish, F., Irvine, B., Beall, E., Yap, P.L., Kolberg, J., Urdea, M.S., 1993. Classification of Hepatitis C virus into six major genotypes and a series of subtypes by phylogenetic analysis of the NS-5 region. J. Gen. Virol. 74, 2391–2399.
Simmonds, P., Bukh, J., Combet, C., Deléage, G., Enomoto, N., Feinstone, S., Halfon, P., Inchauspé, G., Kuiken, C., Maertens, G., Mizokami, M., Murphy, D.G., Okamoto, H., Pawlotsky, J.-M., Penin, F., Sablon, E., Shin-I, T., Stuyver, L.J., Thiel, H.-J., Viazov, S., Weiner, A.J., Widell, A., 2005. Consensus proposals for a unified system of nomenclature of Hepatitis C virus genotypes. Hepatology 42, 962–973.
Sitkoff, D., Sharp, K.A., Honig, B., 1994. Accurate calculation of hydration free energies using macroscopic solvent models. J. Phys. Chem. 98, 1978–1988.
Smith, R.E.T., 2006. Hepatitis C virus therapies. Nat. Rev. Drug Discov. 5, 715–716.
Stote, R.H., Karplus, M., 1995. Zinc binding in proteins and solution: a simple but accurate nonbonded representation. Proteins: Struct. Funct. Bioinform. 23, 12–31.
Tibotec, 2011. http://www.tibotec.com/news/detail.jhtml?action=view&itemname=news_78, accessed Sep 24, 2011.
Tsui, V., Case, D.A., 2000. Theory and applications of the generalized born solvation model in macromolecular simulations. Biopolymers 56, 275–291.
Duan, Y., Wu, C., Chowdhury, S., Lee, M.C., Xiong, G., Zhang, W., Yang, R., Cieplak, P., Luo, R., Lee, T., Caldwell, J., Wang, J., Kollman, P., 2003. A point-charge force field for molecular mechanics simulations of proteins based on condensed-phase quantum mechanical calculations. J. Comput. Chem 24, 1999–2012.
Wang, J., Wolf, R.M., Caldwell, J.W., Kollman, P.A., Case, D.A., 2004. Development and testing of a general amber force field. J. Comput. Chem. 25, 1157–1174.
Welsch, C., Domingues, F.S., Susser, S., Antes, I., Hartmann, C., Mayr, G., Schlicker, A., Sarrazin, C., Albrecht, M., Zeuzem, S., Lengauer, T., 2008. Molecular basis of telaprevir resistance due to V36 and T54 mutations in the NS3/4A protease of the Hepatitis C virus. Genome Biol. 9, R16.
World Health Organization, 2010. Hepatitis C: disease3 burden. World Health Organization, Geneva, Switzerland. http://www.who.int/vaccine_research/diseases/viral_cancers/en/index2.html, accessed Sep 24, 2011.