Insight into the drug resistance mechanisms of GS-9669 caused by mutations of HCV NS5B polymerase via molecular simulation

Di Han [a,b], Huiqun Wang [c], Baerlike Wujieti [a], Beibei Zhang [a], Wei Cui [a,*], Bo-Zhen Chen [a,*]

[a] School of Chemical Sciences, University of Chinese Academy of Sciences, No. 19A, YuQuan Road, Beijing 100049, China
[b] School of Medical Engineering, Xinxiang Medical University, Xinxiang 453003, China
[c] Department of Medicinal Chemistry, School of Pharmacy, Virginia Commonwealth University, 800 E Leigh Street, Richmond, VA 23298, USA

ARTICLE INFO

ABSTRACT

GS-9669 is a non-nucleos(t)ide inhibitor (NNI) binding to the thumb site II of the Hepatitis C virus (HCV) NS5B polymerase and has advanced into phase II trials. To clarify the drug resistance mechanisms of GS-9669 caused by M423T/I/V, L419M, R422K, and I482L mutations of NS5B polymerase (GT1b) and the receptor-ligand interactions during the binding process, a series of molecular simulation methods including molecular dynamics (MD) simulations and adaptive steered molecular dynamics (ASMD) simulations were performed for the wild-type (WT) and six mutant NS5B/GS-9669 complexes. The calculated results indicate that the binding free energies of the mutant systems are less negative than that of the WT system, indicating that these mutations will indeed cause NS5B to produce different degrees of resistance to GS-9669. The mutation-induced drug resistances are mainly caused by the loss of binding affinities of Leu419 and Trp528 with GS-9669 or the formation of multiple solvent bridges. Moreover, the ASMD calculations show that GS-9669 binds to the thumb II sites of the seven NS5B polymerases in distinct pathways without any obvious energy barriers. Although the recognition methods and binding pathways are distinct, the binding processes of GS-9669 with the WT and mutant NS5B polymerases are basically controlled thermodynamically. This study clearly reveals the resistance mechanisms of GS-9669 caused by M423T/I/V, L419M, R422K, and I482L mutations of HCV NS5B polymerase and provides some valuable clues for further optimization and design of novel NS5B inhibitors.

© 2021 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Hepatitis C virus (HCV), an enveloped (+)-strand RNA virus, primarily affects the liver and may lead to liver Cirrhosis and carcinoma [1,2]. The HCV is mainly transmitted to people through blood, mother-to-child, and sexual contact [3,4]. According to the World Health Organization (WHO), approximately 0.4 million people among the 71 million infected with HCV globally died from hepatitis C in 2016 [5].

Prior to the approval of HCV direct-acting antiviral (DAA) therapies, the standard of care (SOC) for treatment of HCV was a combination of pegylated interferon (PEG-IFN) and the antiviral compound ribavirin (RBV) for more than a decade [6]. However, this SOC suffered from poor efficacy and troublesome side effects, such as rash, nausea, anemia, and depression [7]. Fortunately,

thanks to the scientific discoveries related to the structure and replication of HCV [8], the development of DAAs greatly improved the treatment of HCV by using DAAs with PEG-IFN plus RBV [9–12]. Research on DAAs will benefit access to more effective, better-tolerated and even more affordable treatments [13].

With high genetic variability and different geographical distribution, HCV are classified into seven main genotypes (GT1-7), most of which have multiple subtypes (a, b, and so on) [14–16]. The genome of HCV containing approximately 9600 nucleotides can translate to a single open reading frame (ORF) flanked by the 5′- and 3′-untranslated regions (UTR) [17]. The ORF can encode a single large polyprotein precursor which is ultimately cleaved to form ten proteins: three structural proteins (one nucleocapsid protein and two envelope proteins) and seven nonstructural proteins which include two proteins required for virion production (p7 and NS2) as well as five proteins that form the cytoplasmic viral replication complex (NS3, NS4A, NS4B, NS5A, and NS5B) [1,18].

For any virus, the polymerase plays an extremely important role in its life cycle and has been proven to be a very effective target for

* Corresponding authors.
E-mail addresses: cuiwei@ucas.ac.cn (W. Cui), bozhenchen@hotmail.com (B.-Z. Chen).

the development of DAAs [19]. Not surprisingly, HCV NS5B polymerase has become a focus of drug development for the treatment of HCV infection after its discovery [20,21]. NS5B polymerase is a typical RNA-dependent RNA polymerase (RdRp), which is structurally right-hand like and contains thumb, palm, and finger subdomains (Fig. 1a) [22–24]. In recent years, several clinically relevant inhibitors against this polymerase have been successively studied and are divided into two categories classically: nucleos(t)ide inhibitors (NIs) and non-nucleos(t)ide inhibitors (NNIs) [21,25]. The NIs mimic the natural substrates of NS5B polymerase, tackling the active site and terminating the RNA synthesis. In contrast, the NNIs inhibit NS5B polymerase by binding to the allosteric sites (thumb site I or II, or palm site I or II, see Fig. 1a) [16].

Among the reported NS5B inhibitors, GS-9669 (see Fig. 1b), an NNI that binds to the thumb site II, is a promising antiviral drug candidate for the GT1 chronic HCV and has advanced into phase II trials [26–28]. However, resistance variants such as M423T/I/V, L419M, R422K, I482L and so on, were identified to GS-9669 in vitro [26,29], which could reduce its potency. The understanding of the drug resistance mechanism of GS-9669 is very crucial for the design of novel potent agents targeting viral variants. But to date, detailed insight into the resistance mechanism at the atomic level is still unfound. In this work, focusing on the M423T/I/V, L419M, R422K, and I482L mutants which have been found in replicon-based resistance selection experiments with thumb site II inhibitors [26], we systemically investigate the interaction mechanisms of GS-9669 with the wild-type (WT) NS5B polymerase (GT1b) and its resistance variants by using molecular dynamics (MD) simulation, binding free energy calculation, and binding energy decomposition. Additionally, adaptive steered molecular dynamics (ASMD) simulations were also performed to view the details of the receptor-ligand interactions during the binding process. Our calculated results can provide some valuable clues for further design of novel NS5B inhibitors that are less susceptible to drug resistance.

2. Materials and methods

2.1. Initial model preparation

Lacking of the co-crystal structure of NS5B and GS-9669, molecular docking technology is used to construct the complex structure. The co-crystal structure of NS5B$^{T329V, V338A, R544Q}$ polymerase (GT1b) binding with a tri-substituted inhibitor similar with GS-9669 (PDB ID: 4EO6 [27]) was derived from the RCSB Protein Data Bank and only one monomer was retained. In Sybyl-X 2.0 software package [30], the missing residues were added, and GS-9669 was sketched, optimized, and then docked into the thumb site II of NS5B polymerase based on the tri-substituted small molecule. Thus the complex NS5B$^{T329V, V338A, R544Q}$/GS-9669 was obtained. Subsequently, the above engineered mutant residues were re-mutated to the corresponding WT residues. On this basis, the models of the six mutant complexes (M423T/I/V, L419M, R422K, and I482L) were obtained by mutating the WT complex.

The restrained electrostatic potential (RESP) [31] fitting of GS-9669 went through the following process: Initially, optimization and electrostatic potential calculation of GS-9669 were carried out by the Gaussian 09 program [32] at HF/6-31G* level. Then, the RESP charge fitting and the generation of the corresponding force field parameters were realized with the antechamber and parmchk modules in AMBER16 [33], respectively. In the following molecular mechanics (MM) optimizations and the MD simulations, the leaprc.protein.ff14SB [34] AMBER force field was used for the proteins and the general AMBER force field (GAFF) [35] for the ligand. Before the MM optimizations and MD simulations, the missing atoms of NS5B polymerase were added using the tleap module in AMBER16. Each constructed complex was immersed in a cube TIP3P [36] water box, the distance from whose edge to any solute atom was at least 10 Å. Additionally, each system was neutralized with 16 counter ions of Cl$^-$.
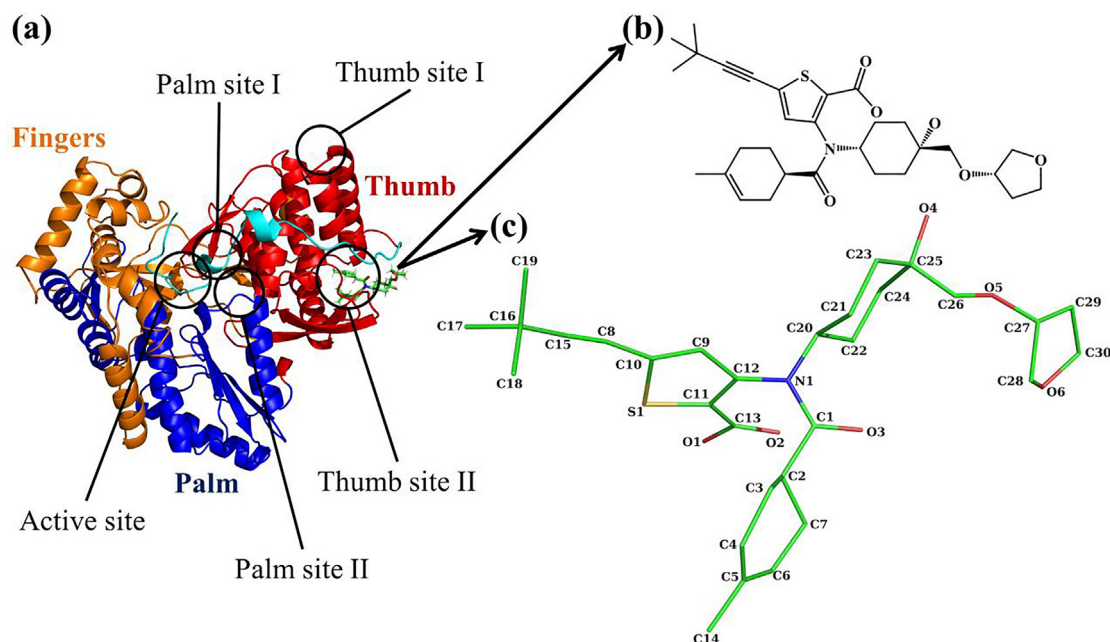

Fig. 1. (a) Cartoon structure of NS5B polymerase (domains thumb, fingers, palm, and the truncated C-terminus are colored red, orange, blue, and cyan, respectively) binding with GS-9669, (b) chemical structure of GS-9669, and (c) atom notations of GS-9669. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

2.2. Molecular minimizations and MD simulations

Initially, each system was minimized via the following three steps using the pmemd.cuda program of AMBER16 running on a GeForce GTX 1080 card. First, the receptor and ligand were restrained, and the water molecules and ions were relaxed for 12,000 cycles (6000 cycles of steepest descent and 6000 cycles of conjugate gradient minimizations). Second, the side chains of protein were optimized with their backbones restrained (6000 cycles of steepest descent and 6000 cycles of conjugate gradient minimizations). Finally, the whole system was minimized without any restraints (6000 cycles of steepest descent and 6000 cycles of conjugate gradient minimizations).

After minimization, the systems were gradually heated up from 0 to 310 K over 350 ps. Then, 60 ns of canonical ensemble MD simulations were carried out under the constant temperature of 310 K using the pmemd.cuda program. During the MD simulations, SHAKE algorithm [37] was used to fix the bonds connected to the hydrogen atoms, and the particle mesh Ewald (PME) method [38] was employed to calculate the long-range coulombic interactions. The integration time step and the nonbonded cutoff were set to 2.0 fs and 10 Å, respectively. Meanwhile, to avoid edge effects, periodic boundary conditions were applied in all calculations. During the sampling process, the coordinates were saved every 1.0 ps. After the MD simulations, 1000 snapshots extracted from the last 20 ns stable MD trajectory at 20 ps interval were used to perform the further binding free energy calculations and binding energy decomposition analyses.

2.3. Binding free energy calculations

The predicted binding free energy of each system was calculated by the mm_pbsa program in AMBER16 according to the following equation [39]:

$\Delta G_{pred} = G_{complex} - G_{protein} - G_{ligand} = \Delta E_{MM} + \Delta G_{GB} + \Delta G_{SA} - T\Delta S$ $= \Delta E_{bind} - T\Delta S$, where ΔE_{MM} is the gas-phase interaction energy between the receptor and the ligand, which includes three parts: ΔE_{ele} (electrostatic energy), ΔE_{vdw} (van der Waals energy), and ΔE_{int} (internal energy); ΔG_{GB} and ΔG_{SA} are the polar and nonpolar contributions to desolvation upon ligand binding, respectively; and $-T\Delta S$ is the conformational entropy change (translation, rotation, and vibration) at temperature T. Here, the ΔG_{GB} (polar desolvation energy) was calculated by the Generalized Born (GB) model proposed by Onufriev et al. (igb = 2) [40] for the MM/GBSA method [41–44], with the dielectric constant values for solvent and solute set to 80 and 1, respectively. The ΔG_{SA} (nonpolar desolvation energy) was estimated by the Laboratoire de Chimie des Polymeres Organiques (LCPO) method [45]: $\Delta G_{SA} = 0.0072 \times \Delta SASA$, where $\Delta SASA$ represents the variation of the solvent accessible surface area. ΔE_{bind} (binding energy) is the sum of ΔE_{MM}, ΔG_{GB}, and ΔG_{SA}. The entropy change ($-T\Delta S$) here was calculated by a recent computationally efficient and numerically reliable approach named interaction entropy (IE) [46]. The IE is defined as $-T\Delta S = KT\ln\langle exp(\Delta E_{pl}^{int}/KT)\rangle$. In the formula, ΔE_{pl}^{int} represents the fluctuation of protein–ligand interaction energy around the average energy over time, and $\Delta E_{pl}^{int} = E_{pl}^{int} - \langle E_{pl}^{int}\rangle$. Here, $\langle E_{pl}^{int}\rangle$ is the average interaction energy between the protein and ligand. In the present work, totally 20,000 snapshots evenly extracted from the last 20 ns were included in the entropy calculation of each system.

2.4. Binding energy decomposition analyses

In order to further understand the detailed protein-inhibitor interactions in the WT and six mutant NS5B/GS-9669 systems, MM/GBSA decomposition analysis of the binding energy was carried out by using the mm_pbsa program in AMBER16. The total

binding energy between each residue and GS-9669 was decomposed into four parts: van der Waals term (ΔE_{vdw}), electrostatic term (ΔE_{ele}), polar desolvation term (ΔG_{GB}), and nonpolar desolvation term (ΔG_{SA}). ΔE_{vdw} and ΔE_{ele} were computed by the sander module in AMBER16. The polar contribution of desolvation (ΔG_{GB}) was calculated by the GB model [47], and the nonpolar contribution of desolvation (ΔG_{SA}) was computed according to SASA determined by the ICOSA method [48].

2.5. Adaptive steered molecular dynamics simulations

Adaptive steered molecular dynamics (ASMD) simulations [49–52], which can provide a more detailed mapping of the force profiles along the binding coordinate, were performed for the WT and six mutant complexes using Amber16. In the ASMD simulations, the predetermined reaction coordinate (the distance between the ligand and the protein) was divided into several stages, with each one having numerous simulations performed in parallel. In between stages, a single trajectory, whose work value was the closest to the Jarzynski average (JA) [53], was selected from simulations. In each stage, the final structure of the selected trajectory was used to initialize the next stage of ASMD trajectories. The potential of mean force (PMF) was calculated with the trajectories selected using the Jarzynski equality (JE) in each stage. Specifically, the pulling direction was determined with the vector of the CA atom of residue 423 and the N1 atom of GS-9669 (Fig. 1c), two atoms being located on a more rigid structure and the line connecting them is nearly parallel to the active pocket. The distance between the above CA and N1 atoms in the equilibrium structure of the complexes was set to the zero point of the distance between the ligand and the protein in each system. The ligand in each system was pulled out by 50 Å with the stretching velocity and force constant set to 10 Å/ns and 7.2 kcal/(mol·Å^2), respectively. The reaction coordinate was split into 10 stages, with each one carrying out 40 simulations for each protein-inhibitor system. The cartoon models were generated using the PyMOL software package [54].

3. Results and discussions

3.1. Stability of the NS5B/GS-9669 complexes during the MD simulations

A total of 60 ns classic MD simulations were conducted under the constant temperature of 310 K for each of the NS5B^{WT}/GS-9669, NS5B^{M423T}/GS-9669, NS5B^{M423I}/GS-9669, NS5B^{M423V}/GS-9669, NS5B^{L419M}/GS-9669, NS5B^{R422K}/GS-9669, and NS5B^{I482L}/GS-9669 complexes. To monitor the convergence of each system and ensure the rationality of the subsequent sampling strategy, the root-mean-square deviations (RMSD) of the polymerase backbone atoms (CA, C, N) and ligands relative to their respective initial structures were calculated, and the results are plotted in Figs. 2 and S1, respectively. The distributions of the RMSD values indicated that all the systems achieve equilibrium after 30 ns, and then remain stable to the end of the simulations. The average RMSD values of the seven complexes from 40 to 60 ns MD simulations are 1.54, 1.57, 1.74, 1.56, 1.87, 1.45, and 1.75 Å, respectively. These results indicate that using the snapshots extracted from the MD trajectories between 40 and 60 ns to conduct the following energy calculations should be both reasonable and reliable [55,56].

### 3.2. Binding free energies predicted by MM/GBSA method

The predicted binding free energies, together with the energy components, are summarized in Table 1. Student's *t*-test calcula-

Table 1
The predicted binding free energies and individual energy terms of the NS5B/GS-9669 complexes calculated by MM/GBSA method (kcal/mol).

System	ΔE_{ele}[a]	ΔE_{vdw}[b]	ΔG_{GB}[c]	ΔG_{SA}[d]	ΔE_{bind}[e]	$-T\Delta S$[f]	ΔG_{pred}[g]
WT	−7.98 ± 6.69	−43.40 ± 3.41	23.69 ± 6.68	−6.68 ± 0.46	−34.37 ± 3.43	17.85 ± 0.25	−16.52 ± 3.44
M423T	−11.14 ± 6.75	−43.21 ± 3.86	26.94 ± 6.62	−6.75 ± 0.44	−34.16 ± 3.68	20.29 ± 0.18	−13.87 ± 3.68
M423I	−9.31 ± 6.95	−42.88 ± 3.73	24.87 ± 7.70	−6.68 ± 0.54	−34.00 ± 3.09	20.02 ± 0.21	−13.98 ± 3.10
M423V	−8.28 ± 6.63	−40.36 ± 3.52	23.00 ± 6.69	−6.50 ± 0.51	−32.13 ± 3.27	21.30 ± 0.08	−10.83 ± 3.27
L419M	−12.55 ± 5.94	−39.68 ± 2.87	26.60 ± 5.94	−5.76 ± 0.27	−31.39 ± 2.74	18.32 ± 0.22	−13.07 ± 2.75
R422K	−10.55 ± 7.99	−46.78 ± 3.55	27.97 ± 7.36	−7.10 ± 0.37	−36.46 ± 3.84	22.67 ± 0.09	−13.79 ± 3.84
I482L	−13.76 ± 6.81	−45.60 ± 3.49	29.68 ± 6.46	−6.93 ± 0.38	−36.60 ± 3.72	21.08 ± 0.29	−15.52 ± 3.73

[a] Electrostatic contribution; [b] van der Waals contribution; [c] the polar contribution of desolvation; [d] the nonpolar contribution of desolvation; [e] the binding energy; [f] T = 310 K and the entropy calculated through IE method; [g] the predicted binding free energy.
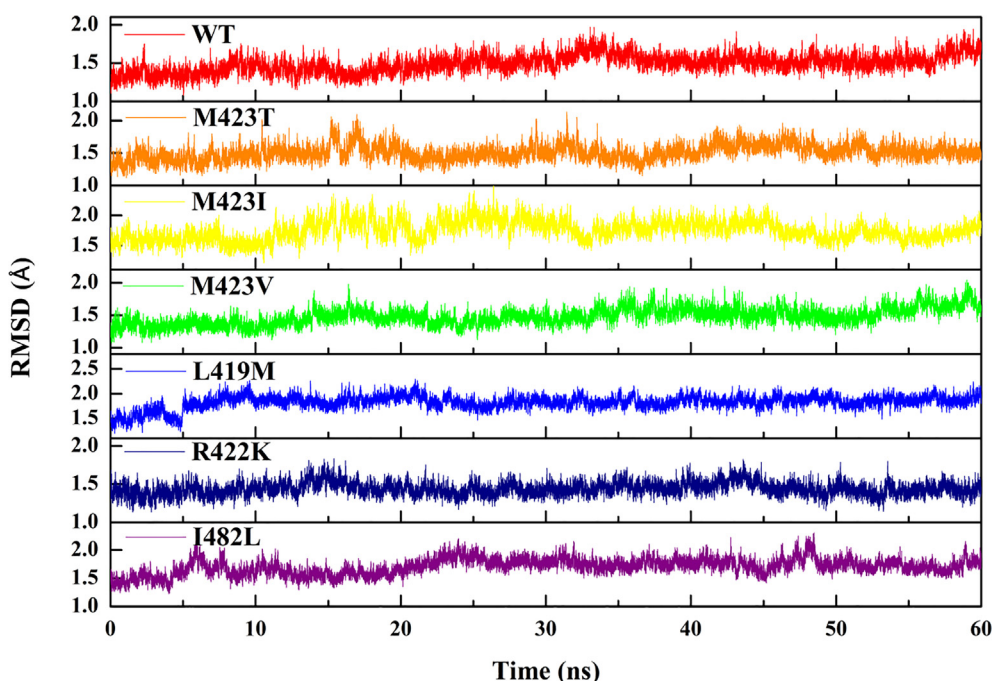
Fig. 2. The root-mean-square deviations (RMSD) of the polymerase backbone atoms (CA, C, N) in the seven systems relative to their respective initial structures.

tions show that the probabilities of no difference for the ΔG_{pred}^{mutant} values towards the ΔG_{pred}^{WT} value are all less than 0.0005, indicating that the relevant data have very significant differences. As can be seen from this table, all the binding energies (ΔE_{bind}) of the M423T (−34.16 kcal/mol), M423I (−34.00 kcal/mol), M423V (−32.13 kcal/mol), and L419M (−31.39 kcal/mol) mutant systems are less negative than that of the WT system (−34.37 kcal/mol). Moreover, the binding energies of these mutation systems increase gradually. The calculation results are consistent with the experimental results [26,29] that these mutations will cause resistance and the levels of resistance increases sequentially. Differently, the binding energies of the R422K (−36.46 kcal/mol) and I482L (−36.60 kcal/mol) mutant systems are more negative than that of the WT system. From the thermodynamic point of view, these two mutations will not cause drug resistance, which is inconsistent with the experimental results. However, it can be seen from the binding free energies (ΔG_{pred}) of the systems (see Table 1) that all the ΔG_{pred} values of the mutant systems (−13.87^{M423T}, −13.98^{M423I}, −10.83^{M423V}, −13.07^{L419M}, −13.79^{R422K}, and −15.52^{I482L} kcal/mol) are less negative than that of the WT system (−16.52 kcal/mol), indicating that these mutations can cause NS5B to produce different levels of resistance to GS-9669. We speculate that the influence of entropies on

these mutants (especially for the R422K and I482L mutations) may be relatively large. It can also be seen from Table 1 that the van der Waals energy term (ΔE_{vdw}) in each system is the main component of the binding free energy. The desolvation energies ($\Delta G_{GB} + \Delta G_{SA}$) of the seven systems are 17.01^{wt}, 20.19^{M423T}, 18.19^{M423I}, 16.5^{M423V}, 20.84^{L419M}, 20.87^{R422K}, and 22.75^{I482L} kcal/mol, indicating that solvation is unfavorable for the binding of GS-9669 and NS5B polymerase. Additionally, in order to identify which energy term has a greater influence on the difference in the binding free energies, the correlation coefficients of the energy components ΔE_{ele}, ΔE_{vdw}, ΔG_{GB}, ΔG_{SA}, and $-T\Delta S$ with respect to ΔG_{pred} were calculated, respectively. The values are 0.16, 0.56, 0.32, 0.36, and 0.37. Obviously, the energy term ΔE_{vdw} should play a key role in distinguishing the binding affinity of the seven systems.

3.3. Comparison of the binding modes of the GS-9669 bound WT and mutant NS5B complexes

To further understand the polymerase/inhibitor recognition patterns and highlight the important residues leading to drug resistance, the total binding energy of each system was decomposed into residue–ligand pairs by the MM/GBSA decomposition
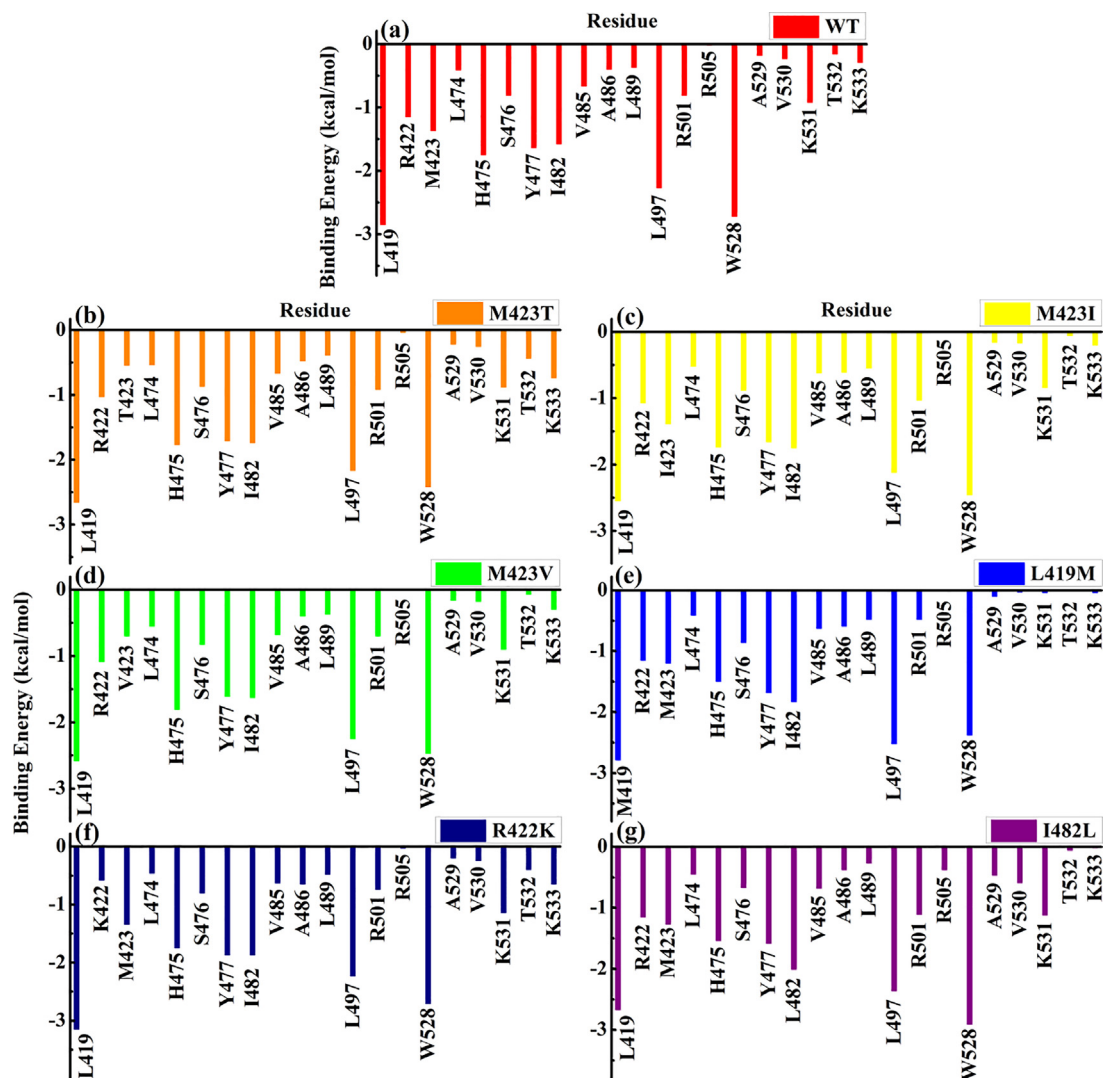
Fig. 3. The binding energy decomposition spectra for the (a) NS5B^{WT}/GS-9669, (b) NS5B^{M423T}/GS-9669, (c) NS5B^{M423I}/GS-9669, (d) NS5B^{M423V}/GS-9669, (e) NS5B^{L419M}/GS-9669, (f) NS5B^{R422K}/GS-9669, and (g) NS5B^{I482L}/GS-9669 systems according to the MM/GBSA decomposition analyses.

Table 2
Visible percentage of hydrogen bonds[a] during the last 20 ns MD simulations between GS-9669 and NS5B polymerases.

System	Donor H	Donor	Acceptor	Occupied(%)	Distance(Å)	Angle(deg)
WT	Tyr477@H	Tyr477@N	GS-9669@O1	19	2.91	151.86
	Ser476@H	Ser476@N	GS-9669@O2	26	2.89	150.42
M423T	Tyr477@H	Tyr477@N	GS-9669@O1	28	2.91	152.75
	Ser476@H	Ser476@N	GS-9669@O2	23	2.89	149.82
M423I	Tyr477@H	Tyr477@N	GS-9669@O1	24	2.92	154.71
	Ser476@H	Ser476@N	GS-9669@O2	27	2.88	148.87
M423V	Tyr477@H	Tyr477@N	GS-9669@O1	22	2.92	153.26
	Ser476@H	Ser476@N	GS-9669@O2	29	2.88	150.52
L419M	Tyr477@H	Tyr477@N	GS-9669@O1	19	2.91	152.68
	Ser476@H	Ser476@N	GS-9669@O2	15	2.88	151.37
R422K	Tyr477@H	Tyr477@N	GS-9669@O1	14	2.92	152.80
	Ser476@H	Ser476@N	GS-9669@O2	27	2.88	149.46
I482L	Tyr477@H	Tyr477@N	GS-9669@O1	17	2.91	153.19
	Ser476@H	Ser476@N	GS-9669@O2	10	2.90	151.66

[a] A hydrogen bond was defined if the donor–acceptor distance was less than 3.00 Å and the donor–donor H-acceptor angle was>135.00°.

Table 3
Solvent bridges during the last 20 ns MD simulations between GS-9669 and the residues of NS5B polymerases.

System	Bridging solute residues		Occupied(%)
WT	Arg501	GS-9669	27
M423T	Arg501	GS-9669	11
M423I	Arg501	GS-9669	23
M423V	Arg501	GS-9669	11
L419M	Arg501	GS-9669	22
R422K	Arg501	GS-9669	21
	Trp528	GS-9669	17
I482L	Arg501	GS-9669	16
	Trp528	GS-9669	31

process. The interaction spectra were plotted in Fig. 3. As can be inferred from the data in this figure, the correlation coefficients between the energy contributions of the key residues in the M423T/I/V, L419M, R422K, and I482L mutant systems and those in the WT system are 0.96, 0.99, 0.98, 0.96, 0.97, and 0.97, respectively. Therefore, GS-9669 in the mutant polymerases should basically remain most of the key interactions in the WT polymerase. Additionally, two other important protein-inhibitor interactions, hydrogen bonds and solvent bridges, were also calculated and summarized in Tables 2 and 3, respectively. It can be seen from Table 2 that the O1 and O2 atoms of GS-9669 (see Fig. 1c) can form hydrogen bonds with Ser476 and Tyr477 in both WT and mutant systems. Therefore, besides Ser476 [19], Tyr477 also plays a key role in anchoring the inhibitor at this site, which is similar to the anchoring mode of VX-222 [57] targeting the thumb II site. Meanwhile, solvent bridging interactions can be formed between the inhibitor and Arg501 in all systems (Table 3). In addition, GS-9669 can also form solvent bridges with Trp528 in both R422K and I482L mutant systems. It can be seen from Fig. 4 that among the above four gatekeeper residues that successfully form hydrogen bonds or solvent bridges with GS-9669, Ser476 and Tyr477 are on one side of the inhibitor, and Arg501 and Trp528 are on the other side. From this point of view, the solvent bridges should be able to help the anchoring hydrogen bonds to better stabilize GS-9669 in the binding pocket, which will prevent the inhibitor from rotating and also lay the foundation for modifying the N-alkyl moiety located at the gate of the active pocket to improve the interaction mode and avoid the drug resistance.

To further validate this conjecture, the root-mean-square fluctuations (RMSF) of the inhibitor's heavy atoms within the last 20 ns MD simulation trajectory for each system was calculated. The results (Fig. S2) show that the RMSF values of the atoms in the N-alkyl moiety are relatively large, indicating that the conformation of this group is indeed flexible in the binding state. Furthermore, it can also be seen from the average structures of GS-9669 during the last 20 ns MD simulations (Fig. S3) that the structures of the N-alkyl moiety (especially for the heterocyclic ether) are twisted [58], which also illustrates that the conformation of this moiety is much flexible. In the subsequent more detailed analyses of binding patterns, we find that the conformations of the thiophene and N-acyl moieties of GS-9669 are basically rigid, and the conformational changes mainly occur in the N-alkyl moiety of the inhibitor.

As observed from the WT system shown in Fig. 4a, the thumb II allosteric pocket dominated by α-helices mainly consists of the residues Leu419, Arg422, Met423, His475, Ser476, Tyr477, Ile482, Leu497, and Trp528. Undoubtedly, the binder to this pocket, GS-9669, will interact with these residues, and our calculations indeed confirm this (Fig. 3a). The above residues surround the thiophene and N-acyl moieties of GS-9669 and the binding energy contributions of them are −2.85, −1.15, −1.37, −1.75, −0.81, −1.64,

−1.58, −2.27, and −2.72 kcal/mol, respectively. In the work of Scott E. Lazerwith et al.[27], to enhance the pharmacokinetic profile, a heterocyclic ether was introduced on this N-alkyl substituent. Our calculations show that the heterocyclic ether is able to adjust its conformation and thus is close to the loop region in space where residues 529–533 are located, interacting with Lys531 (−0.92 kcal/mol) there (see Figs. 3a and 4a). This indicates that the introduction of this group is successful. Overall, GS-9669 binds to the polymerase mainly through hydrophobic, hydrogen bond, and solvent bridge contacts.

To gain a deeper understanding of the residues responsible for the loss of binding affinity in the mutant systems, the energetic differences between the key residues in WT system and every individual mutant system ($\Delta\Delta G = \Delta G_{bind}^{mutant} - \Delta G_{bind}^{WT}$) were calculated and displayed in Fig. 5. The positive values indicate that the residues in the mutant polymerase form weaker interactions with GS-9669 than the corresponding residues in the WT polymerase, while the negative values indicate that the residues in the mutant polymerase form stronger interactions with GS-9669. If ± 0.2 kcal/mol is defined as the threshold, we can find that the M423T mutation directly causes the energy contribution loss of this site residue (see Fig. 5a, $\Delta\Delta G = 0.82$ kcal/mol) in the NS5B^{M423T}/GS-9669 system. This is due to that the side chain of threonine is shorter than that of methionine after the M423T mutation (Fig. 4b), resulting in a weaker interaction with the inhibitor. In addition, we noticed that this mutation also caused a decrease in the energy contribution of Trp528 ($\Delta\Delta G = 0.30$ kcal/mol) and increases in the energy contributions of Thr532 ($\Delta\Delta G = -0.28$ kcal/mol) and Lys533 ($\Delta\Delta G = -0.45$ kcal/mol). Our calculations show that the distance between the group and Trp528 becomes larger and the distances between the group and residues Thr532 and Lys533 become smaller (see Figs. S4a–c). Moreover, this mutation has a certain weakening effect on the strength of the solvent bridge formed between Arg501 and GS-9669 (Table 3). From this, the reason why GS-9669 can still maintain a certain activity against the M423T mutant NS5B polymerase may stem from both the thiophene and the N-acyl moieties rather than from the N-alkyl substituent and the N-acyl group [27]. Therefore, the M423T mutation mainly affects the interaction between the residue at the mutation site and the inhibitor, but failed to cause large changes in the polymerase-inhibitor interactions. This should be the reason that the M423T mutation produced weak drug resistance in experimental detection [26,29].

Unlike the M423T mutant system, from Fig. 5b we can see that the mutation between the two hydrophobic residues of equivalent molecular weight in the M423I mutant system does not cause a significant change in the energy contributions of the residues at the 423 site ($\Delta\Delta G = -0.02$ kcal/mol). Interestingly, this mutation causes the loss of energy contributions of Leu419 ($\Delta\Delta G = 0.30$ kcal/mol) and Trp528 ($\Delta\Delta G = 0.26$ kcal/mol). Further calculations (Figs. S4d–f) show that the distance between Ile423 and the N-acyl moiety of GS-9669 is reduced by 1.35 Å compared to the distance between Met423 and the N-acyl moiety after the M423I mutation, while the distances from the N-acyl moiety to Leu419 and Trp528 are increased by 0.43 Å and 0.14 Å, respectively. These indicate that the bulky side chain of Ile423 can push down the N-acyl moiety of GS-9669 after mutation, which leads to the changes in the energy contributions of the above-mentioned residues. This conformational change of GS-9669 also affects its distances with residues Ala486 and Arg501. Compared to the WT system, the distances from the inhibitor to these two residues in the M423I system are reduced by 0.64 and 1.68 Å (Figs. S4g and h), which is also confirmed by the increases in the energy contributions of these two residues (see Fig. 5b, $\Delta\Delta G^{Ala486} = -0.21$ kcal/mol, $\Delta\Delta G^{Arg501} = -0.22$ kcal/mol). It is worth mentioning that the rotation of the N-alkyl moiety of GS-9669 after mutation should also contribute to the enhancement of the interaction between the inhi-
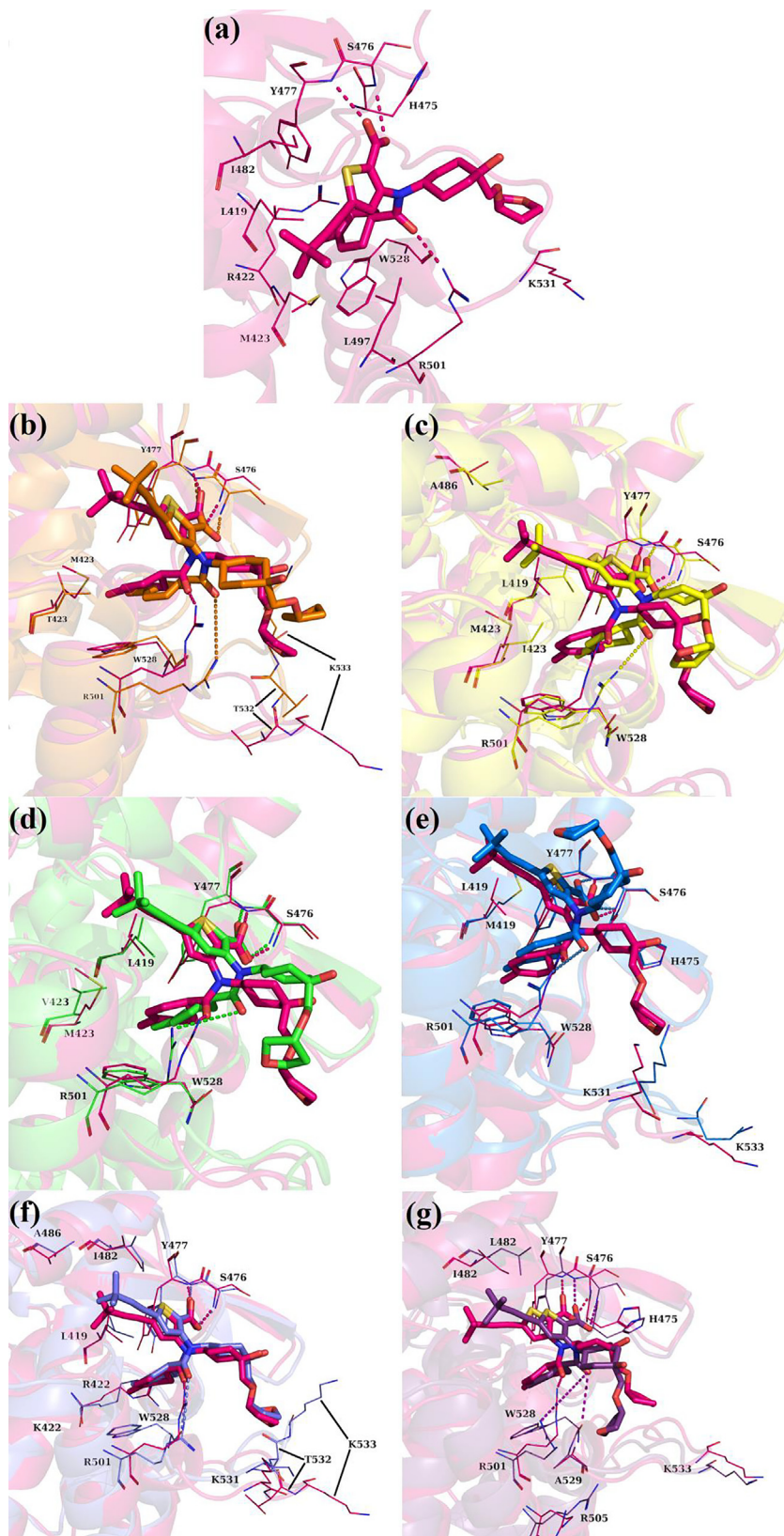
Fig. 4. Binding modes of GS-9669 with key residues in the WT and mutant NS5B polymerases. (a) NS5B^WT/GS-9669; (b) NS5B^WT/GS-9669 (hot pink) versus NS5B^M423T/GS-9669 (orange); (c) NS5B^WT/GS-9669 (hot pink) versus NS5B^M423I/GS-9669 (yellow); (d) NS5B^WT/GS-9669 (hot pink) versus NS5B^M423V/GS-9669 (green); (e) NS5B^WT/GS-9669 (hot pink) versus NS5B^L419M/GS-9669 (blue); (f) NS5B^WT/GS-9669 (hot pink) versus NS5B^R422K/GS-9669 (slate); (g) NS5B^WT/GS-9669 (hot pink) versus NS5B^I482L/GS-9669 (purple). The dashed lines represent hydrogen bonds or solvent bridges. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
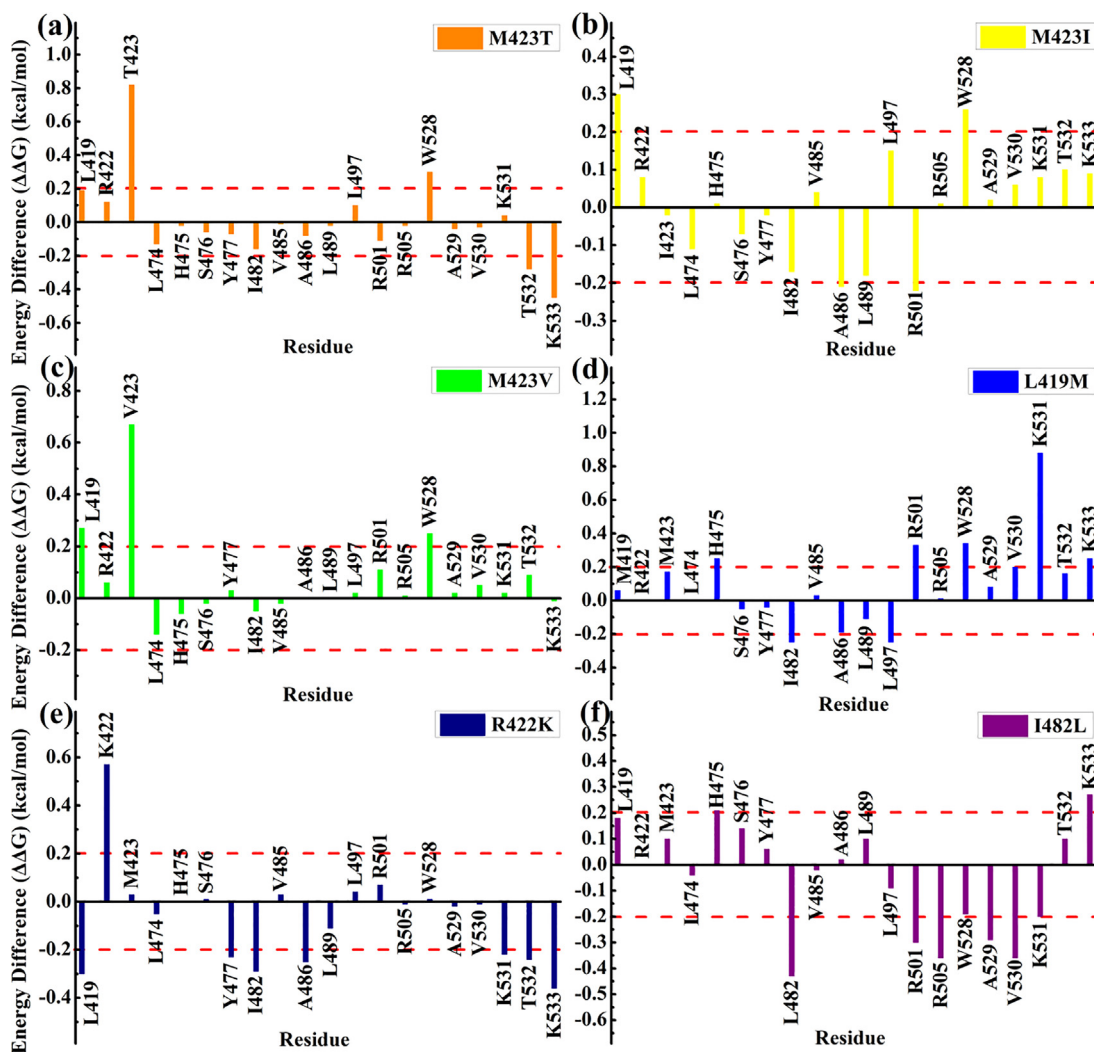
Fig. 5. The energetic difference ($\Delta\Delta G = \Delta G_{bind}^{mutant} - \Delta G_{bind}^{WT}$) spectra between the WT system and the (a) M423T, (b) M423I, (c) M423V, (d) L419M, (e) R422K, or (f) I482L mutant system.

bitor and Arg501. Apart from these, the mutation does not cause other significant changes. It can be seen from the above that the main reason for the resistance of M423I mutation should be that this mutation causes the conformational changes of GS-9669, which led to the reduction of its interactions with Leu419 and Trp528.

To the M423V mutant system, it can be seen from Fig. 5c that the energy contributions of residues 423 and Trp528 decrease by 0.67 and 0.25 kcal/mol after the mutation, which is similar to the situation in the M423T mutant system. Moreover, residues with decreased energy contributions also include Leu419 ($\Delta\Delta G$ = 0.27 kcal/mol). Calculations show that the distances from GS-9669 to residues Val423, Trp528, and Leu419 are increased after the mutation (Figs. S4i–k). As mentioned earlier, the N-alkyl moiety of GS-9669 remains its flexibility after binding to the NS5B polymerase, and it also exists in the M423V mutant system.

From the above discussion, it can be concluded that the RAVs M423T/V directly lead to the decreases in the energy contributions of the residues at the mutation site, thereby disturbing the original conformation of the N-acyl moiety of GS-9669, resulting in the decreases of energy contributions of Trp528 and Leu419. Although the M423I mutation does not cause direct changes in the energy

contributions of the residues located at the mutant site, it disturbs the original conformation of the N-acyl moiety of GS-9669 through a bulky side chain, resulting in the decrease of energy contributions of Leu419 and Trp528. In general, RAVs M423T/I/V can result in decreased interactions between the mutation site residues Leu419/Trp528 and GS-9669. Among them, since Trp528 is located at the entrance of the binding pocket, its interaction with GS-9669 will also be affected by the conformational change of the heterocyclic ether.

For the L419M mutant system, it can be seen from Fig. 5d that the mutation has no significant effect on the interaction between residue 419 and GS-9669 ($\Delta\Delta G$ = 0.06 kcal/mol). It is very interesting that in the L419M mutant system, the orientation of the heterocyclic ether of GS-9669 is rotated by almost 180° compared to that in the WT system (see Fig. 4e). In order to validate this observation, three times independent simulations were performed to this mutant system. The conformations of the inhibitor were almost the same in all of our simulations results. This conformational change results in significant decreases of interactions between the original surrounding residues and the heterocyclic ether of GS-9669. The residues include His475 ($\Delta\Delta G$ = 0.25 kcal/mol), Arg501 ($\Delta\Delta G$ = 0.33 kcal/mol), Trp528 ($\Delta\Delta G$ = 0.34 kcal/mol), Lys531 ($\Delta\Delta G$ = 0.88 kcal/mol), and
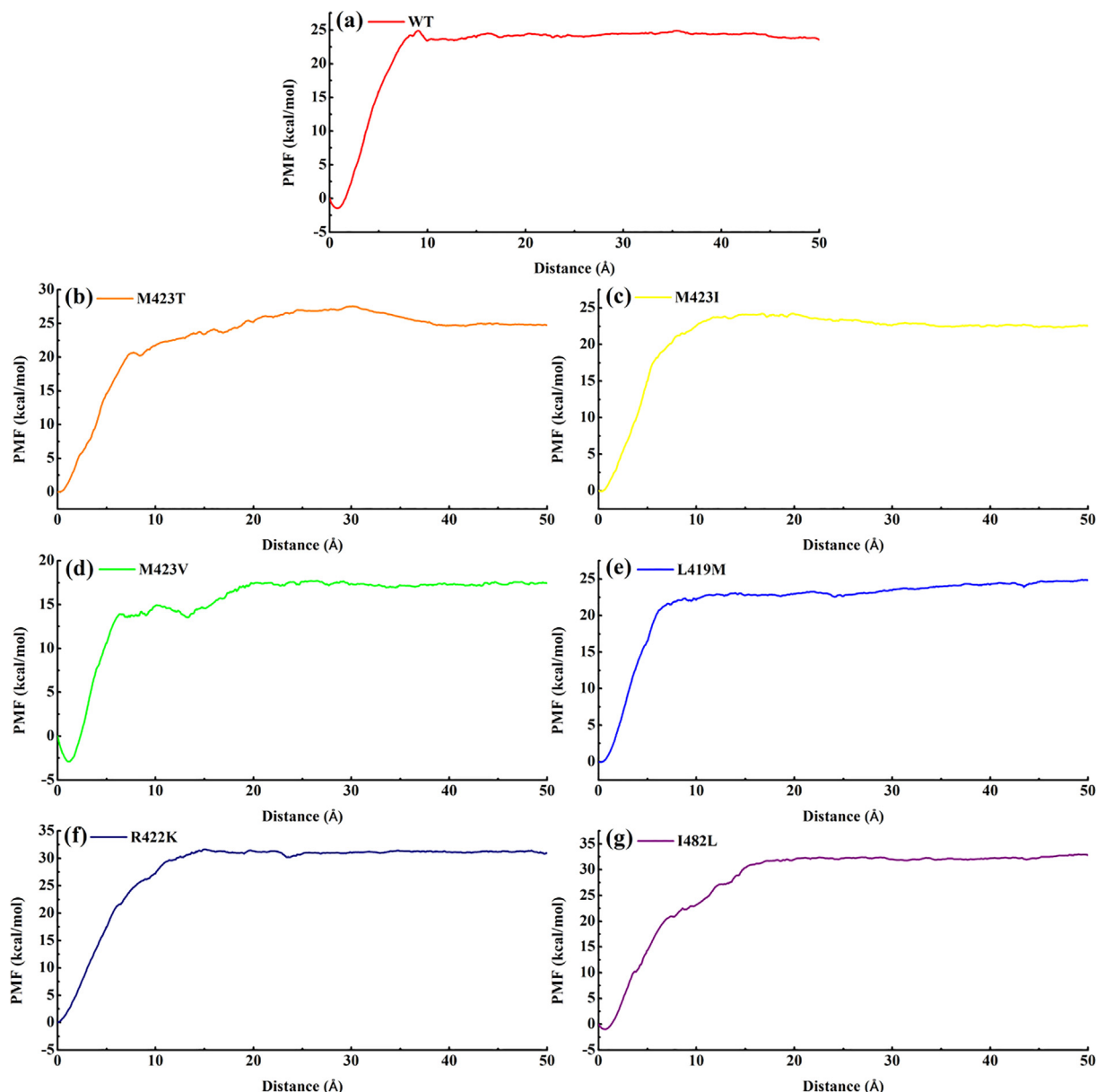
Fig. 6. PMF profiles along the reaction coordinate for the NS5B^{WT}/GS-9669 and NS5B^{mutant}/GS-9669 complexes.

Lys533 ($\Delta\Delta G$ = 0.25 kcal/mol). This may provide a possible explanation why the L419M mutation causes strong drug resistance to GS-9669.

In the R422K mutant system (Fig. 4f), replacement of the larger molecular weight arginine by the smaller molecular weight lysine directly results in loss of energy contribution of residue 422 ($\Delta\Delta G$ = 0.57 kcal/mol). However, the energy contributions of Leu419 ($\Delta\Delta G$ = −0.3 kcal/mol), Tyr477 ($\Delta\Delta G$ = −0.23 kcal/mol), Ile482 ($\Delta\Delta G$ = −0.29 kcal/mol), Ala486 ($\Delta\Delta G$ = −0.25 kcal/mol), Lys531 ($\Delta\Delta G$ = −0.22 kcal/mol), Thr532 ($\Delta\Delta G$ = −0.24 kcal/mol), and Lys533 ($\Delta\Delta G$ = −0.36 kcal/mol) all increase in this system. The calculations show that the distances between the above residues and the inhibitor have become shorter after the mutation (Figs. S4l–r), indicating that the position of GS-9669 in the binding pocket shifts toward Tyr477 after the mutation. Overall, the R422K mutation even makes the polymerase-inhibitor binding energy more negative, but the conformational entropy change of the R422K mutant system is the largest one of these systems (Table 1), which makes the final binding free energy less negative than that

of the WT system, and the mutant polymerase develops drug resistance to GS-9669. The reason may be that more solvent bridges are formed in the R422K system (see Table 3), which causes an increase in the -TΔS term.

From Figs. 4g and 5f, we can see that the difference of the side chains between isoleucine and leucine is able to cause a significant increase in the energy contribution ($\Delta\Delta G$ = −0.43 kcal/mol) in the I482 L mutant system. The distance between Leu482 and GS-9669 in the mutant system is also shorter than that between Ile482 and the inhibitor in the WT system (Fig. S4s). Furthermore, the N-alkyl moiety of the GS-9669 in the I482L mutant system is closer to Ala529 ($\Delta\Delta G$ = −0.29 kcal/mol) than that in the WT system (Fig. S4t), and forms stronger interactions with the neighboring Arg501 ($\Delta\Delta G$ = −0.30 kcal/mol), Arg505 ($\Delta\Delta G$ = −0.36 kcal/mol), and Val530 ($\Delta\Delta G$ = −0.36 kcal/mol), respectively. This conformational change of GS-9669 causes decreases in the energy contributions of His475 ($\Delta\Delta G$ = 0.21 kcal/mol) and Lys533 ($\Delta\Delta G$ = 0.27 kcal/mol). Similar to the case of the R422K system, the I482L mutation makes the polymerase-inhibitor binding energy more

negative. However, also because the -TΔS term becomes more positive (two solvent bridges are formed in this system, see Table 3), the final binding free energy of the I482L mutant system is less negative than that of the WT system. Therefore, the reduction in entropy may be the main reason for the resistance of the I482L mutant system.

In summary, the M423T/V mutation directly attenuates the interactions between GS-9669 and the residue at the mutation site significantly, which mainly causes the conformational change of the N-acyl moiety of GS-9669, resulting the loss of energy contributions of Trp528 and Leu419. The M423I mutation does not affect the interaction between GS-9669 and residues at the mutant site, but it can disturb the original conformation of the N-acyl moiety of GS-9669 through the bulky side chain of Ile423, resulting in the loss of energy contributions of Leu419 and Trp528. Binding with the L419M mutant NS5B polymerase, the orientation of the heterocyclic ether of GS-9669 rotates by almost 180° compared to that in the WT system, which further weakens the binding of the inhibitor with the neighboring residues such as Arg501, Trp528, and Lys531. For the R422K and I482L mutant systems, the decrease in entropy should be the main reason for drug resistance. In addition, the heterocyclic ether moiety introduced to enhance the pharmacokinetic profile is still flexible after binding. It is suggested that further improvement work based on GS-9669 can focus on enhancing the interaction of this moiety with the NS5B polymerase.

3.4. Analyses of binding pathways of the WT and mutant systems by ASMD simulation and PMF calculations

The comparisons and analyses above are based on thermodynamic data. In order to determine whether the drug resistance is related to the kinetic process, we explored the binding processes of GS-9669 with the WT and mutant NS5B polymerases using ASMD simulations and PMF calculations. The results are shown in Fig. 6. It should be noted that the equilibrium distance between GS-9669 and residue 423 of NS5B polymerase in each system was taken as the zero point of the distance.

In the PMF curves, the energy differences between the lowest point (corresponding to the equilibrium state of the system) and the end point (corresponding to the dissociation state of the system, namely the point at 50 Å in the figure) of the systems are -25.02^{wt}, -24.77^{M423T}, -22.64^{M423I}, -20.36^{M423V}, -24.92^{L419M}, -30.94^{R422K}, -33.81^{I482L} kcal/mol, respectively. The energy differences are highly correlated with the binding energy (ΔE_{bind} in Table 1) calculated by MM/GBSA (the correlation coefficient is 0.82). Two calculation methods confirm each other, indicating the rationality of them. Additionally, the energy difference between the highest point and the end point in the PMF curve of
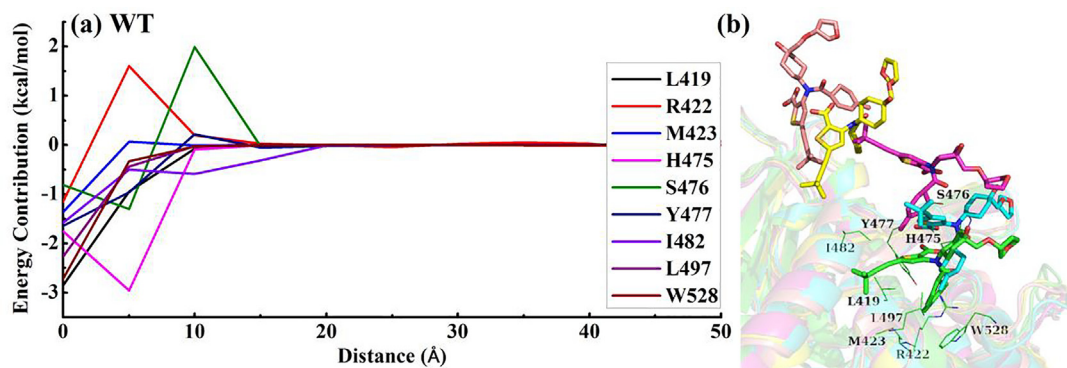

Fig. 7. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the WT system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
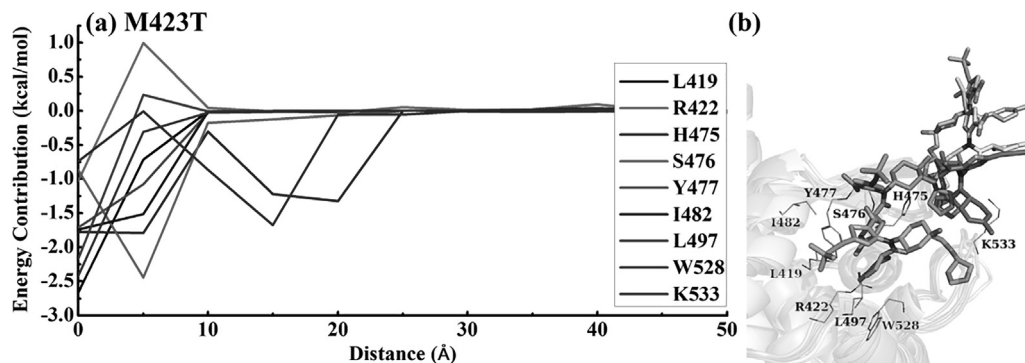

Fig. 8. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the M423T mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

each system is 1.30^{wt}, 2.76^{M423T}, 1.69^{M423I}, 0.27^{M423V}, 0.06^{L419M}, 0.72^{R422K}, 0.12^{I482L} kcal/mol. Except for the WT, M423T, and M423I mutant systems with slightly obvious energy barriers, all other systems have no significant energy barriers. This shows that the binding processes of GS-9669 with WT and mutant NS5B polymerases should be basically controlled thermodynamically, which also indicates that the resistance caused by R422K and I482L mutations should be attributed to the decrease in entropy.

In addition, in order to better understand the details of the interactions between the receptor and ligand during the binding process for each system, the binding energies of the complexes in which the distance from GS-9669 to residue 423 is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å were decomposed, respectively, and the results are shown in Figs. 7a–13a. It can be seen from these figures that the receptor and ligand begin to interact at ~20 Å in each system. According to the principle of micro-reversibility, the binding process has the same pathways as the dissociation process. For convenience of description, the following will be analyzed based on the binding process. By comparing the energy contributions of the residues in the binding process of each system and the binding path of GS-9669, we can find that the binding pathways of GS-9669 with WT and the mutant systems are not the same.

In the process of GS-9669 binding to the WT NS5B polymerase (Fig. 7), first, the *tert*-butyl group of the thiophene moiety of GS-9669 gradually approaches Ile482 at the entrance of the binding

pocket due to hydrophobic interaction (~20 to ~10 Å), then His475 quickly establishes a strong interaction with the thiophene moiety (~10 to ~5 Å). During ~ 15 to ~5 Å, Ser476 at the entrance of the pocket and Arg422 at the bottom of the pocket will successively hinder the binding of GS-9669. The binding energy decomposition results show that at ~ 10 Å, the energy contribution of Ser476 is 1.99 kcal/mol (of which the van der Waals contribution is −1.45 kcal/mol, electrostatic contribution is −0.26 kcal/mol, and polar desolvation energy contribution is as high as 3.7 kcal/mol), and the energy contribution of Arg422 at ~ 5 Å is 1.60 kcal/mol (of which van der Waals contribution is −0.17 kcal/mol, electrostatic contribution is −0.83 kcal/mol, and polar desolvation energy contribution is 2.61 kcal/mol). Therefore, the repelling effect of these two residues on the inhibitor is caused by the desolvation. In addition, during ~ 10 to ~5 Å, the binding strength of GS-9669 with His475 and Ser476 at the entrance of the pocket gradually increases, inducing GS-9669 to adjust the heterocyclic ether to rotate to the crack between Ser476 and Trp528. Eventually, the heterocyclic ether settles in the crack and completes deep bonding. It can be seen that GS-9669 will bind to the WT NS5B polymerase from the direction of Try477 in the thumb II site, and the anchoring and induction of Ile482 and His475 are crucial for the binding of GS-9669.

Unlike the WT system, the heterocyclic ether of GS-9669 will approach the crack between His475 and Lys533 on the binding pocket of the M423T NS5B polymerase (Fig. 8) from the direction
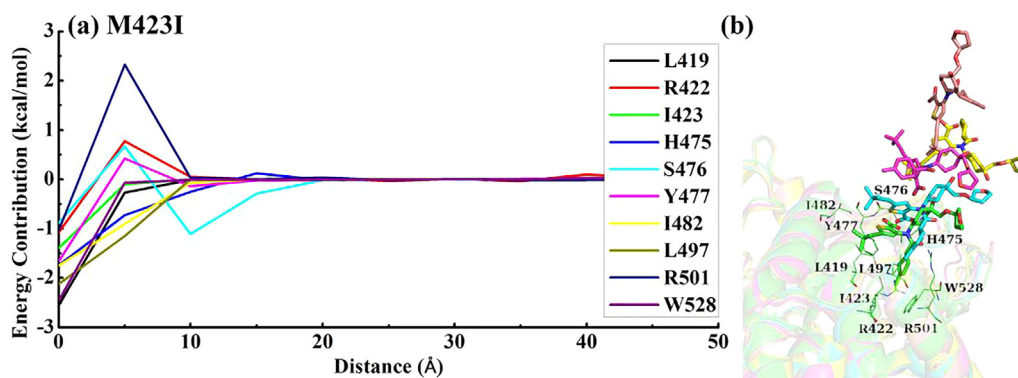

Fig. 9. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the M423I mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
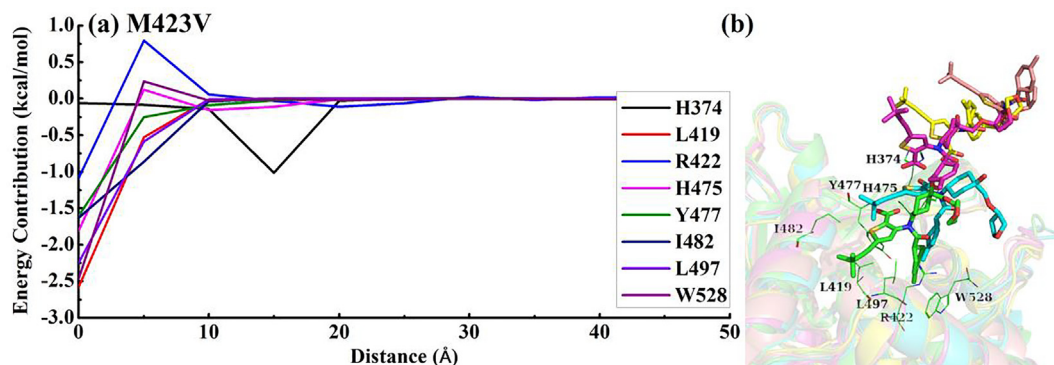

Fig. 10. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the M423V mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Fig. 11. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the L419M mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
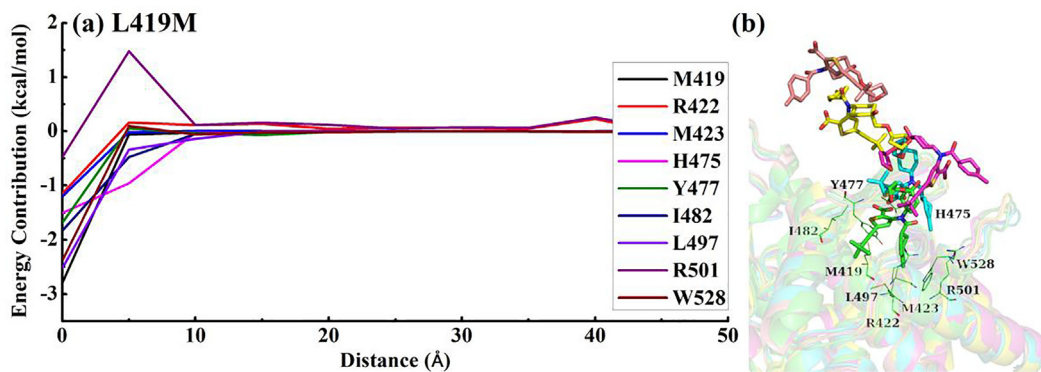

Fig. 12. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the R422K mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
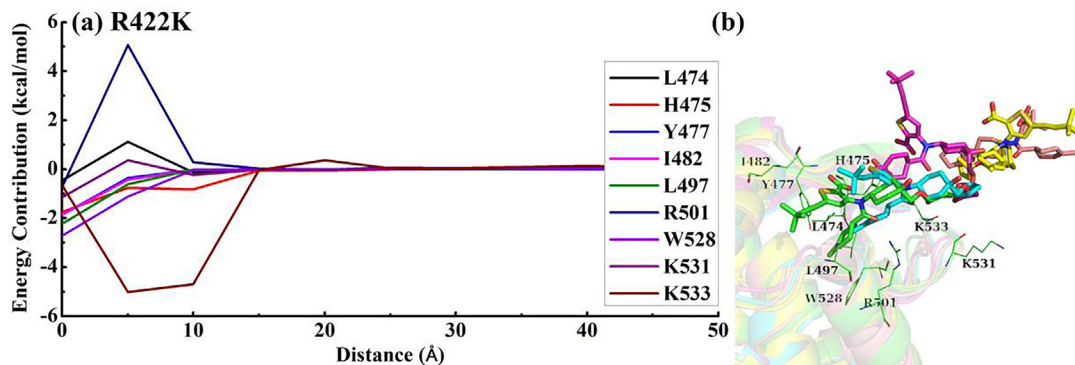

Fig. 13. (a) Binding energy contributions of residues (>1.00 or < −1.00 kcal/mol) in the dissociation processes for the I482L mutant system (the distance of GS-9669 from the equilibrium of the system is 0, 5, 10, 15, 20, 25, 30, 35, 40, 45, and 50 Å, respectively), and (b) the binding path for GS-9669 (the color of the structure at 0, 5, 10, 15, 20 Å is green, cyan, hot pink, yellow, and salmon, respectively). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
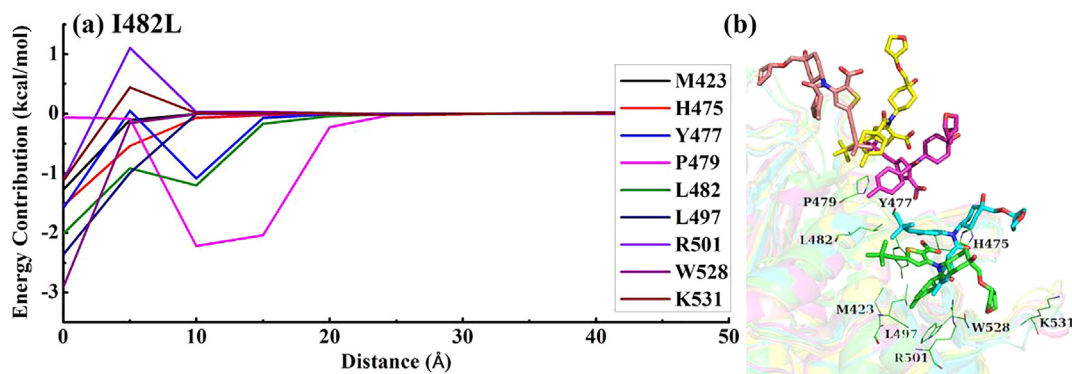
of His475 (~25 to ~20 Å), and then is attracted by Lys533 (~20 to ~10 Å). At this point, the heterocyclic ether of GS-9669 is temporarily fixed in the crack between His475 and Lys533. Then residues such as Ser476, His475, and Ile482 gradually establish interactions with GS-9669, thereby completing the overall binding of the inhi-

bitor (~10 to ~0 Å). During the binding process, Arg422 also has a short-term repulsive effect on GS-9669 (~10 to ~3 Å) due to desolvation (the polar desolvation energy of Arg422 is 1.84 kcal/mol).

For the NS5B^{M423I} polymerase (Fig. 9), GS-9669 firstly approaches Ser476 through the *tert*-butyl group of the thiophene

moiety, and then is induced to bind deeply in a direction almost entirely perpendicular to the entrance plane of the binding pocket. At ∼ 10 to ∼5 Å, GS-9669 needs to overcome the repulsive force generated by Arg422, Ser476, Tyr477, and Arg501 due to polar desolvation (their polar desolvation energies are 1.76, 1.44, 1.49, 5.15 kcal/mol, respectively, see Fig. 9a). This will prevent GS-9669 from binding to the active site, making the NS5B^{M423I} polymerase generate resistance to GS-9669.

In the process of binding to NS5B^{M423V} polymerase (Fig. 10), the thiophene moiety of GS-9669 first establishes interaction with His374 (∼20 to ∼10 Å), and then binds to the binding pocket along the crack between His374 and Trp528. In the process from ∼ 15 to ∼ 5 Å, the affinity of GS-9669 and His374 gradually weakens, and the inhibitor is simultaneously hindered by Arg422, Trp528, and His475 (mainly caused by polar desolvation, and their polar desolvation energies are 1.27, 1.33, 0.81 kcal/mol, respectively), and promoted by residues such as Ile482 and Leu497. The existence of these reverse effects leads to the PMF curve having a local minimum at ∼ 13 Å (Fig. 6d). This may cause the inhibitor to fall into the local energy minimum during the binding process and fail to effectively enter the binding pocket, thereby causing NS5B^{M423V} polymerase to generate resistance to GS-9669.

The binding path of GS-9669 and NS5B^{L419M} polymerase is almost perpendicular to the entrance plane of the active pocket (Fig. 11). In addition to that Arg501 need to overcome the desolvation repulsion during the binding process (the polar desolvation energy of Arg501 is 2.84 kcal/mol), all other key residues promote GS-9669′s binding.

Similar to the case of the M423T mutation system, when GS-9669 binds with NS5B^{R422K} polymerase (∼15 to ∼10 Å), the heterocyclic ether is firstly close to the crack between His475 and Lys533 and establishes stable interactions with these two residues, especially with Lys533 (Fig. 12). Then with the help of Lys533, His475, Trp528 and other residues, the N-acyl moiety of GS-9669 overcomes the electrostatic repulsion and desolvation repulsion of Arg501 (the electrostatic contribution of Arg501 is 0.99 kcal/mol, and the polar desolvation contribution is 5.47 kcal/mol), adjusting its direction and inserting into the binding pocket to complete the binding.

For the NS5B^{I482L} system, the *tert*-butyl group of GS-9669 approaches the binding pocket from where Pro479 is located (Fig. 13), and at the same time establishes obvious interactions with Leu482 and Tyr477 near the entrance of the pocket (∼25 to ∼10 Å). Then the N-acyl moiety of the inhibitor is inserted into the binding pocket with the help of Trp528, Leu497, Leu482, His475 and other residues (∼10 to ∼0 Å). In this process, the inhibitor overcomes the desolvation repulsion of Arg501 and Lys531 (their polar desolvation energies are 1.04 and 0.58 kcal/mol, respectively).

Through the above analyses, it can be found that GS-9669 binds to the thumb II sites of WT and the series of mutant NS5B polymerases in different pathways. For the WT, M423I or I482L mutant NS5B polymerases, GS-9669 starts to bind along the region where Tyr477, Pro479, and Leu482 are located. This region can recognize the *tert*-butyl of GS-9669, and then the inhibitor adjusts its conformation with the help of residues in other regions to complete the binding. For the M423T, M423V, or R422K mutant NS5B polymerases, GS-9669 begins to bind along the crack between His475 and Lys533. This region can recognize both its thiophene moiety (for the M423V mutation) and its heterocyclic ether of its N-alkyl moiety (for M423T and R422K mutation). Then GS-9669 adjusts its conformation by means of residues in other regions to complete the binding. For the L419M mutant NS5B polymerase, GS-9669 completes the binding in a direction almost perpendicular to the entrance plane of the binding pocket. In the above binding processes, GS-9669 will be subjected to the desolvation repulsion of Arg422 or Arg501. Although the recognition method and binding

pathway are different, the kinetic process has little effect on the binding of GS-9669 with WT and mutant NS5B polymerases, and the binding process is basically controlled thermodynamically.

4. Conclusion

In the present study, to clarify the resistance mechanisms of GS-9669 caused by M423T/I/V, L419M, R422K, and I482L mutations of NS5B polymerase (GT1b) and receptor-ligand interactions during the binding process, a serial of computational methods such as molecular dynamics (MD) simulation, binding free energy calculation, binding energy decomposition, and adaptive steered molecular dynamics (ASMD) simulations were performed for the NS5B^{WT}/GS-9669, NS5B^{M423T}/GS-9669, NS5B^{M423I}/GS-9669, NS5B^{M423V}/GS-9669, NS5B^{L419M}/GS-9669, NS5B^{R422K}/GS-9669, and NS5B^{I482L}/GS-9669 complexes. The calculations show that the binding free energies of the M423T, M423I, M423V, L419M, R422K, I482L mutant systems are less negative than that of the WT system, indicating that these mutations will indeed cause NS5B to produce different degrees of resistance to GS-9669.

Further analyses show that the M423T/V mutation attenuates the interactions between GS-9669 and the residue at the mutation site significantly, which disturbs the conformation of the N-acyl moiety of GS-9669, resulting in the loss of the binding affinities of Trp528 and Leu419. The M423I mutation does not affect the interaction between GS-9669 and the mutant site residue, but the steric hindrance of the bulky side chain of Ile423 disturbs the conformation of the N-acyl moiety of GS-9669, resulting in the loss of binding affinities of Leu419 and Trp528. Binding with the L419M mutant NS5B polymerase, the orientation of the heterocyclic ether of GS-9669 is almost rotated by 180° compared to that in the WT system, which further attenuates the binding of the inhibitor to the neighboring residues (such as Arg501, Trp528, and Lys531). For the R422K and I482L mutant systems, the reduction in the entropy of the system caused by the formation of multiple solvent bridges might be the main reason for the drug resistance. Additionally, the heterocyclic ether moiety introduced in order to enhance the pharmacokinetic profile is still flexible after binding. It is suggested that further improvement work based on GS-9669 can focus on enhancing the interaction of this moiety with the NS5B polymerase.

In addition, the PMF curves show that except for the WT and M423T/I mutant systems which have very low energy barriers, there is no obvious energy barrier for the other systems. Further analyses show that GS-9669 binds to the thumb II site of NS5B polymerase in different pathways. Although the recognition methods and binding pathways are distinct, the binding processes of GS-9669 with the WT and mutant NS5B polymerases are basically controlled thermodynamically.

All these results clearly reveal the resistance mechanisms of GS-9669 caused by M423T/I/V, L419M, R422K, and I482L mutations of HCV NS5B polymerase and can provide some valuable clues for further optimization and design of novel NS5B inhibitors.

CRediT authorship contribution statement

Di Han: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization, Project administration. **Huiqun Wang:** Conceptualization, Methodology. **Baerlike Wujieti:** Visualization. **Beibei Zhang:** . **Wei Cui:** Methodology, Software, Formal analysis, Resources. **Bo-Zhen Chen:** Conceptualization, Software, Formal analysis, Resources, Supervision, Project administration, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.04.026.

References

[1] Li DK, Chung RT. Overview of direct-acting antiviral drugs and drug resistance of hepatitis C virus. In: Hepatitis C Virus Protocols. (Springer, 2019) 3–32.

[2] Ganta NM, Rathnakar B, Satyanarayana M, et al. A review on HCV inhibitors: significance of non-structural polyproteins. Eur J Med Chem 2018.

[3] Pradat P, Trepo C. HCV: epidemiology, modes of transmission and prevention of spread. Best Pract Res Clin Gastroenterol 2000;14(2):201–10.

[4] Yeung LT, King SM, Roberts EA. Mother-to-infant transmission of hepatitis C virus. Hepatology 2001;34(2):223–9.

[5] Crawford JJ, Hoeflich KP, Rudolph J. p21-Activated kinase inhibitors: a patent review. 22(3), 293-310).

[6] Alqahtani SA, Sulkowski MS. The Role of Interferon for the Treatment of Chronic Hepatitis C Virus Infection. (2019).

[7] Wang G, Ma J, Jiang L-J et al. Discovery and Development of the Next-Generation HCV NS3 Protease Inhibitor Glecaprevir. (2019).

[8] Nayak A, Pattabiraman N, Fadra N, Goldman R, Kosakovsky Pond SL, Mazumder R. Structure–function analysis of hepatitis C virus envelope glycoproteins E1 and E2. J Biomol Struct Dyn 2015;33(8):1682–94.

[9] De Clercq E. The race for interferon-free HCV therapies: a snapshot by the spring of 2012. Rev Med Virol 2012;22(6):392–411.

[10] Klibanov OM, Vickery SB, Olin JL, Smith LS, Boceprevir WSH. A Novel NS 3/4 Protease Inhibitor for the Treatment of Hepatitis C. Pharmacotherapy. J Hum Pharmacol Drug Ther 2012;32(2):173–90.

[11] Matthews SJ, Lancaster JW. Telaprevir: a hepatitis C NS3/4A protease inhibitor. Clin Ther 2012;34(9):1857–82.

[12] Kanda T, Yokosuka O, Omata M. Treatment of hepatitis C virus infection in the future. Clin Transl Med 2013;2(1):9.

[13] Kwong AD, Perni RB, Graham CS. Development and marketing of INCIVEK (Telaprevir; VX-950): a first-generation HCV protease inhibitor. Combin PEGylated Interferon Ribavirin 2019.

[14] Bartenschlager R, Lohmann V, Penin F. The molecular and structural basis of advanced antiviral therapy for hepatitis C virus infection. Nat Rev Microbiol 2013;11(7):482–96.

[15] Smith DB, Bukh J, Kuiken C, et al. Expanded classification of hepatitis C virus into 7 genotypes and 67 subtypes: updated criteria and genotype assignment web resource. Hepatology 2014;59(1):318–27.

[16] Sorbo MC, Cento V, Di Maio VC, et al. Hepatitis C virus drug resistance associated substitutions and their clinical relevance: Update 2018. Drug Resist Updates 2018;37:17–39.

[17] Venkatesan A. Review on chemogenomic approaches towards hepatitis C viral targets. J Cell Biochem 2019.

[18] Kumthip K, Maneekarn N. The role of HCV proteins on treatment outcomes. Virol J 2015;12(1):217.

[19] Watkins WJ. Evolution of HCV NS5B Non-nucleoside Inhibitors. (2019).

[20] Houghton M. Discovery of the hepatitis C virus. Liver Int 2009;29:82–8.

[21] Eltahla AA, Luciani F, White PA, Lloyd AR, Bull RA. Inhibitors of the hepatitis C virus polymerase; mode of action and resistance. Viruses 2015;7(10):5206–24.

[22] Bressanelli S, Tomei L, Roussel A, et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Proc Natl Acad Sci 1999;96(23):13034–9.

[23] Lesburg CA, Cable MB, Ferrari E, Hong Z, Mannarino AF, Weber PC. Crystal structure of the RNA-dependent RNA polymerase from hepatitis C virus reveals a fully encircled active site. Nat Struct Mol Biol 1999;6(10):937.

[24] Ago H, Adachi T, Yoshida A, et al. Crystal structure of the RNA-dependent RNA polymerase of hepatitis C virus. Structure 1999;7(11):1417–26.

[25] Soriano V, Vispo E, de Mendoza C, et al. Hepatitis C therapy with HCV NS5B polymerase inhibitors. Expert Opin Pharmacother 2013;14(9):1161–70.

[26] Fenaux M, Eng S, Leavitt SA, et al. Preclinical characterization of GS-9669, a thumb site II inhibitor of the hepatitis C virus NS5B polymerase. Antimicrob Agents Chemother 2013;57(2):804–10.

[27] Lazerwith SE, Lew W, Zhang J, et al. Discovery of GS-9669, a thumb site II non-nucleoside inhibitor of NS5B for the treatment of genotype 1 chronic hepatitis C infection. J Med Chem 2013;57(5):1893–901.

[28] Götte M, Feld JJ. Direct-acting antiviral agents for hepatitis C: structural and mechanistic insights. Nat Rev Gastroenterol Hepatol 2016;13(6):338.

[29] Dvory-Sobol H, Voitenleitner C, Mabery E, et al. Clinical and in vitro resistance to GS-9669, a thumb site II nonnucleoside inhibitor of the hepatitis C virus NS5B polymerase. Antimicrob Agents Chemother 2014;58(11):6599–606.

[30] International T. Sybyl-X 2.0. (Ed.^(Eds) (Tripos International St. Louis, MO, 2012).

[31] Bayly CI, Cieplak P, Cornell W, Kollman PA. A well-behaved electrostatic potential based method using charge restraints for deriving atomic charges: the RESP model. J Phys Chem 1993;97(40):10269–80.

[32] Frisch M, Trucks G, Schlegel H, et al. Gaussian 09, rev. Wallingford: Gaussian Inc; 2009.

[33] Case DA, Betz R, Cerutti D et al. AMBER 2016. University of California, San Francisco, 810 (2016).

[34] Maier JA, Martinez C, Kasavajhala K, Wickstrom L, Hauser KE, Simmerling C. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. J Chem Theory Comput 2015;11(8):3696–713.

[35] Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA. Development and testing of a general amber force field. J Comput Chem 2004;25(9):1157–74.

[36] Jorgensen WL, Chandrasekhar J, Madura JD, Impey RW, Klein ML. Comparison of simple potential functions for simulating liquid water. J Chem Phys 1983;79(2):926–35.

[37] Ryckaert J-P, Ciccotti G, Berendsen HJ. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J Comput Phys 1977;23(3):327–41.

[38] Darden T, York D, Pedersen L. Particle mesh Ewald: an N· log (N) method for Ewald sums in large systems. J Chem Phys 1993;98(12):10089–92.

[39] Kollman PA, Massova I, Reyes C, et al. Calculating structures and free energies of complex molecules: combining molecular mechanics and continuum models. Acc Chem Res 2000;33(12):889–97.

[40] Onufriev A, Bashford D, Case DA. Modification of the generalized Born model suitable for macromolecules. J Phys Chem B 2000;104(15):3712–20.

[41] Wang E, Sun H, Wang J, et al. End-point binding free energy calculation with MM/PBSA and MM/GBSA: strategies and applications in drug design. Chem Rev 2019;119(16):9478–508.

[42] Wang E, Weng G, Sun H, et al. Assessing the performance of the MM/PBSA and MM/GBSA methods. 10. Impacts of enhanced sampling and variable dielectric model on protein–protein interactions. PCCP 2019;21(35):18958–69.

[43] Weng G, Wang E, Wang Z, et al. HawkDock: a web server to predict and analyze the protein–protein complex based on computational docking and MM/GBSA. Nucleic Acids Res 2019;47(W1):W322–30.

[44] Sun H, Duan L, Chen F, et al. Assessing the performance of MM/PBSA and MM/GBSA methods. 7. Entropy effects on the performance of end-point binding free energy calculation approaches. PCCP 2018;20(21):14450–60.

[45] Weiser J, Shenkin PS, Still WC. Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). J Comput Chem 1999;20(2):217–30.

[46] Duan L, Liu X, Zhang JZ. Interaction entropy: a new paradigm for highly efficient and reliable computation of protein–ligand binding free energy. J Am Chem Soc 2016;138(17):5722–8.

[47] Onufriev A, Bashford D, Case DA. Exploring protein native states and large-scale conformational changes with a modified generalized born model. Proteins Struct Funct Bioinf 2004;55(2):383–94.

[48] Gohlke H, Kiel C, Case DA. 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 2003;330(4):891–913.

[49] Ozer G, Quirk S, Hernandez R. Adaptive steered molecular dynamics: Validation of the selection criterion and benchmarking energetics in vacuum. J Chem Phys 2012;136(21):06B606.

[50] Ozer G, Keyes T, Quirk S, Hernandez R. Multiple branched adaptive steered molecular dynamics. J Chem Phys 2014;141(6):08B606_601.

[51] Bureau HR, Merz Jr DR, Hershkovits E, Quirk S, Hernandez R. Constrained unfolding of a helical peptide: Implicit versus explicit solvents. PLoS ONE 2015;10(5).

[52] Bureau HR, Hershkovits E, Quirk S, Hernandez R. Determining the energetics of small β-sheet peptides using adaptive steered molecular dynamics. J Chem Theory Comput 2016;12(4):2028–37.

[53] Jarzynski C. Nonequilibrium equality for free energy differences. Phys Rev Lett 1997;78(14):2690.

[54] DeLano WL. The PyMOL molecular graphics system. http://www.pymol.org, (2002).

[55] Wang H, Guo C, Chen B-Z, Ji M. Computational study on the drug resistance mechanism of HCV NS5B RNA-dependent RNA polymerase mutants V494I, V494A, M426A, and M423T to Filibuvir. Antiviral Res 2015;113:79–92.

[56] Wang H, Cui W, Guo C, Chen B-Z, Ji M. Molecular modeling study on the drug resistance mechanism of NS5B polymerase to TMC647055. Biochem Cell Biol 2016;94(2):147–58.

[57] Xue W, Jiao P, Liu H, Yao X. Molecular modeling and residue interaction network studies on the mechanism of binding and resistance of the HCV NS5B polymerase mutants to VX-222 and ANA598. Antiviral Res 2014;104:40–51.

[58] Case DA, Betz R, Cerutti D et al. AMBER 2016 reference manual. University of California: San Francisco, CA, USA, 1-923 (2016).