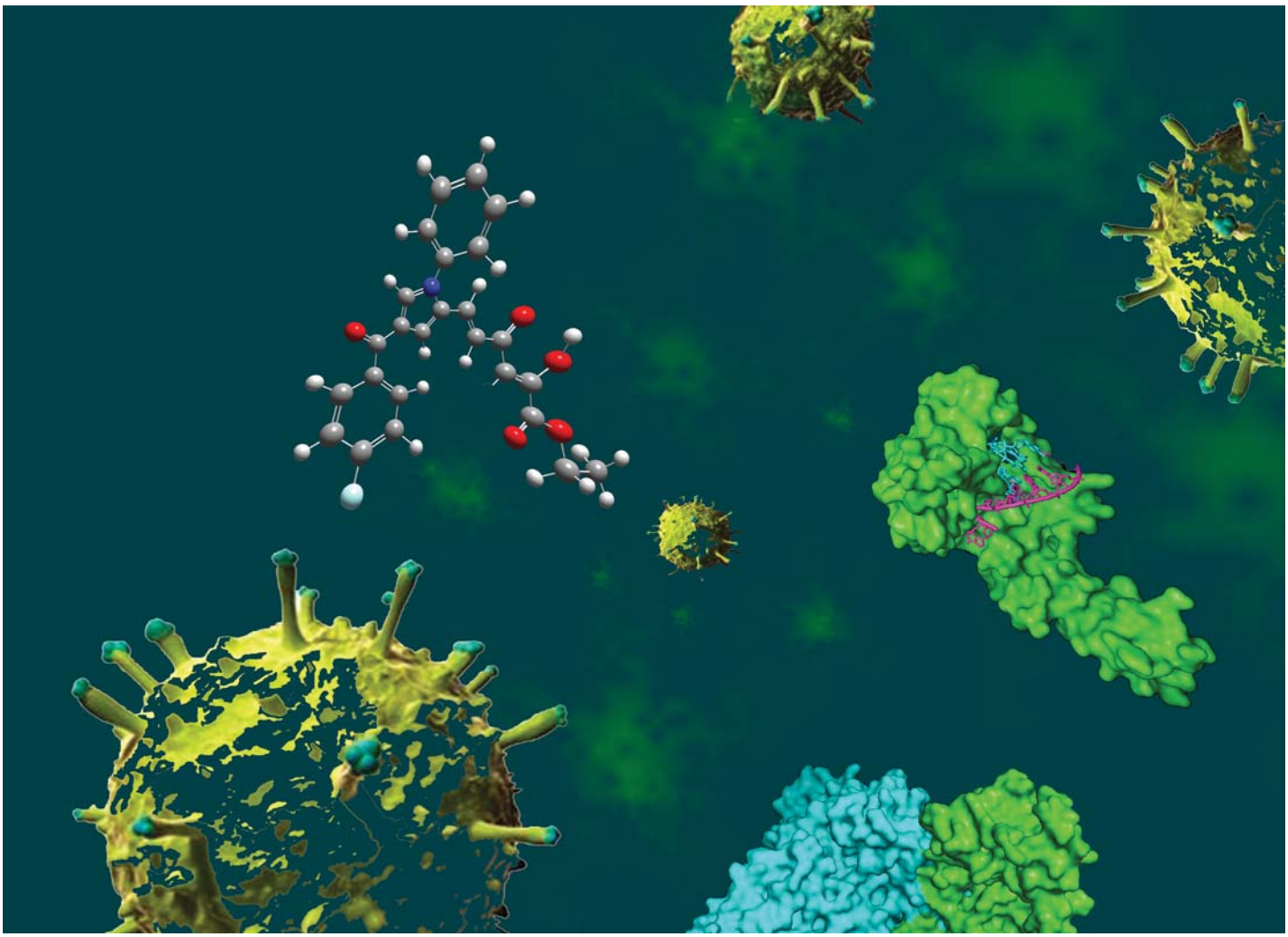

Showcasing research from the Group of Assoc. Prof. Weiwei Xue and Prof. Feng Zhu at Chongqing University and Zhejiang University, China

Prediction of the binding mode and resistance profile for a dual-target pyrrolyl diketo acid scaffold against HIV-1 integrase and reverse-transcriptase-associated ribonuclease H

The Innovative Drug Research and Bioinformatics (IDRB) group is working on the binding mechanism identification and structure-based design of multi-target drugs for complex diseases, such as mood disorders and HIV infection. Multi-target drugs have the benefits of rapid onset and/or higher efficacy for complex disease treatment. This computational work predicts the binding mode and potential resistance profile of the dual-target pyrrolyl diketo acid scaffold against HIV-1 integrase (IN) and the reverse-transcriptase-associated ribonuclease H (RNase H) active sites.

As featured in:

See Weiwei Xue, Feng Zhu *et al.*, *Phys. Chem. Chem. Phys.*, 2018, **20**, 23873.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/pccp

Registered charity number: 207890

PCCP

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2018, **20**, 23873

Prediction of the binding mode and resistance profile for a dual-target pyrrolyl diketo acid scaffold against HIV-1 integrase and reverse-transcriptase-associated ribonuclease H†

Fengyuan Yang,[ab] Guoxun Zheng,[ab] Tingting Fu,[ab] Xiaofeng Li,[ab] Gao Tu,[ab] Ying Hong Li,[ab] Xiaojun Yao,[c] Weiwei Xue *[a] and Feng Zhu *[ab]

The rapid emergence of drug-resistant variants is one of the most common causes of highly active antiretroviral therapeutic (HAART) failure in patients infected with HIV-1. Compared with the existing HAART, the recently developed pyrrolyl diketo acid scaffold targeting both HIV-1 integrase (IN) and reverse transcriptase-associated ribonuclease H (RNase H) is an efficient approach to counteract the failure of anti-HIV treatment due to drug resistance. However, the binding mode and potential resistance profile of these inhibitors with important mechanistic principles remain poorly understood. To address this issue, an integrated computational method was employed to investigate the binding mode of inhibitor JMC6F with HIV-1 IN and RNase H. By using per-residue binding free energy decomposition analysis, the following residues: Asp64, Thr66, Leu68, Asp116, Tyr143, Gln148 and Glu152 in IN, Asp443, Glu478, Trp536, Lys541 and Asp549 in RNase H were identified as key residues for JMC6F binding. And then computational alanine scanning was carried to further verify the key residues. Moreover, the resistance profile of the currently known major mutations in HIV-1 IN and 2 mutations in RNase H against JMC6F was predicted by *in silico* mutagenesis studies. The results demonstrated that only three mutations in HIV-1 IN (Y143C, Q148R and N155H) and two mutations in HIV-1 RNase H (Y501R and Y501W) resulted in a reduction of JMC6F potency, thus indicating their potential role in providing resistance to JMC6F. These data provided important insights into the binding mode and resistance profile of the inhibitors with a pyrrolyl diketo acid scaffold in HIV-1 IN and RNase H, which would be helpful for the development of more effective dual HIV-1 IN and RNase H inhibitors.

Received 22nd March 2018,
Accepted 12th June 2018

DOI: 10.1039/c8cp01843j

rsc.li/pccp

Introduction

Human immunodeficiency virus 1 (HIV-1) remains one of the most devastating health problems all over the world, which leads to 70 million infections and 36 million deaths.[1–3] Current therapeutic regimens for HIV-1 infection are primarily based on highly active antiretroviral therapy (HAART).[4–6] This therapy usually combines multiple drugs aiming at various targets such as HIV-1 protease, reverse transcriptase (RT) and integrase (IN) as well as the host chemokine receptor (CCR5).[7–9] In particular, both RT and IN play essential roles in viral replication,[10–12] and three integrase strand transfer inhibitors (INSTI; raltegravir (RAL), elvitegravir (EVG) and dolutegravir (DTG), as shown in Fig. 1) and a dozen RT inhibitors (RTIs) have been approved by the U.S. Food and Drug Administration (FDA).[13–15] In spite of the clinical success of HAART, several factors (extensive drug resistance,[16–18] severe adverse effects,[19] limited number of efficacious combinations,[20] drug–drug interactions[21] and the requirement for lifelong patient compliance[22]) have seriously affected its efficacy. Therefore, the discovery of anti-HIV agents of novel mechanisms other than drug combination is urgently needed.[23–25]

HIV-1 IN consists of a highly conserved DDE motif (Fig. 2A) that chelates with two Mg^{2+} ions in its catalytic core domain (CCD).[26] Since the catalytic site of the ribonuclease H domain (RNase H) in HIV-1 RT shared a very similar topology with that of the CCD in IN(Fig. 2B),[11] these 2 proteins are reported as a potential pair for discovering multi-target drugs.[27–30]

[a] *Innovative Drug Research and Bioinformatics Group, School of Pharmaceutical Sciences and Collaborative Innovation Center for Brain Science, Chongqing University, Chongqing 401331, China. E-mail: xueww@cqu.edu.cn*
[b] *Innovative Drug Research and Bioinformatics Group, College of Pharmaceutical Sciences, Zhejiang University, Hangzhou, Zhejiang 310058, China. E-mail: zhufeng@zju.edu.cn, zhufeng.ns@gmail.com*
[c] *State Key Laboratory of Applied Organic Chemistry and Department of Chemistry, Lanzhou University, Lanzhou 730000, China*

† Electronic supplementary information (ESI) available. See DOI: 10.1039/c8cp01843j

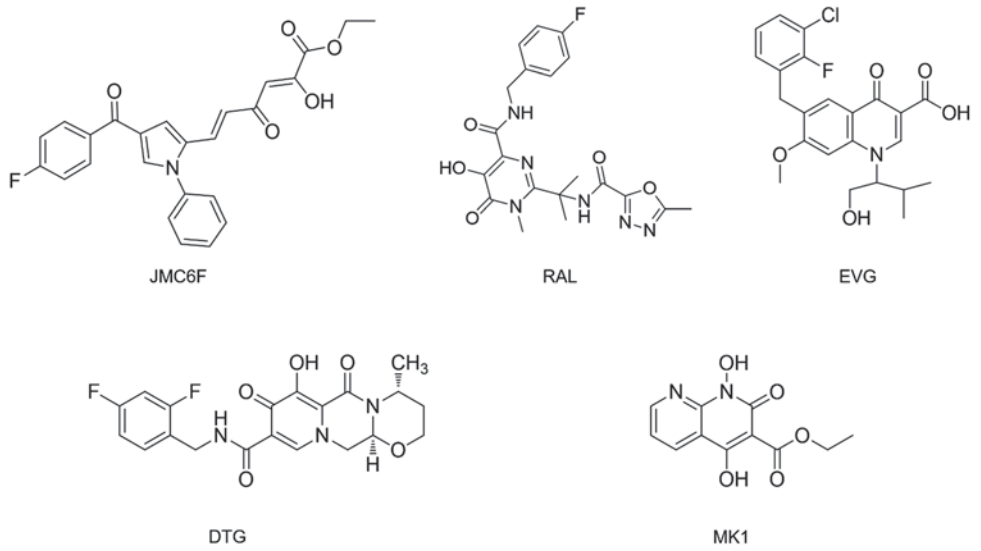

Fig. 1 Chemical structures of inhibitors studied in this work.

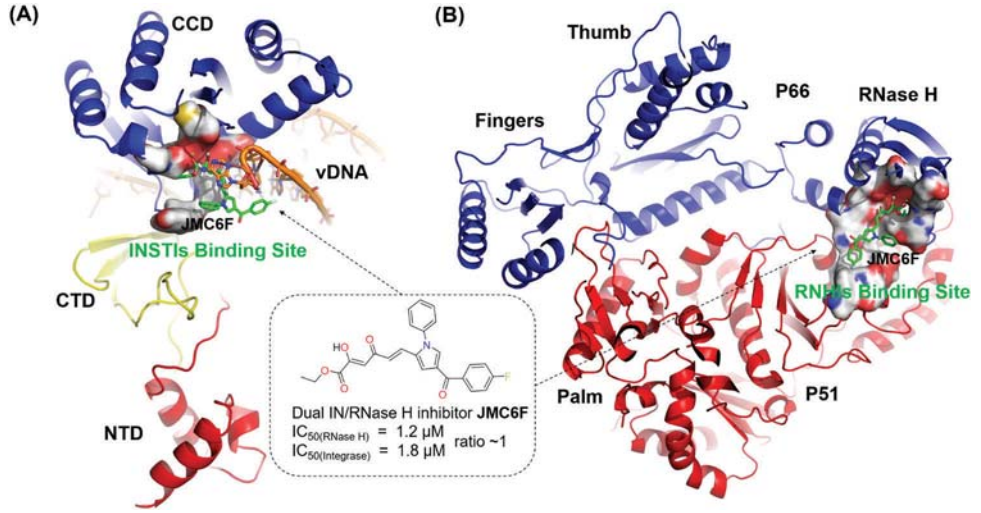

Fig. 2 Schematic representation of active sites of IN and RNase H binding with JMC6F. (A) HIV-1 IN/vDNA model. (B) X-ray crystal structure of HIV-1 reverse transcriptase. The proteins are represented as red and blue cartoons and vDNA is displayed as orange sticks. The INSTI and RHI binding sites are shown at the surface and the JMC6F is represented as green sticks.

The design of dual-action drugs with an anti-IN/RNase H pharmacological profile has therefore emerged as a promising strategy for treating HIV-1 infection.[31–35] Different from the traditional approved RTIs,[36,37] this new strategy blocks the catalytic activity of RT by inhibiting the RNase H domain (RHI),[38] which is beneficial for suppressing viral resistance, reducing toxic side effects and ameliorating drug–drug interactions.[39–41] 4 or more molecular scaffolds based on this strategy have been discovered, including hydroxytropolone,[42] 2-hydroxyisoquinoline-1,3(2*H*,4*H*)-dione,[43] madurahydroxylactone[44] and diketo derivatives.[38] Pyrrolyl diketo acid derivatives showed good effectiveness against the RAL resistant HIV-1 strain[35] and the most recently discovered dual-action inhibitor (JMC6F in Fig. 2), a pyrrolyl diketo acid derivative, showed good correlation between IN and RNase H inhibition.[45] Because JMC6F is reported to be one of the best dual inhibitors of this class,[38,45] it can be selected as a representative of drugs with an anti-IN/RNase H pharmacological profile.

However, the molecular basis of (1) why the pyrrolyl diketo acid derivatives can simultaneously bind to IN and RNase H and (2) how these currently available drug-resistant mutations for FDA approved INSTIs and RTIs affect the affinity of this dual-action scaffold remains elusive. Molecular docking has provided some information on the interaction between the diketo derivatives and their corresponding target,[35,46] but more detailed information such as dynamic and quantitative identification of the interactions between key residues and drugs is still not available. This information is especially valuable when considering the fact that the mutation of the key residues is the main reason leading to the resistance of anti-HIV drugs.[47]

With the success of computational methods, such as molecular dynamics (MD) simulation and the binding free energy calculation tool, in structure-based drug design,[48–53] it is of great interest to evaluate the application of these methods in order to address the elusive molecular basis discussed above. The MD simulations can provide plentiful dynamic structural information about protein–ligand interactions.[53–55] Based on MD trajectories, several approaches can be applied to predict binding free energies, including free energy perturbation (FEP), thermodynamic integration (TI), and molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) methods.[56] Compared to FEP or TI, due to the good balance between the speed and accuracy of the MM/PBSA method it has been widely applied in various protein–ligand systems,[57] including the prediction of the binding mode and mutation-induced drug resistance profile of antiviral drugs to their targets (like influenza neuraminidase,[58] HCV RNA polymerase[59] and HIV protease[60]). Herein, an integrated computational strategy was therefore used to identify the binding mode of JMC6F in HIV-1 IN and RT-associated RNase H at the atomic level. First, a homology model of full-length HIV-1 IN was constructed based on the crystal structure of the IN from the prototype foamy virus (PFV).[61] Then, molecular docking and MD simulations were employed to investigate the binding of JMC6F in HIV-1 IN and RNase H. Third, the simulation models constructed in this study were fully validated from multiple perspectives. Finally, *in silico* site-directed mutagenesis and binding free energy calculations were applied to assess the resistance profile of JMC6F based on the known resistance mutations in HIV IN or RNase H. In summary, this study revealed the resistance profile of the most promising dual inhibitor of HIV IN and RNase H, which could facilitate the discovery of new anti-HIV drugs of this class with an improved resistance profile.

Materials and methods

Homology modeling of HIV-1 IN and construction of the HIV-1 intasome complex

Homology modeling of the HIV-1 IN full-length structure was performed using *Prime* in Maestro. The sequence of HIV-1 IN was extracted from the GenBank (accession number: AAC83551.1). The crystal structure of PFV IN (PDB ID: 3OY9[61]) was used as a model template. Before the construction of the homology model, sequence alignment of HIV/PFV IN and viral U5 DNA was done and adjusted manually to obtain reasonable matching. Then, the stereochemical quality of the models was evaluated using the Ramachandran plot of *PROCHECK*.[62] Compared with the free HIV-1 IN, INSTIs preferentially bind to the intasome complex (IN, vDNA and Mg^{2+}).[63] The construction of the HIV-1 intasome complex was done using PyMOL. First, the 19-base-pair mimic of the pre-processed terminal vDNA from the structure of PFV IN (PDB ID: 3OY9[61]) and the Mg^{2+} ions from the inhibited structure of PFV IN (PDB ID: 3S3M[64]) were introduced into the HIV-1 IN model. Then, vDNA from the structure of the PFV intasome was mutated based on the sequence alignment of vDNA between HIV and PFV.[61] In addition, three FDA approved INSTIs (RAL, EVG and DTG) from the inhibited structures of PFV IN (PDB ID: 3OYA,[65] 3L2W[61] and 3S3M[64]) were fitted into the HIV-1 intasome.

Molecular Docking of JMC6F into HIV-1 IN and RNase H

Prior to docking JMC6F into HIV-1 IN and RNase H, cross-docking was performed on the co-crystal structures of PFV IN in complex with RAL and DTG (PDB ID: 3OYA[65] and 3S3M[64]) and HIV-1 RNase H in complex with MK1 and MK2 (PDB ID: 3LP0[32] and 3LP1[32]) by Glide with standard precision (SP). In order to imitate the real biological environment of inhibition, Mg^{2+} ions were substituted for Mn^{2+} in the reported co-crystal structures (PDB ID: 3LP0 and 3LP1) in this study. In the process of cross-docking, the structures of RAL, DTG, MK1 and MK2 were preprocessed by *LigPrep* using *OPLS-2005* force field,[66] and the ionized state was assigned by *Epik* at a pH value of 7.0 ± 2.0. *OPLS-2005* is a popular force field used to perform molecular docking, and was successfully adopted previously for revealing the binding mechanisms between HIV protein targets (IN or RT) and their corresponding inhibitors.[35,67,68] Moreover, the *OPLS3 force field*[69] has emerged as the most recent force field developed by Schrödinger with significantly enhanced performance. In this study, to be consistent with previous studies,[35,67,68] *OPLS-2005 force field* is applied. Co-crystal structures were prepared by adding hydrogen atoms, assigning partial charges and protonation states, and minimizing the structure using the *Protein Preparation Wizard* module in *Maestro*. And the docking grid box was defined by centering on the corresponding ligand in each complex using the *Receptor Grid Generation* module in *Maestro*. In order to ensure that there are no alternative binding sites with higher (similar) affinity, additional calculations were performed using the SiteMap code included by Schrödinger and the results are shown in Table S1 (ESI†). As shown in Table S1 (ESI†), seven potential binding sites were found and JMC6F more favorably binds at S1. Therefore, the S1 site was found with the highest docking performance which corresponds to the site we chose in this study. In molecular docking, 300 poses were generated during the initial phase of the docking procedure and all the poses were chosen for energy minimization by 100 steps of conjugate gradient minimization. Fig. S1 (ESI†) illustrated the superposition of the cross-docking poses of RAL, DTG, MK1 and MK2 onto their corresponding co-crystallized poses. The RMSD values between the docking pose and crystal pose of DTG, RAL, MK1 and MK2 were calculated (2.71, 3.40, 1.63 and 1.14, respectively). As shown, the RMSD values of RAL and DTG were higher than those of MK1 and MK2. This was mainly due to the flexible linker of RAL and DTG. But the docking pose at the catalytic site was quite similar. These results indicated that the Glide SP docking procedure is able to produce poses that are in close agreement with the crystal structures.

After validation of the docking procedure, the same parameter settings in cross-docking were applied to dock the prepared JMC6F into the INSTI binding site of IN and the RHI binding site of RT. The 3D structure of JMC6F was constructed with the

Table 1 The wide type and corresponding mutant simulation systems in this study

Systems	Wide type	Mutations
IN-JMC6F	JMC6F docked into the HIV-1 IN model	T66A/I/K, E92Q, E138A/K, G140C/S, Y143C/H/R, S147G, Q148H/K/R, N155H
IN-DTG	DTG in the crystal structure (PDB ID: 3S3M) fitted into the HIV-1 IN model	T66A/I/K, E92Q, E138A/K, G140C/S, Y143C/H/R, S147G, Q148H/K/R, N155H
IN-RAL	RAL in the crystal structure (PDB ID: 3OYA) fitted into the HIV-1 IN model	T66A/I/K, E92Q, E138A/K, G140C/S, Y143C/H/R, S147G, Q148H/K/R, N155H
IN-EVG	EVG in the crystal structure (PDB ID: 3L2W) fitted into the HIV-1 IN model	T66A/I/K, E92Q, E138A/K, G140C/S, Y143C/H/R, S147G, Q148H/K/R, N155H
RNase H-JMC6F	JMC6F docked into the crystal structure (PDB ID: 3LP0)	Y501R, Y501W
RNase H-MK1	Crystal structure (PDB ID: 3LP0)	ND[a]

[a] ND, not determined in this study.

building tool in *Maestro*. Docking grid boxes in IN and RNase H were defined by centering on DTG and MK1, respectively. The docking poses with the highest docking score were chosen for further MD simulation.

MD simulation

System setup. Six wild type (WT) and sixty-six mutant type systems were constructed for MD simulation (Table 1). Before the simulation, each system was prepared using the LEaP module of *AMBER14 package.* The AMBER force field ff14SB[70] was applied for describing proteins, vDNA as well as Mg^{2+} ions. The force field parameters for optimizing ligands were created using the *Antechamber program*,[71] using the *General AMBER Force Field*[72] and *restrained electrostatic potential* partial charge.[73] The geometric optimization and the electrostatic potential calculation of ligands were performed at the HF/6-31G* level of the Gaussian09 suite. Finally, each system was immersed into a rectangular periodic box of pre-equilibrated TIP3P water with at least 10 Å distance around the complexes. Finally, appropriate numbers of sodium counter ions were added to maintain the electroneutrality of the simulation system.

Equilibration and production runs. All MD simulations were performed using GPU-accelerated PMEMD integrated in *AMBER14 package.* For each simulation, a sophisticated protocol (minimization, heating, and equilibration) was followed. Initially, the prepared system was subjected to energy minimization through 2 steps. The first step is to apply harmonic restraints with a force constant of 10 kcal $(\text{mol Å}^2)^{-1}$ to solute atoms, and the second step is to allow all atoms to move freely. In each step, energy minimization was performed by the steepest descent method for the first 3000 steps and by the conjugated gradient method for the subsequent 2000 steps. After energy minimization, a 50 ps *NVT* ensemble heating process was applied to gradually increase the temperature from 0 to 300 K and the density of the system to 1 g cm^{-3} in this step. Then, to equilibrate the system, three MD equilibrations of 50 ps at 300 K were performed with decreased restraint weights (5.0, 1.0 and 0.1 kcal mol^{-1} Å^{-2}). Finally, a 50 ns production run in the *NPT* ensemble was carried out without any restraint on the simulation system. During MD simulations, periodic boundary conditions were employed and the direct space interaction was calculated by using the particle-mesh Ewald (PME) method[74] with a long range electrostatic interaction (cutoff = 12.0 Å). All bonds involving hydrogen atoms were constrained with the SHAKE algorithm[75] allowing an integration time step of 2 fs.

Thermodynamic analysis

Binding free energy calculation. The mm_pbsa.pl script was used to calculate the binding free energy (ΔG_{calc}) by molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) methods.[76–79] The free energy perturbation (FEP) method[80–82] could also be applied to perform qualitative predictions in this analysis. To be consistent with previous publications,[83,84] MM/PBSA methods were adopted for analysis in this study. For each simulation system, a total of 500 snapshots were taken from the last equilibrium trajectory with equal intervals. And for each snapshot, ΔG_{calc} was calculated by:

$$\Delta G_{\text{calc}} = \Delta E_{\text{vdW}} + \Delta E_{\text{ele}} + \Delta G_{\text{PB}} + \Delta G_{\text{SA}} \quad (1)$$

where ΔE_{vdW} and ΔE_{ele} represent van der Waals and electrostatic energy changes in the gas phase. ΔG_{PB} and ΔG_{SA} represent polar and non-polar solvent interaction energy changes. ΔG_{PB} is calculated as the description of the PB equation, with dielectric constants for the solute and the solvent set to 1 and 80,[85] respectively. While ΔG_{SA} is estimated by the solvent-accessible surface area (SASA) determined by a water probe radius of 1.4 Å and the surface tension constant γ was set to 0.0072 kcal $(\text{mol Å}^{-1})^{-1}$.[86]

Per-residue binding free energy decomposition. To quantitatively evaluate the contribution of each residue for drug's binding, ΔG_{calc} was decomposed at a per-residue basis, including the van der Waals term ($\Delta E_{\text{vdW}}^{\text{per-residue}}$), the electrostatic term ($\Delta E_{\text{ele}}^{\text{per-residue}}$), the polar term ($\Delta G_{\text{PB}}^{\text{per-residue}}$) and the nonpolar term ($\Delta G_{\text{SA}}^{\text{per-residue}}$) for ligands and each residue, as shown in eqn (2):

$$\Delta G_{\text{calc}}^{\text{per-residue}} = \Delta E_{\text{vdW}}^{\text{per-residue}} + \Delta E_{\text{ele}}^{\text{per-residue}} + \Delta G_{\text{PB}}^{\text{per-residue}} + \Delta G_{\text{SA}}^{\text{per-residue}} \quad (2)$$

Computational alanine scanning (CAS). CAS[87] was performed to verify the key residues of IN and RNase H contributing significantly to JMC6F ($\geq$0.5 kcal mol^{-1}). The whole process includes the generation of mutated snapshots and the calculation of the binding free energy difference between the WT and mutant (MUT) complex. Firstly, 500 snapshots were collected from the last 20 ns WT trajectory. The alanine

mutation was generated by truncating the selected mutation residue at Cγ and by replacing Cγ with a hydrogen atom at a distance of 1.09 Å from Cβ along the direction of the Cγ–Cβ bond.[87] And the topology files with alanine mutations were regenerated by the LEaP module in *AMBER14* and the MM/PBSA method[76–78] was used to calculate the relative binding free energy ($\Delta\Delta G_{CAS}$) defined by the difference between WT and MUT complexes shown as below:

$$\Delta\Delta G_{CAS} = \Delta G_{MUT} - \Delta G_{WT} \quad (3)$$

where ΔG_{WT} and ΔG_{MUT} refer to the binding free energy of the WT and MUT complexes, respectively.

In silico site-directed mutagenesis

Collecting the reported drug resistant mutations in HIV-1 IN and RNase H. A total of 16 IN and 2 RNase H mutations were collected from the *Stanford HIV Drug Resistance Database*[88] and the experiment previously reported by Arion *et al.*[89] to validate the MD simulation models was used in this study. The resistance profile of JMC6F was further analyzed based on these previously reported resistance mutations.

In silico site-directed mutagenesis was performed to validate the accuracy of the constructed model and to predict the drug resistance profile of JMC6F. Based on the representative structure extracted from the equilibrated WT trajectories, a total of 48 mutant complexes (Table 1) were generated using PyMOL and prepared by LEaP in *AMBER14*. For each system, the MD simulation (10 ns) and binding free energy calculations were performed following the same way as described in previous sections.

Results and discussion

Modeled structure of the full-length HIV-1 IN

The resolving of HIV-1 intasome complexed with inhibitors has been hampered by many technical hurdles, such as the low solubility of IN.[90] Herein, the homology model of HIV-1 IN was built based on the crystal structure of PFV IN.[61] The sequence alignment of PFV/HIV-1 IN and viral U5 DNA is shown in Fig. S2 (ESI†). From the sequence alignment, the overall identity between HIV-1 IN and PFV IN was up to ~50%, showing the high degree of homology and highly conserved DDE motif (Asp64, Asp116 and Glu152). The homology model of HIV-1 IN represents a strand transfer conformation for inhibitor binding. The overall architecture of the HIV-1 IN model comprised three domains including NTD (residues 1–46), CCD (residues 56–202) and CTD (residues 220–272) connected with each other by linkers (Fig. 2A). The stereochemical quality of the model was evaluated using Ramachandran plot analysis in PROCHECK.[62] As shown in Fig. S3 (ESI†), 98.2% residues were found in the allowed regions, and 83.5% were in the favored regions indicating that the homology mode was reliable.

Initial poses of JMC6F in the binding sites of HIV-1 IN and RNase H

The inhibitor JMC6F was found active to inhibit both HIV-1 IN and RT associated RNase H.[23] Therefore, JMC6F was docked into the INSTI binding site of HIV-1 IN[91] and RT associated RNase H.[92] The docking poses of JMC6F in HIV-1 IN and RNase H were similar to the co-crystal structures of DTG in PFV IN[64] and MK1 in HIV-1 RNase H,[64] respectively. Since JMC6F, DTG and MK1 shared a similar structure portion (Fig. 1) chelation occurs with two Mg^{2+} ions at the IN and RNase H active site.[11] The superposition of poses of JMC6F with DTG in PFV IN (PDB ID: 3S3M[64]) and MK1 in HIV-1RT (PDB ID: 3LP0[32]) is displayed in Fig. 3. As shown in Fig. 3, JMC6F occupied the active sites of HIV-1 IN and RNase H with the DKA moiety chelating with Mg^{2+}. The initial poses of JMC6F presented in docking could render a prediction of protein–ligand interactions. The initial pose obtained from the docking procedure was subjected to MD simulation to explore the detailed structural and energetic properties for protein–ligand recognition.

MD simulation and stability of the systems. For each system, 50 ns explicit solvent MD simulation was carried, and the fluctuation of simulation trajectory was monitored by root-mean-square deviation (RMSD) relative to the initial structure. The RMSD plots of backbone heavy atoms of proteins and vDNA, heavy atoms of ligands and backbone atoms of the binding site (the residues within 5 Å around the ligand) for each system are shown in Fig. S4 (ESI†). For HIV-1 IN simulation systems (Fig. S4A–D, ESI†),

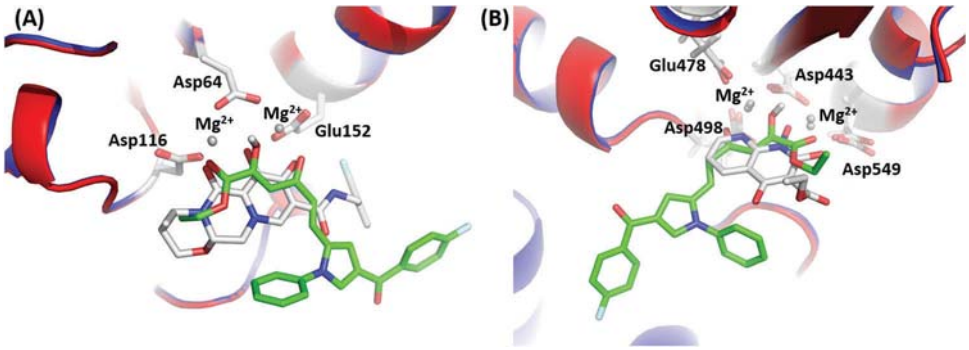

Fig. 3 The close view of docking poses of JMC6F at the active site of HIV-1 IN and RNase H. (A) The conformational imposition of JMC6F (green) with DTG (white) at the active site of HIV-1 IN (B) The conformational imposition of JMC6F (green) with MK1 (white) at the active site of HIV-1 RNase H. The conserved catalytic residues of IN and RNase H are also displayed as white sticks.

the RMSD values of the binding site and ligand were within ~2 Å and ~1.5 Å respectively, while the RMSD values of the overall protein and vDNA were within ~8 Å and ~6 Å. In order to further depict the changes of conformation, RMSFs of the HIV-1 integrase in the simulation trajectory of the JMC6F-IN system were calculated as shown in Fig. S5 (ESI†). And this indicates that the fluctuation was rather higher in the flexible loop regions (residues 15–48, residues 60–72, residues 135–148, residues 200–250), especially the loop region (residues 140–149) at the binding site, which is required for IN function as described previously in a reported experiment.[61] By combining RMSD and RMSF analyses, it was observed that the large fluctuation of the overall protein was caused by the flexible linkers[93] in HIV-1 IN as shown in Fig. S5 (ESI†). For bound HIV-1 RNase H complexes (Fig. S4E and F, ESI†), the overall proteins, binding sites and ligands showed low fluctuations within a RMSD value of ~3 Å. In summary, the results indicated that 50 ns MD simulation was sufficient to obtain an equilibrated simulation trajectory. Thus, the last 20 ns equilibrated trajectory was used for the following structural and energetic analysis.

Validation of the models constructed by MD simulation

Correlation of the binding affinity between the simulation and experimental results. Based on the 500 snapshots extracted from the last 20ns MD simulation, the MM/PBSA method[76] was used to analyze the binding free energy (ΔG_{calc}) and the results are listed in Table 2. The ΔG_{calc} for the six WT systems were −21.89 kcal mol^{-1} (IN-JMC6F), −17.30 kcal mol^{-1} (RNase H-JMC6F), −17.50 kcal mol^{-1} (RNase H-MK1), −32.83 kcal mol^{-1} (IN-DTG), −31.34 kcal mol^{-1} (IN-RAL) and −28.22 kcal mol^{-1} (IN-EVG). And the detailed energy components of the binding free energy are shown in Table S2 (ESI†). Both van der Waals and electrostatic contributions play key roles in JMC6F binding with IN and RNase H. And it is interesting to note that the electronic contribution is higher when JMC6F binds with IN, while the van der Waals contribution is lower when JMC6F binds with RNase H, which may contribute to the balanced binding affinity of JMC6F against HIV-1 IN and RNase H (ratio ~ 1).[38] In addition, $\Delta G_{exp} = RT\ln IC_{50}$ was used to convert the experimental IC_{50} values[32,92,94] into the binding free energies (ΔG_{exp}). As shown in Fig. S6A (ESI†), the relative binding free energy of the $\Delta\Delta G_{calc}$ correlated very well (R^2 = 0.95) with the experimental value ($\Delta\Delta G_{exp}$). The good correlation coefficient between $\Delta\Delta G_{cal}$ and $\Delta\Delta G_{exp}$ can be good evidence for verifying the resulting model from MD simulation.[47]

Table 2 The calculated and experimentally estimated binding free energies of JMC6F, DTG, EVG, RAL and MK1 binding to the wide type HIV-1 IN and RNase H, respectively. (ΔG is in kcal mol^{-1} and the IC_{50} value is in μM)

Drugs	Targets	ΔG_{calc} [a]	$\Delta\Delta G_{calc}$ [b]	ΔG_{exp} [c]	$\Delta\Delta G_{exp}$ [b]	IC_{50} [d]
JMC6F	IN	−21.89	10.94	−8.08	3.62	1.2000
	RNase H	−17.30	15.53	−7.84	3.86	1.8000
MK1	RNase H	−17.50	15.33	−7.58	4.12	2.8000
DTG	IN	−32.83	0.00	−11.70	0.00	0.0027
EVG	IN	−28.22	4.61	−11.20	0.50	0.0060
RAL	IN	−31.34	1.49	−11.60	0.10	0.0033

[a] Calculated binding free energy by the MM/PBSA method in this work. [b] Binding free energy variation $\Delta\Delta G_{calc}$ values were calculated using the IN-DTG complex as a reference. [c] Estimated binding free energy ΔG_{exp} based on IC_{50} values using $\Delta G_{exp} = RT\ln(IC_{50})$. [d] Experimental IC_{50} value from the reported work in ref. 32, 38, 94 and 95.

Resistance profile identified by *in silico* site-directed mutagenesis analysis. To further validate the model constructed in the study, the resistance profile associated with 16 single point mutations of the approved INSTIs (RAL, EVG and DTG) (Table 1) was investigated by *in silico* site mutagenesis analysis. It is noted that the resistance profile of the 16 mutations from experimental results[95] was reproduced well (R^2 = 0.85) by the calculated $\Delta\Delta G_{cal}$ (Fig. S6B and Table S3, ESI†). More detailed information of the binding affinity is listed in Tables S4–S6 (ESI†). As the *in silico* site-directed mutagenesis analysis was based on the WT model from MD simulation in this work, the good correlation further indicated that the resulting models from MD simulation were capable of predicting the mutation-induced drug resistance profile.[47]

Comparison of MD refined and cryo-EM solved HIV-1 IN structures. It is important to indicate that the cryo-EM structure of the core tetrameric HIV-1 strand transfer complex (STC) binding without inhibitors was solved[96] at the end of this work. Solving the HIV-1 IN cryo-EM structure provided an opportunity for validating the model from our MD simulation. Thus, a structure comparison between the refined HIV-1 IN and the cryo-EM HIV-1 IN was conducted. As shown in Fig. S7 (ESI†), the overall distribution of NTD, CCD, CTD and the linkers matched pretty well. The RMSDs of the backbone atoms for NTD, CCD, CTD and the residues at the binding site between the MD refined and cryo-EM solved HIV-1 IN structures were 3.21 Å, 2.26 Å, 3.08 Å and 1.40 Å, respectively. These suggested that the HIV-1 IN model was reliable.

MD simulated snapshots of JMC6F in complex with HIV-1 IN and RNase H. After validating the reliability of the simulation models, the binding mechanism of JMC6F with HIV-1 IN and RNase H was analyzed by visualizing the MD-simulated snapshots and the identification and verification of the key residues contribute to JMC6F binding with IN and RNase H. The representative snapshots of JMC6F in complex with HIV-1 IN and RNase H over the equilibrated MD trajectories were extracted. And a close view of the interactions is depicted in Fig. 4A1 and B1. Fig. 4A1 shows that the diketo acid group of JMC6F is oriented towards the two Mg^{2+} ions, chelating with both the ions and residues Asp64, Asp116, Glu148 and Glu152. The interaction with the two Mg^{2+} ions is important in the development of INSTIs.[97] The phenyl ring linked in the nitrogen atoms of the pyrrole ring exhibited a hydrophobic interaction with Tyr143. Moreover, Tyr143 has been proven to be critical in the binding of RAL or other IN inhibitors with HIV-1 IN.[68] The 4-F-benzoyl moiety linked at the 4 position of the pyrrole ring extended into the space of A17 and C16. Meanwhile, as shown in Fig. 4B1, the diketo acid group of JMC6F interacted with the Mg^{2+} ions at the RNase H active site and residues D443 and D549. The two metal ion mechanism was reported as the key component of RHIs.[32] The pyrrole ring was pointing outside the active site and involved in hydrophobic interactions with residues Trp535, Trp820 and Trp980 (Fig. 4B1).

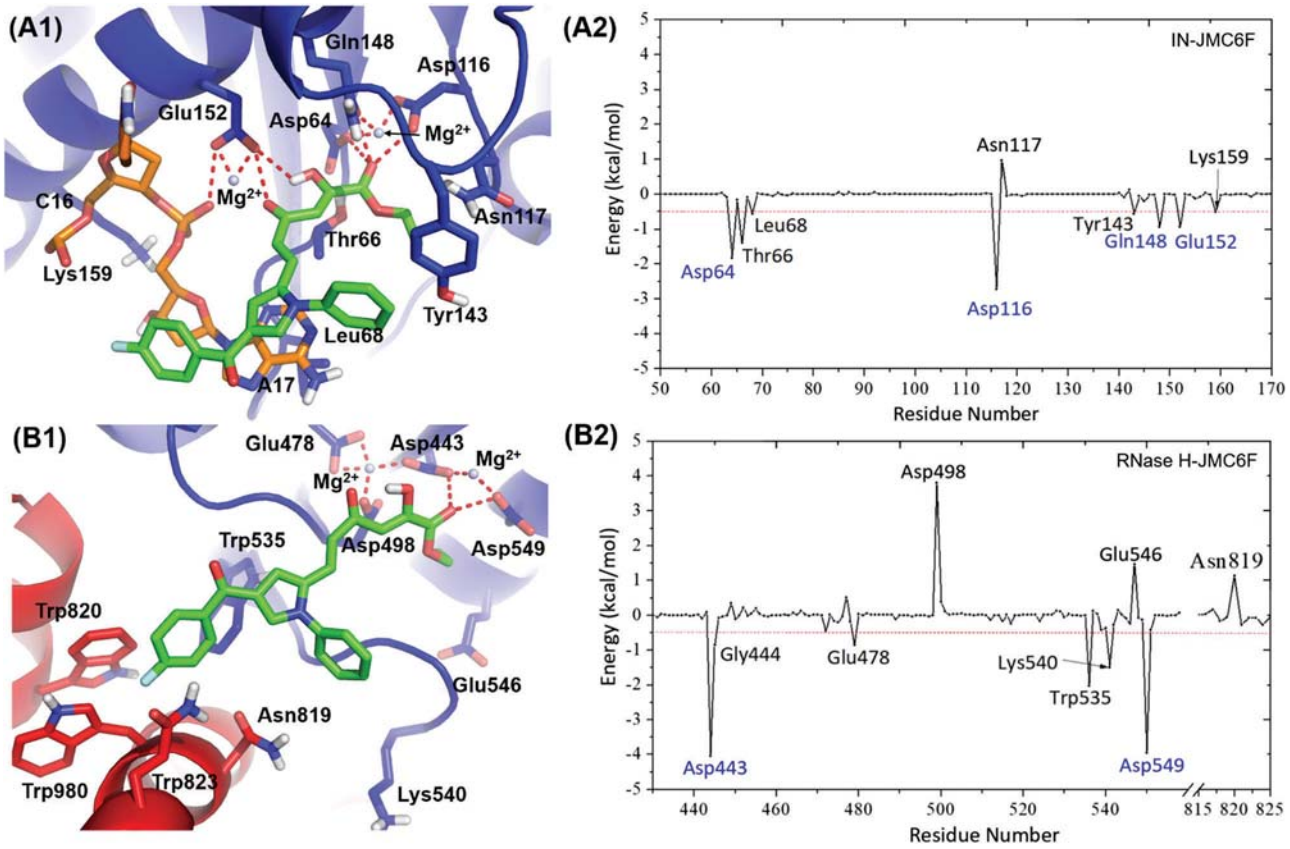

Fig. 4 (A1 and B1) The average structures of the MD simulations of JMC6F interacted with the key residues of the binding pocket of the wide type IN and RNase H. The proteins are shown as cartoons in blue and red color. The side chains of the active site are shown as a stick drawing (red and blue) and JMC6F is represented as a stick drawing (green). The chelating interaction of JMC6F with IN and RNase H is shown as red dash lines. (A2 and B2) The energetic contributions of each individual amino acid residue of the wide type IN and RNase H binding with JMC6F the wide type IN-JMC6F complex. Residues chelating with JMC6F and Mg^{2+} are labelled in blue color.

Identification and verification of the key residues contributing to JMC6F inhibition. Based on equilibrated MD simulation trajectories, the energy contribution of each residue for the inhibition of JMC6F was calculated. As shown in Fig. 4A2 and B2, residues Asp64, Thr66, Leu68, Asp116, Tyr143, Gln148 and Glu152 at the INSTI binding site of HIV-1 IN and residues Asp443, Gly444, Trp535, Lys540 and Asp549 at the RHI binding site of HIV-1 RNase H were identified as key residues ($\geq$0.5 kcal mol^{-1}) for the inhibition of JMC6F. The computational alanine scanning (CAS) method[87] has been widely applied to verify the reliability of key residues identified in protein–ligand complexes by MD simulations.[98–100] Therefore, the identified seven HIV-1 IN residues (Asp64, Thr66, Asp116, Thr143, Gln148, Glu152 and Lys159) and five HIV-1 RNase H residues (Asp443, Glu478, Trp535, Lys540 and Asp549) were further analyzed by CAS calculations (Fig. 5 and Table S7, ESI†). By comparing Fig. 4 and 5, it was obvious that the per-residue binding free energy and CAS results were consistent with each other. Besides, as shown in Table S7 (ESI†) and Fig. 5, in all the simulated complexes, residues Asp64, Leu68, Asp116, Thr143, Lys159, Asp443, Glu478 and Asp549 were found to be hotspots for JMC6F.

Shared binding modes of JMC6F in HIV-1 IN and RNase H

The extracted structures as well as the identified key residues inferred the similarity of the binding mode of JMC6F with

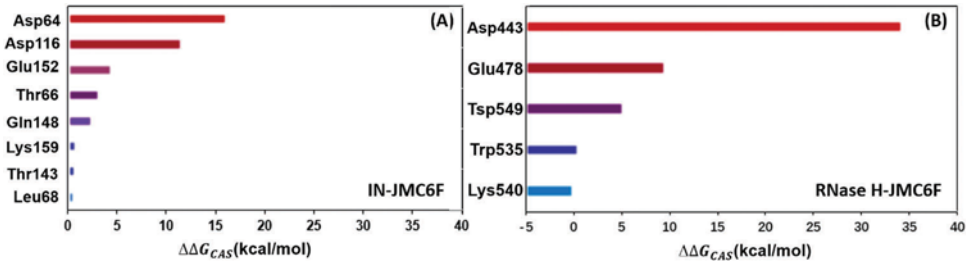

Fig. 5 The computational alanine scanning mutagenesis results for (A) IN-JMC6F and (B) RNase H-JMC6F.

HIV-1 IN and RNase H. The similar binding mode shared by the two complexes was then generalized and is schematically represented in Fig. 6. As illustrated, the shared binding mode could be characterized with two pharmacophores: the diketo acid group chelated with Mg^{2+} ions at the enzyme catalytic sites which was coordinated by the DDE motif (site I) of IN and RNase H, while the pyrrole substitutions were exposed to a hydrophobic cavity (site II and site III) out of the catalytic sites. Compared with the binding site of IN, the RNase H binding site is largely hydrophobic, which apparently affects the interactions of JMC6F with the two targets.

Resistance profile of JMC6F identified based on the available mutations

Though three INSTIs (RAL, EVG and DTG) of HIV-1 IN were approved by the U.S. FDA,[101] the clinical drug resistance as well as the cross resistance between these drugs due to the emergence of mutations could not be ignored.[102] Moreover, combination of G140S/Q148H/N155H substitutions was associated with the reduced susceptibility to the second-generation inhibitor DTG.[91] The major goal of developing dual inhibitors targeting both HIV-1 IN and RNase H is expected to reduce the chance of selecting drug resistant viruses.[33,103–105] However, the sensitivity of JMC6F based on the available mutations remains not well understood.

In this study, *in silico* site-directed mutagenesis analysis of MD simulations based on binding free energy (ΔG_{cal}) calculations[106–110] was employed to predict sensitivity of JMC6F based on known major single mutations (Table 1). The resistance profile of certain residues towards inhibitors can be reflected by the relative binding free energy ($\Delta\Delta G_{cal}$) before and after *in silico* mutation on the corresponding residues.[84,111,112] The predicted drug resistance profile based on currently major single mutations in HIV-1 IN against the inhibition of JMC6F is given in Fig. 7. More detailed information is given in Table S8 (ESI†). As shown, the resistance profile of JMC6F was markedly different from that of the three approved INSTIs. The majority (81.3%) of the mutations were predicted to confer no resistance to JMC6F, except that three mutations show varying degrees of resistance to JMC6F. The Y143C and Q148R mutations (FC was 3.6 and 4.64, respectively) showed moderate resistance to JMC6F. These two mutants also presented resistance to RAL (FC was 3.20 and 47.00, respectively) and EVG (FC was 1.50 and 240, respectively), but no single mutations showed resistance to DTG. Moreover, the N155H substitution led to a high reduction in affinity to JMC6F with an FC of 34.46. In order to reveal the resistance mechanism of the three mutants, we further calculated the energy contribution of each residue, and the results and extracted representative structures depict how mutants affect inhibition (Fig. S8 and S9, ESI†).

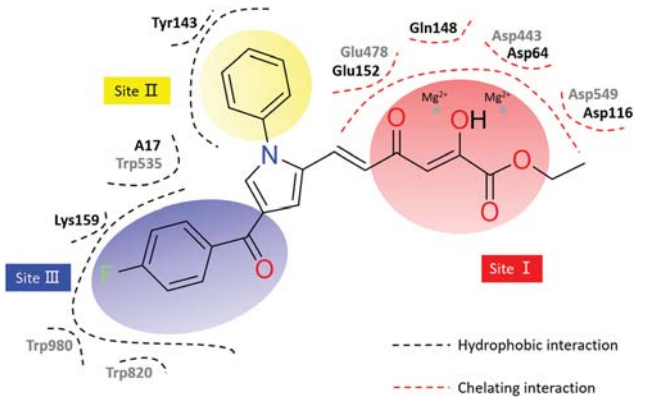

Fig. 6 The binding mode shared by IN-JMC6F and RNase H-JMC6F. The identified hydrophobic interactions and chelating interactions are depicted in black and red dashed lines, respectively. The red color (site I) indicated the chemical moiety chelating with magnesium ions and residues in the vicinity, while the yellow and blue color (site II and site III) represented the chemical moiety forming hydrophobic interactions with residues near the active site. The residues and the nucleotide of the IN/vDNA system are labeled in dark (Asp64, Asp116, Tyr143, Gln148, Glu152, Lys159, and A17), and the residues of RNase H were labeled in gray (Asp443, Glu478, Trp535, Asp549, Trp820, Glu823, and Trp980).

Compared to the wild type system Fig S8A and S9A (ESI†), the three mutants caused obvious impacts on residues which contribute to the inhibition of JMC6F in the wide type. As shown in Fig. S8B (ESI†), the mutation on the residue 143 caused a reduction in the contribution of Leu66, Tyr143 and Lys159. This mutation caused the loss of hydrophobic interactions between the phenyl ring of JMC6F and Tyr143 in the wide type system as shown in Fig. S9B (ESI†). As shown in Fig. S8C and S9C (ESI†), a mutation on the Glu148 caused a reduction in the contribution of Leu68, Asp116, Tyr143, Gln148 and Lys159. And the representative structure showed that the hydrogen bond between Gln148 and JMC6F in the wide type system was lost. In addition, the residue Tyr143 shifts away from the phenyl ring linked in the nitrogen atoms of JMC6F causing the loss of contribution to hydrophobic interaction. As shown in Fig. S8D and S9D (ESI†), a mutation on the residue 155 caused a reduction in the contribution of Asp64, Leu68, Asp116 and Try143. The results shown in Fig. S8D and S9D (ESI†) indicated that these mutants all caused subtle conformational changes at the binding site and the 3′ end of the viral DNA in comparison with the wide type system. It is worth noting that the three mutants all caused an obvious decrease in the energy contribution and conformation changes of Tyr143. These findings indicated that the main cause of resistance may be due to the conformation changes in the binding site, especially in the flexible loop region (residues 140–149). Thus, based on our results and previous studies,[35,46] the key to overcoming drug resistance is to design optimized ligands with substituted groups in the pyrrolyl ring which will strengthen the interaction with the residues at the binding site. Furthermore, even 3 single mutations caused resistance to JMC6F, and it is encouraging that JMC6F was active against most RAL and EVG resistance caused by single mutations.

Finally, despite that no RHIs were approved for anti-HIV clinical therapy, two known mutations on Y501 in RNase H (an important residue for the inhibition of RNase H activity by RHIs)[89,92] were selected to predict the resistance profile of JMC6F. As shown in Table 3, $\Delta\Delta G_{cal}$ values from *in silico* Y501W and Y501R mutations were 1.47 kcal mol^{-1} and 2.67 kcal mol^{-1},

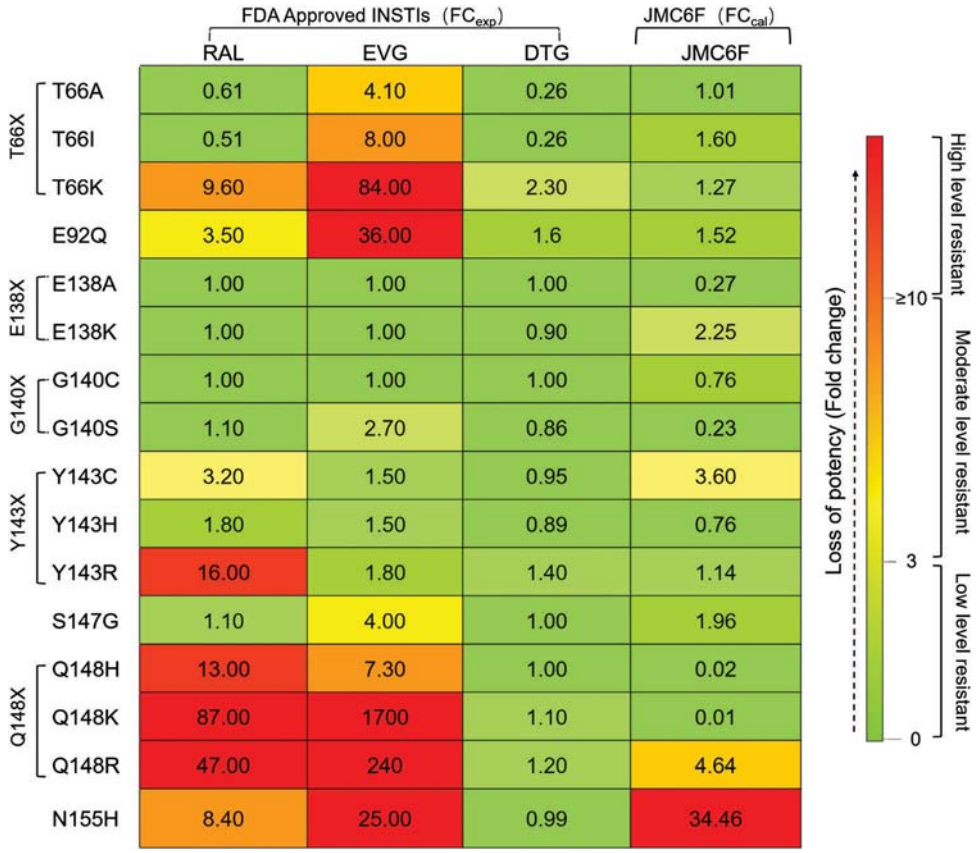

Fig. 7 Resistance profile for compounds toward HIV-1 IN mutants. FC_{exp}, the fold changes calculated in experiments.[95] FC ≤ 3-fold that of the wild type were considered to be resistant, which is shown in green. 10-fold > FC > 3-fold were considered to be moderately resistant, which is shown in yellow. FC ≥ 10-fold were considered to be highly resistant, which is shown in orange. FC_{calc}, fold change values of potency were derived from the equation $\Delta\Delta G_{calc} = RT\ln(FC_{potency})$.

Table 3 The calculated binding free energies of JMC6F binding with mutated HIV-1 RNase H (all units are kcal mol^{-1})

Drug	Target	Mutants	ΔG_{calc} [a]	ΔG_{calc} [b]
JMC6F	RNase H	WT	−17.30	0.00
		Y501R	−15.83	1.47
		Y501W	−14.63	2.67

[a] Calculated binding free energy by the MM/PBSA method in this work. [b] $\Delta\Delta G_{calc}$ is defined as the change of binding free energy using the wild type as a reference.

respectively. The relatively large decrease of binding free energy induced by Y501W and Y501R mutations implies their potential resistance to JMC6F.

Conclusion

In the present study, the binding mode and resistance profile of a novel dual inhibitor JMC6F targeting HIV-1 IN and RNase H were analyzed by computational methods. The models from MD simulations were verified by the good correlation (R^2 = 0.95) of the relative binding free energy between calculated and experimental values, the good reproducibility of the *in silico* site-directed mutagenesis study (R^2 = 0.85) and the structural comparison of the model with the very recently reported HIV-1 IN cryo-EM. By using integrated computational methods, seven residues (Asp64, Thr66, Asp116, Thr143, Gln148, Glu152 and Lys159) of HIV-1 IN and five residues (Asp443, Glu478, Trp535, Lys540 and Asp549) of HIV-1 RNase H were identified as key residues which contributed to the inhibition of JMC6F binding. Thus, a shared binding mode of JMC6F with IN and RNase H was summarized in terms of the diketo acid group chelating with Mg^{2+} ions at the enzyme catalytic sites which was coordinated by the DDE motif (site I) of IN and RNase H, while the pyrrole substitutions were exposed to a hydrophobic cavity (site II and site III) out of the catalytic sites. The drug resistance profile of JMC6F based on the currently known major single mutations in HIV-1 IN and RNase H suggested only three mutations in HIV-1 IN (Y143C, Q148R and N155H) and the key to overcoming drug resistance is to design optimized ligands with substituted groups in the pyrrolyl ring which will strengthen the interaction with residues at the binding site. Finally, two mutations in HIV-1RNase H (Y501R and Y501W) lead to a reduction in JMC6F potency, implying their potential role in providing resistance to JMC6F. The observations obtained in this study will be useful for the structure-based design of more efficient HIV-1 IN and RNase H dual targeting inhibitors.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (21505009), the Precision Medicine Project of the National Key Research and Development Plan of China (2016YFC0902200), the Innovation Project on Industrial Generic Key Technologies of Chongqing (cstc2015zdcy-ztzx120003) and the Fundamental Research Funds for Central Universities (10611CDJXZ238826, CDJZR14468801, and CDJKXB14011).

References

1 H. Kelly, H. A. Weiss, Y. Benavente, S. de Sanjose and P. Mayaud, *Lancet HIV*, 2018, **5**, e45–e58.
2 L. F. Johnson, M. T. May, R. E. Dorrington, M. Cornell, A. Boulle, M. Egger and M. A. Davies, *PLoS Med.*, 2017, **14**, e1002468.
3 F. Zhu, C. Qin, L. Tao, X. Liu, Z. Shi, X. Ma, J. Jia, Y. Tan, C. Cui, J. Lin, C. Tan, Y. Jiang and Y. Chen, *Proc. Natl. Acad. Sci. U. S. A.*, 2011, **108**, 12943–12948.
4 M. Touma, L. D. Rasmussen, R. Martin-Iguacel, F. N. Engsig, N. B. Staerke, M. Staerkind, N. Obel and M. Ahlström, *Clin. Epidemiol.*, 2017, **9**, 385–392.
5 O. A. Uthman, J. B. Nachega, J. Anderson, S. Kanters, E. J. Mills, F. Renaud, S. Essajee, M. Doherty and L. Mofenson, *Lancet HIV*, 2017, **4**, e21–e30.
6 F. Zhu, B. Han, P. Kumar, X. Liu, X. Ma, X. Wei, L. Huang, Y. Guo, L. Han, C. Zheng and Y. Chen, *Nucleic Acids Res.*, 2010, **38**, D787–D791.
7 C. Herrera, N. Armanasco, J. Garcia-Perez, P. Ziprin, N. Olejniczak, J. Alcami, J. Nuttall and R. J. Shattock, *AIDS*, 2016, **30**, 1015–1025.
8 S. Serrano-Villar, T. Sainz, Z. M. Ma, N. S. Utay, T. W. Chun, S. Mann, A. D. Kashuba, B. Siewe, A. Albanese, P. Troia-Cancio, E. Sinclair, A. Somasunderam, T. Yotter, S. G. Deeks, A. Landay, R. B. Pollard, C. J. Miller, S. Moreno and D. M. Asmuth, *PLoS Pathog.*, 2016, **12**, e1005381.
9 F. Zhu, C. J. Zheng, L. Y. Han, B. Xie, J. Jia, X. Liu, M. T. Tammi, S. Y. Yang, Y. Q. Wei and Y. Z. Chen, *Curr. Mol. Pharmacol.*, 2008, **1**, 213–232.
10 Y. Wu, *Retrovirology*, 2004, **1**, 13.
11 K. Das, S. E. Martinez, R. P. Bandwar and E. Arnold, *Nucleic Acids Res.*, 2014, **42**, 8125–8137.
12 F. Zhu, X. X. Li, S. Y. Yang and Y. Z. Chen, *Trends Pharmacol. Sci.*, 2018, **39**, 229–231.
13 P. E. Sax, A. Pozniak, M. L. Montes, E. Koenig, E. DeJesus, H. J. Stellbrink, A. Antinori, K. Workowski, J. Slim, J. Reynes, W. Garner, J. Custodio, K. White, D. SenGupta, A. Cheng and E. Quirk, *Lancet*, 2017, **390**, 2073–2082.
14 P. K. Quashie, R. D. Sloan and M. A. Wainberg, *BMC Med.*, 2012, **10**, 34.
15 F. Zhu, L. Han, C. Zheng, B. Xie, M. T. Tammi, S. Yang, Y. Wei and Y. Chen, *J. Pharmacol. Exp. Ther.*, 2009, **330**, 304–315.
16 L. Menendez-Arias, *Antiviral Res.*, 2013, **98**, 93–120.
17 S. Y. Rhee, T. Liu, J. Ravela, M. J. Gonzales and R. W. Shafer, *Antimicrob. Agents Chemother.*, 2004, **48**, 3122–3126.
18 C. Delaugerre, R. Rohban, A. Simon, M. Mouroux, C. Tricot, R. Agher, J. M. Huraux, C. Katlama and V. Calvez, *J. Med. Virol.*, 2001, **65**, 445–448.
19 T. Hawkins, *Antiviral Res.*, 2010, **85**, 201–209.
20 E. De Clercq, *Int. J. Antimicrob. Agents*, 2009, **33**, 307–320.
21 N. R. Srinivas, *Clin. Pharmacol. Ther.*, 2018, **103**, 570.
22 S. N. Kudalkar, J. Beloor, E. Quijano, K. A. Spasov, W. G. Lee, J. A. Cisneros, W. M. Saltzman, P. Kumar, W. L. Jorgensen and K. S. Anderson, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, E802–E811.
23 E. De Clercq and G. Li, *Clin. Microbiol. Rev.*, 2016, **29**, 695–747.
24 B. Li, J. Tang, Q. Yang, S. Li, X. Cui, Y. Li, Y. Chen, W. Xue, X. Li and F. Zhu, *Nucleic Acids Res.*, 2017, **45**, W162–W170.
25 F. Zhu, X. H. Ma, C. Qin, L. Tao, X. Liu, Z. Shi, C. L. Zhang, C. Y. Tan, Y. Z. Chen and Y. Y. Jiang, *PLoS One*, 2012, **7**, e39782.
26 Y. Goldgur, F. Dyda, A. B. Hickman, T. M. Jenkins, R. Craigie and D. R. Davies, *Proc. Natl. Acad. Sci. U. S. A.*, 1998, **95**, 9150–9154.
27 T. Fu, G. Zheng, G. Tu, F. Yang, Y. Chen, X. Yao, X. Li, W. Xue and F. Zhu, *ACS Chem. Neurosci.*, 2018, **9**, 1492–1502.
28 H. Yang, C. Qin, Y. H. Li, L. Tao, J. Zhou, C. Y. Yu, F. Xu, Z. Chen, F. Zhu and Y. Z. Chen, *Nucleic Acids Res.*, 2016, **44**, D1069–D1074.
29 F. Zhu, Z. Shi, C. Qin, L. Tao, X. Liu, F. Xu, L. Zhang, Y. Song, X. Liu, J. Zhang, B. Han, P. Zhang and Y. Chen, *Nucleic Acids Res.*, 2012, **40**, D1128–D1136.
30 Y. H. Li, C. Y. Yu, X. X. Li, P. Zhang, J. Tang, Q. Yang, T. Fu, X. Zhang, X. Cui, G. Tu, Y. Zhang, S. Li, F. Yang, Q. Sun, C. Qin, X. Zeng, Z. Chen, Y. Z. Chen and F. Zhu, *Nucleic Acids Res.*, 2018, **46**, D1121–D1127.
31 R. Di Santo, *Curr. Med. Chem.*, 2011, **18**, 3335–3342.
32 H. P. Su, Y. Yan, G. S. Prasad, R. F. Smith, C. L. Daniels, P. D. Abeywickrema, J. C. Reid, H. M. Loughran, M. Kornienko, S. Sharma, J. A. Grobler, B. Xu, V. Sardana, T. J. Allison, P. D. Williams, P. L. Darke, D. J. Hazuda and S. Munshi, *J. Virol.*, 2010, **84**, 7625–7633.
33 F. Esposito and E. Tramontano, *Antiviral Chem. Chemother.*, 2014, **23**, 129–144.
34 G. Zheng, W. Xue, P. Wang, F. Yang, B. Li, X. Li, Y. Li, X. Yao and F. Zhu, *Sci. Rep.*, 2016, **6**, 26883.
35 A. Corona, F. S. di Leva, G. Rigogliuso, L. Pescatori, V. N. Madia, F. Subra, O. Delelis, F. Esposito, M. Cadeddu, R. Costi, S. Cosconati, E. Novellino, R. di Santo and E. Tramontano, *Antiviral Res.*, 2016, **134**, 236–243.
36 L. Tian, M. S. Kim, H. Li, J. Wang and W. Yang, *Proc. Natl. Acad. Sci. U. S. A.*, 2018, **115**, 507–512.
37 L. Tao, F. Zhu, F. Xu, Z. Chen, Y. Y. Jiang and Y. Z. Chen, *Pharmacol. Res.*, 2015, **102**, 123–131.

38 L. Pescatori, M. Metifiot, S. Chung, T. Masoaka, G. Cuzzucoli Crucitti, A. Messore, G. Pupo, V. N. Madia, F. Saccoliti, L. Scipione, S. Tortorella, F. S. Di Leva, S. Cosconati, L. Marinelli, E. Novellino, S. F. Le Grice, Y. Pommier, C. Marchand, R. Costi and R. Di Santo, *J. Med. Chem.*, 2015, **58**, 4610–4623.
39 Y. Mehellou and E. De Clercq, *J. Med. Chem.*, 2010, **53**, 521–538.
40 L. Tao, F. Zhu, C. Qin, C. Zhang, F. Xu, C. Y. Tan, Y. Y. Jiang and Y. Z. Chen, *Nat. Biotechnol.*, 2014, **32**, 979–980.
41 Y. H. Li, P. P. Wang, X. X. Li, C. Y. Yu, H. Yang, J. Zhou, W. W. Xue, J. Tan and F. Zhu, *PLoS One*, 2016, **11**, e0165737.
42 J. Didierjean, C. Isel, F. Querre, J. F. Mouscadet, A. M. Aubertin, J. Y. Valnot, S. R. Piettre and R. Marquet, *Antimicrob. Agents Chemother.*, 2005, **49**, 4884–4894.
43 M. Billamboz, F. Bailly, M. L. Barreca, L. De Luca, J. F. Mouscadet, C. Calmels, M. L. Andréola, M. Witvrouw, F. Christ, Z. Debyser and P. Cotelle, *J. Med. Chem.*, 2008, **51**, 7717–7730.
44 C. Marchand, J. A. Beutler, A. Wamiru, S. Budihas, U. Mollmann, L. Heinisch, J. W. Mellors, S. F. Le Grice and Y. Pommier, *Antimicrob. Agents Chemother.*, 2008, **52**, 361–364.
45 G. Cuzzucoli Crucitti, M. Metifiot, L. Pescatori, A. Messore, V. N. Madia, G. Pupo, F. Saccoliti, L. Scipione, S. Tortorella, F. Esposito, A. Corona, M. Cadeddu, C. Marchand, Y. Pommier, E. Tramontano, R. Costi and R. Di Santo, *J. Med. Chem.*, 2015, **58**, 1915–1928.
46 V. Poongavanam, N. S. H. N. Moorthy and J. Kongsted, *RSC Adv.*, 2014, **4**, 38672–38681.
47 G. F. Hao, G. F. Yang and C. G. Zhan, *Drug Discovery Today*, 2012, **17**, 1121–1126.
48 Y. Li, S. Xuan, Y. Feng and A. Yan, *Drug Discovery Today*, 2015, **20**, 435–449.
49 E. R. Gray, J. C. Brookes, C. Caillat, V. Turbe, B. L. J. Webb, L. A. Granger, B. S. Miller, L. E. McCoy, M. El Khattabi, C. T. Verrips, R. A. Weiss, D. M. Duffy, W. Weissenhorn and R. A. McKendry, *ACS Infect. Dis.*, 2017, **3**, 479–491.
50 W. Xue, F. Yang, P. Wang, G. Zheng, Y. Chen, X. Yao and F. Zhu, *ACS Chem. Neurosci.*, 2018, **9**, 1128–1140.
51 W. Xue, P. Wang, G. Tu, F. Yang, G. Zheng, X. Li, X. Li, Y. Chen, X. Yao and F. Zhu, *Phys. Chem. Chem. Phys.*, 2018, **20**, 6606–6616.
52 B. Li, J. Tang, Q. Yang, X. Cui, S. Li, S. Chen, Q. Cao, W. Xue, N. Chen and F. Zhu, *Sci. Rep.*, 2016, **6**, 38881.
53 X. Liu, D. Shi, S. Zhou, H. Liu, H. Liu and X. Yao, *Expert Opin. Drug Discovery*, 2018, **13**, 23–37.
54 W. Xue, P. Wang, B. Li, Y. Li, X. Xu, F. Yang, X. Yao, Y. Z. Chen, F. Xu and F. Zhu, *Phys. Chem. Chem. Phys.*, 2016, **18**, 3260–3271.
55 F. Y. Yang, T. T. Fu, X. Y. Zhang, J. Hu, W. W. Xue, G. X. Zheng, B. Li, Y. H. Li, X. J. Yao and F. Zhu, *Mol. Simul.*, 2017, **43**, 1089–1098.
56 J. Wang, P. Morin, W. Wang and P. A. Kollman, *J. Am. Chem. Soc.*, 2001, **123**, 5221–5230.
57 C. Wang, D. Greene, L. Xiao, R. Qi and R. Luo, *Front. Mol. Biosci.*, 2017, **4**, 87.
58 L. Li, Y. Li, L. Zhang and T. Hou, *J. Chem. Inf. Model.*, 2012, **52**, 2715–2729.
59 K. H. Barakat, J. Law, A. Prunotto, W. C. Magee, D. H. Evans, D. L. Tyrrell, J. Tuszynski and M. Houghton, *J. Chem. Inf. Model.*, 2013, **53**, 3031–3043.
60 D. W. Wright, B. A. Hall, O. A. Kenway, S. Jha and P. V. Coveney, *J. Chem. Theory Comput.*, 2014, **10**, 1228–1241.
61 S. Hare, S. S. Gupta, E. Valkov, A. Engelman and P. Cherepanov, *Nature*, 2010, **464**, 232–236.
62 R. A. Laskowski, M. W. MacArthur, D. S. Moss and J. M. Thornton, *J. Appl. Crystallogr.*, 1993, **26**, 283–291.
63 A. S. Espeseth, P. Felock, A. Wolfe, M. Witmer, J. Grobler, N. Anthony, M. Egbertson, J. Y. Melamed, S. Young, T. Hamill, J. L. Cole and D. J. Hazuda, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 11244–11249.
64 S. Hare, S. J. Smith, M. Metifiot, A. Jaxa-Chamiec, Y. Pommier, S. H. Hughes and P. Cherepanov, *Mol. Pharmacol.*, 2011, **80**, 565–572.
65 S. Hare, A. M. Vos, R. F. Clayton, J. W. Thuring, M. D. Cummings and P. Cherepanov, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 20057–20062.
66 G. A. Kaminski and R. A. Friesner, *J. Phys. Chem. B*, 2001, **105**, 6474–6487.
67 J. Kankanala, K. A. Kirby, F. Liu, L. Miller, E. Nagy, D. J. Wilson, M. A. Parniak, S. G. Sarafianos and Z. Wang, *J. Med. Chem.*, 2016, **59**, 5051–5062.
68 W. Xue, J. Qi, Y. Yang, X. Jin, H. Liu and X. Yao, *Mol. BioSyst.*, 2012, **8**, 2135–2144.
69 E. Harder, W. Damm, J. Maple, C. Wu, M. Reboul, J. Y. Xiang, L. Wang, D. Lupyan, M. K. Dahlgren, J. L. Knight, J. W. Kaus, D. S. Cerutti, G. Krilov, W. L. Jorgensen, R. Abel and R. A. Friesner, *J. Chem. Theory Comput.*, 2016, **12**, 281–296.
70 V. Hornak, R. Abel, A. Okur, B. Strockbine, A. Roitberg and C. Simmerling, *Proteins*, 2006, **65**, 712–725.
71 J. Wang, W. Wang, P. A. Kollman and D. A. Case, *J. Mol. Graphics Modell.*, 2006, **25**, 247–260.
72 J. Wang, R. M. Wolf, J. W. Caldwell, P. A. Kollman and D. A. Case, *J. Comput. Chem.*, 2004, **25**, 1157–1174.
73 C. I. Bayly, P. Cieplak, W. D. Cornell and P. A. Kollman, *J. Phys. Chem. B*, 1993, **97**, 10269–10280.
74 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**, 10089–10092.
75 M. Springborg and B. Kirtman, *J. Chem. Phys.*, 2007, **126**, 104107–104112.
76 P. A. Kollman, I. Massova, C. Reyes, B. Kuhn, S. Huo, L. Chong, M. Lee, T. Lee, Y. Duan, W. Wang, O. Donini, P. Cieplak, J. Srinivasan, D. A. Case and T. E. Cheatham, *Acc. Chem. Res.*, 2000, **33**, 889–897.
77 L. Xu, H. Sun, Y. Li, J. Wang and T. Hou, *J. Phys. Chem. B*, 2013, **117**, 8408–8421.
78 H. Sun, Y. Li, S. Tian, L. Xu and T. Hou, *Phys. Chem. Chem. Phys.*, 2014, **16**, 16719–16729.
79 J. Xu, P. Wang, H. Yang, J. Zhou, Y. Li, X. Li, W. Xue, C. Yu, Y. Tian and F. Zhu, *BioMed Res. Int.*, 2016, **2016**, 2509385.

80 S. Jo, W. Jiang, H. S. Lee, B. Roux and W. Im, *J. Chem. Inf. Model.*, 2013, **53**, 267–277.
81 D. Cappel, M. L. Hall, E. B. Lenselink, T. Beuming, J. Qi, J. Bradner and W. Sherman, *J. Chem. Inf. Model.*, 2016, **56**, 2388–2400.
82 J. Heimdal and U. Ryde, *Phys. Chem. Chem. Phys.*, 2012, **14**, 12592–12604.
83 Q. Chen, J. K. Buolamwini, J. C. Smith, A. Li, Q. Xu, X. Cheng and D. Wei, *J. Chem. Inf. Model.*, 2013, **53**, 3297–3307.
84 W. Xue, X. Jin, L. Ning, M. Wang, H. Liu and X. Yao, *J. Chem. Inf. Model.*, 2013, **53**, 210–222.
85 W. Rocchia, E. Alexov and B. Honig, *J. Phys. Chem. B*, 2001, **105**, 6507–6514.
86 D. Sitkoff, K. A. Sharp and B. Honig, *J. Phys. Chem. B*, 1994, **98**, 1978–1988.
87 I. Massova and P. A. Kollman, *J. Am. Chem. Soc.*, 1999, **121**, 8133–8143.
88 R. W. Shafer, *J. Infect. Dis.*, 2006, **194**(Suppl 1), S51–S58.
89 D. Arion, N. Sluis-Cremer, K. L. Min, M. E. Abram, R. S. Fletcher and M. A. Parniak, *J. Biol. Chem.*, 2002, **277**, 1370–1374.
90 L. Feng, R. C. Larue, A. Slaughter, J. J. Kessl and M. Kvaratskhelia, *Curr. Top. Microbiol. Immunol.*, 2015, **389**, 93–119.
91 I. Malet, E. Thierry, M. Wirden, S. Lebourgeois, F. Subra, C. Katlama, E. Deprez, V. Calvez, A. G. Marcelin and O. Delelis, *J. Antimicrob. Chemother.*, 2015, **70**, 2870–2880.
92 A. Corona, A. Schneider, K. Schweimer, P. Rosch, B. M. Wohrl and E. Tramontano, *Antimicrob. Agents Chemother.*, 2014, **58**, 4086–4093.
93 W. Xue, H. Liu and X. Yao, *J. Comput. Chem.*, 2012, **33**, 527–536.
94 B. A. Johns, T. Kawasuji, J. G. Weatherhead, T. Taishi, D. P. Temelkoff, H. Yoshida, T. Akiyama, Y. Taoda, H. Murai, R. Kiyama, M. Fuji, N. Tanimoto, J. Jeffrey, S. A. Foster, T. Yoshinaga, T. Seki, M. Kobayashi, A. Sato, M. N. Johnson, E. P. Garvey and T. Fujiwara, *J. Med. Chem.*, 2013, **56**, 5901–5916.
95 M. Kobayashi, T. Yoshinaga, T. Seki, C. Wakasa-Morimoto, K. W. Brown, R. Ferris, S. A. Foster, R. J. Hazen, S. Miki, A. Suyama-Kagitani, S. Kawauchi-Miki, T. Taishi, T. Kawasuji, B. A. Johns, M. R. Underwood, E. P. Garvey, A. Sato and T. Fujiwara, *Antimicrob. Agents Chemother.*, 2011, **55**, 813–821.
96 D. O. Passos, M. Li, R. Yang, S. V. Rebensburg, R. Ghirlando, Y. Jeon, N. Shkriabai, M. Kvaratskhelia, R. Craigie and D. Lyumkis, *Science*, 2017, **355**, 89–92.
97 L. Krishnan, X. Li, H. L. Naraharisetty, S. Hare, P. Cherepanov and A. Engelman, *Proc. Natl. Acad. Sci. U. S. A.*, 2010, **107**, 15910–15915.
98 Z. Tian, J. Liu and Y. Zhang, *Sci. Rep.*, 2016, **6**, 22336.
99 F. Dapiaggi, S. Pieraccini and M. Sironi, *Mol. BioSyst.*, 2015, **11**, 2152–2157.
100 O. Keskin, B. Ma and R. Nussinov, *J. Mol. Biol.*, 2005, **345**, 1281–1294.
101 X. Z. Zhao, S. J. Smith, D. P. Maskell, M. Metifiot, V. E. Pye, K. Fesen, C. Marchand, Y. Pommier, P. Cherepanov, S. H. Hughes and T. R. Burke Jr., *J. Med. Chem.*, 2017, **60**, 7315–7332.
102 D. J. McColl and X. Chen, *Antiviral Res.*, 2010, **85**, 101–118.
103 S. Distinto, E. Maccioni, R. Meleddu, A. Corona, S. Alcaro and E. Tramontano, *Curr. Pharm. Des.*, 2013, **19**, 1850–1859.
104 P. K. Quashie, T. Mesplede and M. A. Wainberg, *Curr. Opin. Infect. Dis.*, 2013, **26**, 43–49.
105 P. Wang, F. Yang, H. Yang, X. Xu, D. Liu, W. Xue and F. Zhu, *Biomed. Mater. Eng.*, 2015, **26**(Suppl 1), S2233–S2239.
106 R. Chachra and R. C. Rizzo, *J. Chem. Theory Comput.*, 2008, **4**, 1526–1540.
107 P. Wang, X. Zhang, T. Fu, S. Li, B. Li, W. Xue, X. Yao, Y. Chen and F. Zhu, *ACS Chem. Neurosci.*, 2017, **8**, 1416–1428.
108 P. Wang, T. Fu, X. Zhang, F. Yang, G. Zheng, W. Xue, Y. Chen, X. Yao and F. Zhu, *Biochim. Biophys. Acta*, 2017, **1861**, 2766–2777.
109 C. Y. Yu, X. X. Li, H. Yang, Y. H. Li, W. W. Xue, Y. Z. Chen, L. Tao and F. Zhu, *Int. J. Mol. Sci.*, 2018, **19**, E183.
110 G. Zheng, W. Xue, F. Yang, Y. Zhang, Y. Chen, X. Yao and F. Zhu, *Phys. Chem. Chem. Phys.*, 2017, **19**, 28885–28896.
111 W. Xue, P. Jiao, H. Liu and X. Yao, *Antiviral Res.*, 2014, **104**, 40–51.
112 W. Xue, D. Pan, Y. Yang, H. Liu and X. Yao, *Antiviral Res.*, 2012, **93**, 126–137.

