Alternative HIV-1 Non-Nucleoside Reverse Transcriptase Inhibition: Drugging the p51 subunit

Kwok-Fong Chan[1,#], Chinh Tran-To Su[1,#], Alexander Krah[1], Ser-Xian Phua[1], Peter J. Bond[1], Samuel Ken-En Gan[1,2,3,*]

[1]Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), Singapore

[2]p53 Laboratory, A*STAR, Singapore

[3]Experimental Drug Development Centre, A*STAR, Singapore

[#] Both authors contributed equally to the work

*Corresponding author: samuelg@bii.a-star.edu.sg

Bioinformatics Institute, A*STAR

30 Biopolis Street, #07-01 Matrix

Singapore 138671

Tel: +65 6407 0584 (S. K-E. G.)

KEYWORDS: HIV, novel p51 drug target, NNRTIs

ABSTRACT

HIV resistance continues to raise the need for alternative drug targets. Since Reverse Transcriptase (RT) is unique and critical for the virus life cycle, it is a rational target that is likely to have less host side effects. By screening for alternative HIV-1 RT inhibitors, we found that two out of forty compounds from the NCI Diversity Set V could inhibit the RT catalytic activity. Experimental and computational structural analyses showed that one of the two inhibitory compounds binds to a novel location in the HIV-1 RT p51 subunit, revealing a new druggable site and opening opportunities for a new class of non-nucleoside RT Inhibitors that may inhibit HIV-1 RT allosterically. This is the first report of potential inhibitors targeting HIV-1 RT p51.

INTRODUCTION

Human immunodeficiency virus (HIV) remains one of the major global pandemics, with millions infected worldwide[1]. Current therapies including pre-exposure prophylaxis (PrEP) for prevention and cocktail antiretroviral therapy (ART) for treatment target stages of the HIV life cycle: Protease Inhibitors (PIs to interfere with Protease binding to its substrate Gag during viral maturation[2]); Reverse Transcriptase Inhibitors (RTIs) to inhibit viral DNA production[3]; Integrase Inhibitors to block viral DNA insertion into the host genome[4]; and many others that include chemokine receptor antagonists[5], attachment inhibitors[6], and pharmacokinetic enhancers[7].

While integrase inhibitors are currently widely accepted and recommended by WHO[8] as the first line of treatment, there remains potential safety issues for pregnant women living. On the other hand, the RTI drug class, comprising of both the nucleoside (NRTI) and non-nucleoside (NNRTI) inhibitors, are safer and affordable. RTIs target the Reverse Transcriptase (RT) enzyme that is unique to viruses, making it a safer target with reduced chance of toxicity, and hence more suitable than integrase inhibitors for children, pregnant women, or individuals treated tuberculosis with rifampicin. Recently, a two-drug switch regimen[9] that involves a NNRTI (Rilpivirine) and PI (cobicistat-boosted darunavir[9]) elicits the combination as an effective therapy. In addition to WHO-recommended use of RTIs[8] in the standard triple-inhibitor ART, such finding of beneficial two-drug regimens demonstrate the importance of RTIs in the control of HIV.

Both NRTI and NNRTI are used in ART, each exerting inhibition via different mechanisms. NRTIs mimic nucleotides and bind to the nucleotide binding site to competitively block the reverse transcription[10], leading to termination of the DNA elongation process[11]. In contrast, NNRTIs bind to a site located near to the polymerase active site, disrupting the structural alignment of the deoxyribonucleotide triphosphates (dNTPs)

and template/primer substrates at the RT "primer grip", thus inducing effects on the polymerase active site distally[12] and inhibiting reverse transcription[13]. Unlike NRTIs, which are constrained to mimicking nucleotides to directly interfere with cellular replication machinery[14], NNRTIs can vary in structure, generally do not exhibit off-target effects, and have been found to non-competitively cause conformational changes in the RT structure[15]. Therefore, NNRTI drug resistant mutations typically occur at the less conserved drug-binding pocket[16].

HIV-1 RT is a heterodimer of the p66 and p51 subunits. The larger p66 subunit contains five subdomains: fingers, palm, thumb, connection, and RNaseH. Despite having an identical sequence to that of p66 excluding the RNaseH subdomain, the p51 subunit forms a more compact structure[17] with no inhibitors targeting the subunit in the clinic to date.

In our search for other potential inhibitors against HIV-1 RT, we experimentally screened forty chemicals from a diverse set of chemical compounds using the RT-PCR assay, finding two potential inhibitors. Utilizing computational docking and molecular dynamics simulations, our analyses revealed the likely binding mechanisms of the two compounds to the RT structure, where one compound bound to the p51 subunit, and the other near to the original known NNRTI binding site. Our findings thus demonstrate a novel druggable site in the RT p51 and a potential inhibitor scaffold for novel intervention.

RESULTS

Screening for compounds that inhibit HIV-1 RT activity

GAPDH housekeeping gene PCR was used as a cDNA synthesis control given its constitutive expression[18]. RT-PCR screening of 40 compounds from the NCI/DTP Diversity Set V at 300 μM showed that two compounds, i.e. NSC48443 and NSC127133 (subsequently referred to as compounds 1 and 2, respectively) inhibited HIV-1 RT cDNA synthesis (Figure 1). A non-inhibitor (i.e. NSC6145, referred to as "ctrl cmpd") from the preliminary RT-PCR screening results was used as a negative control. Compounds 1 and 2 were found to partially inhibit HIV-1 RT, based on independent triplicate measurements, at 150 and 300 μM respectively (Figure 1), and PCR band intensities were observed to decrease at higher concentrations of the inhibitor candidates.

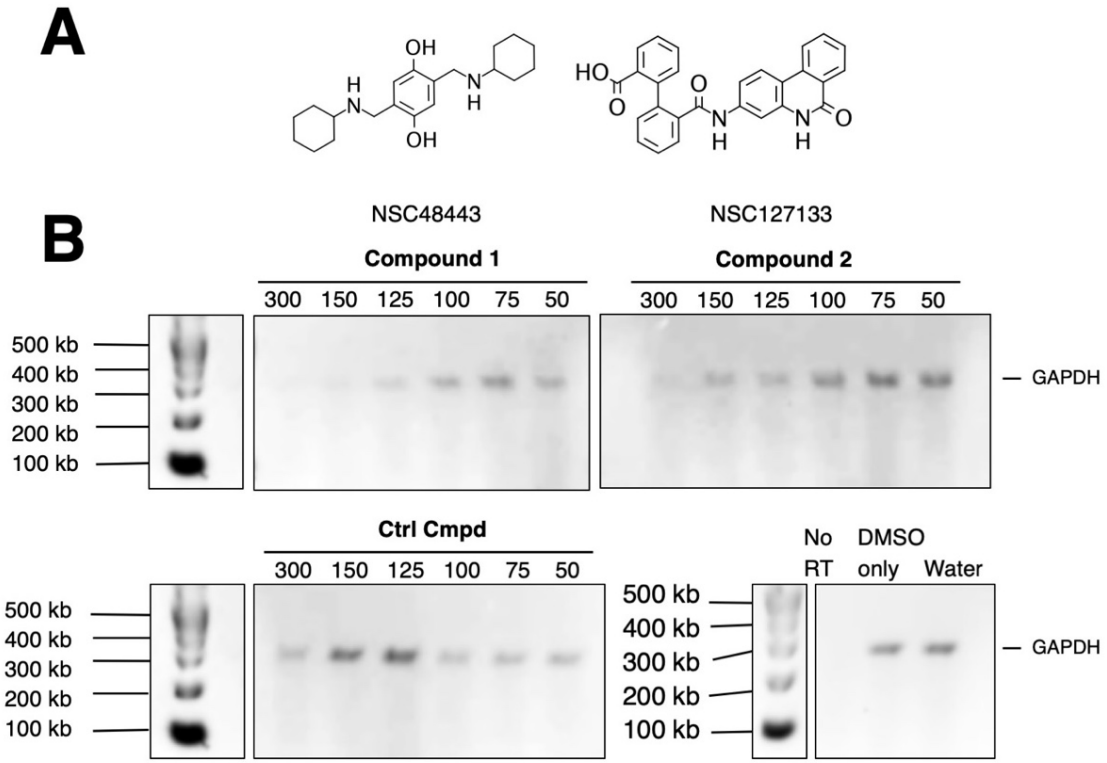

Figure 1. The two compounds that inhibited HIV-1 RT function. (**A**) 2D chemical structures of compounds 1 (NSC48443) and 2 (NSC127133). (**B**) RT-PCR using HIV-1 RT assay,

representative results of triplicate experiments on GAPDH gene with the two identified compounds. The experiments were performed at various concentrations (50 - 300 μM) of the two potential inhibitors with the controls.

Identifying binding sites of the inhibitory compounds on HIV-1 RT

Our functional assays using p51 and p66 subunits separately on their own did not show detectable bands for the p51 alone (Figure 2A). On the other hand, bands were detected for the p66-only reactions in the case of compound 2 and the control compound. However, there was no band detected in the case of compound 1 added to the p66 only experiment (Figure 2A). We also found that inhibition of enzyme activity did not occur during the GAPDH PCR step, as the PCR bands were still observed when the compounds were added only after the cDNA synthesis step (Figure 2B). The findings showed that compound 1, but not compound 2, inhibited the RT p66 subunits, suggesting that compound 2 does not bind to p66.

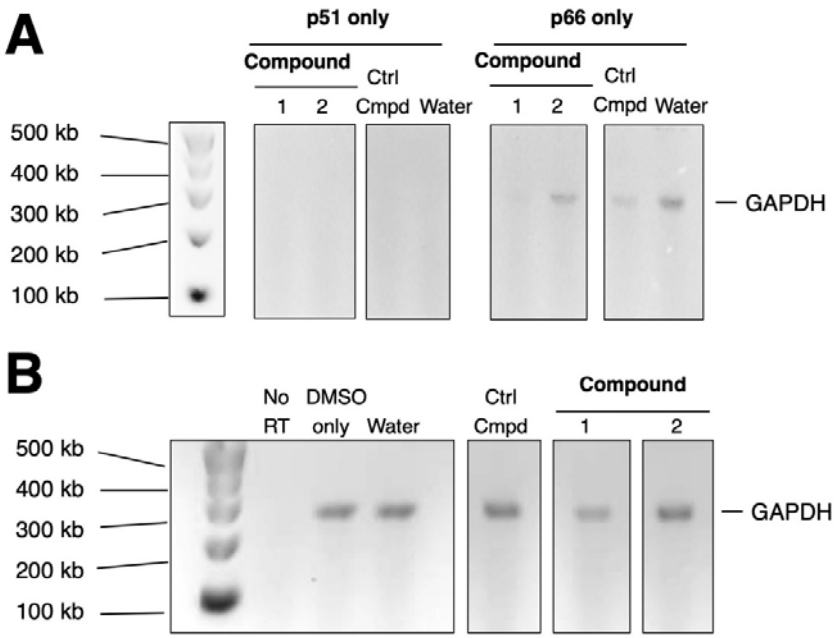

Figure 2. Agarose gel electrophoresis results of RT-PCR on GAPDH gene products on a 2% agar gel. Water, DMSO incubated reaction mixture and a randomly picked non-inhibitor, NSC6145 ('ctrl cmpd') were used as negative controls. **(A)** RT-PCR results of compounds on

HIV-1 RT p51 or p66 subunits separately. **(B)** Results of post-cDNA synthesis stage in the cases of the two compound and the respective controls. All agarose gel electrophoresis results were analysed with GelApp[19] and bands were found to be within the expected size of ~300bp.

To identify the binding sites of compound 1 and 2 on the HIV-1 RT structure heterodimer of the p66 and p51 subunits, blind docking experiments were performed using Autodock Vina[20], followed by molecular dynamics (MD) simulations with various initial bound conformations of compounds 1 and 2 as inputs (Supplementary Figure S1). Compound 1 bound to two different sites on the HIV-1 RT, with one site on the p66 subunit close to a loop located on p51 comprising of S134 to P140, and the other possible site on the p51 subunit. Compound 2 was observed to only bind to one site on the p51 subunit (Figure 3A).

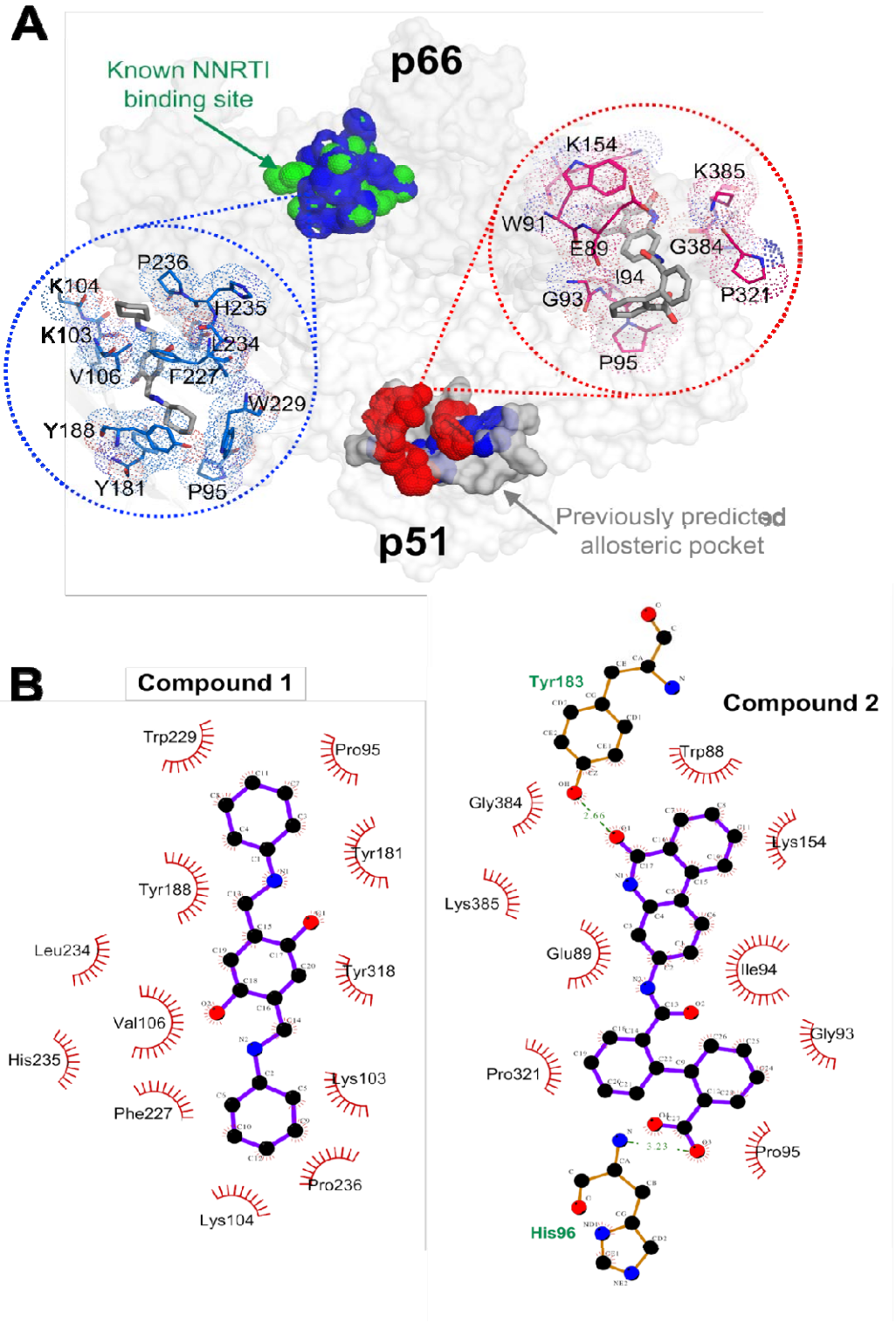

Figure 3. Structural analyses of the HIV-1 RT and the two inhibitory compounds. **(A)** Binding sites of compound 1 (blue surface and lines) and of compound 2 (red surface and lines) on HIV-1 RT. One of the two binding sites for compound 1 overlaps with the known

NNRTI binding pocket[21] (green, on p66), whereas the other overlaps with our previously predicted allosteric pocket[21] (grey surface on p51), on which compound 2 was also found to bind (red). The two chemical compounds are presented in grey sticks within the binding pockets. **(B)** 2D presentation of the interactions between the compounds and their binding sites are shown, constructed using LigPlot+[22].

Determining the favourable binding sites of the two compounds

Structural analyses of the RT showed that compound 1 occupied the site on p66 with a more favourable binding energy than the binding site on p51 (Supplementary Table S1). It was also observed that compound 1 maintained a more stable bound conformational state at the p66 site than at the p51 site (Figure 4A, showing the more stable center-of-mass distances between the binding site and the compound). With more hydrophobic contacts but fewer hydrogen bonds detected between compound 1 and its p66 binding site than in the p51 binding site (Supplementary Table S2), it is likely for hydrophobic interactions to be the dominant binding force to p66.

Hydrophobic contacts were observed between L100, K103, V106, Y181, Y188, P225, F227, L234, P236 and Y318 with less prominent interactions between P95, S105 and W229. Unstable/weak hydrogen bonds were detected between compound 1 and residue K101 (Supplementary Figure S2).

On the other hand, compound 2 was found to maintain stable binding to the p51 subunit on three out of six successful setups for the MD simulations – each setup has a unique initial compound 2 conformation (Figure 3 and Supplementary Figure S1). In these three conformations, compound 2 exhibited low binding energies to the same site on p51, with the lowest in conformation 3 (Supplementary Table S3), thus the binding site for conformation 3 was selected. Compared to the binding site of compound 1 on p66 subunit,

the binding site of compound 2 on p51 is more solvated and noticeably polar due to several polar/charged nearby residues H96, K154, and K385. It was further affirmed by the consistent trends of energy contributions (with increasing interior dielectric constants, shown in Supplementary Table S3) that the binding might be dominated by electrostatic interactions. We observed that the site residue Y183 established strong hydrogen bonds with compound 2 whereas some others (W88, E89, G93, I94, P95, K154, P157 and K183) made hydrophobic contacts with compound 2 (Supplementary Figure S2).

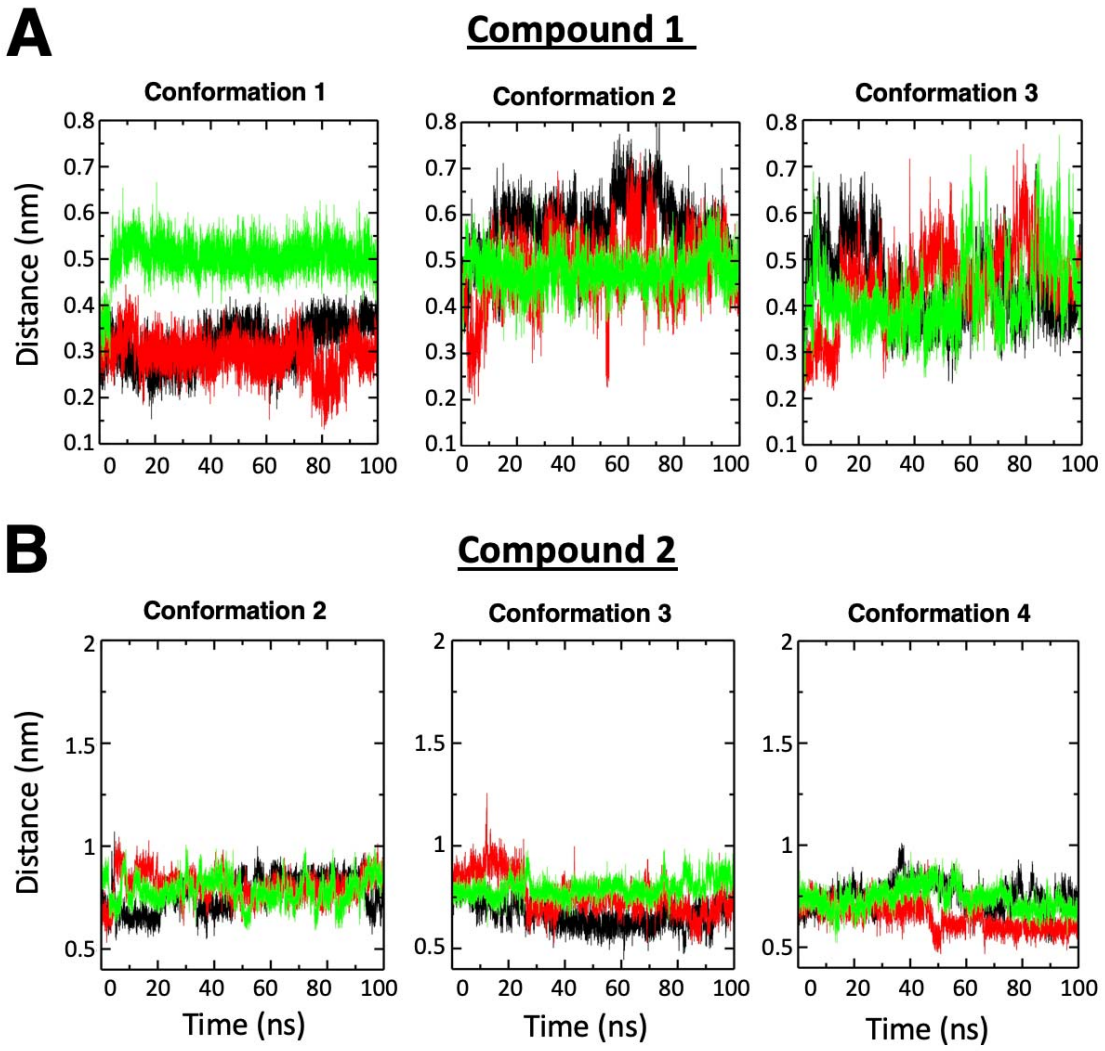

Figure 4. The center-of-mass distances between ligand and the respective binding site on RT. Data are shown for **(A)** compound 1 and **(B)** compound 2, calculated using independent

triplicates (black, green and red) of 100 ns trajectories. For compound 2, only the results of the three setups that exhibited stable ligand binding (i.e. conformation 2, 3, and 4) are shown. The remaining data are shown in Supplementary Figure S2.

DISCUSSION

We set out to search for novel HIV-1 inhibitors by experimentally screening 40 compounds of the NCI Diversity Set V and shortlisted two compounds that could inhibit the HIV-1 RT activity in reverse transcription assays. Given that RT is unique to viruses and absent in mammalian cells, the potential inhibitors are unlikely to cause side effects (unpublished limited mice experiments showed no toxicity effects at dosages up to 7 mg/kg for both compounds).

Our initial blind docking results found two potential binding sites (on p66 and p51 subunits) for compound 1 (Figure 3). Structural analyses demonstrated that compound 1 bound energetically favourably to the binding site on p66, which was also supported by our experimental functional assay when compound 1 inhibited the activity of p66 homodimers (Figure 2A). Compound 1 was found to bind to the p66 site predominantly via hydrophobic interactions. On the other hand, compound 2 bound to the p51 subunit, driven predominantly by electrostatic interactions with its more polar pocket (Supplementary Table S2).

It was computationally shown that both compounds interact with the RT p51 subunit (with less pronounced propensity for compound 1 at its second binding site in p51), distant from the polymerase site located on p66 (Figure 3A), suggesting an underlying allosteric mechanism. Interestingly, compound 2 established stable hydrogen bonds with residue Y183 of the p51 aside from other contacts (Figure 3B). To date, there are no known mutations reported to occur at this residue of p51[23], suggesting conservation and the possibility for compound 2 to be used against HIV-1 variants with resistance against known NNRTIs.

It should be noted that compound 2 also interacted weakly (e.g. via conformations 5 and 6, as shown in Supplementary Figure S2) with another site on the p51 subunit with a similar binding mode (data not shown) to those at the other dominant site; nonetheless, this affirmed that compound 2 would target p51 rather than p66.

We found the first binding site on p51 for compound 2 overlapped with our previously predicted allosteric pocket[21]. Since the structural fold of p51 differs from that of p66 (e.g. overall RMSD ~4.93 Å, excluding the RNase H region), this further supports that the main inhibition site of compound 2 is located on p51. Previous studies[24, 25] showed that p66 or p51 homodimers of HIV-1 RT exhibit RT activity, albeit at a reduced level. We were unable to detect p51 homodimer catalytic activity in our assay, but given the lack of inhibition on p66 homodimers, there is good evidence that the main activity would be that of p51, supporting our computational analysis (Figure 2A). Given that the DNA polymerase active site is on the p66 subunit, compound 2 is a potential novel type of NNRTI acting via p51, thus explaining why its inhibitory effect occurs at a relatively higher concentration of 300 μM compared to the direct action of compound 1 on p66 at 150 μM, albeit also a relatively high concentration.

Given the high concentrations required for the inhibitory effect, the two compounds are not likely to be effective inhibitors in themselves. However, they provide scaffolds and a structural description of key interactions that can guide the modification and development of new inhibitors. Most importantly, compound 2 confirmed the possibility of targeting the p51 subunit at a novel site that we previously identified computationally[21].

In addition, our structural conservation analysis revealed that the compound 2 binding region on HIV-1 p51 is conserved across multiple RTs (Figure 5 and Supplementary Data 1), highlighting its significance. Since viruses can develop cross-resistance to multiple drugs[21, 26], and Y183 has yet to have a reported clinical mutation, compound 2 serves as a potentially

important guide for targeting drug-resistant HIV-1 variants, and with the conservation of the site across RTs – a possible broad-spectrum RTI as well. By taking a holistic approach[27] in studying the whole structure of the target proteins and looking for common sites and how resistance develops, as demonstrated for HIV-1 Gag[28], HIV-1 Protease[26], and antibodies[29-33], the search for common druggable conserved sites across protein families can add to the success of developing broad-spectrum therapeutics.

In conclusion, we have found two potential inhibitory leads using detailed computational interaction data validated by experiments for HIV-1 RT inhibition, which could serve as scaffolds for new HIV-1 RT drugs and even potential broad-spectrum RTIs. Concomitantly, we also demonstrated the possibility of inhibiting HIV-1 RT by targeting the p51 subunit.

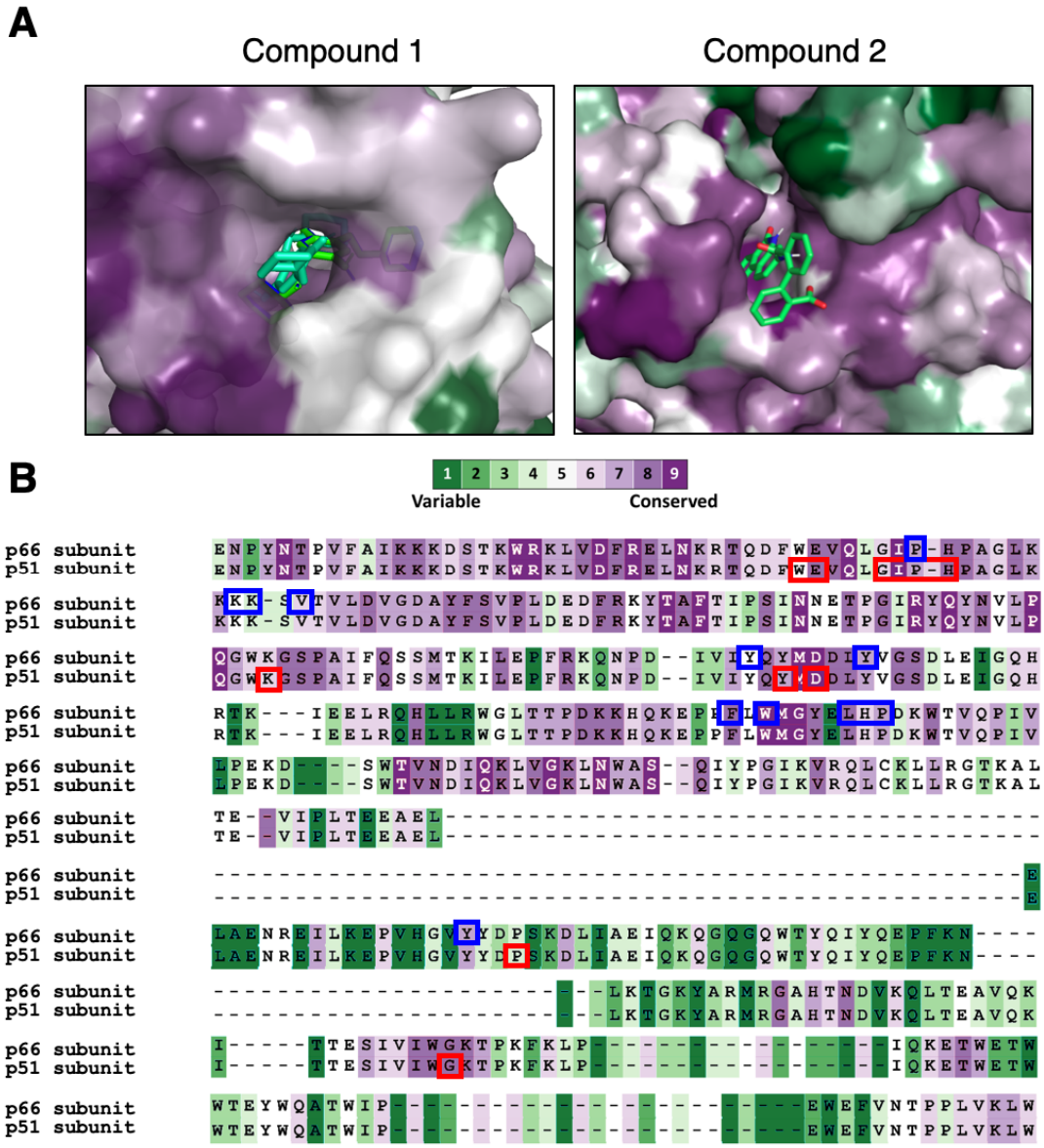

Figure 5. The two binding sites on HIV-1 RT that are highly conserved. **(A)** The binding sites on p66 and p51 complexed with compound 1 and 2, respectively, are shown in surface representation (coloured according to the level of conservation among the RT families). The two potential inhibitory compounds are shown as sticks. **(B)** Sequence alignment of the binding sites on p66 and p51, for compound 1 (blue boxes) and compound 2 (red boxes). The conservation colour scheme is the same as in (A). An augmented reality presentation of the

binding pockets could be viewed using our “APD AR Holistic Review” App available on both Google and Apple app stores (for more details, see Poh et al. [34])

METHODS

RNA Extraction

RNA was extracted from EXPI293F (Invitrogen, Cat no. A14527) cells using TRIzol Reagent (ThermoFisher Scientific, Cat. No.: 15596018) according to the manufacturer's recommendations. The RNA was treated with RNase-free DNase I (Roche, Cat. No.: 04716728001) and stored at -80 °C.

Library Reagent Preparation

NCI Diversity Set V chemical compounds from the National Cancer Institute (NCI) Developmental Therapeutic Program's Open Compound Repository, NIH (http://dtp.cancer.gov) were reconstituted into 1 mM concentrations in dimethyl sulfoxide (DMSO) and stored at -20 °C. The chemicals were initially screened at 300 μM, and potential hit compounds were subsequently diluted to lower concentrations: 150 μM, 30 μM, 3 μM and 0.3 μM.

cDNA synthesis Reverse Transcription (HIV-1)

All the library compounds were first heated to 37 °C for 10 min for better solubility prior to addition into the RT reaction mix. For HIV-1 RT, SinoBiological HIV-1 Reverse Transcriptase p51 (isolate HXB2, Cat No.: 40244-V07E) and p66 (isolate HXB2, Cat No.: 40244-V07E1) subunits were both used to form the RT complex for cDNA synthesis. The 20 μL cDNA synthesis mix is as follows: 0.5 μL of random hexamers (ThermoFisher Scientific, Cat No.: SO142, 100μM), 0.5 μL of Oligo d(T)s (ThermoFisher Scientific, Cat No.: SO131, 100 μM), 1.0 μL of Ribolock RNase Inhibitor (ThermoFisher Scientific, Cat No.: EO0381, 40 U/μL), 1.0 μL of dNTP mixture (1st Base, Cat. No.: BIO-5120, 10 mM of each dNTP), 4.0 μL of 5 × RT Buffer (250 mM Tris-HCl pH8.3, 375mM KCl, 15mM $MgCl_2$, 0.1M DTT),

2 μL of DEPC-Treated Water, 0.5 μL of RT p51 (0.17 μg/μL), 1.0 μL of RT p66 (0.33 μg/μL), 6.0 μL of candidate inhibitors and 3.0 μg of DNase-treated RNA. The reaction mixes were prepared on ice, followed by incubation in ProFlex 3x 32-Well PCR System (Applied Biosystems) at 25 °C for 18 min, followed by 37 °C for 1 hour, and RT termination at 85 °C for 5 min. Reaction mixtures containing either water or DMSO only without RNA were used as negative controls.

Polymerase Chain Reaction (PCR) Gene Amplification

The GAPDH PCR reactions were setup in 27.0 μL reactions, containing 13.5 μL of GoTaq Green Master Mix (Promega, Cat No.: M7123), 1.5 μL of Human GAPDH Forward Primer (5''–AGAAGGCTGGGGCTCATTTG–3'), 1.5 μL of Human GAPDH Reverse Primer (5'–AGGGGCCATCCACAGTCTTC–3') and 2 μL of cDNA template. Thermocycler conditions were set at 95 °C (2 min), 20 cycles of 95 °C (30 s), 50 °C (30 s), and 72 °C (1 min 30 s). The final extension step was performed at 72 °C for 7 min. Agarose gel electrophoresis analysis were carried out on the PCR products using 2% agar gels and GAPDH amplified products. Band sizes were confirmed using GelApp[19].

Structural Docking

The chemical structures of the two compounds (NSC48443 and NSC127133 from the NCI/DTP Diversity Set V) were retrieved from PubChem library, CID: 241217 and 278037 (particularly from the ligand code 6B3 in this entry) for compounds 1 and 2, respectively. Wild type HIV-1 RT structure (PDB: 3T19) was processed as previously described[21].

We performed rigid blind docking using the program Autodock Vina[20] to find potential binding sites in the HIV-1 RT structure for the two identified compounds. We next performed focused docking with flexible side chains of residues involved in the binding sites (as

receptor) derived from the initial blind docking. The compound bound-conformations that were significantly different and had high docking scores were selected for further analysis and subsequent MD simulations.

MD simulations of HIV-1 RT complexes.

We simulated the whole HIV-1 RT p66/p51 with the selected compound conformations using GROMACS (version 2018)[35], with the CHARMM36m[36] forcefield for the protein, the CHARMM[37] general force field for the ligand, and the TIP3P[38] water model. Nose-Hoover [39] and Parrinello-Rahman[40] thermostat and barostat were used to maintain constant temperature and pressure at 300 K and 1 bar, respectively. The particle mesh Ewald (PME) method was used for electrostatics, with a real-space cut-off of 12 Å. A 12 Å cut-off was used to calculate the van der Waal's potential, switching the energy function after 10 Å. Bonds involving hydrogen atoms were restrained using the LINCS[41] algorithm. We carried out a 20 ns equilibration, restraining the protein backbone atoms, followed by unrestrained simulations for 100 ns in triplicate.

Conventional MD simulation analysis

Hydrophobic contacts were calculated using a 4.0 Å cut-off between the ligand heavy atoms and protein carbon atoms. The number of water molecules around the ligand was determined using a cut-off of 3.5 Å between the ligand heavy atoms and water oxygen atoms. Hydrogen bonds were calculated between donor and acceptors of the ligand and the protein using a cut-off of 3.5 Å, with a maximum angle between donor – hydrogen atom and acceptor of 30°. To determine stable positions of the ligand in each binding site, we calculated distances between the centre of mass of the ligand and of the binding site residues, which were defined via the LigPlot+ program[22].

Binding energy calculation

The molecular mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method was applied to calculate the binding energies of the compounds and their corresponding binding sites on the HIV-1 RT subunits, using *g_mmpbsa*[42]. Various internal dielectric constants of the solute[43-45] (ε_{in}), were used to estimate the polarizability of the binding sites, e.g. $\varepsilon_{in} = 2$ (hydrophobic), $\varepsilon_{in} = 8$ (polar), or $\varepsilon_{in} = 20$ (highly charged). The external dielectric constant ε_{out} was set to 80 (for water). The calculations were performed using the last 50 ns of each MD simulation trajectory for all independent replicas.

Analysis of structural conservation among RT proteins

Multiple sequence alignment of the HIV-1 RT sequences with 152 retrovirus RT sequences (retrieved from NCBI RefSeq Databases) was first performed using ClustalW[46]. The resulting alignment was then subjected to the ConSurf server[47, 48] with default parameters to study the conservation of the regions of interest across the RT families.

DATA AVAILABILITY

All data generated or analysed during this study are included in this published article (and its Supplementary Information files)

REFERENCES

1. WHO, Number of people (all ages) living with HIV | Estimates by country, http://apps.who.int/gho/data/view.main.22100WHO?lang=en, 2019).
2. C. S. Adamson, *Molecular Biology International*, 2012, **2012**.
3. M. A. Parniak and N. Sluis-Cremer, in *Advances in Pharmacology*, Elsevier, 2000, vol. 49, pp. 67-109.
4. Y. Pommier, A. A. Johnson and C. Marchand, *Nature Reviews Drug discovery*, 2005, **4**, 236.
5. K. Princen and D. Schols, *Cytokine & Growth Factor Reviews*, 2005, **16**, 659-677.
6. K. Qian, S. L. Morris-Natschke and K. H. Lee, *Medicinal Research Reviews*, 2009, **29**, 369-393.
7. K. B. Larson, K. Wang, C. Delille, I. Otofokun and E. P. Acosta, *Clinical Pharmacokinetics*, 2014, **53**, 865-872.
8. WHO, *Updated recommendations on first-line and second-line antiretroviral regimens and post-exposure prophylaxis and recommendations on early infant diagnosis of HIV: interim guidelines: supplement to the 2016 consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection*, World Health Organization, 2018.
9. F. Maggiolo, N. Gianotti, L. Comi, E. Di Filippo, L. Fumagalli, S. Nozza, L. Galli, D. Valenti, M. Rizzi and A. Castagna, *Journal of Antimicrobial Chemotherapy*, 2020, DOI: 10.1093/jac/dkaa018.
10. Y. Cheng, G. E. Dutschman, K. F. Bastow, M. Sarngadharan and R. Ting, *Journal of Biological Chemistry*, 1987, **262**, 2187-2189.
11. P. Huang, D. Farquhar and W. Plunkett, *Journal of Biological Chemistry*, 1992, **267**, 2817-2822.
12. Y. Hsiou, J. Ding, K. Das, A. Clark Jr, S. Hughes and E. Arnold, *Structure*, 1996, **4**, 853-860.
13. Q. Xia, J. Radzio, K. S. Anderson and S. C. Nicolas, *Protein Science*, 2007, **16**, 1728-1737.
14. M. J. Young, *Frontiers in Molecular Biosciences*, 2017, **4**, 74.
15. J. M. Seckler, M. D. Barkley and P. L. Wintrode, *Biophysical Journal*, 2011, **100**, 144-153.
16. C. F. Silberstein, F. Gago, M. Santoro, C. Gori, V. Svicher, F. Rodríguez Barrios, R. d'Arrigo, M. Ciccozzi, A. Bertoli and A. d. A. Monforte, *Journal of Virology*, 2005, **79**, 10718-10729.
17. S. G. Sarafianos, B. Marchand, K. Das, D. M. Himmel, M. A. Parniak, S. H. Hughes and E. Arnold, *Journal of Molecular Biology*, 2009, **385**, 693-713.
18. T. Suzuki, P. J. Higgins and D. R. Crawford, *Biotechniques*, 2000, **29**, 332-337.
19. J. Z. Sim, P. V. Nguyen, H. K. Lee and S. K. E. Gan, *Nature Methods Application Notes*, 2015.
20. O. Trott and A. J. Olson, *Journal of computational chemistry*, 2010, **31**, 455-461.
21. R. Z. H. Chiang, S. K. E. Gan and C. T. T. Su, *Bioscience Reports*, 2018, **38**, BSR20171113.
22. R. A. Laskowski and M. B. Swindells, *Journal of Chemical Information and Modeling*, 2011, **51**, 2778-2786.
23. A. M. Wensing, V. Calvez, F. Ceccherini-Silberstein, C. Charpentier, H. F. Günthard, R. Paredes, R. W. Shafer and D. D. Richman, *Top Antivir Med*, 2019, **27**, 111-121.
24. X. Zheng, G. A. Mueller, M. J. Cuneo, E. F. Derose and R. E. London, *Biochemistry*, 2010, **49**, 2821-2833.

25. N. G. Sharaf, Z. Xi, R. Ishima and A. M. Gronenborn, *Proteins: Structure, Function, and Bioinformatics*, 2017, **85**, 2191-2197.
26. C. T. T. Su, W. L. Ling, W. H. Lua, Y. X. Haw and S. K. E. Gan, *BMC Bioinformatics*, 2016, **17**, 500.
27. S. X. Phua, K. F. Chan, Chinh T. T. Su, J. J. Poh and Samuel K. E. Gan, *Bioscience Reports*, 2019, **39**, BSR20181958.
28. C. T. T. Su, C. K. Kwoh, C. S. Verma and S. K. E. Gan, *Journal of Biomolecular Structure and Dynamics*, 2017, DOI: 10.1080/07391102.2017.1417160, 1-12.
29. W. H. Lua, C. T. T. Su, J. Y. Yeo, J. J. Poh, W. L. Ling, S. X. Phua and S. K. E. Gan, *Journal of Allergy and Clinical Immunology*, 2019.
30. W. L. Ling, W. H. Lua, J. J. Poh, J. Y. Yeo, D. P. Lane and S. K. E. Gan, *Frontiers in Immunology*, 2018, **9**.
31. W. H. Lua, W. L. Ling, J. Y. Yeo, J. J. Poh, D. P. S. Lane and S. K. E. Gan, *Scientific Reports*, 2018, **8**, 718.
32. C. T. T. Su, W. L. Ling, W. H. Lua, J. J. Poh and S. K. E. Gan, *Scientific Reports*, 2017, **7**.
33. C. T. T. Su, W. H. Lua, W. L. Ling and S. K. E. Gan, *Antibodies*, 2018, **7**, 20.
34. J. J. Poh, S. X. Phua, K. F. Chan and S. K.-E. Gan, *Scientific Phone Apps and Mobile Devices*, 2018, **4:4**.
35. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess and E. Lindahl, *SoftwareX*, 2015, **1**, 19-25.
36. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller and A. D. MacKerell, Jr., *Nat Methods*, 2017, **14**, 71-73.
37. K. Vanommeslaeghe, E. Hatcher, C. Acharya, S. Kundu, S. Zhong, J. Shim, E. Darian, O. Guvench, P. Lopes, I. Vorobyov and A. D. Mackerell, Jr., *Journal of computational chemistry*, 2010, **31**, 671-690.
38. W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *The Journal of Chemical Physics*, 1983, **79**, 926-935.
39. D. J. Evans and B. L. Holian, *The Journal of Chemical Physics*, 1985, **83**, 4069-4074.
40. M. Parrinello and A. Rahman, *Journal of Applied Physics*, 1981, **52**, 7182-7190.
41. B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *Journal of Computational Chemistry*, 1997, **18**, 1463-1472.
42. R. Kumari, R. Kumar and A. Lynn, *Journal of Chemical Information and Modeling*, 2014, **54**, 1951-1962.
43. S. Genheden and U. Ryde, *Proteins: Structure, Function, and Bioinformatics*, 2012, **80**, 1326-1342.
44. T. Yang, J. C. Wu, C. Yan, Y. Wang, R. Luo, M. B. Gonzales, K. N. Dalby and P. Ren, *Proteins*, 2011, **79**, 1940-1951.
45. C. N. Schutz and A. Warshel, *Proteins: Structure, Function, and Bioinformatics*, 2001, **44**, 400-417.
46. J. D. Thompson, D. G. Higgins and T. J. Gibson, *Nucleic Acids Research*, 1994, **22**, 4673-4680.
47. H. Ashkenazy, E. Erez, E. Martz, T. Pupko and N. Ben-Tal, *Nucleic Acids Research*, 2010, **38**, W529-W533.
48. G. Celniker, G. Nimrod, H. Ashkenazy, F. Glaser, E. Martz, I. Mayrose, T. Pupko and N. Ben-Tal, *Israel Journal of Chemistry*, 2013, **53**, 199-206.

Acknowledgements

This research was funded by A*STAR core fund. The forty compounds tested were provided by National Cancer Institute (NCI) Developmental Therapeutic Program's Open Compound Repository, NIH. We thank Ling Wei Li, Lua Wai Heng, Poh Jun Jie and Yeo Joshua for their help in performing independent replicates of the RT-PCR experiments.

Author Contributions

K.F.C. and S.X.P. performed the RT-PCR experiments. A.K. performed the molecular simulations. K.F.C., C.T.T.S., A.K., P.J.B, and S.K.E.G. analysed the results and wrote the manuscript. C.T.T.S. and S.K.E.G. conceived and supervised the study. All authors read and approved the final version of the manuscript.

Competing Interests

The authors declare no competing interests.