Inhibiting HIV-1 and MMLV Reverse Transcriptase: The potential of an Allosteric Broad-Spectrum Inhibitor

Kwok-Fong Chan[1], Ser-Xian Phua[1], Chinh Tran-To Su[1,*], Samuel Ken-En Gan[1,*]

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

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

Bioinformatics Institute, A*STAR

30 Biopolis Street, #07-01 Matrix

Singapore 138671

Tel: +65 6478 8301 (C. T-T. S.)

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

KEYWORDS: HIV-1 Reverse Transcriptase; MMLV Reverse Transcriptase; Drug Discovery; Broad-spectrum inhibitor; Cross-Reactive

ABSTRACT

Since drug resistance mutations in HIV-1 have been increasingly reported to resist most of the current drug repertoire in the antiretroviral therapy, there is a demand for new drugs. In this study, we focused on the viral enzyme Reverse Transcriptase (RT) as it is a good drug target given its absence in non-viruses. Through a reverse transcription assay screen, we found two out of forty compounds from the NCI Diversity Set V to inhibit HIV-1 RT activity. The less potent compound also inhibited MMLV RT. Molecular docking, structural conservation and binding pocket analyses suggested similar binding mechanisms of the dual inhibitor to the targets, implying that the phenylbenzoic scaffold may be potentially used to design broad-spectrum inhibitors against multiple Reverse Transcriptase enzymes from multiple viruses.

INTRODUCTION

Human Immunodeficiency Virus (HIV) infection remains one of the major global epidemics today with millions infected[1]. The current cocktail antiretroviral therapy (ART) targets various stages of the viral life cycle, such as using Protease Inhibitors (PIs to interfere 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 PI is the most used drug class in ART due to lower rates of induced drug resistance, it is a significant challenge in terms of treatment cost despite being the recommended first line of treatment by WHO[8]. On the other hand, the RTI drug class (including nucleoside (NRTI) and non-nucleoside (NNRTI) inhibitors) within ART is more affordable, and is safer since it targets RT that is unique to viruses, making it more suitable than PIs for a subset of the treatment population (e.g. children, pregnant women, or individuals treated tuberculosis with rifampicin).

The two RTI categories: NRTIs and NNRTIs inhibit RT function by different mechanisms. NRTIs bind to the nucleotide binding site and block the reverse transcription by competing with dNTPs[9], terminating the DNA elongation process[10] and thereby viral replication. In contrast, NNRTIs bind to a pocket located near the polymerase active site, and change the structural alignment of the dNTPs and template/primer substrates at the RT "primer grip", thus inducing distal effects on the polymerase active site[11], consequently inhibiting reverse transcription[12]. Unlike NRTIs, which are known to interfere with cellular replication machinery[13], NNRTIs vary in structure, generally do not exhibit off-target effects, and have been found to non-competitively allosterically cause conformational changes in the RT

structure[14]. To resist NNRTIs, HIV mutates at the drug-binding pocket[15] that is less conserved.

In the battle against viral drug resistance, broad spectrum antivirals with high barrier to resistance have been proposed as an alternative solution to inhibit multiple virus families as they bind to binding sites with common features and conserved structural sites[16-18]. For example, broad spectrum antiviral inhibitors[19-21] that block multiple viral families protease activity (such as Xenotropic murine leukemia virus-related virus, Human T-lymphotropic virus, and Feline immunodeficiency viruses) were also blocked HIV-1 protease. In the case of RT, although several RT broad spectrum inhibitors were found against various HIV-1 strains[22-24], there have been so far none against across multiple retrovirus families. Thereby, targeting conserved regions across all viral RT may also be effective within drug resistant variants.

HIV-1 RT is a heteromeric complex, made up of the p66 and p51 subunits. The larger p66 subunit contains five subdomains: fingers, palm, thumb, connection and RNaseH subdomains. Despite having an identical sequence to that of p66 excluding the RNaseH subdomains, the p51 subunit forms a more compact structure[25]. To date, no inhibitors that target the p51 subunit are used in the clinics.

Given that WHO guidelines recommends the use of NNRTIs, a guide to NNRTI drug selection usage and novel allosteric druggable sites in HIV-1 RT[26] was proposed previously which may alleviate drug resistance. Furthering this, we experimentally screened a diverse set of chemical compounds using the RT-PCR assay. In addition to computational molecular docking, structural conservation and inhibitor-binding pocket analyses, we further studied the mechanisms of how the inhibitors bound to the RT structures, finding a possible scaffold for broad spectrum RT inhibitor design that may be effective across mutant variants of HIV-1 RT and that of other viruses.

RESULTS

Two novel compounds that could inhibit RT activity

GAPDH housekeeping gene PCR was used as cDNA synthesis target given its high and consistent expression[27]. Preliminary 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 (thereafter named compound 1 and 2, respectively) inhibited HIV-1 RT cDNA synthesis (Figure 1). The inhibition did not occur during the GAPDH PCR step (Supplementary Figure S1A) given that bands were still observed when the inhibitors were only added after the cDNA synthesis. A non-inhibitor from the preliminary RT-PCR screen (NSC6145) was randomly picked as a negative control for the experiments. Compounds 1 and 2 were found to partially inhibit HIV-1 RT in independent triplicates at 150 and 300 μM respectively (Figure 1) and PCR products were observed to decrease at higher concentrations.

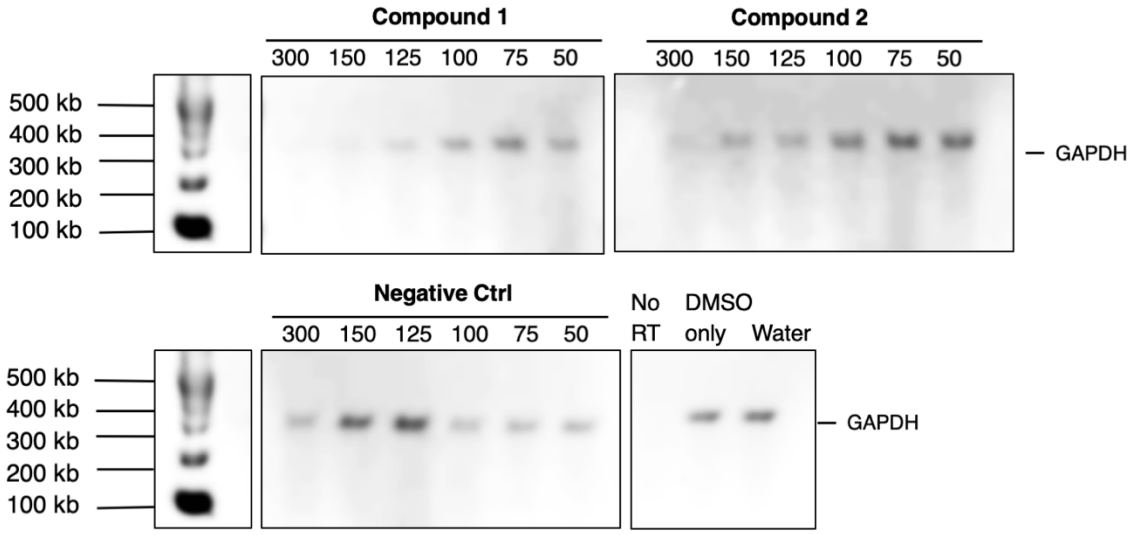

Figure 1. RT-PCR (HIV-1 RT) representative results of triplicate experiments on GAPDH gene with the two identified compounds. The experiments were performed at various concentrations (50 - 300 μM) of the two inhibitor compounds with the negative controls.

Broad-spectrum RT inhibitors

In order to determine if the two compounds could function as broad-spectrum RT inhibitors, we performed similar RT-PCR functional assays on another RT enzyme (i.e. Moloney Murine Leukemia Virus (MMLV) RT available in our lab). Among the two compounds, only compound 2 was found to inhibit the MMLV RT activity at concentrations above 400 μM (Figure 2).

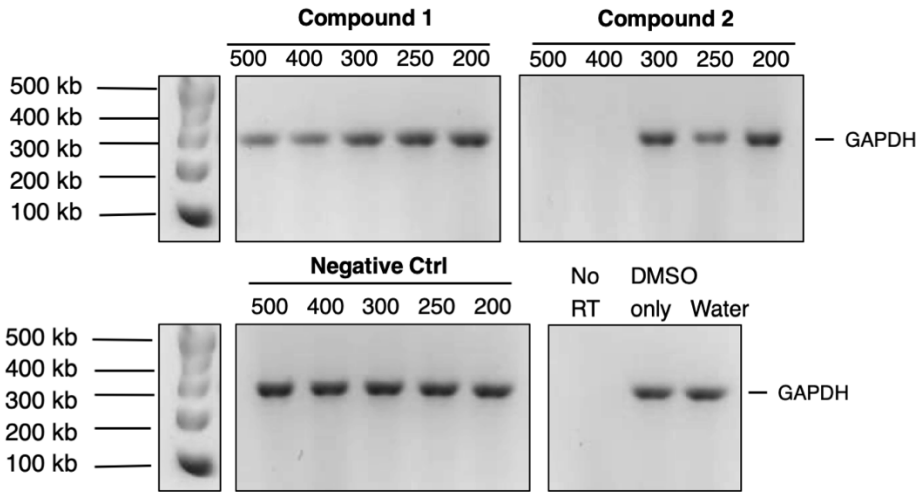

Figure 2. MMLV RT cDNA Synthesis PCR assays of the two compounds at various concentrations using MMLV RT. Results shown here is representative of triplicate experiments.

We next performed computational blind dockings using the Achilles Blind Docking Server [28,29]. The dockings were individually performed using compound 2 structure (as ligand) to both the RT structures (as receptors) without specifying the binding regions. It was shown that compound 2 bound to several regions in HIV-1 RT, mostly on the p51 domain, as well as specific regions in MMLV RT (Figure 3A). Among the binding sites on both HIV-1 and MMLV RT, we found that one particular site shared structural similarity between HIV-1 p51 domain and MMLV RT where compound 2 conformers were highly populated (~11.17% and ~9.22%, respectively, of the total conformer sampling). Docking affinity scores of the

compound 2 conformers to the HIV-1 RT p51 and MMLV RT binding sites were also found among the top highly ranked results in both HIV-1 (rank 2) and MMLV (rank 7) (Supplementary Table S5).

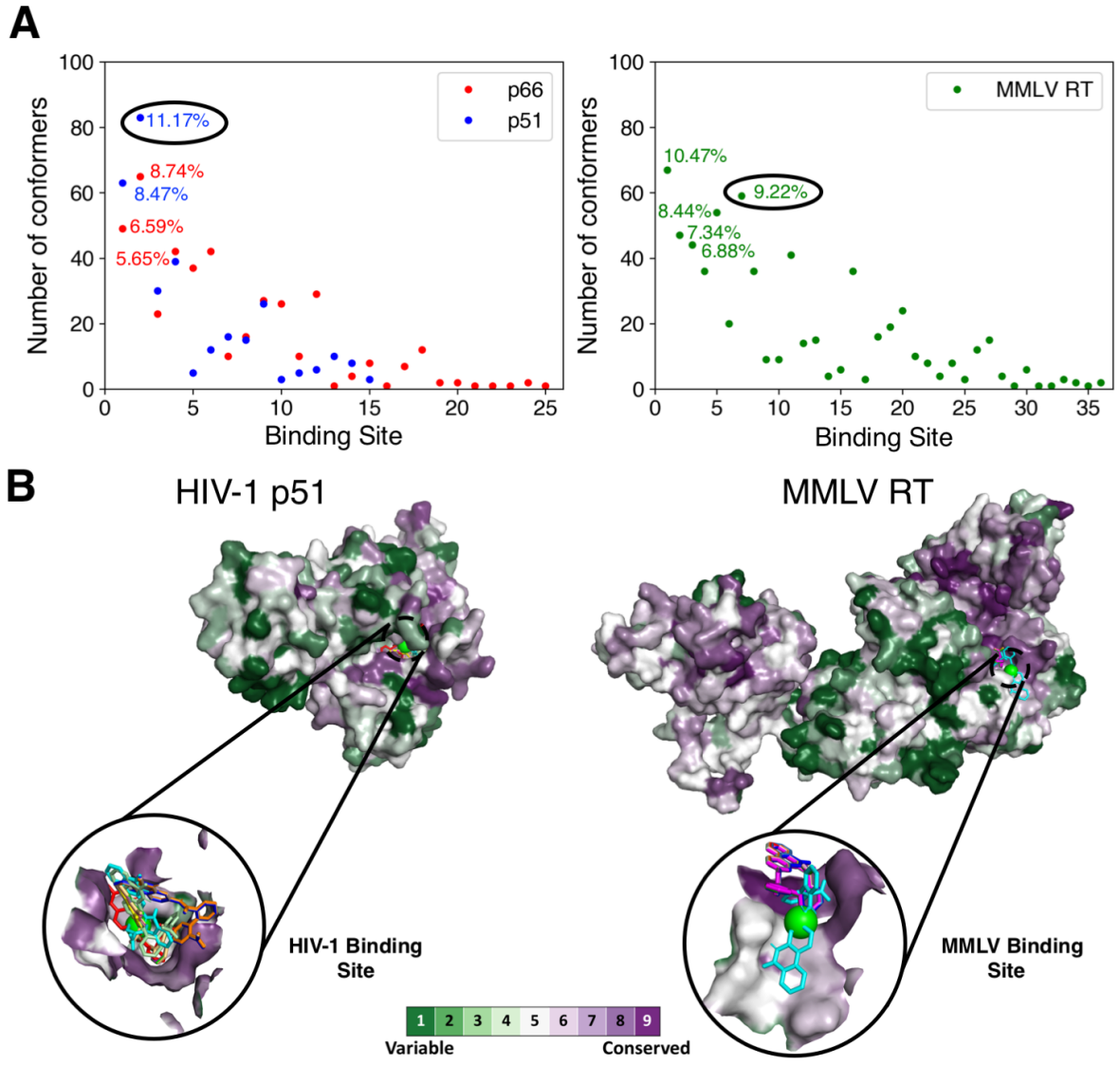

Figure 3. Computational structural analysis of compound 2 bindings to HIV-1 and MMLV RT regions. (A) Sampling distribution of the binding conformers in both the RT structures, with % indicating a few top binding regions that compound 2 populated at high frequency. The binding sites of interest are circled in black. (B) Analysis of sequence and structural conservation of the common binding regions using the ConSurf[30,31] server indicates similar binding mechanisms of the compound 2 to both the viral RT structures. An animated

presentation of the MMLV and HIV-1 RT-inhibitor complexes could be viewed using the "APD AR Holistic Review App" available on both Google and Apple app stores (for more details, see Poh et al.[32]).

In addition, we found that the common binding site on both HIV-1 and MMLV RT were conserved among the RT protein family (Figure 3B, Supplementary Figure S2 and S3). It was also shown that compound 2 has a phenylbenzoic acid group that was likely to be involved in these bindings to both the regions that contain a predominantly charged region adjacent to a hydrophobic patch in both the HIV-1 p51 and MMLV RT (Supplementary Figure S4). This phenylbenzoic acid group might be crucial to the binding of compound 2 (but absent in the 39 other compounds tested) to both the RT regions.

DISCUSSION

We set out to search for novel HIV-1 NNRTI alternatives by experimentally screening 40 compounds of the NCI Diversity Set V. We identified two compounds that could inhibit the HIV-1 RT activity, and of the two compounds, the less potent compound 2 also inhibited MMLV RT function.

Computational structural analyses suggested a common binding site for compound 2 on both the RT structures that was conserved among the RT protein families as well as across viruses infecting different species, thereby suggesting the scaffolds for novel broad-spectrum RT inhibitor design. Both the binding regions in HIV-1 p51 and MMLV RT were found to be conserved among the RT families and they shared similar binding mechanism with compound 2 conformers, particularly via the phenylbenzoic acid group. Therefore, the phenylbenzoic acid group could be used as a scaffold template for the design of potent inhibitors against retroviral RT. Furthermore, our structural conservation analysis revealed that the common binding region on MMLV RT and HIV-1 p51 was also present across multiple Reverse Transcriptases (Supplementary Data 1).

In addition to p51 domain binding, compound 2 was also found to interact with p66 (at lower frequency, Figure 3A) away from the polymerase site in our blind docking analysis, suggesting the involvement of allostery. Since the structural folds of p51 differed from that of p66 subunits (e.g. overall RMSD ~4.93 Å, excluding the RNase H region) but shared similar structural characteristics on MMLV RT, it was likely that the main inhibition site by the compound 2 was located on p51 of HIV-1 RT.

Studies[33,34] showed that p66 or p51 homodimers of HIV-1 RT remained functional, albeit at reduced activity. Thus, it is likely that compound 2 might also be able to inhibit the p66 homodimer formation/activity. In our assay, we did not find the PCR reactions with the p51 alone to produce detectable bands after PCR, even in the negative control (Supplementary

Figure S1B). However, faint bands were detected for the p66 alone in the case of compound 2 (and also of the negative control), but there was none in the case of compound 1 (Supplementary Figure S1B). Such a result strongly supports the computational analysis that compound 2 was unlikely to bind to p66 but instead displayed its effect by binding to p51, whereas the inhibition of compound 1 was likely to be narrowed to p66 subunit. From this observation, the predominant inhibition of compound 2 on p51 may explain why compound 2 was only able to inhibit at much higher concentrations than compound 1 (Figure 1A). Given that the active site of RT is located on p66 for the HIV-1, the inhibition likely involves an allosteric mechanism.

With possible inhibition via the phenylbenzoic acid group on HIV-1 RT, there is great promise to further modify and design broad-spectrum RTIs, and possible safer ones with lower toxicity against multiple viral RTs and mutant HIV-1 RTs. The conservation of the site suggests its crucial role across viruses, thus making it an attractive target for drug-resistant HIV-1 RT variants. Just as viruses can develop cross-resistance to multiple drugs[26,35], it is possible to have cross-inhibition against multiple viral RT. With the potential of allosteric inhibitors, rational drug design can be more diverse beyond the scaffold of nucleosides in NRTIs and be targeted against specific RTs at different levels while maintaining the broad-spectrum activity. To do this, it would be necessary to take a holistic approach[36] in studying the whole structure of the target proteins. Such an approach has been shown to be important in the studies of HIV-1 Gag[37], HIV-1 Protease[35], and antibodies[38-42], in the search for allosterically important regions that if conserved, would be likely broad-spectrum druggable targets.

In conclusion, we found two potential inhibitors from our experimental screen that inhibited HIV-1 RT. One of the inhibitors demonstrated cross-inhibition to MMLV RT at a similar conserved binding site across viral RTs, and likely inhibited HIV-1 RT via its binding to the

p51 subunit, against which no known clinical drugs target. We propose that this compound could serve as a scaffold for inhibitor design as a broad-spectrum RTI, and certainly in the case of HIV-1, one of the first p51 targeting inhibitors.

METHODS

RNA Extraction

RNA was extracted from overnight cultured 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), was reconstituted into 1 mM concentrations with 100% dimethyl sulfoxide (DMSO) and stored at -20 °C. The chemicals were initially screened at 300 μM, and subsequently diluted (for potential hit compounds) in varying concentrations (150 μM, 30 μM, 3 μM and 0.3 μM).

cDNA synthesis Reverse Transcription (HIV-1 & MMLV)

All the library compounds were first heated to 37 °C for 10 min for better solubility prior to addition into 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 mix was 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, RT termination at 85 °C for 5 min. Reaction mixtures containing either water or DMSO only without RNA and inhibitors were used as negative controls.

For MMLV RT, Tetro Reverse Transcription kit (Bioline, Cat. No.: BIO 65050) was used in the cDNA synthesis as performed using HIV-1 RT.

Polymerase Chain Reaction (PCR) Gene Amplification

The total volume of the GAPDH PCR reaction mixture was set up 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). Final extension step was performed at 72 °C for 7 min. Agarose Gel Electrophoresis was carried out on the PCR products using a 2% agar gel and GAPDH amplified products were analysed and confirmed using GelApp[43].

Structural Analysis & Blind Docking

The chemical structure of compound 2 (NSC127133 from the Diversity Set V) was retrieved from the National Cancer Institute Developmental Therapeutic Program (NCI/DTP). The 2D structures were then converted into 3D structures using Open Babel[44]. Wild type HIV-1 RT

structure (PDB: 3T19, preprocessed as previously[26]) and MMLV RT structure (PDB: 5DMQ) were used.

Missing residues were first added into the MMLV RT using Modeller[45] and Swiss-Pdb Viewer[46] and followed by energy minimization using UCSF Chimera[47]. The blind docking calculations were performed using HIV-1 RT and MMLV RT structures as receptors and the converted compound 2 structure as ligand using Achilles Blind Docking Server[28,29] (available at http://bio-hpc.eu/software/blind-docking-server/), which have demonstrated successful prediction of binding poses in various academic studies[48,49] and industrial projects[28]. We selected the binding regions on the receptors with the most populated ligand bound-conformers for analysis. To minimize possible confounding factors (e.g. disrupting the HIV-1 RT structural stability by interfering the interactions between p66 and p51 domains), we excluded the binding regions at the HIV-1 RT p66/p51 interface.

Analysis of structural conservation among RT proteins

Multiple sequence alignment of the HIV-1 and MMLV RT sequences with 152 retrovirus RT sequences (retrieved from NCBI RefSeq Databases) was first performed using ClustalW[50]. The resulting alignment was then subjected to the ConSurf server[30,31] with default parameters to study the conservation of the regions of interest across the RT families.

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> (2018).

2 Adamson, C. S. Protease-mediated maturation of HIV: inhibitors of protease and the maturation process. *Molecular Biology International* **2012** (2012).

3 Parniak, M. A. & Sluis-Cremer, N. in *Advances in Pharmacology* Vol. 49 67-109 (Elsevier, 2000).

4 Pommier, Y., Johnson, A. A. & Marchand, C. Integrase inhibitors to treat HIV/AIDS. *Nature Reviews Drug discovery* **4**, 236 (2005).

5 Princen, K. & Schols, D. HIV chemokine receptor inhibitors as novel anti-HIV drugs. *Cytokine & Growth Factor Reviews* **16**, 659-677 (2005).

6 Qian, K., Morris-Natschke, S. L. & Lee, K. H. HIV entry inhibitors and their potential in HIV therapy. *Medicinal Research Reviews* **29**, 369-393 (2009).

7 Larson, K. B., Wang, K., Delille, C., Otofokun, I. & Acosta, E. P. Pharmacokinetic enhancers in HIV therapeutics. *Clinical Pharmacokinetics* **53**, 865-872 (2014).

8 WHO. *Consolidated guidelines on the use of antiretroviral drugs for treating and preventing HIV infection: recommendations for a public health approach*. (World Health Organization, 2016).

9 Cheng, Y., Dutschman, G. E., Bastow, K. F., Sarngadharan, M. & Ting, R. Human immunodeficiency virus reverse transcriptase. General properties and its interactions with nucleoside triphosphate analogs. *Journal of Biological Chemistry* **262**, 2187-2189 (1987).

10 Huang, P., Farquhar, D. & Plunkett, W. Selective action of 2', 3'-didehydro-2', 3'-dideoxythymidine triphosphate on human immunodeficiency virus reverse transcriptase and human DNA polymerases. *Journal of Biological Chemistry* **267**, 2817-2822 (1992).

11 Hsiou, Y. *et al*. Structure of unliganded HIV-1 reverse transcriptase at 2.7 Å resolution: implications of conformational changes for polymerization and inhibition mechanisms. *Structure* **4**, 853-860 (1996).

12 Xia, Q., Radzio, J., Anderson, K. S. & Nicolas, S. C. Probing nonnucleoside inhibitor-induced active-site distortion in HIV-1 reverse transcriptase by transient kinetic analyses. *Protein Science* **16**, 1728-1737 (2007).

13 Young, M. J. Off-target effects of drugs that disrupt human mitochondrial DNA maintenance. *Frontiers in Molecular Biosciences* **4**, 74 (2017).

14 Seckler, J. M., Barkley, M. D. & Wintrode, P. L. Allosteric suppression of HIV-1 reverse transcriptase structural dynamics upon inhibitor binding. *Biophysical Journal* **100**, 144-153 (2011).

15 Silberstein, C. F. *et al*. High sequence conservation of human immunodeficiency virus type 1 reverse transcriptase under drug pressure despite the continuous appearance of mutations. *Journal of Virology* **79**, 10718-10729 (2005).

16 Kim, Y. *et al*. Broad-spectrum antivirals against 3C or 3C-like proteases of picornaviruses, noroviruses, and coronaviruses. *Journal of Virology* **86**, 11754-11762 (2012).

17 Patick, A. *et al*. In vitro antiviral activity of AG7088, a potent inhibitor of human rhinovirus 3C protease. *Antimicrobial Agents and Chemotherapy* **43**, 2444-2450 (1999).

18 De Palma, A. M. *et al*. Potential use of antiviral agents in polio eradication. *Emerging Infectious Diseases* **14**, 545 (2008).

19 Lee, T. *et al*. Analysis of the S3 and S3′ subsite specificities of feline immunodeficiency virus (FIV) protease: development of a broad-based protease inhibitor efficacious against FIV, SIV, and HIV in vitro and ex vivo. *Proceedings of the National Academy of Sciences* **95**, 939-944 (1998).

20 Kuhnert, M., Steuber, H. & Diederich, W. E. Structural basis for HTLV-1 protease inhibition by the HIV-1 protease inhibitor indinavir. *Journal of Medicinal Chemistry* **57**, 6266-6272 (2014).

21 Li, M. *et al*. Structural and biochemical characterization of the inhibitor complexes of xenotropic murine leukemia virus-related virus protease. *The FEBS Journal* **278**, 4413-4424 (2011).

22 Famiglini, V. *et al*. New indolylarylsulfones as highly potent and broad spectrum HIV-1 non-nucleoside reverse transcriptase inhibitors. *European Journal of Medicinal Chemistry* **80**, 101-111 (2014).

23 Campiani, G. *et al*. Pyrrolobenzoxazepinone derivatives as non-nucleoside HIV-1 RT inhibitors: further structure– activity relationship studies and identification of more potent broad-spectrum HIV-1 RT inhibitors with antiviral activity. *Journal of Medicinal Chemistry* **42**, 4462-4470 (1999).

24 Romero, D. L. *et al*. Targeting delavirdine/atevirdine resistant HIV-1: identification of (alkylamino) piperidine-containing bis (heteroaryl) piperazines as broad spectrum HIV-1 reverse transcriptase inhibitors. *Journal of Medicinal Chemistry* **39**, 3769-3789 (1996).

25 Sarafianos, S. G. *et al*. Structure and function of HIV-1 reverse transcriptase: molecular mechanisms of polymerization and inhibition. *Journal of Molecular Biology* **385**, 693-713 (2009).

26 Chiang, R. Z. H., Gan, S. K. E. & Su, C. T. T. A computational study for rational HIV-1 non-nucleoside reverse transcriptase inhibitor selection and the discovery of novel allosteric pockets for inhibitor design. *Bioscience Reports* **38**, BSR20171113 (2018).

27 Suzuki, T., Higgins, P. J. & Crawford, D. R. Control selection for RNA quantitation. *Biotechniques* **29**, 332-337 (2000).

28 Achilles Blind Docking Server (Bioinformatics and High Performance Computing (BIO-HPC) Research group, 2017).

29 Linares, S. I., Pérez, S. H., Cecilia, J. M. & García, J. M. High-throughput parallel blind virtual screening using BINDSURF. *BMC Bioinformatics* **13**, S13 (2012).

30 Ashkenazy, H., Erez, E., Martz, E., Pupko, T. & Ben-Tal, N. ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Research* **38**, W529-W533 (2010).

31 Celniker, G. *et al*. ConSurf: using evolutionary data to raise testable hypotheses about protein function. *Israel Journal of Chemistry* **53**, 199-206 (2013).

32 Poh, J. J., Phua, S. X., Chan, K. F. & Gan, S. K. E. *Scientific Phone Apps and Mobile Devices* **4:4** (2018).

33 Zheng, X., Mueller, G. A., Cuneo, M. J., Derose, E. F. & London, R. E. Homodimerization of the p51 subunit of HIV-1 reverse transcriptase. *Biochemistry* **49**, 2821-2833, doi:10.1021/bi902116z (2010).

34 Sharaf, N. G., Xi, Z., Ishima, R. & Gronenborn, A. M. The HIV-1 p66 homodimeric RT exhibits different conformations in the binding-competent and-incompetent NNRTI site. *Proteins: Structure, Function, and Bioinformatics* **85**, 2191-2197 (2017).

35 Su, C. T. T., Ling, W. L., Lua, W. H., Haw, Y. X. & Gan, S. K. E. Structural analyses of 2015-updated drug-resistant mutations in HIV-1 protease: an implication of protease inhibitor cross-resistance. *BMC Bioinformatics* **17**, 500 (2016).

36 Phua, S. X., Chan, K. F., Su, Chinh T. T., Poh, J. J. & Gan, Samuel K. E. Perspective: The promises of a holistic view of proteins—impact on antibody engineering and drug discovery. *Bioscience Reports* **39**, BSR20181958, doi:10.1042/BSR20181958 (2019).

37 Su, C. T. T., Kwoh, C. K., Verma, C. S. & Gan, S. K. E. Modeling the full length HIV-1 Gag polyprotein reveals the role of its p6 subunit in viral maturation and the effect of non-cleavage site mutations in protease drug resistance. *Journal of Biomolecular Structure and Dynamics*, 1-12, doi:10.1080/07391102.2017.1417160 (2017).

38 Lua, W. H. *et al*. Role of the IgE variable heavy chain in FcεRIα and superantigen binding in allergy and immunotherapy. *Journal of Allergy and Clinical Immunology* (2019).

39 Ling, W. L. *et al*. Effect of VH–VL Families in Pertuzumab and Trastuzumab Recombinant Production, Her2 and FcγIIA Binding. *Frontiers in Immunology* **9**, doi:10.3389/fimmu.2018.00469 (2018).

40 Lua, W. H. *et al*. The effects of Antibody Engineering CH and CL in Trastuzumab and Pertuzumab Recombinant Models: Impact on antibody production and antigen-binding. *Scientific Reports* **8**, 718 (2018).

41 Su, C. T. T., Ling, W. L., Lua, W. H., Poh, J. J. & Gan, S. K. E. The role of Antibody Vκ Framework 3 region towards Antigen binding: Effects on recombinant production and Protein L binding. *Scientific Reports* **7** (2017).

42 Su, C. T. T., Lua, W. H., Ling, W. L. & Gan, S. K. E. Allosteric Effects between the Antibody Constant and Variable Regions: A Study of IgA Fc Mutations on Antigen Binding. *Antibodies* **7**, 20 (2018).

43 Sim, J. Z., Nguyen, P. V., Lee, H. K. & Gan, S. K. E. GelApp: Mobile gel electrophoresis analyser. *Nature Methods Application Notes* (2015).

44 O'Boyle, N. M. *et al*. Open Babel: An open chemical toolbox. *Journal of Cheminformatics* **3**, 33 (2011).

45 Sali, A. & Blundell, T. L. Comparative protein modelling by satisfaction of spatial restraints. *Journal of Molecular Biology* **234**, 779-815 (1993).

46 Guex, N. & Peitsch, M. C. SWISS-MODEL and the Swiss-Pdb Viewer: an environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723 (1997).

47 Pettersen, E. F. *et al*. UCSF Chimera—a visualization system for exploratory research and analysis. *Journal of Computational Chemistry* **25**, 1605-1612 (2004).

48 Sinicropi, M. S. *et al*. N-thioalkylcarbazoles derivatives as new anti-proliferative agents: synthesis, characterisation and molecular mechanism evaluation. *Journal of Enzyme Inhibition and Medicinal Chemistry* **33**, 434-444 (2018).

49 Carrasco, C. J. P. *et al*. Labelling Herceptin with a novel oxaliplatin derivative: a computational approach towards the selective drug delivery. *Journal of Molecular Modeling* **20**, 2401 (2014).

50 Thompson, J. D., Higgins, D. G. & Gibson, T. J. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Research* **22**, 4673-4680 (1994).

Acknowledgements

This research was funded by the Bioinformatics Institute 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. K.F.C. performed the computational dockings and analysis. K.F.C., C.T.T.S. and S.K.E.G. analyzed 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 financial interests.