Article

Homology modeling of clinically-relevant rilpivirine-resistant HIV-RT variants identifies novel rilpivirine analogs with retained binding affinity against NNRTI-resistant HIV mutations

Charissa Luk[1,2], Jeslyn Wu[1,3], Edward Njoo[1]

[1]Department of Chemistry, Biochemistry, & Physics, Aspiring Scholars Directed Research Program, Fremont, CA
[2]Bishop O'Dowd High School, Oakland, CA
[3]Mission San Jose High School, Fremont, CA

SUMMARY

Human immunodeficiency virus (HIV), which affects tens of millions of individuals worldwide, can lead to acquired immunodeficiency syndrome (AIDS). Though there is currently no cure for HIV, the development of small molecule antiretroviral agents has greatly improved the prognosis of infected individuals, especially in developed countries. In particular, compounds such as rilpivirine have been developed as non-nucleoside reverse transcriptase inhibitors (NNRTIs), which allosterically target the reverse transcriptase enzyme in the retrovirus. However, mutations to the reverse transcriptase enzyme threaten to undermine the efficacy of this class of antivirals. Recent advances in computational biophysical modeling enable the structural analyses of such variants without the complications of obtaining crystal structures. Here, we employ homology modeling and molecular docking towards the identification of novel rilpivirine analogs that retain high binding affinity to clinically relevant rilpivirine-resistant mutations of the HIV reverse transcriptase enzyme.

INTRODUCTION

Human immunodeficiency virus (HIV), which can progress to acquired immunodeficiency syndrome (AIDS) without antiretroviral therapies, affects about 38 million people worldwide as of 2019 (1). When a person progresses to AIDS, the virus has caused a CD4+ T cell count lower than 200 cells per cubic millimeter of blood as compared to a normal count of 500 to 1600, resulting in a compromised immune system in the host (2). This loss of the immune system's function increases the host's susceptibility to opportunistic illnesses, cancers, and infectious diseases that would otherwise be well controlled.

While to this day there is no cure for HIV-infected individuals, the development of small molecule therapeutics has advanced antiretroviral therapy (ART), and this is advanced enough to allow infected individuals to live a normal life without developing AIDS (3). One specific drug class used for HIV treatment is that of non-nucleoside reverse transcriptase inhibitors (NNRTIs), which allosterically

inhibit the reverse transcriptase (RT) enzyme of HIV (4). This inhibition of the RT takes advantage of HIV's reliance on RT to complete the transformation of viral RNA into DNA compatible with the human system, thereby limiting the virus's ability to replicate within the host. Because RT is essential in the HIV life cycle, it is a key target for drug discovery. Rilpivirine, a second generation diarylpyrimidine (DAPY) NNRTI, has demonstrated superior potency and toxicology profiles compared to previous NNRTI's, and was approved for clinical use in the United States in 2011 (5–7) **(Figure 1a)**. Like other NNRTI's, rilpivirine binds to the allosteric binding pocket in reverse transcriptase, where it induces a conformational change in the protein's structure that prevents reverse transcription **(Figure 1b)**.

However, the emergence and prevalence of drug-resistant variants of HIV pose a threat to the continued success of antiretroviral therapy (8, 9). Mainly, mutations in the allosteric NNRTI binding pocket of HIV-1 RT have the potential to dramatically decrease binding affinity and antiretroviral efficacy of rilpivirine and other NNRTIs, therefore necessitating the continued development and discovery of novel chemical entities which retain activity against viral variants (10).

Molecular docking allows for predicted binding affinity, and through homology modeling, it is possible to model known mutant structures. Previously, we reported the identification of five top analogs of rilpivirine through a high-throughput virtual screen, wherein we described the results of screening the library of 2,4-diarylpyrimidines structurally analogous to rilpivirine (11). Here, we employ homology modeling to model the most common clinically-relevant rilpivirine-resistant HIV variants and to predict binding affinity of our reported hit compounds against. Additional, further analysis provides structural basis for the retention of high binding affinity of our compounds in the presence of key clinical variants through molecular docking experiments. We hypothesized that altering the 4-substituted aniline fragments on our analogs may remove dependence of binding affinity on key residue-ligand interactions, resulting in the retention of high binding affinity to the allosteric pocket of HIV-RT. Through this, we report the identification of a number of novel diarylpyrimidine (DAPY) small molecules that, by computer modeling, might
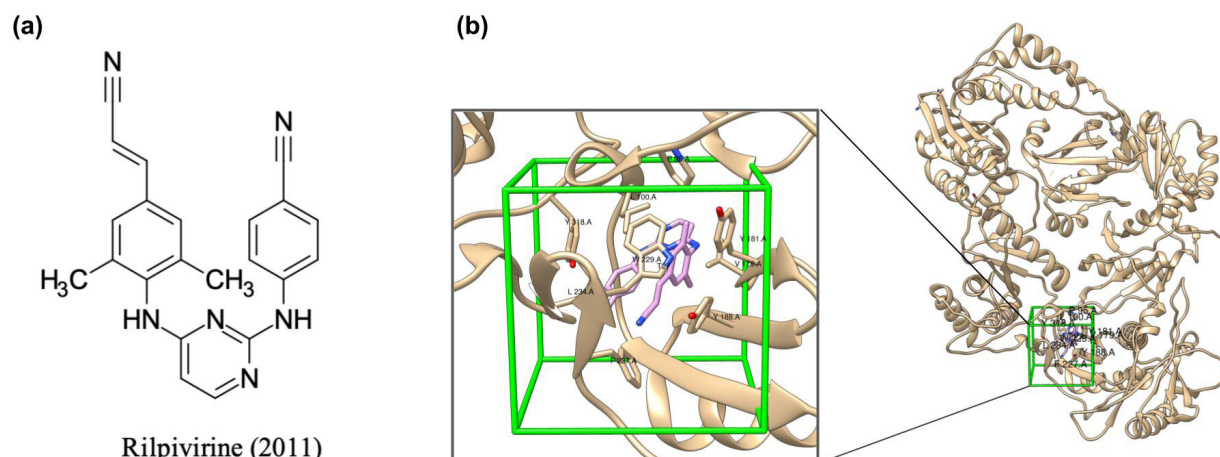
Figure 1: Rilpivirine is a diarylpyrimidine (DAPY) non-nucleoside reverse transcriptase inhibitor that binds to and allosterically inhibits the reverse transcriptase enzyme in the human immunodeficiency virus (HIV). (a) Chemical structure of rilpivirine, a diaryl pyrimidine (DAPY) second generation NNRTI approved for clinical use by the FDA in 2011; (b) Crystal structure binding pose (PDB:3MEE) of HIV-1 Reverse Transcriptase in Complex with TMC278 (Rilpivirine). Key residues in the HIV-1 reverse transcriptase allosteric NNRTI- binding pocket include K101, W229, Y181, and Y188.
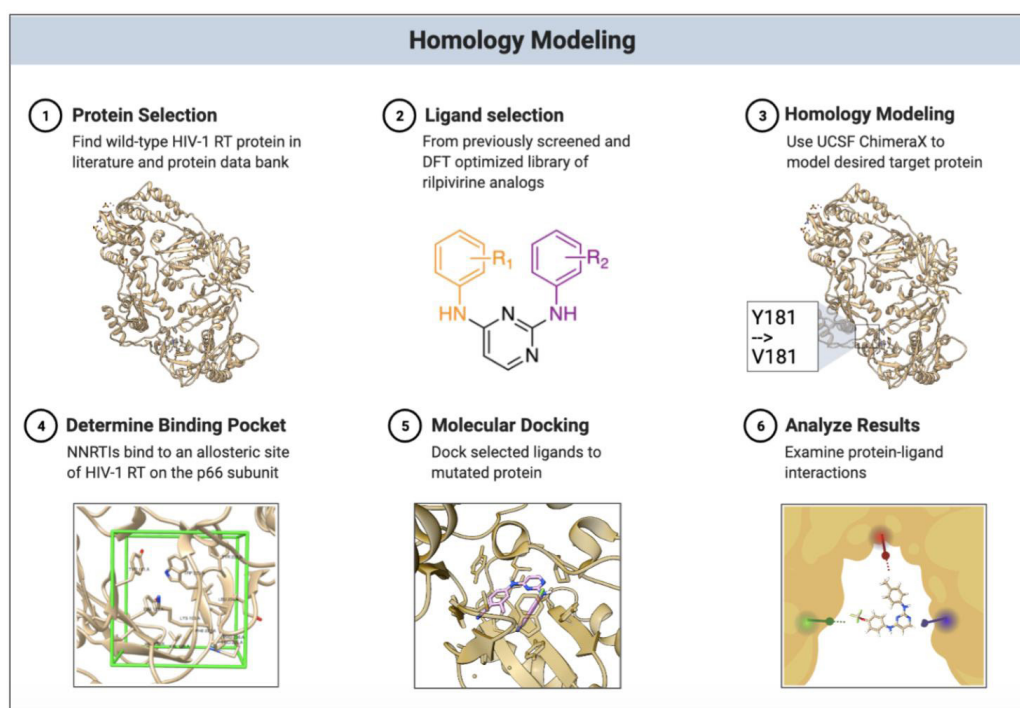

Figure 2: Homology modeling workflow for modeling HIV-1 RT variants. 1) An unliganded HIV-1 RT crystal structure of rilpivirine bound to a representative reverse transcriptase of HIV [PDB:3MEE (13)] was chosen. 2) Top performing analogs from our previously screened library were selected. 3) Different clinically-relevant rilpivirine-resistant mutations were modeled using UCSF ChimeraX. 4) The docking gridbox location was determined based on where rilpivirine is known to bind, and the size was determined based on the size of our analogs. 5) Molecular docking was used to determine the binding affinity of the analogs to mutated proteins. 6) Binding scores and poses of the analogs were analyzed to provide justification.

provide structural leads against rilpivirine-resistant HIV-RT mutants.

RESULTS

We utilized homology modeling to screen clinically relevant NNRTI-resistant variants **(Figure 2)**. A total of 23 structural variants of RT were visualized, including 16 which are associated with decreased rilpivirine efficacy (K101E/P, E138A/G/K/Q/R, V179L, Y181C/I/V, Y188L, H221Y, F227C, and M230I/L) and 5 common double and triple mutants

(L100I/K103N, L100I/K103R/V179D, K103N/Y181C, V106A/F227L, and V106A/Y181C) (10, 12) **(Figure 3)**. We calculated binding affinities from molecular docking experiments using our previously described gridbox parameters (11) based on the crystal structure of HIV-RT liganded with rilpivirine **(Figure 4a)**. The crystal structure of wild-type HIV RT (PDB: 3MEE) (13) acted as the base scaffold wherewith each variant was separately constructed via single-residue substitution homology models. We used the top five analogs identified in our previously reported HTVS: 4G, 7D, 10D, 7B, and 4E **(Figure 4b)**, which all exhibited binding affinities comparable to rilpivirine against wild type HIV-1 RT **(Figure 5a)** (11).

K103N and K103N/Y181C Variants

Mutations at K103 cause the binding affinity of rilpivirine to drop significantly to -7.9 kcal/mol. We found that our analogs experience loss in binding affinity as well, though not as drastic. Notably, analog 7D performs the best, with a binding affinity of -10.8 kcal/mol. The K103N/Y181C double mutant significantly drops the binding affinity of rilpivirine and our analogs to between -9.0 and -8.2 kcal/mol. The decrease in binding affinity of rilpivirine is due to both the loss of the interaction with the positively charged lysine and loss of the aromatic tyrosine ring. The decrease in the binding affinity of our analogs is due to both K103N and Y181C, but mostly due to the Y181C variant, as the loss of the tyrosine residue aromatic ring eliminates the halogen-π interaction which our analogs strongly rely on. The Y181C variant also eliminates critical π-π interactions for rilpivirine and our analogs.

L100I, L100I/K103N, and L100I/K103R/V179D Variants

L100 is a nonpolar, aliphatic residue and forms hydrophobic interactions with the amino nitrile on rilpivirine. Interestingly, the L100I variant causes the binding affinity of analog 4G to be superior to rilpivirine when docked to WT RT. Leucine's side chain does not protrude out as much as isoleucine's side chain in the allosteric binding pocket. This mutation increases the distance between the leucine residue at position 100 and rilpivirine in the L100I variant, which causes weaker hydrophobic interactions. For reference, the hydrophobic interaction distance between the nitrogen at position two on rilpivirine's pyrimidine ring and WT L100 is 3.1 Å, whereas the interaction distance between the same nitrogen on rilpivirine

to the L100I variant increases to 3.5 Å, resulting in a weaker hydrophobic interaction **(Figure 5b)**.

It is important to note that L100I mutations rarely occur on their own; rather, they occur in combination with K103N, which causes a 10-fold reduced susceptibility to rilpivirine (12). Against the L100I/K103N variant, the binding affinity of rilpivirine is decreased to -8.1 kcal/mol and the binding affinities for our analogs range from -8.9 to -10.4 kcal/mol. Interestingly, in variants that contain the K103N mutation, analog 7D consistently has a superior binding affinity than rilpivirine or our other analogs. The trifluoromethoxy on 7D forms C-H---X interactions with isoleucine 100 at a distance of 3.3 Å, and the terminal methyl (-CH3) group on 7D forms electrostatic interactions with mutant K103N at a distance of 3.0 Å. For rilpivirine, L100I has hydrophobic interactions with the amino nitrile at a distance of 4.4 Å, and the amino nitrile interacts with arginine 103 at a distance of 4.9 Å **(Figure 5c)**. The combination of RT associated mutation L100I/K103R/V179D is strongly associated with decreased rilpivirine efficacy, consistent with our findings that rilpivirine's binding affinity decreases to -9.8 kcal/mol. Most of our analogs retain their binding affinity, as their binding is not largely dependent on mutations at the L100 or V179 position.

Y181C/I/V Variants

Y181 in WT HIV-1 RT forms π-π stacking with the aromatic ring of our analogs and halogen-π interactions with the trifluoromethoxy group of our analogs. Thus, mutations at this position, especially Y181V, are observed to be detrimental to the binding affinity of rilpivirine, which is -4.9 kcal/mol, and to our analogs, which range from -5.6 to -4.7 kcal/mol. This is consistent with the reported 10- to 15-fold decrease in rilpivirine susceptibility seen in patients with the Y181V mutation (10). The loss of the aromatic tyrosine ring in mutations eliminates potential for critical π-π stacking interactions. It also eliminates the halogen-π interaction with the trifluoromethoxy group and tyrosine 181 critical to the success of our analogs.

Additionally, mutant Y181V caused significantly lower binding affinity for rilpivirine and our analogs compared to mutant Y181I, in which our analogs and rilpivirine retained most of their binding affinity.

Mutant Y181C is reported to reduce rilpivirine susceptibility

Residue	L	K	K	V	E	V	Y	Y	H	F	M
Position	100	101	103	106	138	179	181	188	221	227	230
Mutated Residue(s)	I	E	N	A	A	D	C	L	Y	C	I
		P	R		G	L	I				L
					K		V				
					Q						
					R						

Figure 3: List of clinically-relevant rilpivirine-resistant point mutations that were modeled.

3-fold, consistent with the docking results for rilpivirine, which was -10.8 kcal/mol (12). Our analogs had binding affinities similar to that of rilpivirine to Y181C. Despite the loss of π-π interactions, the trifluoromethoxy on our analogs is able to form orthogonal multipolar interactions with the sulfur on cysteine 181, with the distance from the closest fluorine measuring 3.2 Å. Although weak, this fits into criteria for fluorine–sulfur contacts (2.8–3.4 Å) in experimentally observed protein structures (14, 15). It is important to note that the binding pose of rilpivirine and our analogs are oriented similarly spatially, with each R-group forming similar interactions with nearby residues, such as C181, **(Figure. 5d),** showing that these similar interactions confer similar reduced binding affinities to the Y181C mutant.

Y188L Variant

The Y188L mutant is known to decrease rilpivirine susceptibility 6-fold, which is in line with our findings that this mutant causes rilpivirine to have a reduced binding affinity of -8.3 kcal/mol, while analog 7B retains the best binding affinity at -10.3 kcal/mol (10). Analog 7B is the only analog without a trifluoromethoxy group, as it does not rely as strongly
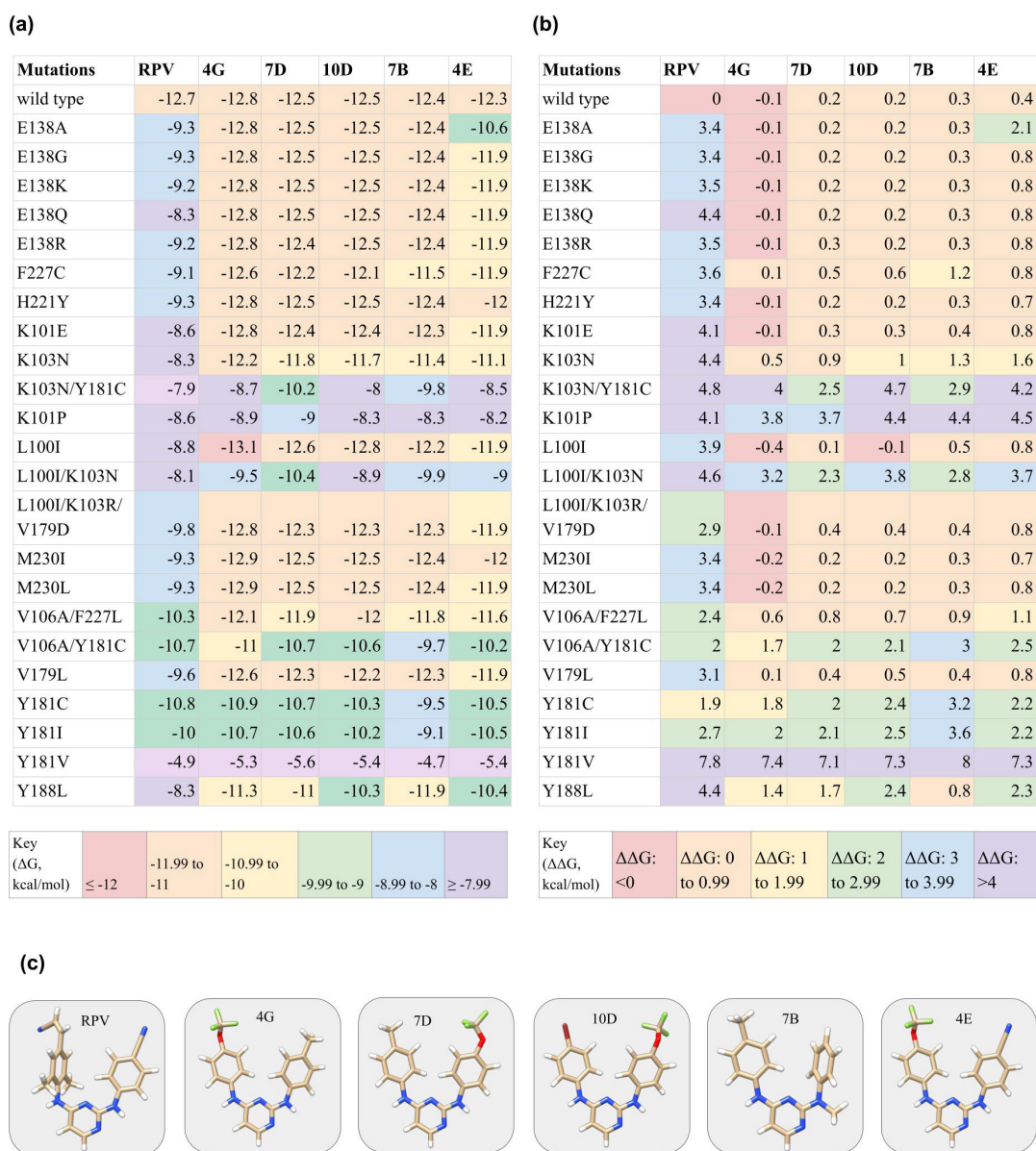
(a)

Mutations	RPV	4G	7D	10D	7B	4E
wild type	-12.7	-12.8	-12.5	-12.5	-12.4	-12.3
E138A	-9.3	-12.8	-12.5	-12.5	-12.4	-10.6
E138G	-9.3	-12.8	-12.5	-12.5	-12.4	-11.9
E138K	-9.2	-12.8	-12.5	-12.5	-12.4	-11.9
E138Q	-8.3	-12.8	-12.5	-12.5	-12.4	-11.9
E138R	-9.2	-12.8	-12.4	-12.5	-12.4	-11.9
F227C	-9.1	-12.6	-12.2	-12.1	-11.5	-11.9
H221Y	-9.3	-12.8	-12.5	-12.5	-12.4	-12
K101E	-8.6	-12.8	-12.4	-12.4	-12.3	-11.9
K103N	-8.3	-12.2	-11.8	-11.7	-11.4	-11.1
K103N/Y181C	-7.9	-8.7	-10.2	-8	-9.8	-8.5
K101P	-8.6	-8.9	-9	-8.3	-8.3	-8.2
L100I	-8.8	-13.1	-12.6	-12.8	-12.2	-11.9
L100I/K103N	-8.1	-9.5	-10.4	-8.9	-9.9	-9
L100I/K103R/ V179D	-9.8	-12.8	-12.3	-12.3	-12.3	-11.9
M230I	-9.3	-12.9	-12.5	-12.5	-12.4	-12
M230L	-9.3	-12.9	-12.5	-12.5	-12.4	-11.9
V106A/F227L	-10.3	-12.1	-11.9	-12	-11.8	-11.6
V106A/Y181C	-10.7	-11	-10.7	-10.6	-9.7	-10.2
V179L	-9.6	-12.6	-12.3	-12.2	-12.3	-11.9
Y181C	-10.8	-10.9	-10.7	-10.3	-9.5	-10.5
Y181I	-10	-10.7	-10.6	-10.2	-9.1	-10.5
Y181V	-4.9	-5.3	-5.6	-5.4	-4.7	-5.4
Y188L	-8.3	-11.3	-11	-10.3	-11.9	-10.4

Key (ΔG, kcal/mol)	≤ -12	-11.99 to -11	-10.99 to -10	-9.99 to -9	-8.99 to -8	≥ -7.99

(b)

Mutations	RPV	4G	7D	10D	7B	4E
wild type	0	-0.1	0.2	0.2	0.3	0.4
E138A	3.4	-0.1	0.2	0.2	0.3	2.1
E138G	3.4	-0.1	0.2	0.2	0.3	0.8
E138K	3.5	-0.1	0.2	0.2	0.3	0.8
E138Q	4.4	-0.1	0.2	0.2	0.3	0.8
E138R	3.5	-0.1	0.3	0.2	0.3	0.8
F227C	3.6	0.1	0.5	0.6	1.2	0.8
H221Y	3.4	-0.1	0.2	0.2	0.3	0.7
K101E	4.1	-0.1	0.3	0.3	0.4	0.8
K103N	4.4	0.5	0.9	1	1.3	1.6
K103N/Y181C	4.8	4	2.5	4.7	2.9	4.2
K101P	4.1	3.8	3.7	4.4	4.4	4.5
L100I	3.9	-0.4	0.1	-0.1	0.5	0.8
L100I/K103N	4.6	3.2	2.3	3.8	2.8	3.7
L100I/K103R/ V179D	2.9	-0.1	0.4	0.4	0.4	0.8
M230I	3.4	-0.2	0.2	0.2	0.3	0.7
M230L	3.4	-0.2	0.2	0.2	0.3	0.8
V106A/F227L	2.4	0.6	0.8	0.7	0.9	1.1
V106A/Y181C	2	1.7	2	2.1	3	2.5
V179L	3.1	0.1	0.4	0.5	0.4	0.8
Y181C	1.9	1.8	2	2.4	3.2	2.2
Y181I	2.7	2	2.1	2.5	3.6	2.2
Y181V	7.8	7.4	7.1	7.3	8	7.3
Y188L	4.4	1.4	1.7	2.4	0.8	2.3

Key (ΔΔG, kcal/mol)	ΔΔG: <0	ΔΔG: 0 to 0.99	ΔΔG: 1 to 1.99	ΔΔG: 2 to 2.99	ΔΔG: 3 to 3.99	ΔΔG: >4

(c)

RPV   4G   7D   10D   7B   4E

Figure 4:Docking results from homology modeling virtual screening experiments. (a) Heat map of computational binding affinities (reported in kcal/mol) from molecular docking experiments against key HIV-RT mutants, color coded by binding affinity values, (b) relative change in free energy based on rilpivirine's binding affinity to WT, and (c) Density Functional Theory (DFT)-optimized 3D structures of rilpivirine and five top analogs identified from our previously published high throughput virtual screen

on halogen-π interactions with tyrosine 188 like our other analogs. Instead, the aromatic ring of 7B is able to form C-H---π interactions with leucine 188 at a distance of 3.5 Å and retain most of its binding affinity **(Figure 5e)**.

E138A/G/K/Q/R and K101E/P Variants

In wild type RT, E138, a negatively charged residue, engages K101 through electrostatic attractions, which is the basis for stabilizing tertiary structure of the allosteric binding

pocket (13). However, in variants wherein E mutated to alanine or glycine, which are short, non-polar residues; or arginine or lysine, which are positively charged; or glutamine, which is polar but not anionic at physiological pH, the interaction between E138 and K101 is lost, distorting the binding pocket shape. This loss of structure greatly decreases the binding affinity of rilpivirine from -12.7 kcal/mol against WT to between -8.3 and -9.3 kcal/mol against these variants, and this is consistent with clinical data that suggests that these
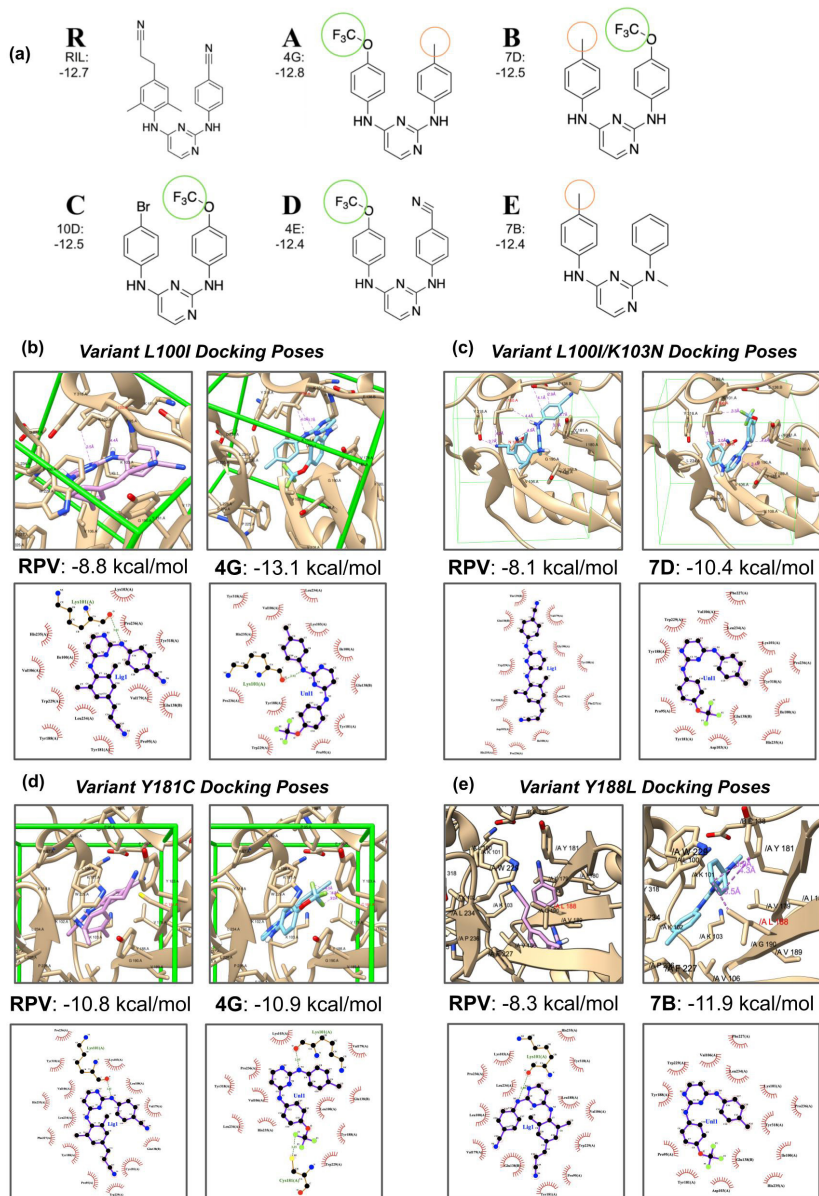
Figure 5: Docking poses of rilpivirine and key analogs against select homology-modeled structural variants of HIV-RT. (a) Chemical structure of rilpivirine and identified analogs screened in this study. (b) Binding pose of rilpivirine compared to top analog 4G in binding pocket of the L100I variant. (c) Binding pose of rilpivirine compared to top analog 7D in binding pocket of the L100I/K103N variant. (d) Binding pose of rilpivirine compared to top analog 4G in binding pocket of the Y181C variant. (e) Binding pose of rilpivirine compared to top analog 7B in binding pocket of Y188L variant. (f) 2D ligand plot of binding pose of rilpivirine compared to top analog 4G in binding pocket of the L100I variant. (g) 2D ligand plot of binding pose of rilpivirine compared to top analog 7D in binding pocket of the L100I/K103N variant. (h) 2D ligand plot of binding pose of rilpivirine compared to top analog 4G in binding pocket of the Y181C variant. (i) binding pose of rilpivirine compared to top analog 7B in binding pocket of Y188L variant.

mutants are resistant to treatment by rilpivirine, as they reduce rilpivirine susceptibility at least two- to three-fold (12). Our analogs, however, retain the same binding affinity as when docked to WT, as they do not rely strongly on the structure of the binding pocket in the inner E138/K101 region and instead rely mostly on halogen-π interactions with aromatic residues **(Figure 5f).**

Mutation of K101 to a glutamic acid residue causes a loss of electrostatic interaction with E138 and the two glutamic acid residues repulse each other, resulting in a decreased binding affinity of -8.6 kcal/mol which is consistent with clinical data where K101E mutation often causes 2.5-3 fold reduced susceptibility to rilpivirine (10). Mutation of K101 to proline, which is a nonpolar residue, also causes a loss in binding affinity to -8.3 kcal/mol. Our analogs are able to maintain their binding affinity, again due to the reliance on halogen-π interactions, rather than relying on binding pocket confirmation from the E138/K101 electrostatic interactions like rilpivirine.

DISCUSSION

A general trend observed is that mutations from aromatic amino acids to aliphatic or nonpolar amino acids cause a significant decrease in binding affinity. Mutations resulting in amino acids with shorter nonpolar side chains cause worse binding affinity. Furthermore, loss of an aromatic ring, especially the aromatic rings in Y181 and Y188, is the most detrimental for NNRTI binding affinity. Specifically, Y181C and Y188L mutations cause steric hindrances between rilpivirine and the binding site and loss of critical π-π interactions between the aromatic rings in rilpivirine and the side chain of tyrosine. In the Y181C/I/V mutants, which are some of the most common NNRTI-resistant mutations, our analogs have binding affinities similar to those of rilpivirine and do not perform better, due to the loss of the proximal aromatic ring of tyrosine 181, which is crucial for halogen-π bonding between the fluorinated analogs and the aromatic ring. Against mutant Y188L, analog 7B retains most of its binding affinity. The residues E138 in the p51 subunit and K101 in the p66 subunit determine the tertiary structure of the NNRTI binding pocket through electrostatic forces. These residues also form electrostatic attractions between the p51 and p66 subunits, stabilizing the tertiary structure of RT (16). Against E138A/G/K/Q/R mutations, rilpivirine decreases from -12.7 kcal/mol to about -9.3 kcal/mol, whereas the analogs retain their binding affinity. Binding of rilpivirine is dependent on the hydrogen bonding to the peptide backbone, and the binding pocket becomes more distorted without the electrostatic interactions between E138 and K101, causing decreased binding affinity. Our analogs are able to retain their binding affinity even when the electrostatic interactions are lost because they largely rely on halogen-π interactions with tryptophan 229 and tyrosine 181 instead of relying on amino nitrile bonding like rilpivirine. Lysine is a positively charged amino acid, and when K101 mutates to either K101E/G, which are not positively charged,

it causes a loss in electrostatic interactions between K101 and E138. Our analogs have a binding affinity significantly better than rilpivirine against the L100I mutation. However, L100I rarely occurs alone; it mostly occurs in combination with K103N, causing over 10-fold reduced susceptibility to rilpivirine (12). Both L100 and K103 have hydrophobic, fatty tails, which enable hydrophobic interactions with the rilpivirine cyanobenzene. The K103N mutant loses the aliphatic tail, which is important for binding affinity. Analog 7D retains better binding affinity than rilpivirine against double mutant L100I/K103N and single mutant K103N, suggesting that hydrophobic interactions with K103 are not overly significant in 7D's binding pose. The other analogs, especially rilpivirine, still faced decreased binding affinity in this mutant. Moreover, we found that mutations at the Y181 position to valine resulted in a greater loss in binding affinity as compared to Y181L mutations. Valine has a shorter side chain compared to isoleucine, which could explain the more significant decrease in binding affinity of the Y181V variant.

The number of NNRTI-resistant mutants from prolonged usage of NNRTIs poses a growing threat to the efficacy of antiretroviral therapy. Thus, the continued design and development of novel antiretroviral small molecule therapeutics is paramount to addressing the future of HIV/AIDS. With the capabilities of computational tools available, we can rapidly screen large compound libraries, which is more time and cost-effective than in vitro or in vivo screening. Using biophysical simulations of homology modeling and molecular docking, we have effectively modeled structural variants in RT without having to obtain crystallographic x-ray structures. It should be noted that molecular docking has limitations when predicting binding poses and binding affinity that are consistent with in vivo results; however, it is still useful for rapidly screening potential drugs. AutoDock Vina was found to have the best scoring power in ranking the binding affinities of docked poses.

While the results presented herein demonstrate the identification of and provide a structural rationale for hit compounds that retain high binding affinity to several rilpivirine-resistant HIV variants, whether the trends observed in binding affinity correspond with retention of antiretroviral activity remain to be seen. To this end, laboratory syntheses of our top hit compounds are currently underway, and in the future, we hope this will provide information about the in vitro functional efficacy of the compounds described. Moreover, since the emergence of new viral variants is a continued threat in a variety of viral diseases, we envision that such computational workflows might be similarly employed towards the development of new antiviral compounds to fight such variants.

MATERIALS AND METHODS
Modeling, Design, and Molecular Mechanics Preoptimization

We systematically created and modeled each

diarylpyrimidine (DAPY) analog using Avogadro, an open-source molecular modeling software package (17). All chemical entities screened in this study were initially optimized by molecular mechanics using the UFF94 force field at 10,000 steps, as per our previous report (11).

Density Functional Theory (DFT)

We created input files for rigorous quantum-mechanical optimization through Avogadro (17). The geometries of each structure were thermodynamically minimized via density functional theory (DFT) through ORCA, an ab initio quantum molecular modeling software, using a B3LYP functional and def2-SVP basis set with a continuum solvation model (CPCM) in water (18). All DFT calculations were performed on a Dell PowerEdge 710 server with a 24 core Intel Xeon X5660 processor at 2.80 GHz and 32GB RAM.

Homology Modeling

We used the structure of the HIV-1 Reverse Transcriptase in complex with TMC278 [PDB:3MEE] as the WT reference sequence from which variants were modeled using ChimeraX (13, 22, 23). We modeled mutated residues from the backbone-dependent Dunbrack 2010 rotamer library. The models were then checked for clashes between nearby amino acids. HIV-1 RT mutations were taken from the 2019 Update of the Drug Resistance Mutations in HIV-1 (10). Homology modeling was performed on the ChimeraX version 1.1 (2020-09-09) suite on a Lenovo ThinkCentre computer with a 4-core Intel I5 processor at 3.20 GHz and 8 GB RAM.

Molecular Docking

In concordance with our previous report, with a batch script submission to AutoDock Vina, we docked each optimized structure to the allosteric binding pocket of RT to predict binding affinities of our designed analogs (11, 19, 20). For this, we used default genetic algorithm settings, gridbox spacing, and scoring functions. We used the unliganded structure of rilpivirine bound to a representative reverse transcriptase of HIV [PDB:3MEE] as the receptors and internal standard for comparison. We set the center atom of rilpivirine as the center of a 16 by 16 by 16 Å grid box **(Figure 1b)** (13). We extracted predicted binding modes as a value of free energy of binding (ΔG) in kcal/mol. We visualized the final binding poses in Chimera before the final docking positions and protein-ligand interactions were analyzed to extract both predicted binding thermodynamics and the structural basis for such results (Fig. 5) (21). Molecular docking experiments were performed on a Dell PowerEdge 710 server with a 24 core Intel Xeon X5660 processor at 2.80 GHz and 32GB RAM.

REFERENCES

1. "The Global HIV/AIDS Epidemic." HIV.gov, 7 July 2020, www.hiv.gov/hiv-basics/overview/data-and-trends/global-statistics.
2. "What Are HIV and AIDS?" HIV.gov, 18 June 2020, www.hiv.gov/hiv-basics/overview/about-hiv-and-aids/what-are-hiv-and-aids.
3. HIV Treatment: The Basics. (n.d.). Retrieved January 07, 2021, from https://hivinfo.nih.gov/understanding-hiv/fact-sheets/hiv-treatment-basics
4. Sharp, Paul M, and Beatrice H Hahn. "The evolution of HIV-1 and the origin of AIDS." *Philosophical transactions of the Royal Society of London.* Series B, Biological sciences vol. 365,1552, 2010: 2487-94. doi:10.1098/rstb.2010.0031
5. Usach, Iris et al. "Non-nucleoside reverse transcriptase inhibitors: a review on pharmacokinetics, pharmacodynamics, safety and tolerability." *Journal of the International AIDS Society* vol. 16,1, 2013: 1-14. doi:10.7448/IAS.16.1.18567
6. Azijn, Hilde et al. "TMC278, a next-generation nonnucleoside reverse transcriptase inhibitor (NNRTI), active against wild-type and NNRTI-resistant HIV-1." *Antimicrobial agents and chemotherapy* vol. 54, no. 2, 2010: pp. 718-27. doi:10.1128/AAC.00986-09
7. Molina, Jean-Michel, et al. "Rilpivirine versus efavirenz with tenofovir and emtricitabine in treatment-naive adults infected with HIV-1 (ECHO): a phase 3 randomised double-blind active-controlled trial." *The Lancet*, vol. 378, 9787, 2011: pp. 238-246.
8. Buscher, April, et al. "Impact of antiretroviral dosing frequency and pill burden on adherence among newly diagnosed, antiretroviral-naive HIV patients." *International journal of STD & AIDS* vol. 23, no. 5, 2012, pp. 351-5. doi:10.1258/ijsa.2011.011292
9. Minuto, Joshua J, and Richard Haubrich. "Etravirine: a second-generation NNRTI for treatment-experienced adults with resistant HIV-1 infection." *Future HIV therapy* vol. 2, no. 6, 2008, pp. 525-537. doi:10.2217/17469600.2.6.525
10. Wensing, Annemarie, et al. "2019 update of the drug resistance mutations in HIV-1." *Topics in Antiviral Medicine*, vol. 27, 3, 2019: 111–121. https://pubmed.ncbi.nlm.nih.gov/31634862
11. Wu, Jeslyn, et al. "Design and in silico screening of analogs of rilpivirine as novel non-nucleoside reverse transcriptase inhibitors (NNRTIs) for antiretroviral therapy." *Journal of Emerging Investigators*, vol. 2, 2021:

online.

12. "NNRTI Resistance Notes." *NNRTI Resistance Notes - HIV Drug Resistance Database,* hivdb.stanford.edu/dr-summary/resistance-notes/NNRTI/.

13. Lansdon, Eric B., et al. "Crystal structures of HIV-1 reverse transcriptase with etravirine (TMC125) and rilpivirine (TMC278): implications for drug design." *Journal of medicinal chemistry* vol. 53, no. 10, 2010, pp. 4295-9. doi:10.1021/jm1002233

14. Pollock, Jonathan, et al. "Rational Design of Orthogonal Multipolar Interactions with Fluorine in Protein–Ligand Complexes." *Journal of Medicinal Chemistry,* vol. 58, no. 18, 2015, pp. 7465–7474., doi:10.1021/acs.jmedchem.5b00975.

15. Bauer, Matthias R., et al. "Harnessing Fluorine–Sulfur Contacts and Multipolar Interactions for the Design of p53 Mutant Y220C Rescue Drugs." *ACS Chemical Biology,* vol. 11, no. 8, 2016, pp. 2265–2274., doi:10.1021/acschembio.6b00315.

16. Smith, Steven J., et al. "Rilpivirine Analogs Potently Inhibit Drug-Resistant HIV-1 Mutants." *Retrovirology,* vol. 13, no. 1, 2016, doi:10.1186/s12977-016-0244-2.

17. Hanwell, Marcus D., et al. "Avogadro: an advanced semantic chemical editor, visualization, and analysis platform." *Journal of Cheminformatics,* vol. 4, no. 17, 2012, doi:10.1186/1758-2946-4-17.

18. Neese, Frank, et al. "The ORCA Program System." *WIREs Computational Molecular Science*, vol. 2, no. 1, 2011, pp. 73–78. doi:10.1002/wcms.81.

19. Ashok, Bhavesh. "Bhaveshashok/AutoDockVina-BatchSubmission." *GitHub,* 2020, github.com/bhaveshashok/AutoDockVina-BatchSubmission.

20. Trott, Oleg, and Arthur J Olson. "AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading." *Journal of computational chemistry,* vol. 31, no. 2, 2010, pp. 455-61. doi:10.1002/jcc.21334

21. Pettersen, Eric F. et al. "UCSF Chimera--a visualization system for exploratory research and analysis." *Journal of computational chemistry,* vol. 25, no. 13, 2004, pp. 1605-12. doi:10.1002/jcc.20084

22. Pettersen, Eric F., et al. "UCSF ChimeraX: Structure visualization for researchers, educators, and developers." *Protein Science* 30.1 (2021): 70-82.

23. Goddard, Thomas D., et al. "UCSF ChimeraX: Meeting modern challenges in visualization and analysis." *Protein Science* 27.1 (2018): 14-25.