MedChemComm

Accepted Manuscript

This article can be cited before page numbers have been issued, to do this please use: V. Poongavanam, A. Corona, C. Svendsen, L. Scipione, N. Grandi, F. Pandolfi, R. Di Santo, R. Costi, F. Esposito, E. Tramontano and J. Kongsted, *Med. Chem. Commun.*, 2018, DOI: 10.1039/C7MD00600D.

ROYAL SOCIETY OF CHEMISTRY

rsc.li/medchemcomm

View Article Online
DOI: 10.1039/C7MD00600D

Structure-guided approach identifies a novel class of HIV-1 Ribonuclease H inhibitors: Binding mode insights through magnesium complexation and site-directed mutagenesis studies

Vasanthanathan Poongavanam[1], Angela Corona[2], Casper Steinmann,[1] Luigi Scipione[3], Nicole Grandi[2], Fabiana Pandolfi[3], Roberto Di Santo[3], Roberta Costi[3], Francesca Esposito, [2] Enzo Tramontano[2,4*] and Jacob Kongsted[1*]

[1] Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, DK-5230 Odense M, Denmark

[2] Department of Life and Environmental Sciences, University of Cagliari, Italy

[3] Dipartimento di Chimica e Tecnologie del Farmaco, Istituto Pasteur-Fondazione Cenci Bolognetti, "Sapienza" Università di Roma, Roma, Italy

[4] Istituto di Ricerca Genetica e Biomedica, Consiglio Nazionale delle Ricerche (CNR), Monserrato(CA), Italy.

Running title. Identification of novel RNase H inhibitors by *in silico* screening

Keywords. HIV-1, RNase H, *in silico* screening, magnesium complexation, mutagenesis, allosteric inhibition

View Article Online
DOI: 10.1039/C7MD00600D

ABSTRACT

HIV persistent infection requires a life-long treatment and among the 2.1 million new HIV infections that occur every year there is an increased rate of transmitted drug-resistant mutations. This fact requires a constant and timely effort in order to identify and develop new HIV inhibitors endowed with innovative mechanisms. The HIV-1 Reverse Transcriptase (RT) associated Ribonuclease H (RNase H) is the only viral encoded enzymatic activity that still lacks an efficient inhibitor despite the fact that it is a well-validated target whose functional abrogation compromises viral infectivity. Identification of new drugs is a long and expensive process that can be speeded up by *in silico* methods. In the present study, a structure-guided screening is coupled with a similarity-based search on the Specs database to identify a new class of HIV-1 RNase H inhibitors. Out of 45 compounds selected for experimental testing, 15 of these inhibited the RNase H function below 100 μM with three hits possessing IC_{50} values <10 μM. The most active compound **AA** inhibits HIV-1 RNase H with an IC_{50} of 5.1 μM, and possesses Mg-independent mode of inhibition. Site-directed mutagenesis studies provide valuable insight into the binding mode of newly identified compounds, for instance, compound **AA** involves extensive interactions with a lipophilic pocket formed by Ala502, Lys503, Trp (406, 426 and 535) and polar interactions with Arg557 and with the highly conserved RNase H primer-grip residue Asn474. The structural insights obtained from this work provide the bases for further lead optimization.

View Article Online
DOI: 10.1039/C7MD00600D

INTRODUCTION

The increasing resistance to current therapeutic agents for HIV drug regiment remains a major issue for effective acquired immune deficiency syndrome (AIDS) therapy.[1, 2] It has recently been estimated (2016) that more than 36.7 million people are living with HIV, and among them only 18.2 million are on antiretroviral therapy and nearly 1.1 million people died in 2015.[1, 3, 4] Furthermore, 2.1 million new HIV infections occur every year. Although AIDS related mortality has been significantly reduced since the introduction of highly active antiretroviral therapy (HAART), (35% since 2005), most of the currently marketed drugs that are being used today for HIV therapy are prone to viral resistance.[5, 6] The number of newly infected people/year is constant and among them there is an overall increase of transmitted drug resistance mutations (TDRMs) detected in antiretroviral treatment–naive patients.[2, 7] The resistance, which often results from lack of compliance, can relay to multiple types of drug, and can dramatically affect the outcome of HAART,[8, 9] therefore, there is a strong need to search for non-traditional chemo-targets for anti-HIV drug developments.[10-12] Inhibition of the HIV-1 reverse transcriptase associated ribonuclease H (RNase H) function provides a novel target for anti-HIV chemotherapy, since it is the only viral encoded enzymatic function for which no inhibitor went in clinical development so far.[13-15] Although it has been shown that inactivation of the RNase H function by amino acidic substitution lead to not infectious virions[16] and its selective inhibition completely blocks viral replication.[17, 18]

RNase H is one of the two catalytically active domains of the p66 subunit of HIV-1 reverse transcriptase, which is essential for viral genome replication by performing the hydrolysis of the viral RNA strand in the RNA/DNA hybrid during the retro-transcription process.[19-22] In order to carry out the hydrolytic process, water molecules (act as nucleophiles) and magnesium ions (Mg^{2+}) (initiate the deprotonation of water) coordinating with the highly conserved DDDE motif (Asp443, Asp498, Asp549 and Glu478) residues of RNase H are essential.[23] The topology of the HIV-1 RNase H domain is well characterized from different studies.[22, 24, 25] The active site is small and slightly hydrophobic and consists of a highly electronegative region, which is primarily due to the DDDE motif coordinated with the two catalytically active magnesium ions (**Figure 1**).[22, 26]

In the past years, a large number of RNase H inhibitors have been reported but none

View Article Online
DOI: 10.1039/C7MD00600D

of these inhibitors have yet reached the market. [14, 15, 18, 20, 27-31] RNase H inhibitors can be classified into two categories based on the mode of action (i) compounds that block the magnesium ions availability to the hydrolytic process (the so called active-site inhibitors), such as diketo,[17, 32] N-hydroximide (e.g. NHQD, cf. Figure 1),[33] tropolone (e.g. β-thujaplicinol, cf. Figure 1), [34] pyrido-pyrimidinone, [30] N-hydroxy-naphthyridinone, [35] thiocarbamates, [27] and 5-nitro-furan-2-carboxylic acid, [36] and (ii) allosteric inhibitors that impair the efficiency of the RNase H activity by binding outside the catalytic domain, inducing conformational changes that affect a proper substrate binding at the RNase H active site. This second class includes vinylogous ureas (e.g. NSC727447, cf. Figure 1), [37] acylhydrazones (e.g. BHMP07, cf. Figure 1), [29, 38, 39] propenones, [25] isatines, [40] anthraquinones.[41] Both classes of compounds present critical issues that hamper the process of drug-development. On the one hand Mg^{2+} coordination provides active-site inhibitors a strong and stable interaction with the catalytic domain, as shown for the last approved innovation for HIV-1 treatment, the integrase (IN) strand transfer inhibitors[42], and as it has been proved for RNase H active site inhibitors RDS1759.[17] However, both cases show that the Mg binding mechanism itself lacks specificity towards the other polynucleotidyl transferases (and indeed it is deeply investigated in order to look for IN and RNase H dual inhibitors [43-45]), and that the metal-binding moiety needs to be flanked by additional fragments able to specifically interact with the molecular environment in order to achieve sufficient potency and specificity.[46] On the other hand, it has been proved that compounds that bind to allosteric non-conserved regions, as the non-nucleoside RT-inhibitors, are more prone to selection of resistant strains.[2] Therefore, in order to identify effective inhibitors with lower possibility of toxicity and drug resistance selection, there is a need to target also the most conserved peripheral regions of the HIV-1 RNase H domain, adjoining to allosteric pockets that are reported to differ in location in humans and viruses. [47, 48]

View Article Online
DOI: 10.1039/C7MD00600D

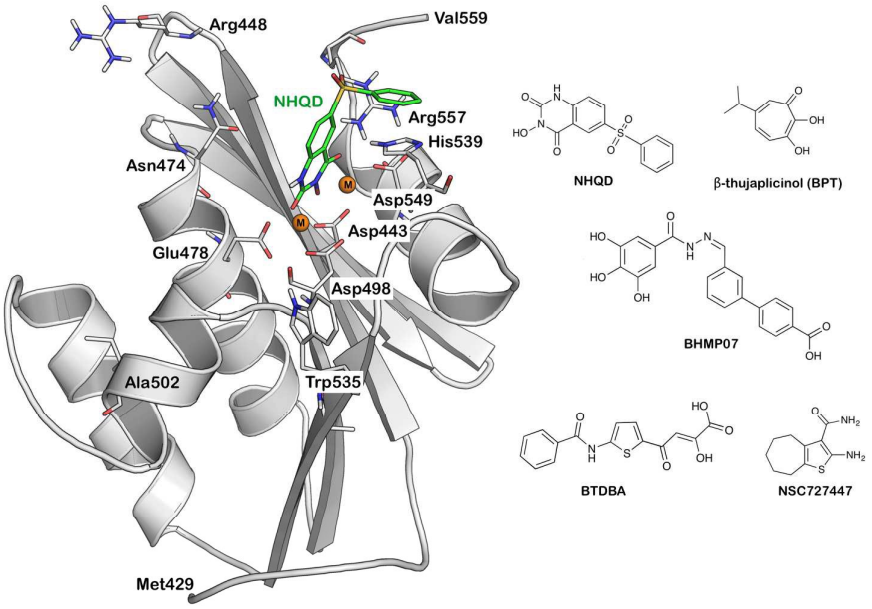


Figure 1. The HIV-RT associated RNase H structural topology is shown with a bound ligand NHQD (green, ball and stick model). The catalytically important residues and magnesium ions (orange sphere) have been highlighted. 2D structures of representative molecules belong to the active site and allosteric RNase H inhibition.

Recently, we proved that the lateral chains of prototype DKA inhibitors designed and synthesized by Costi et al.[43, 44] arrange within the active site binding the highly conserved residues Asn474, Gln475 and Tyr501 of the HIV-1 RNase H.[17]

These amino acids are part of the RNase H primer grip, so called because it extensively interacts with the DNA strain of the RNA/DNA heteroduplex, allowing the RNA to be in the right orientation to be cleaved by the active site. Mutations in the primer grip strongly impair viral replication[49] by compromising RNase H specificity[50] and efficiency.[17] Compound BHMP07 (cf. Figure 1) binds to an allosteric pocket centered on Gln507 of p66 and composed of the residues Leu425, Leu422, and Asn404 residues, was also reported to interact with Asn474 (H-bond), Gln500 (H-bond), Tyr501 (π-π stack), Asp498 (H-bond), Gly444 (H-bond) and Asp443 (H-bond) residues.[47] This observation was further supported by a virtual enrichment study.[48] Therefore, highly conserved residues within the RNase H domain represent a promising target for development of efficient RNase H inhibitors, and this effort can be optimized using computer-aided drug design techniques.

With the availability of a large number of high-resolution 3D crystal structures of drug target proteins, and deposits of a large number of ligand data in public databases,

View Article Online
DOI: 10.1039/C7MD00600D

both ligand and structure-guided *in silico* approaches are extensively utilized to identify new ligands with affinity for a wide range of protein targets.[51-53] Recently, a ligand-based virtual screening (VS) strategy was used on HIV-1 RT for identification of dual inhibitors (polymerase and RNase H activity), [29] leading to identification of four compounds that inhibited the RNase H function in the low micromolar range, and the binding mode was further confirmed by a site-directed mutagenesis approach thus confirming the effectiveness of this method for identification of novel RNase H inhibitors. Along this line, we have developed various ligand- and structure-based virtual screening models such as machine learning methods, docking, shaped based evaluation, pharmacophore modeling and QM-based calculations.[54, 55] From a structure-based docking approach, we found about 80-90 % of accuracy in terms of "enrichment" compared to other methods. In continuation, in the present study, we used virtual screening models from docking followed by shape-based methods to screen the Specs database (www.specs.net) to identify new chemotypes able to bind within the RNase H domain and inhibit the RNase H function. The best-ranked compounds from *in silico* screening have been selected for in vitro testing. Among the 45 compounds screened for RNase H inhibition, three compounds were showed IC_{50} values below 10 μM. In addition, the binding mode of these compounds has been investigated by site-directed mutagenesis and "induced-fit" molecular modeling obtaining insights for the further lead optimization.

RESULTS AND DISCUSSION

1. Structure-Based Virtual Screening

We have previously shown that effective virtual screening of RNase H inhibitors from large chemical databases could be achieved using the combination of docking and QM-based refinement calculations.[54, 55] In order to identify a novel chemo-type for RNase H inhibition and to validate previously developed computational methods, the best models were used to screen the Specs database (containing 277325 drug-like compounds for purchase) for HIV-1 RNase H inhibition screening (Figure 2). After application of general filters for ADMET property score and reactive groups, 343518 molecules were docked within a computational model of the HIV-1 RT associated RNase H domain.[22] A set of 1205 compounds was obtained at the end of the docking-based virtual screening, and these compounds were subsequently used for QM-based refinement calculation based on density functional theory (DFT) calculations as

View Article Online
DOI: 10.1039/C7MD00600D

previously described [55] (*Supporting information S1*). The best-ranked 180 compounds from the screening were sorted for further inspection. To select a diverse set of structures for the biochemical assays, these compounds were clustered according to the structural similarity using the Canvas software from the Schrodinger suite.[56] In total, 50 individual clusters were defined containing between 1 and 15 compounds and each cluster accounts for at least 75% structural similarity. In the next step, each individual cluster was visually inspected and the previously known compounds were removed from the hit list, which includes NSC727447, tropolone, pyrido-pyrimidinone and tropolone scaffolds (*The list of compounds used for clustering is provided in S2*). At the end, 25 structurally diverse compounds, with the best scores (structures are provided in the S3) were chosen and purchased from the chemical vendor www.specs.net to be tested against the HIV RT-associated RNase H function in enzymatic assays.

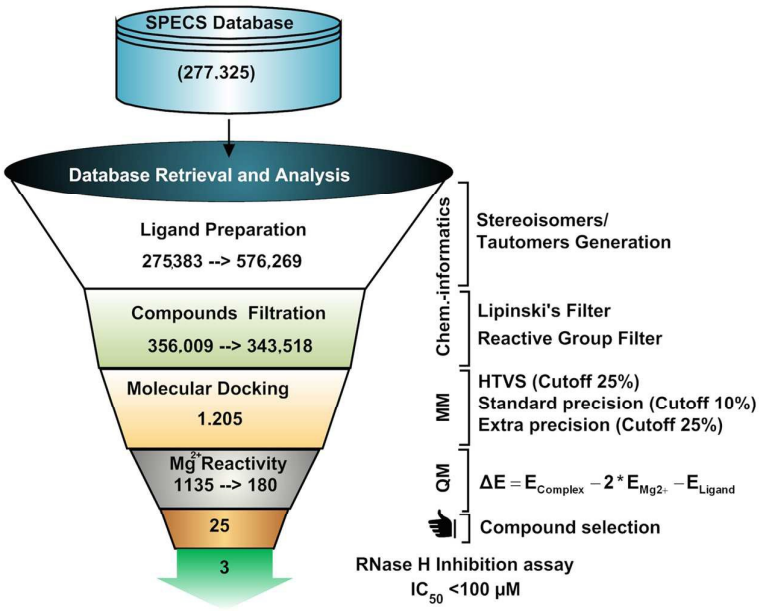


Figure 2. Overall workflow of structure-based virtual screening strategies applied.

2. Inhibition of HIV-1 RT associated RNase H function

The selected compounds were tested against HIV-1 RT associated RNase H function (Table 1). Out of 25 compounds tested (A-Y), 3 compounds inhibited the RNase H activity below an IC_{50} value of 100 μM and compound **A** showed an IC_{50} value of 9.35 μM. Notably, none of these compounds has previously been reported as an

View Article Online
DOI: 10.1039/C7MD00600D

inhibitor for RNase H, however, compound **A**, share structural similarities to hydantoin molecules described as HIV RNase H inhibitors in the patent in 2005 (US 20100063118) [57], and compound **B** shares very similar structural patterns with BHMP07 (cf. Figure 3). These interesting similarities confirm the robustness of the VS methods in identifying potential RNase H inhibitors. Interestingly, among the active compounds, compound **C**, that exhibited a modest IC_{50} value (61.25 μM), presents a scaffold that has never been reported before in relation to RNase H modulator (Figure 3). Thus, further investigation on this scaffold could be interesting. To find more similar compounds, which could potentially be used to understand the structure-activity relationship (SAR) for lead optimization, all three compounds **A**, **B** and **C** were used as query molecules for a similarity based further search.

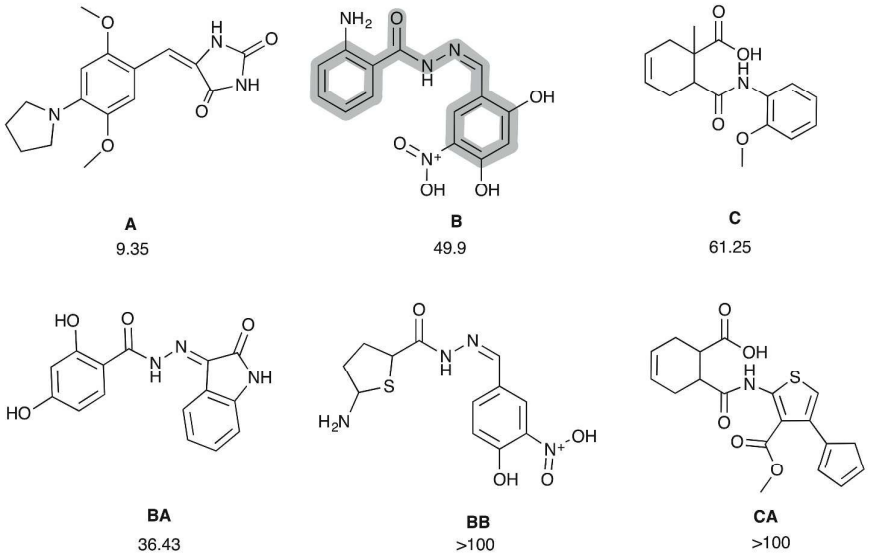


Figure 3. Initial hit molecules (A, B, C) from the first screening and compound "B" is highlighted regions where it shares common structural pattern as BHMP07. Compounds AQ, AR and as found as analogs of B and C by the shaped-based screening (see text for detail).

Similarity based search and structure activity relationships (SAR)

Similarity searches assume that molecules with similar physicochemical properties also have similar biological activity. This approach not only finds similar molecules or analogous but is also able to identify completely new chemo-types but having similar pharmacophoric points.[58] In this study, a combination of shape or pharmacophore-based (vROCS) [59] and electrostatic-based similarity (EON)[60] was

View Article Online
DOI: 10.1039/C7MD00600D

used on three active compounds (**A**, **B** and **C**) in the Specs database, identifying the 20 best-scored compounds to be tested against the HIV RT-associated RNase H function in enzymatic assays. Among them 17 molecules were analogs of compound **A** with structures varying greatly in terms of length and also functional groups (Figure 4). Two compounds **BA** and **BB** were found as analogs of **B**, and only one **CA** was found as an analogue of compound **C** (cf. Figure 3).

Out of 20 hits, 12 compounds inhibited the RNase H function with IC_{50} values ranging from 5 to 91 μM (cf. Table 1, compound **AA-AP**). Interestingly, the better inhibitors were found among **A** analogs, with compound **AA** showing an IC_{50} value of 5.17 μM. The **A** analogs showed indeed a great variety of structures in terms of length and also functional groups, however, there is a clear pattern in the scaffold among the tested compounds, as hydantoin-like molecules that all belongs to benzylidene hydantoins, a group less explored.[57] Indeed, the structure of the compounds in the **A** series can be fragmented into three parts (moieties): a head (Part **a**), a linker (Part **b**) and a lipophilic group (Part **c**). Part **a** is either an imidazolidinedione or a thiazolidinedione moiety. In some cases one of the two carbonyl groups of the azolidindione ring can be replaced by a thiocarbonyl or an imino moiety (Figure 4).

Part **b** consists of a more or less complex aromatic linker: aromatic or heteroaromatic ring substituted with hydroxyl, alkyloxy or amino groups, halogens or again aryl rings. If the linker is composed of the biaryl fragments, such as the case of compounds **AB**, **AC**, **AD**, **AG**, the compound exhibits stronger inhibition compared to shorter compounds **AH**-**AO** (the sole exception is AA discussed below).

Part **c** significantly modulates the RNase H function and is mostly composed of groups such as arylhalides or nitroaryl or carbamoylphenyl groups. The approximate distance of part **a**, **b** and **c** is found to be 7.5 Å (*benzylideneimidazolidinedione*), 3.5 Å (*linker*), 4.5 Å (*active region*), respectively. Overall, the observation suggests that parts **a**, and **b** are essential for the activity, while part **c** greatly modulates the activity against the RNase H function, depending on the nature of the substituent: more lipophilic is the substituent more active is the inhibitor (compare derivatives **AA**, **AB**, **AE** with the less potent **AG** derivative).

View Article Online
DOI: 10.1039/C7MD00600D

Among the **B** analogues, the results obtained from the shape-based screening e.g., **BA** (IC_{50} = 36.43 μM) and **BB** (IC_{50} = >100 μM) (Table 1, cf. Figure 2), reveal that small changes in the structure leads to a large difference in inhibition e.g., compound **B** (IC_{50} = 49.9 μM) and **BB**. The major structural difference between these compounds lies in the presence of hydrogen bond donor e.g., hydroxyl or amino groups at the phenyl ring. Indeed, it has previously been shown that hydroxyl groups interact directly to the negatively charged residues at the active site.[39] Moreover, the decrease of molecular density of the molecule (molecular density is molecular weight divided by vdw_vol (amu/Å^3)) to an increasing inhibitory activity e.g. **BHMP07** (0.795) > **BA** (0.817) > **B** (0.837) > **BB** (0.861) (the molecular density of each molecule is provided in the parenthesis). This observation is particularly interesting in relation to design of new compounds in this series. Finally, only one analog **CA** of compound **C** was unable to inhibit the RNase H function up to 100 μM concentration.

View Article Online
DOI: 10.1039/C7MD00600D

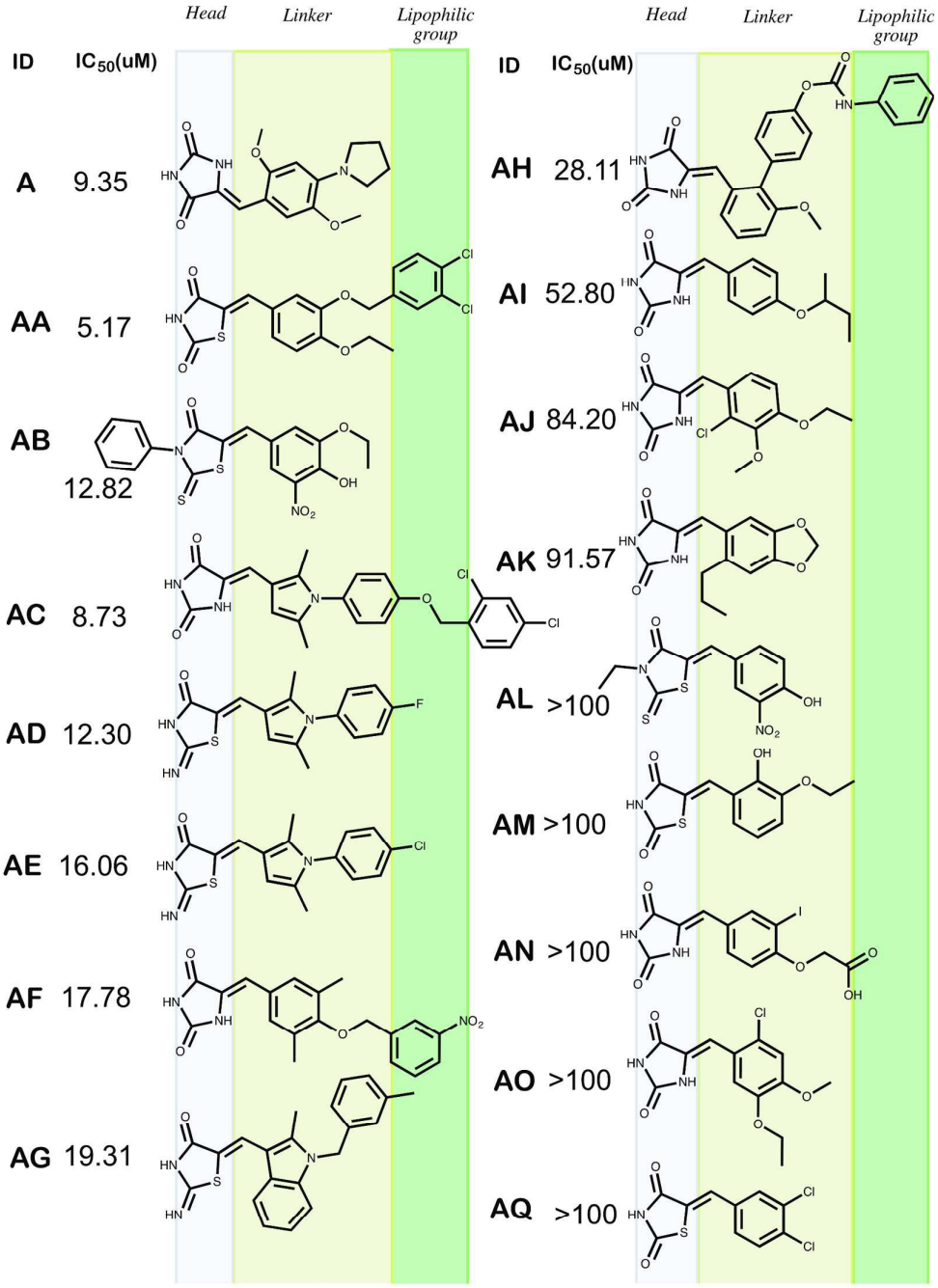


Figure 4. Comparison of structure and activity of **A** analogs.

Binding mode of inhibitors

In order to acquire more information about the binding mode of the compounds a further investigation was performed with a particular focus on the four "hits" identified**, A, AB**, and **AC** comparing their binding mode with BHMP07 (an allosteric inhibitor) and N-hydroxy quinazolinedinone (a prototype inhibitor active binding the

View Article Online
DOI: 10.1039/C7MD00600D

active site) (NHQD, PDB ID: 3QIO). [22] Since there is no crystal structure with BHMP07 bound to the RNase H domain, we compared the docking pose of BHMP07 from our docking methodology as previously reported. [39]

In the case of NHQD, for instance, His539 and Mg^{2+} interact with one of the three oxygen atoms of NHQD and the neighboring two oxygen atoms interact with the other Mg^{2+} (1) ion and bound water molecules (Figure 5, panel A). Differently, the binding pose of BHMP07 matched the previously determined binding pose, [39] as having hydrogen bonds with Asn474, Gln500, Asp498, Gly444, Asp433 and π-π stacking with Tyr501 (Figure 5, panel A).

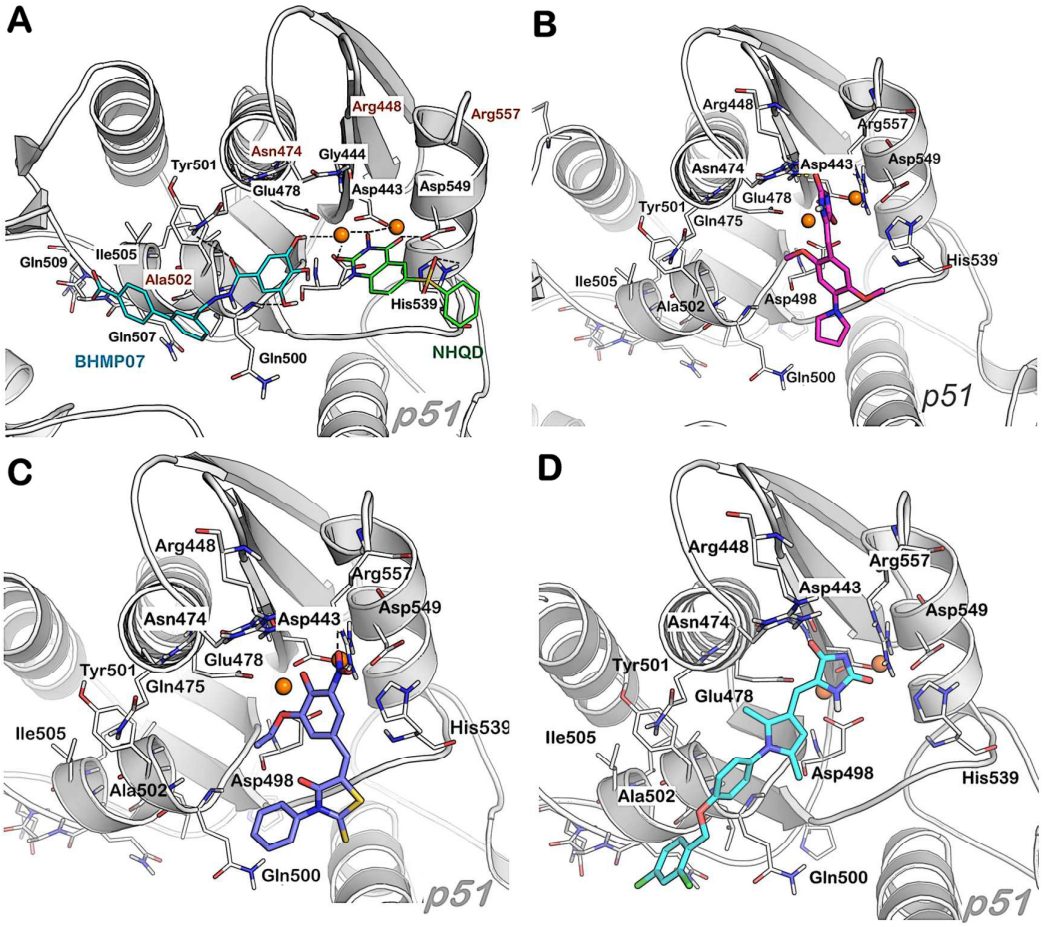


Figure 5. Comparison of the binding mode of known RNase H inhibitors, in a X-ray crystal, and important residues are highlighted. **Panel A,** N-hydroxy quinazolinedione inhibitor (NHQD, active site binder, green) and BHMP07 (allosteric site binder, cyan). **Panel B,** compound **A** (magenta), **Panel C** compound **AA** (blue), **Panel D,** compound **AB** (cyan).

View Article Online
DOI: 10.1039/C7MD00600D

For the imidazolidinedione scaffold, the *head* portion of all the inhibitors lies in close vicinity to negatively charged residues (Asp433 or Asp498) and interacts with Asn474. Moreover, a negatively charged nitrogen atom in the imidazolidinedione ring (e.g. compound **A**, panel B) is partially exposed to the magnesium ions and Arg557 as we have observed for the active site inhibitor NHQD. Overall, the *head* portion of inhibitors is exposed or interacts with Arg557 (ionic or π interaction), the polar interactions are found either with Gly444, Asn474, Asp498, Gln500 or His539 residues. The *linker* of the "hits" is primarily exposed to hydrophobic residues such as Ala502, Lys503, Trp (406, 426, 535) and Tyr501, thereby confirming that a linker composed as described above with one or two aromatic portions properly substituted, can provide additional interactions contributing to a stronger inhibition compared to the compounds that have substitution with halo-groups or ends in this region.

The most important part of these hits is an "*active region*" fragment (preferably with *arylhalides)*. Residues that lie in this region are more hydrophobic in nature e.g., Phe346, Trp406, Tyr406, Tyr405, Ala502 and Ile505. In addition, residues such as Gln507 and Pro421 also significantly interact with most active compounds. Some of these residues are located between the polymerase domain and RNase H domain of reverse transcriptase.

Compounds in the same scaffold such as **A** and **AA** showed a slightly different binding mode compared to the rest of the compounds (**Figure 5** panels B and C) since the negatively charged nitrophenol moiety of AA has high affinity towards the magnesium ions and this ionic interaction pulls the whole compound towards the active site. Since these ligands are small in size, the ionic interactions dominate the overall binding pose. Moreover, the hydrophobic residue Trp535 strongly interacts with the phenyl substituent at position 3 of the thioxothiazolidine of **AA** and with the pyrrolidine ring of **A**, respectively. The similar binding mode was observed for hydrazone derivatives **AQ** and **B** (Figure S3), in which favorable interactions of the inhibitors with Arg557 (ionic/π interaction), Gly444 (H-bonding) and Ala446 (H-bonding) were found.

3. Investigation of Magnesium-Complexation

In order to investigate the potential importance for interaction between the active compounds and Mg^{2+} ions, spectrophotometric complexation studies were carried on

View Article Online
DOI: 10.1039/C7MD00600D

the selected compounds **A**, **AA**, **AB**, **AC**, **AD**, **AE**, **AF** and **AG.** Titration with $MgCl_2$ did not induce significant changes in the UV-Vis spectra of compounds **AA**, **AC**, **AD**, **AE**, **AF** and **AG**, thereby indicating a lack of interaction of these molecules with Mg^{2+} (Figures 6 panels A and C, and *Supplementary material, S4 to S8*)

For compound **A**, we observed an overall increase in absorbance across the analyzed UV-Vis spectrum that could reasonably be attributed to an increase in the solubility due to the presence of magnesium ion rather than to a real complexation (Figure 6B).

Differently, titration of compound **AB** with $MgCl_2$ produced an increase of the absorbance at 282 and 392 nm (hyperchromic effect) with a slight shift towards higher wavelengths (hypsochromic effects); furthermore a disappearance of the 494 nm maxima was observed. The titration spectra present two isosbestic points: the first at 407 nm appears at lower magnesium concentration, the second at 449 nm appears at higher magnesium concentration. This type of behavior suggests the presence of two subsequent complexation equilibria between **AB** and magnesium ions (Figure 5D), supporting the hypothesis of the involvement of the interaction with the cofactor Mg^{2+} in the binding to the protein, and in agreement with the docking model.

This finding is very interesting since it is the first time that the inhibition of metalloenzymes is reported for the phenol with flanking nitro and ethoxy group fragment. Compounds containing the nitrophenol fragment have been previously been investigated as inhibitors of HIV-1 integrase (IN)[61], showing that the nitrophenol moiety is not sufficient to achieve inhibition of IN . Indeed this is in agreement with the much lower activity of compound **B** (IC_{50} = 49 μM) and **BB** (IC_{50} = >100 μM) that contain a similar nitrophenol moiety, but lack the ethoxy group, and are, therefore not able to coordinate the Mg as compound **AB** does. The phenol with flanking nitro and ethoxy groups fragment is structurally closely related to the more investigated 5-nitrovanillin (CHEBI ID: 48385), that, on its side, has never been investigated for such property. Therefore, this finding could represent a new insight for the development of the **AB** scaffold.

View Article Online
DOI: 10.1039/C7MD00600D

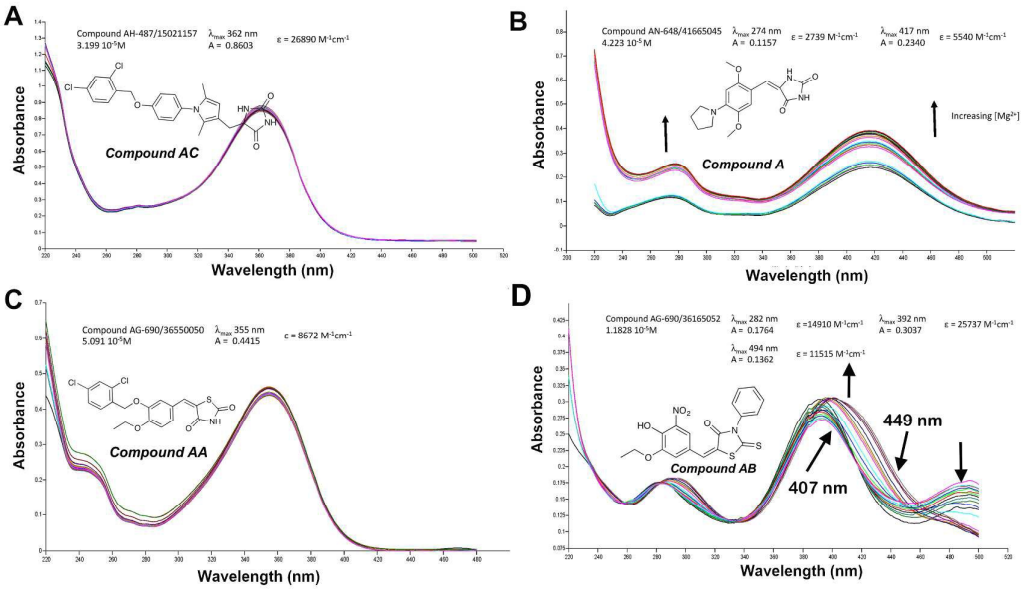


Figure 6. Titration graph of inhibitor with Mg^{2+}; Panel A: compound **AC**, Panel B: compound **A**, Panel C**:** Compound **AA** and Panel D: compound **AB**.

4. Site-directed mutagenesis

In order to experimentally verify the binding model suggested by the computational studies, the most active compounds **A**, **AA** and **AB** were chosen to perform a site-directed mutagenesis, determining the independent impact of several amino acid substitutions on the potency of the compounds to inhibit RNase H function. First, we assessed the impact of the substitution of Asn474, a primer grip residue, with Ala. Then we attempted to reduce the space available for the binding of the "*active region*" fragment introducing a phenylalanine instead of Ala502. Moreover, we attempted to disrupt the interactions engaged by the *head* portion by mutating Arg557 into Ala. A further substitution, R448A, as previously identified specifically affecting the activity of DKA active site inhibitors[17] was included as negative control. The active-site inhibitor β-thujaplicinol (BTP) [62] was used as a control (Figure 7 and Table 2). To understand the binding mode of the compounds used in the site-directed mutagenesis studies, molecular docking based on the "induced-fit" approach was utilized to mutated RT RNase H structure models.

View Article Online
DOI: 10.1039/C7MD00600D

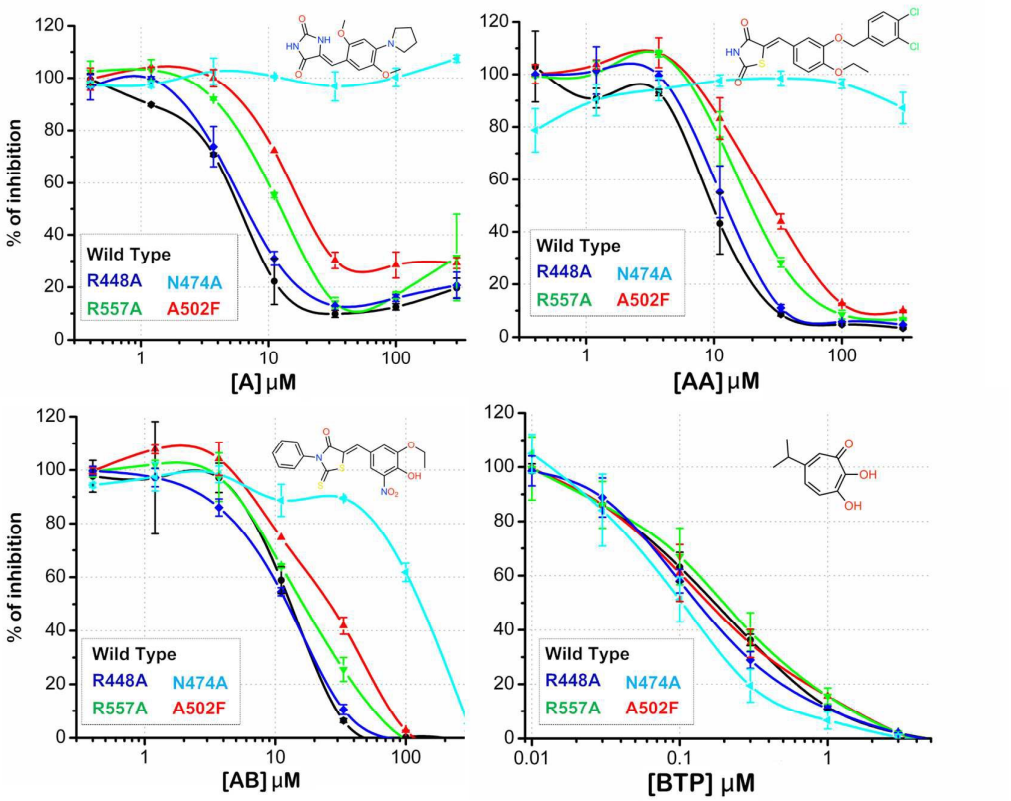


Figure 7. The comparison of the binding mode of compounds A, AA, AB and BPT in the wild type and various clinically relevant mutants.

Initially, the binding mode of BPT was investigated in four mutants (i.e. A502F, R557A, R4448A, N474A) and wild-type (WT). As shown in **Table 2**, there is no significant change in the IC_{50} of BPT in the mutants as compared to the WT. The "induced-fit" analysis also reveals that BPT possesses similar binding mode in all mutants compared to WT (cf. *Supporting information, S9*). BPT has shown to inhibit RNase H function through metal chelation. It is evident from the various binding poses that two of the three-hydroxyl groups strongly interact with one of the magnesium ions in the active site. We observe similar metal interaction pattern for compound **AB (**cf. Figure 4C**)** that was indeed the only one to show the isosbestic point and the shift proper to an actual Mg coordination. When assessed against the mutated enzymes, **AB** showed a lower impact of N474A substitution on its potency of inhibition when compared to the other compounds, a slight increase in IC_{50} when assessed against A502F and no significant effect of R557A substitution, proving that **AB** inhibits RNase H through metal chelation.

View Article Online
DOI: 10.1039/C7MD00600D

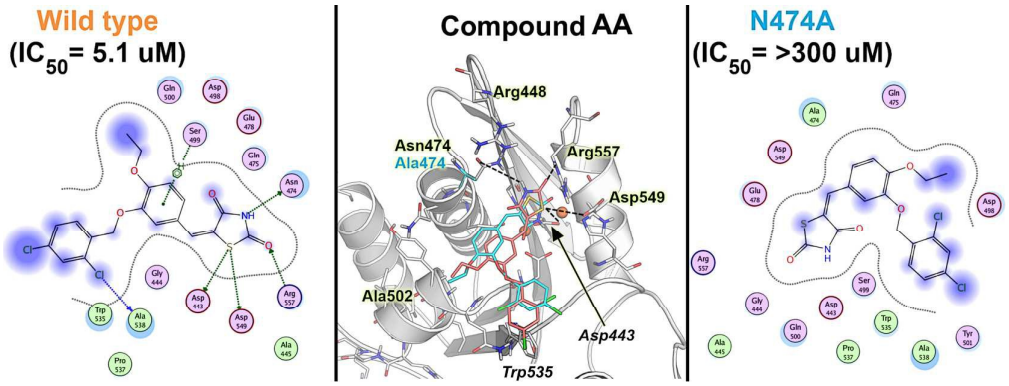


Figure 8. The comparison of various binding modes of compound **AA** in the RNase H active site, wild type is shown in orange and N474A mutant shown in cyan. Important residues are highlighted.

The other two compounds, **AA** and **A**, showed a similar behavior when tested against the mutated enzymes, with more pronounced variations in IC_{50} in the case of compound **AA**. Compound **AA** totally lost its potency of inhibition when assessed against N474A (>59 fold as compared to the WT) in accordance with the "induced-fit" calculation that clearly shows the different binding interaction of **AA** in N474A-substituted and WT RT (**Figure 8**). Although the binding mode looks quite similar in both cases, the binding pose in N474A is slightly shifted, leading to a significantly weakening of its key interactions with residues such as Ala538, Arg557, Asp549, Asp443, Asn474 and Ser499. Compound **AA** showed a significant loss in potency also in the case of A502F (5 fold) and R557A 2.5 fold, proving to have multiple and strong interactions within the RNase H domain, and, as expected, no change was seen for R448A. The overall results show that compound **AA** has a binding mode that differs significantly from the one reported for NHQD, [22] BHMP07, [35] and DKAs, [30] providing interesting perspectives for the hit optimization.

Furthermore, in order to confirm if the scaffolds are suitable of improvements without becoming toxic, and to check if they were able to inhibit HIV-1 replication in cell-based assays, compounds **A**, **AA**, and **AB** were tested for cytotoxicity and inhibition of HIV-1 replication (data not shown). The results showed that compounds **A**, **AA**, and **AB** are not toxic up to 40 μM. Unfortunately, at the highest concentration tested they were not able to inhibit viral replication, however, the structural insights gained from the present work will drive the scaffolds improvements in order to reach the antiviral activity.

View Article Online
DOI: 10.1039/C7MD00600D

CONCLUSIONS

In the effort to search for novel inhibitors of HIV-1 RT associated RNase H function we have identified 4 scaffolds **A, AA, AB** and **AC** that inhibited the RNase H function in the low micromolar range. Binding affinity studies, computational investigation of the binding modes, magnesium-complexation studies and a site-directed mutagenesis revealed that the most active compound **AA** inhibit HIV-1 RNase H with an IC_{50} of 5.1 μM, and showed a Mg-independent mode of action that involves multiple interactions within the RNase H domain, including the one with the highly conserved residue Asn474. The molecular determinants identified in this study of the binding mode will serve as a good starting point for further ongoing lead optimization.

MATERIALS AND METHODS

In silico screening protocol

Specs compounds (www.specs.net) from the ZINC database[63] were retrieved due to the fact that these compounds are commercially available for experimental screening. The purity of the final compounds (hits) used in the biochemical screening was verified by the chemical vendor using LC-MS and/or ^{1}H NMR and showed at least 95% purity. Out of 277.324 compounds, 1943 compounds were removed due to missing stereo chemical information using the KNIME workflow.[64] The remaining, 275.381 compounds were imported into the OMEGA (a 3D conformation generating tool) tool from the OpenEye software. [65] By default OMEGA reports multiple conformers for all compounds using the MMFF94S force field, however, in the present study only a single i.e., the lowest energy conformer was used.

Compound preprocessing

3D coordinates of all the compounds were imported into the Maestro module[66] available in the Schrödinger suite for further structural filtration and virtual screening process. The virtual screening (VS) workflow implemented in the Schrödinger suite was used. This includes ligand pre- and post-preparations [66-69] (e.g. generation of tautomers, minimization etc.), Lipinski filter, reactive group filters and docking workflow. In each step of the VS workflow, a significant reduction in the number of compounds was achieved. [67]

View Article Online
DOI: 10.1039/C7MD00600D

Homology modeling and structure-based screening

A computational model of the HIV-1 RT associated RNase H domain was built from an X-ray crystal structure (resolution of 1.4 Å) downloaded from the Protein Data Bank (PDB ID: 3QIO). [22] Missing residues were added using the Swiss-Model. [70] Subsequently, the atomic coordinate of the protein was imported into the Maestro module[66] implemented in Schrödinger Suite and the protein was further optimized (e.g., adding hydrogen atoms, assigning correct bond orders, building di-sulfide bonds and replacing crystal bound Mn^{2+} ions with catalytically active Mg^{2+} ions) using the Protein Preparation Wizard.[69] The docking experiments were performed using the Glide module[67] of the Schrödinger Suite as described in the previous section. Before the compound screening, the protein structure and the docking protocol were assessed by reproducing the bound conformation of ligands (3-hydroxy-6-(phenylsulfonyl)quinazoline-2,4(1H,3H)-dione) by docking. [54] In this study, water molecules at the active site beyond 3 Å from the bound ligand were deleted.

Shape-based similarity screening

The rapid overlay of chemical structures (ROCS)[59] and electrostatic (EON)[60] screening tools from the OpenEye [71] were used for shape and electrostatics based structural similarity searches. In both methods, the specs database is screened against the shape and electrostatic potential of a query molecule against a reference molecule. A detailed account of the theory behind these methods can be found elsewhere. [53] Once the overlap is optimized, the shape or electrostatics similarity is computed using Tanimoto Combo (ranges 0-2, a high score indicates the higher shape and pharmacophore similarity).

Structural clustering

For compound clustering, a set of 166 MACCS structural keys was computed, subsequently, the hierarchical clustering was employed using the Tanimoto similarity (TS) as method of metric (TS_{AB} =c/a+b-c, a= bits set to 1 in molecule A, b = bits set to 1 in molecule B, c = number of 1 bits common to molecule A and B).

Binding affinity prediction

Prediction of binding affinities for ligands in lead identification and optimization is a major challenge in computer-aided drug design. In this study, MM-GBSA was used to validate the binding mode of a series of ligands that binds to the protein. A detailed

View Article Online
DOI: 10.1039/C7MD00600D

account of the theory behind this method can be found elsewhere. [72, 73] To capture the "induced-fit" effect, protein-ligand complexes from docking experiments were further energy optimized using the Prime MM-GBSA with an implicit solvation model. During the optimization, both protein and ligands were treated as flexible for instance 5Å region of the protein around the ligand was treated as flexible.

HIV-DNA polymerase-independent RNase H activity determination

HIV RT-associated RNase H activity was measured as described, [32] briefly in 100 µL reaction volume containing 50 mM Tris-HCl buffer pH 7.8, 6 mM $MgCl_2$, 1 mM dithiothreitol (DTT), 80 mM KCl, 0.25 µM hybrid RNA/DNA 5'-GAUCUGAGCCUGGGAGCU-Fluorescin-3' (HPLC, dry, QC: Mass Check) (available from Metabion) 5'-Dabcyl-AGCTCCCAGGCTCAGATC-3' (HPLC, dry, QC: Mass Check), increasing concentrations of inhibitor, whose dilution were made in water, and different amount of enzymes according to a linear range of dose-response curve: 20 ng of WT RT, 37.5 ng R448A, 62.5 ng N474A; 100ng A502F; 75 ng R557A. The reaction mixture was incubated for 1 h at 37 °C, stopped by addition of EDTA and products were measured with a multi-label counter plate reader Victor 3 (Perkin Elmer model 1420-051) equipped with filters for 490/528 nm (excitation/emission wavelength).

Site-directed Mutagenesis

Amino acid substitutions were introduced into the p66 HIV-1 RT subunit of a HIV-1 RT group M subtype B coded in a p6HRT-prot plasmid using the QuikChange mutagenesis kit by following the manufacturer's instructions (Agilent Technologies Inc., Santa Clara, CA). Heterodimeric RT was expressed essentially and purified as described. [74]

Magnesium complexation study

Complexation studies on were carried on the following compounds: AN-648/41665045(**A**); AG-690/36550050 (**AA**); AG-690/36165052 (**AB**); AH-487/15021157 (**AC**); AN-698/40677529 (**AD**); AN-698/40677530 (**AE**); AH-487/40935633 (**AF**); AH-487/41035472(**AG**). The effects of the magnesium ions were evaluated by mean of spectrophotometric method, using a Perkin Elmer lambda 40 UV-Vis spectrophotometer and Hellma quartz cuvette with 1cm optical path. Magnesium chloride 1M solution were purchased from Sigma-Aldrich (Milano, Italy) and was diluted with absolute ethanol to obtain 13 stock solutions ranging from 4 10^{-5}

View Article Online
DOI: 10.1039/C7MD00600D

to 0.2 M. Each studied compound was dissolved in 50-100 mL of absolute ethanol and the final concentration ranges between 3×10^{-5} and 8×10^{-5}M; the obtained solutions were maintained 5 minutes in a ultrasonic bath. Each solution (3 mL) was placed in a cuvette and the UV-Vis spectra was recorded between 220 and 500 or 600 nm using ethanol as reference. Thereafter, small volumes (10-15 μL) of the appropriate $MgCl_2$ ethanolic stock solution were added both in the sample and in the reference cuvettes and the UV-Vis spectra were repeated; the Mg^{2+} concentration in the solutions was increased from 0 to 100-200 fold with respect to the studied compound, in about twenty consecutive increments. Each experiment was triplicate.
Due to the scarce solubility of compound AH-487/40935633 (**AF**) in absolute ethanol was not possible to prepare a solution of known concentration; for this reason only three spectra were recorded, one in the absence and two in presence of a great excess of Mg^{2+}. The titration graphs for the selected compounds are reported in the supplementary material.

Acknowledgements

This work has been supported by Department of Physics, Chemistry and Pharmacy, University of Southern Denmark, Denmark. The authors are thankful to ChemAxon and OpenEye scientific software for providing a free academic license. Authors thank the Italian MIUR for financial support (PRIN 2010, 2010W2KM5L_003).

Conflicts of interest

The authors declare no competing interest.

Author Information

VP and AC contributed equally to this manuscript.

Present address

VP: Department of Chemistry - BMC, Uppsala University, SE-75123, Uppsala, Sweden;

CS: Department of Chemistry and Bioscience, Aalborg University, Aalborg, Denmark.

Author contributions

VP, JK and ET conceived and designed the experiments; VP and CS performed computational modeling, AC, NG, FE performed RNase H inhibition experiments, LS, RC and FP performed chelation experiments. VP AC RDS ET JK analyzed the

View Article Online
DOI: 10.1039/C7MD00600D

data; VP ET JK RDS contributed reagents/materials/analysis tools. The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Supporting Information

SF1: Comparison of frequency of compounds used for chelation calculation (n=1135, blue bar) and compounds selected from chelation calculation (n=180, red bar); **SF2**: List of compounds used for cluster analysis and its chelation score; **SF3**: Compounds used in the first screening and the hits selected for second is highlighted; **SF4**: Comparison of schematic representation of protein-ligand interactions of compound B and AQ; **SF5-9**: UV□Vis Spectra of compound AF, AE, AA, AC, AD; **SF10**: Comparison of binding modes of BPT in various site-directed mutagenesis structural models with schematic representation of protein □ ligand interactions; **SF11**: Spectral data of compounds screened against HIV-1 RT RNase H.

References

1. A. S. Fauci, *Nat Med*, 2003, **9**, 839-843.
2. R. K. Gupta, J. Gregson, N. Parkin, H. Haile-Selassie, A. Tanuri, L. Andrade Forero, P. Kaleebu, C. Watera, A. Aghokeng, N. Mutenda, J. Dzangare, S. Hone, Z. Z. Hang, J. Garcia, Z. Garcia, P. Marchorro, E. Beteta, A. Giron, R. Hamers, S. Inzaule, L. M. Frenkel, M. H. Chung, T. de Oliveira, D. Pillay, K. Naidoo, A. Kharsany, R. Kugathasan, T. Cutino, G. Hunt, S. Avila Rios, M. Doherty, M. R. Jordan and S. Bertagnolio, *Lancet Infect Dis*, 2017, DOI: 10.1016/S1473-3099(17)30702-8.
3. E. De Clercq, *Curr Med Chem*, 2001, **8**, 1543-1572.
4. UNAIDS report, on the global AIDS epidemic 2012: Join UN program on HIV/AIDS 2012.
5. E. De Clercq, *BBA-Mol Basis Dis*, 2002, **1587**, 258-275.
6. M. M. Zdanowicz, *Am J Pharm Educ*, 2006, **70**.
7. N. A. Margot, P. Wong, R. Kulkarni, K. White, D. Porter, M. E. Abram, C. Callebaut and M. D. Miller, *J Infect Dis*, 2017, **215**, 920-927.
8. A. Schneider, A. Corona, I. Sporing, M. Jordan, B. Buchholz, E. Maccioni, R. Di Santo, J. Bodem, E. Tramontano and B. M. Wohrl, *Nucl Acids Res*, 2016, **44**, 2310-2322.
9. J. Darbyshire, *Drugs*, 1995, **49**, 1-3.
10. FDA, Antiretroviral drugs used in the treatment of HIV infection, http://www.fda.gov, (accessed December, 2014).
11. C. Flexner, *Nat Rev Drug Discov*, 2007, **6**, 959-966.
12. W. C. Greene, Z. Debyser, Y. Ikeda, E. O. Freed, E. Stephens, W. Yonemoto, R. W. Buckheit, J. A. Este and T. Cihlar, *Antivir Res*, 2008, **80**, 251-265.
13. A. Corona, F. Esposito and E. Tramontano, *Future Virol*, 2014, **9**, 445-448.
14. T. Ilina, K. Labarge, S. G. Sarafianos, R. Ishima and M. A. Parniak, *Biology*, 2012, **1**, 521-541.

View Article Online
DOI: 10.1039/C7MD00600D

15. E. Tramontano and R. Di Santo, *Curr Med Chem*, 2010, **17**, 2837-2853.
16. O. Schatz, F. Cromme, T. Naas and D. Lindemann, *Gene Regulation and AIDS*. 1990, portfolio, 293-404.
17. A. Corona, F. S. Di Leva, S. Thierry, L. Pescatori, G. C. Crucitti, F. Subra, O. Delelis, F. Esposito, G. Rigogliuso, R. Costi, S. Cosconati, E. Novellino, R. Di Santo and E. Tramontano, *Antimicrob Agents Chemother*, 2014, **58**, 6101-6110.
18. E. Tramontano, F. Esposito, R. Badas, R. Di Santo, R. Costi and P. La Colla, *Antivir Res*, 2005, **65**, 117-124.
19. J. F. Davies, 2nd, Z. Hostomska, Z. Hostomsky, S. R. Jordan and D. A. Matthews, *Science*, 1991, **252**, 88-95.
20. D. M. Himmel, K. A. Maegley, T. A. Pauly, J. D. Bauman, K. Das, C. Dharia, A. D. Clark, K. Ryan, M. J. Hickey, R. A. Love, S. H. Hughes, S. Bergqvist and E. Arnold, *Structure*, 2009, **17**, 1625-1635.
21. D. M. Himmel, S. G. Sarafianos, S. Dharmasena, M. M. Hossain, K. McCoy-Simandle, T. Ilina, A. D. Clark, J. L. Knight, J. G. Julias, P. K. Clark, K. Krogh-Jespersen, R. M. Levy, S. H. Hughes, M. A. Parniak and E. Arnold, *ACS Chem Biol*, 2006, **1**, 702-712.
22. E. B. Lansdon, Q. Liu, S. A. Leavitt, M. Balakrishnan, J. K. Perry, C. Lancaster-Moyer, N. Kutty, X. H. Liu, N. H. Squires, W. J. Watkins and T. A. Kirschberg, *Antimicrob Agents Chemother*, 2011, **55**, 2905-2915.
23. O. Schatz, F. V. Cromme, T. Naas, D. Lindemann, J. Mous and S. F. J. Legrice, *Adv Ap Biot*, 1990, **7**, 293-303.
24. M. T. Christen, L. Menon, N. S. Myshakina, J. Ahn, M. A. Parniak and R. Ishima, *Chem Biol Drug Des*, 2012, **80**, 706-716.
25. R. Meleddu, V. Cannas, S. Distinto, G. Sarais, C. Del Vecchio, F. Esposito, G. Bianco, A. Corona, F. Cottiglia, S. Alcaro, C. Parolin, A. Artese, D. Scalise, M. Fresta, A. Arridu, F. Ortuso, E. Maccioni and E. Tramontano, *ChemMedChem*, 2014, **9**, 1869-1879.
26. V. Poongavanam, J. M. H. Olsen and J. Kongsted, *Integr Biol-UK*, 2014, **6**, 1010-1022.
27. M. Di Grandi, M. Olson, A. S. Prashad, G. Bebernitz, A. Luckay, S. Mullen, Y. B. Hu, G. Krishnamurthy, K. Pitts and J. O'Connell, *Bioorg Med Chem Lett*, 2010, **20**, 398-402.
28. J. Didierjean, C. Isel, F. Querre, J. F. Mouscadet, A. M. Aubertin, J. Y. Valnot, S. R. Piettre and R. Marquet, *Antimicrob Agents Chemother*, 2005, **49**, 4884-4894.
29. S. Distinto, F. Esposito, J. Kirchmair, M. C. Cardia, M. Gaspari, E. Maccioni, S. Alcaro, P. Markt, G. Wolber, L. Zinzula and E. Tramontano, *Eur J Med Chem*, 2012, **50**, 216-229.
30. T. A. Kirschberg, M. Balakrishnan, N. H. Squires, T. Barnes, K. M. Brendza, X. Chen, E. J. Eisenberg, W. Jin, N. Kutty, S. Leavitt, A. Liclican, Q. Liu, X. Liu, J. Mak, J. K. Perry, M. Wang, W. J. Watkins and E. B. Lansdon, *J Med Chem*, 2009, **52**, 5781-5784.
31. M. Nowotny, S. A. Gaidamakov, R. Ghirlando, S. M. Cerritelli, R. J. Crouch and W. Yang, *Mol Cell*, 2007, **28**, 513-513.

View Article Online
DOI: 10.1039/C7MD00600D

32. R. Costi, M. Metifiot, S. Chung, G. C. Crucitti, K. Maddali, L. Pescatori, A. Messore, V. N. Madia, G. Pupo, L. Scipione, S. Tortorella, F. S. Di Leva, S. Cosconati, L. Marinelli, E. Novellino, S. F. J. Le Grice, A. Corona, Y. Pommier, C. Marchand and R. Di Santo, *J Med Chem*, 2014, **57**, 3223-3234.

33. K. Klumpp, J. Q. Hang, S. Rajendran, Y. L. Yang, A. Derosier, P. W. K. In, H. Overton, K. E. B. Parkes, N. Cammack and J. A. Martin, *Nucl Acids Res*, 2003, **31**, 6852-6859.

34. S. R. Budihas, I. Gorshkova, S. Gaidamakov, A. Wamiru, M. K. Bona, M. A. Parniak, R. J. Crouch, J. B. McMahon, J. A. Beutler and S. F. J. Le Grice, *Nucl Acids Res*, 2005, **33**, 1249-1256.

35. H. P. Su, Y. W. Yan, G. S. Prasad, R. F. Smith, C. L. Daniels, P. D. Abeywickrema, J. C. Reid, H. M. Loughran, M. Kornienko, S. Sharma, J. A. Grobler, B. Xu, V. Sardana, T. J. Allison, P. D. Williams, P. L. Darke, D. J. Hazuda and S. Munshi, *J Virol*, 2010, **84**, 7625-7633.

36. H. Fuji, E. Urano, Y. Futahashi, M. Hamatake, J. Tatsumi, T. Hoshino, Y. Morikawa, N. Yamamoto and J. Komano, *J Med Chem*, 2009, **52**, 1380-1387.

37. S. M. Chung, M. Wendeler, J. W. Rausch, G. Beilhartz, M. Gotte, B. R. O'Keefe, A. Bermingham, J. A. Beutler, S. X. Liu, X. W. Zhuang and S. F. J. Le Grice, *Antimicrob Agents Chemother*, 2010, **54**, 3913-3921.

38. K. A. Kirby, B. Marchand, Y. T. Ong, T. P. Ndongwe, A. Hachiya, E. Michailidis, M. D. Leslie, D. V. Sietsema, T. L. Fetterly, C. A. Dorst, K. Singh, Z. Wang, M. A. Parniak and S. G. Sarafianos, *Antimicrob Agents Chemother*, 2012, **56**, 2048-2061.

39. Q. G. Gong, L. Menon, T. Ilina, L. G. Miller, J. Ahn, M. A. Parniak and R. Ishima, *Chem Biol Drug Des*, 2011, **77**, 39-47.

40. S. M. Chung, J. T. Miller, B. C. Johnson, S. H. Hughes and S. F. J. Le Grice, *J Biol Chem*, 2012, **287**, 4066-4075.

41. F. Esposito, T. Kharlamova, S. Distinto, L. Zinzula, Y. C. Cheng, G. Dutschman, G. Floris, P. Markt, A. Corona and E. Tramontano, *FEBS J*, 2011, **278**, 1444-1457.

42. F. Esposito and E. Tramontano, *Antiviral Chem Chemother*, 2014, **23**, 129-144.

43. R. Costi, M. Metifiot, F. Esposito, G. Cuzzucoli Crucitti, L. Pescatori, A. Messore, L. Scipione, S. Tortorella, L. Zinzula, E. Novellino, Y. Pommier, E. Tramontano, C. Marchand and R. Di Santo, *J Med Chem*, 2013, **56**, 8588-8598.

44. G. Cuzzucoli Crucitti, M. Metifiot, L. Pescatori, A. Messore, V. N. Madia, G. Pupo, F. Saccoliti, L. Scipione, S. Tortorella, F. Esposito, A. Corona, M. Cadeddu, C. Marchand, Y. Pommier, E. Tramontano, R. Costi and R. Di Santo, *J Med Chem*, 2015, **58**, 1915-1928.

45. A. Corona, F. S. di Leva, G. Rigogliuso, L. Pescatori, V. N. Madia, F. Subra, O. Delelis, F. Esposito, M. Cadeddu, R. Costi, S. Cosconati, E. Novellino, R. di Santo and E. Tramontano, *Antiviral Res*, 2016, **134**, 236-243.

46. V. Summa, A. Petrocchi, F. Bonelli, B. Crescenzi, M. Donghi, M. Ferrara, F. Fiore, C. Gardelli, O. Gonzalez Paz, D. J. Hazuda, P. Jones, O. Kinzel, R. Laufer, E. Monteagudo, E. Muraglia, E. Nizi, F. Orvieto, P. Pace, G.

View Article Online
DOI: 10.1039/C7MD00600D

Pescatore, R. Scarpelli, K. Stillmock, M. V. Witmer and M. Rowley, *J Med Chem*, 2008, **51**, 5843-5855.

47. A. Corona, T. Masaoka, G. Tocco, E. Tramontano and S. F. J. Le Grice, *Future Med Chem*, 2013, **5**, 2127-2139.
48. A. K. Felts, K. LaBarge, J. D. Bauman, D. V. Patel, D. M. Himmel, E. Arnold, M. A. Parniak and R. M. Levy, *J Chem Inf Model*, 2011, **51**, 1986-1998.
49. J. G. Julias, M. J. McWilliams, S. G. Sarafianos, E. Arnold and S. H. Hughes, *Proc Natl Acad Sci USA,* 2002, **99**, 9515-9520.
50. J. W. Rausch, D. Lener, J. T. Miller, J. G. Julias, S. H. Hughes and S. F. Le Grice, *Biochemistry*, 2002, **41**, 4856-4865.
51. V. Poongavanam, V. Namasivayam, M. Vanangamudi, H. Al Shamaileh, R. N. Veedu, J. Kihlberg and N. A. Murugan, *Wiley Interdisciplinary Reviews: Comput Mol Sci*, 2017, 8, e1328.
52. G. Schneider and H. J. Bohm, *Drug Discov Today*, 2002, **7**, 64-70.
53. P. Vasanthanathan, J. Lastdrager, C. Oostenbrink, J. N. M. Commandeur, N. P. E. Vermeulen, F. S. Jorgensen and L. Olsen, *MedChemComm*, 2011, **2**, 853-859.
54. V. Poongavanam and J. Kongsted, *Plos One*, 2013, **8**, e73478.
55. V. Poongavanam, C. Steinmann and J. Kongsted, *Plos One*, 2014, **9**, e98659.
56. Canvas. Schrödinger, LLC, New York, NY, 2014
57. M. W. Olson, G. M. Di and A. Prashad, US Patent No. WO2005090316A1 2005.
58. E. Naylor, A. Arredouani, S. R. Vasudevan, A. M. Lewis, R. Parkesh, A. Mizote, D. Rosen, J. M. Thomas, M. Izumi, A. Ganesan, A. Galione and G. C. Churchill, *Nat Chem Biol*, 2009, **5**, 220-226.
59. J. A. Grant, M. A. Gallardo and B. T. Pickup, *J Comput Chem*, 1996, **17**, 1653-1666.
60. S. W. Muchmore, A. J. Souers and I. Akritopoulou-Zanze, *Chem Biol Drug Des*, 2006, **67**, 174-176.
61. R. Costi, R. D. Santo, M. Artico, S. Massa, R. Ragno, R. Loddo, M. La Colla, E. Tramontano, P. La Colla and A. Pani, *Bioorg Med Chem*, 2004, **12**, 199-215.
62. S. M. Chung, D. M. Himmel, J. K. Jiang, K. Wojtak, J. D. Bauman, J. W. Rausch, J. A. Wilson, J. A. Beutler, C. J. Thomas, E. Arnold and S. F. J. Le Grice, *J Med Chem*, 2011, **54**, 4462-4473.
63. J. J. Irwin, T. Sterling, M. M. Mysinger, E. S. Bolstad and R. G. Coleman, *J Chem Inf Model*, 2012, **52**, 1757-1768.
64. R. B. Michael, C. Nicolas, D. Fabian, R. G. Thomas, K. Tobias, tter, M. Thorsten, O. Peter, T. Kilian and W. Bernd, *SIGKDD Explor. Newsl.*, 2009, **11**, 26-31.
65. P. C. D. Hawkins, A. G. Skillman, G. L. Warren, B. A. Ellingson and M. T. Stahl, *J Chem Inf Model*, 2010, **50**, 572-584.
66. Maestro, Schrödinger, LLC, New York, NY, 2014.
67. R. A. Friesner, J. L. Banks, R. B. Murphy, T. A. Halgren, J. J. Klicic, D. T. Mainz, M. P. Repasky, E. H. Knoll, M. Shelley, J. K. Perry, D. E. Shaw, P. Francis and P. S. Shenkin, *J Med Chem*, 2004, **47**, 1739-1749.

View Article Online
DOI: 10.1039/C7MD00600D

68. LigPrep, Schrödinger, LLC, New York, NY, 2014.
69. G. M. Sastry, M. Adzhigirey, T. Day, R. Annabhimoju and W. Sherman, *J Comput Aid Mol Des*, 2013, **27**, 221-234.
70. K. Arnold, L. Bordoli, J. Kopp and T. Schwede, *Bioinformatics*, 2006, **22**, 195-201.
71. OpenEye. (2014) OpenEye Scientifc Software, CR Sanda Fe, NM 87508, USA (https://http://www.eyesopen.com).
72. P. D. Lyne, M. L. Lamb and J. C. Saeh, *J Med Chem*, 2006, **49**, 4805-4808.
73. J. Du, H. J. Sun, L. L. Xi, J. Z. Li, Y. Yang, H. X. Liu and X. J. Yao, *J Comput Chem*, 2011, **32**, 2800-2809.
74. A. Corona, R. Meleddu, F. Esposito, S. Distinto, G. Bianco, T. Masaoka, E. Maccioni, L. Menendez-Arias, S. Alcaro, S. F. J. Le Grice and E. Tramontano, *Plos One*, 2016, **11**.

Figure legends

Figure 1. The topology of HIV-RT associated RNase H is shown with a bound ligand NHQD (green, ball and stick model). The catalytically important residues and magnesium ions (orange sphere) have been highlighted. 2D structures of representative molecules belong to the active site and allosteric RNase H inhibition.

View Article Online
DOI: 10.1039/C7MD00600D

Figure 2. Overall workflow of structure-based virtual screening strategies applied.

Figure 3. Initial hit molecules (A, B, C) from the first screening and compound "B" is highlighted regions where it shares common structural pattern as BHMP07. Compounds AQ, AR and as found as analogs of B and C by the shaped-based screening (see text for detail).

Figure 4. Comparison of structure and activity of **A** analogs.

Figure 5. Comparison of the binding mode of known RNase H inhibitors, in a X-ray crystal, and important residues are highlighted. Panel A, N-hydroxy quinazolinedione inhibitor (NHQD, active site binder, green) and BHMP07 (allosteric site binder, cyan). Panel B, compound **A** (magenta), Panel C compound **AA** (blue), Panel D, compound **AB** (cyan).

Figure 6. Titration graph of inhibitor with Mg^{2+}; Panel A: compound **AC**, Panel B: compound **A**, Panel C: Compound **AA** and Panel D: compound **AB**.

Figure 7. The comparison of the binding mode of compounds A, AA, AB and BPT in the wild type and various clinically relevant mutants.

Figure 8. The comparison of various binding modes of compound **AA** in the RNase H active site, wild type is shown in orange and N474A mutant shown in cyan. Important residues are highlighted.

Table 1.
In silico prediction scores and biological effects of VS hits on HIV-1 RT-associated RNase H activity

ID	Vendor code	Cluster	Chelation[#]	XP [#]	Mol Wt (Da)	IC_{50} (μM) [a]
A[1]	AN-648/41665045	36	-21.17	-7.68	317.34	9.35 ± 0.007

View Article Online
DOI: 10.1039/C7MD00600D

B[1]	AG-205/10367010	34	-51.14	-9.91	316.27	49.90 ± 9.7
C[1]	AL-281/41935302	21	-42.49	-6.86	289.33	61.25 ± 2.7
D[1]	AK-968/41922459	1	-11.03	-8.38	295.77	>100 (73.40)
E[1]	AK-968/12713114	3	-34.61	-8.01	241.33	>100 (76.28)
F[1]	AQ-099/42181915	4	-59.34	-8.57	286.10	>100 (84.45)
G[1]	AP-853/43445377	8	-13.66	-9.67	275.29	>100 (82.70)
H[1]	AF-399/07629003	10	-22.42	-9.18	228.21	>100 (89.07)
I[1]	AE-848/08809049	11	-23.03	-7.36	369.17	>100 (69.14)
J[1]	AN-648/41701113	11	-13.91	-6.86	348.38	>100 (77.90)
K[1]	AQ-149/42447764	13	-87.78	-7.27	299.35	>100 (81.80)
L[1]	AQ-149/43372378	13	-25.62	-7.19	319.40	>100 (77.55)
M[1]	AK-968/41922490	27	-22.30	-8.13	190.24	>100 (83.52)
N[1]	AI-204/31719028	28	-37.80	-8.45	237.28	>100 (76.50)
O[1]	AJ-292/41685670	28	-35.68	-9.83	309.34	>100 (74.50)
P[1]	AN-329/43248835	29	-54.79	-8.92	304.34	>100 (83.70)
Q[1]	AO-080/42837941	30	-42.83	-9.65	305.37	>100 (86.00)
R[1]	AK-968/41923774	42	-33.40	-9.04	322.34	>100 (80.00)
S[1]	AO-433/42007915	49	-27.79	-7.54	371.43	>100 (80.41)
T[1]	AF-399/41368457	45	-34.22	-9.34	335.36	>100 (80.00)
U[1]	AG-205/32708023	31	-12.43	-7.79	359.34	>100 (59.43)
V[1]	AG-690/33090008	39	-35.05	-7.49	313.35	>100 (87.24)
W[1]	AK-968/41922494	43	-15.19	-9.05	340.14	>100 (94.56)
X[1]	AK-968/40355672	26	-18.26	-8.15	236.31	>100 (70.44)
Y[1]	AO-080/43441633	30	-35.71	-8.50	374.44	>100 (80.25)
AA[2]	AG-690/36550050	36	-	-	424.30	5.17 ±0.56
AB[2]	AG-690/36165052	36	-	-	402.45	12.82 ± 0.68
AC[2]	AH-487/15021157	36	-	-	456.33	8.73 ± 1.84
AD[2]	AN-698/40677529	36	-	-	315.37	12.30 ± 1.92
AE[2]	AN-698/40677530	36	-	-	331.83	16.06 ± 0.68
AF[2]	AH-487/40935633	36	-	-	591.09	17.78 ± 4.13
AG[2]	AH-487/41035472	36	-	-	331.37	19.31 ± 1.05
AH[2]	AH-487/41457445	36	-	-	367.36	28.11 ± 4.66
AI[2]	AH-487/40935863	36	-	-	260.29	52.80 ± 3.43
AJ[2]	AN-988/14610013	36	-	-	266.68	84.20 ± 15.50
AK[2]	AH-487/41449196	36	-	-	276.25	91.57 ± 11.92
AL[2]	AN-698/40704420	36	-	-	294.29	>100 (65.5)[b]
AM[2]	AG-690/40144774	36	-	-	265.29	>100 (60.0)[b]
AN[2]	AH-487/40935550	36	-	-	388.11	>100 (67.5)[b]
AO[2]	AH-487/41449786	36	-	-	296.71	>100 (70.0)[b]
AP[2]	AG-690/40124767	36	-	-	274.13	>100 (80.5)[b]
AQ[2]	AK-968/41022607	36	-	-	279.25	>100 (89.5)[b]
BA[2]	AN-329/40081786	34	-	-	297.27	36.43 ± 5.83
BB[2]	AF-399/15030247	34	-	-	321.32	>100 (55.0)[b]
CA[2]	AK-968/15608063	21	-	-	391.47	>100 (93.0)[b]

View Article Online
DOI: 10.1039/C7MD00600D

[#] Chelation and docking scores are in kcal mol^{-1} unit;
[a] compound concentration required to reduce the HIV-1-RT associated RNase H activity by 50%; The table reports the average and standard deviation of three independent experiments.

[b] Percentage of control activity in the presence of 100 μM concentration of compound;
[1]First screening;
[2]Secondary screening (Shape based search).

Table. 2

Effect of selected compounds on HIV-1 RT associated RNase H function of mutated RTs

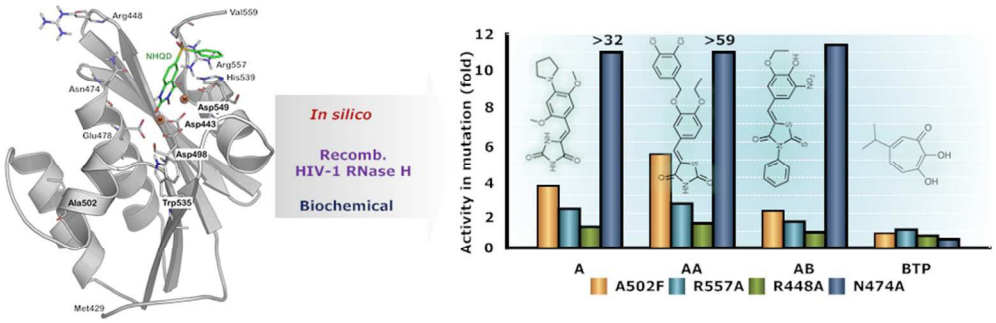
	A		AA		AB		BTP	
	IC_{50} (µM) [a]	Fold[b]	**IC_{50} (µM)** [a]	Fold[b]	**IC_{50} (µM)** [a]	Fold[b]	**IC_{50} (µM)** [a]	Fold[b]
Wild	9.35 ± 0.73	-	5.1 ± 1.89	-	12.82 ± 1.05	-	0.20± 0.02	-
A502F	32.52 ± 2.83	3.5***	26.72 ± 3.14	5.3***	26.35 ± 1.01	2.1****	0.17 ± 0.03	0.85
R557A	20.54 ± 1.26	2.2***	12.53 ± 1.36	2.5**	18.61 ± 0.42	1.5**	0.21 ± 0.05	1.05
R448A	10.96 ± 1.12	1.20	7.01 ± 0.78	1.40	11.28 ± 0.64	0.9	0.14 ± 0.02	0.70
N474A	>300	>32****	>300	>59****	146 ± 12	11.4****	0.10 ± 0.01	0.50

The table reports the average and standard deviation of three independent experiments.

[a] Concentration required to inhibit HIV-1 RT-associated RNase H activity by 50% obtained by three independent experiments (reported as average ± standard deviation). p value < 0.05 (*); p value <0.01 (**); p value <0.001 (***); p value <0.0001(****).

[b] Fold of increase with respect to wt RT.

View Article Online
DOI: 10.1039/C7MD00600D



85x28mm (300 x 300 DPI)