BL-17A/2022G642

Sticklac-Derived Natural Compounds Inhibiting RNase H Activity of HIV-1 Reverse Transcriptase

Tyuji Hoshino [1,*], Yuma Ito [1], Huiyan Lu [1]
[1] Graduate School of Pharmaceutical Sciences, Chiba University,
1-8-1 Inohana, Chuo-ku, Chiba 260-8675, Japan

1 Introduction

Human immunodeficiency virus type 1 (HIV-1) has three viral enzymes. One is reverse transcriptase (RT), which generates single-strand DNA from the viral RNA gene. Another is integrase, which incorporates the generated DNA into the host gene. The other is HIV-1 protease, which processes the precursor peptides into mature viral proteins. Currently, approved drugs are available for these enzymes. Highly active antiretroviral therapy (HAART) effectively suppresses viral replication and the emergence of mutants, and it currently constitutes a standard treatment for HIV-1 infectious diseases. Despite the marked efficacy of HAART, cytotoxicity and resistance to drugs used in HAART sometimes limit long-term chemotherapy only with the currently available drugs. The discovery of a novel class of chemicals having a mechanism of action that is different from those of approved drugs is important to establish a new anti-viral therapy. Hence, there is a continuing need for the development of novel antiviral agents with new inhibitory profiles.

Reverse transcriptase is a unique protein because it has multi-enzymatic functions. One function is polymerase activity to generate genomic DNA. Another function is RNase H activity. RNase H acts on the viral RNA hybridized to DNA and hydrolyzes the phosphodiester bond of viral RNA. Reverse transcriptase is a heterodimer consisting of p51 and p66 subunits. The RNase H active site is far apart from the polymerase domain in RT. Although many HIV-1 drugs have been produced, there is no approved inhibitor for RNase H activity. Inhibitors blocking the viral enzymatic activity are effective in chemotherapy, and chemicals blocking RNase H activity are expected to complement the currently available anti-HIV-1 drugs. Therefore, RT-associated RNase H activity is an attractive target for the development of a novel class of antiviral drugs.

Computational analysis is indispensable for screening hit chemicals and for rational drug design [1]. In this work, we applied in silico screening to a collection of 400 natural compounds [2] to select candidates for an experimental assay. The inhibitory potency was evaluated by an in vitro enzymatic assay and a viral cell-based assay. X-ray crystal analysis revealed the binding modes of three hit compounds. Molecular dynamics (MD) simulations suggested the stability of the binding modes. A potent chemical modification of the hit compounds is discussed with molecular mechanics (MM) computations.

2 Experiment

The collection of natural compounds consisted of 400 chemicals derived from plants and fungi. Since some natural products were rare, the compounds were screened to select candidates for an experimental assay by a computational method. The narrowing down of the compounds for the assay was cost-effective and time-saving for experiments. Docking simulations were carried out for the individual compounds to generate binding poses to the target domain. The calculation model of the target domain was built from the crystal structure, PDB code: 7XIS, which was a complex structure of an HIV-1 RNase H recombinant protein and a nitro-furan-based inhibitor [3]. AutoDock Vina was utilized for docking simulation. Five binding poses were generated for each compound, and 2,000 binding poses were thus predicted in total. All of the 2,000 binding poses were optimized by molecular mechanics calculations [4]. Binding scores were obtained for the optimized poses. The binding pose with the best score was selected from the five predicted ones for each compound. Then the top 30 compounds were selected on the basis of the calculated binding scores.

A recombinant protein for crystallography was a conjugation of the HIV-1 RT-associated RNase H domain and the RNase H enzyme of E. coli. The HIV-1 RNase H domain is called p15, and a p15 partial protein shows no enzymatic activity. The conjugation of a helix of E. coli RNase H into p15 recovers the enzymatic activity. The recombinant protein, p15Ec, has been utilized to investigate the binding structure of an RNase H inhibitor in X-ray crystal analysis.

The recombinant protein was expressed in an E. coli Rosetta2 strain transformed with the pET50 vector in which p15Ec is coded in a fusion form with a soluble tag. The culture medium containing the transformed E. coli was incubated at 28℃ overnight after induction with 0.2 M isopropyl-β-D-thiogalactopyranoside (IPTG). The fusion protein was purified by a Co-affinity column. The soluble tag on the N-terminal side was cleaved by human rhinovirus 3C (HRV-3C) protease, followed by removal of the soluble tag and the HRV-3C protease with Ni-NTA

resin. The purified p15Ec protein was finally obtained by gel filtration.

The whole HIV-1 RT was used for the assay of compound inhibitory activity. The RT was expressed by a plasmid vector, RT69A, in which F160S and C280S amino mutations were introduced to promote crystallization. The recombinant RT was produced by bacterial expression, and a Co-affinity column was used for purification. The purified RT was dialyzed to remove imidazole from the buffer and then incubated with HRV-3C protease to cleave an N-terminal 6×His-tag. The HRV-3C protease and the cleaved 6×His-tag protein were separated by Ni-NTA resin. The RT protein was stored at -20 °C with a buffer of 50 mM Tris-HCl at pH 7.5, 200 mM NaCl, and 50% (v/v) glycerol.

The purified recombinant protein, p15Ec, was concentrated at 10 mg/mL for crystallization. The protein was crystallized by mixing with a precipitant solution of 100 mM MES, 10 mM $ZnSO_4$, 1 mM $MnCl_2$, and 26% (v/v) polyethylene glycol (PEG) 600 at pH6.6. The crystal was grown by the sitting drop vapor diffusion method at 18°C. The co-crystal was obtained by soaking a protein crystal in a solution containing each inhibitory compound. A solution containing 40% (v/v) 1,6-hexanediol was used as a cryo-protectant. X-ray diffractions were acquired at BL-17A beamline of KEK Photon-Factory (PF), Japan. The phase was solved by the molecular replacement method, in which a crystal structure of PDB code: 7XIS was used as a search model [2]. The diffraction data were indexed, collected, and merged by XDS. Molecular replacement and structure refinement were performed by Phenix 40 and visualized by Coot.

3 Results and Discussion

A search was performed with 400 natural compounds for active chemicals interacting with the RNase H domain by an in silico screening technique. A combination of docking simulations by AutoDock Vina and MM calculations by in-house software was carried out to reduce the number of compounds for an experimental assay. In the first step of the virtual screening, each of the 400 compounds was docked to the calculation model built from a crystal structure of HIV-1 RNase H recombinant protein, PDB code: 7XIS. The docking simulation generated 5 binding poses for each of 400 compounds. Hence, 2,000 binding poses were generated in total. In the second step, atom geometries of all of the 2,000 binding poses were optimized by MM calculations. The MM calculations gave the binding scores for the respective optimized poses. The best binding score among the 5 generated poses was selected for each compound. Then the top 30 compounds in ranking of the binding scores were selected from the 400 compounds.

The inhibitory activities of the selected compounds were evaluated by an enzymatic assay with the FRET technique. Initially, the suppression of RNase H activity was measured at a constant concentration of 50 μM for every compound. The measurement was performed twice with comparison to the control without any compounds. The measurements suggested that three compounds derived from sticklac suppressed RNase H enzymatic activity. A common chemical structure was observed in the three identified compounds, in which a hydroxy phenyl group was connected to an anthraquinone backbone. Compound 1 is laccaic acid A, in which an acetamido-ethyl is attached to the phenyl. Compounds 2 and 3 are laccaic acids C and E, in which the acetylamino-ethyl is replaced by amino-carboxy-ethyl and amino-ethyl, respectively. The inhibitory activity was evaluated in a two-fold serial dilution for the compound concentration (Figure 1). Compound 2 showed the best 50 % inhibitory concentration (IC_{50}) value among the three compounds. While the IC_{50} value of compound 2 was 8.1 μM, the values were more than 40 μM for compounds 1 and 3.

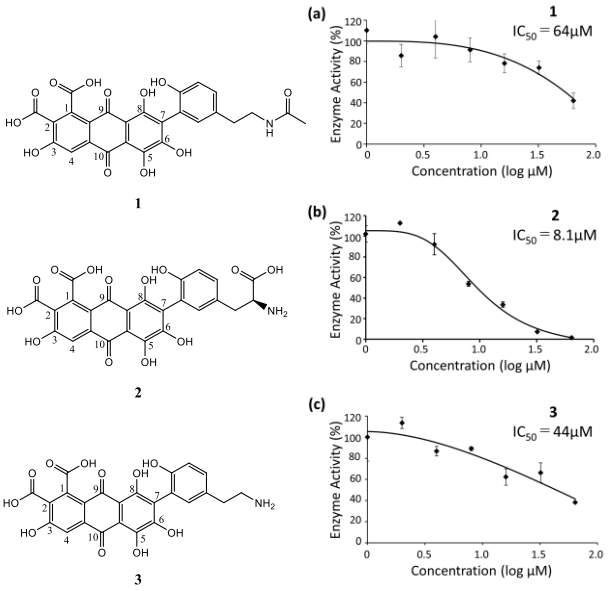

Fig. 1: Chemical structures of the hit natural compounds (left). Enzymatic assay of inhibitory potency of the identified compounds for HIV-1 RNase H activity (right).

The binding structures of the three inhibitory compounds to the RNase H active domain were investigated by X-ray crystal analysis. A recombinant protein, p15Ec,15 was used for the crystal analysis. Single crystals of p15Ec were obtained by the vapor diffusion sitting drop method. Compounds 1-3 were introduced into the crystals by the soaking technique. The co-crystals diffracted X-rays in the range of 1.9 - 2.2 Å in resolution. The space group of all of the crystals was $P4_12_12$, and one p15Ec-compound complex was detected in an asymmetric unit. Two

metal ions were observed at the active site for all of the crystal structures. The divalent metal ions are required to functionalize the RNase H activity. Mn2+ was employed for the divalent metals in crystallography to enforce the coordination bonds with acidic residues. One of the Mn2+ ions was held by Asp23, Glu58, and Asp78, while the other was held by Asp23 and Asp139 (Figure 2). These acidic residues, Asp23, Glu58, Asp78 and Asp139, in p15Ec correspond to Asp443, Glu478, Asp498, and Asp594 of the RNase H domain of HIV-1 RT. The lengths of the coordination bonds from these four acidic residues to metal ions ranged from 2.0 Å to 2.4 Å. Therefore, two Mn2+ ions are kept at the appropriate positions for activating the enzymatic function. An Fo-Fc omit map was depicted to determine the binding poses of compounds at the enzymatic active site (Figure 3).

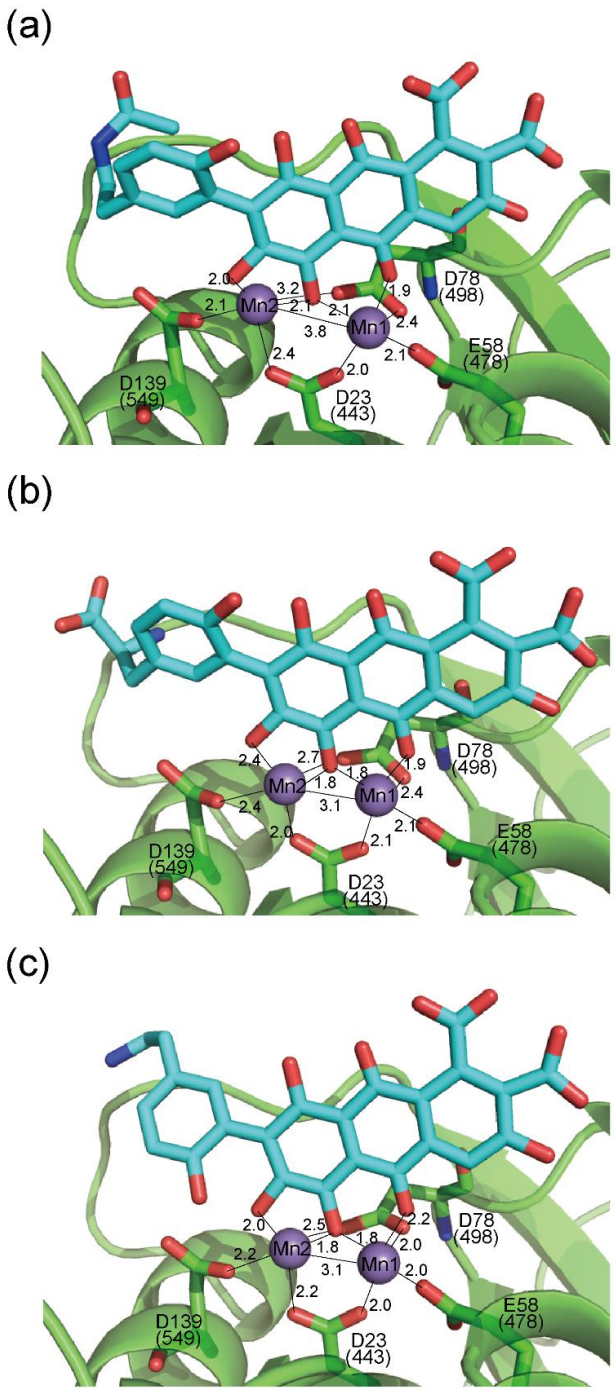

Fig. 2: Binding structures of the identified natural compounds at the active site of p15Ec in crystallographic analysis. (a)-(c) correspond to compounds 1-3, respectively.

The anthraquinone backbone was positioned at the active center in the binding structure of compound 1 (Figure 2(a)). Two Mn2+ ions and the anthraquinone atoms were aligned on the same plane. Two hydroxy groups at the 5th- and 6th-positions and one ketone at the 10th-position were coordinated with the Mn2+ ions. The 3rd-position hydroxy group made a hydrogen bond with Ser79 Oγ. The 1st-position

carboxylic group faced the solvent. The hydroxyphenyl connected to the 7th-position was exposed to the solvent, and the contribution to the compound binding affinity to the target was little. The terminal acetylamino occupied a space near His129 and Lys130. Cβ methyl of Ala128 and Nη amine of Arg147 made a CH-π and an NH-π interaction with the anthraquinone backbone.

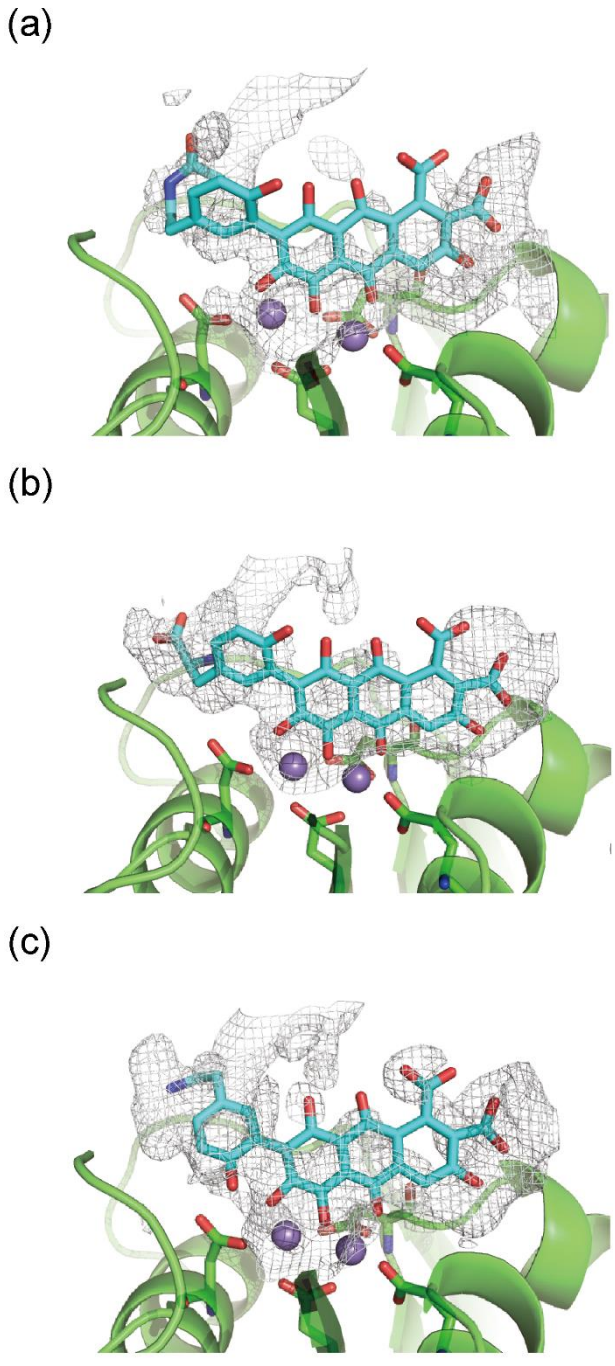

Fig. 3: Fo-Fc omit map (grey mesh) at 2.0 σ for the crystal structure of p15Ec and inhibitor complex. (a) - (c) corresponds to compounds 1-3. Inhibitory compounds are depicted in the stick representation colored cyan (carbon), red (oxygen), and blue (nitrogen). A manganese ion is shown by sphere (purple).

The anthraquinone backbone of compound 2 was also bound to the active center of p15Ec (Figure 2(b)). The anthraquinone atoms and two Mn2+ ions were on the same plane. The 5th-position hydroxy group formed two coordination bonds with the respective Mn2+ ions. Dual coordination of the 5th-position hydroxy was observed in the binding modes of both compounds 1 and 2. The 2nd-position carboxylic group interacted with the main chain nitrogen of Gln80. While the 7th-position hydroxyphenyl was exposed to the solvent, the phenyl aromatic ring attracted the side chain of Ala128 through a CH-π interaction. The terminal amine group loosely interacted with the main chain oxygen of His129. The Ala128 Cβ methyl and Arg147 Nη amine made CH-π and NH-π interactions with the anthraquinone backbone.

The anthraquinone plane was aligned to two Mn2+ ions also in the crystal structure of compound 3 (Figure 2(c)), in which the 5th- and 6th-position hydroxy and the 10th-position ketone groups were coordinated to Mn2+ ions. The 5th-position hydroxy showed dual coordination with two Mn2+ ions. The hydroxyphenyl was exposed to the solvent, and the terminal amine group interacted with the His129 main chain nitrogen via a water molecule. The Ala128 Cβ and the Arg147 Nη amine formed CH-π and NH-π interactions with compound 3.

In the binding poses of the three compounds, the directions of the backbone are consistent. As shown in a previous study on isoquinoline-base inhibitors,30 two opposite directions are possible at the HIV-1 RNase H active domain. Since the binding modes were identical among the compounds, the bulky hydroxy phenyl group was a key factor in determining the direction of the compound association to the RNase H active site.

Anthraquinone is a backbone of laccaic acids identified as RNase H inhibitors in this work. Several anthraquinone derivatives were already reported to block RNase H activity [5], and it was shown that attaching two hydroxy groups to both aromatic rings of anthraquinone promoted the inhibitory activity. The inhibitory activity was weak only when small substituents were attached to the anthraquinone backbone. However, the activity was markedly enhanced by connections of large chemical groups such as glucose. Sennosides A and B were found to inhibit RNase H activity at an IC_{50} value of 2 μM [6]. Two anthraquinones are connected by reducing the carbonyl groups on the center cyclic ring. A monosaccharide, glucose, is attached to one anthraquinone, and a hydroxy-methyl-cyclohexane-tetraol is bound to the other.

4 Conclusion

Natural compounds, laccaic acids A, C, and E, were identified to be inhibitors for HIV-1 RT-associated RNase H activity. Laccaic acids are natural products extracted from sticklac, which is a secretion of the insect Coccus laccae (Laccifer lacca Kerr). The identified compounds are water-soluble red pigments. Therefore, laccaic acids have high potentials in many applications such as cosmetic ingredients, coloring of food, oil painting, and dyeing textiles. Laccaic acids possess an anthraquinone backbone, and two carboxy and four hydroxy groups are attached to the backbone. X-ray crystal analysis clarified that two hydroxy groups and one carbonyl group were coordinated with two metal ions at the active site. Laccaic acid C showed the best inhibitory activity among the three laccaic acids. Modification was attempted and the binding affinity was predicted by MM calculations. Substituting carboxymethyl for amino-carboxy-ethyl at the terminal moiety gave a good binding score. The findings for the binding structure and chemical modification will be useful for the development of potent anthraquinone-based inhibitors for HIV-1 RNase H activity.

Acknowledgement

X-ray diffractions were acquired at the Photon Factory, Tsukuba, Japan (proposal no. 2022G642). Calculations were performed at the Research Center for Computational Science, Okazaki, Japan and at the Information Technology Center of the University of Tokyo. A part of this work was supported by a grant for Scientific Research from the Japan Society for the Promotion of Science.

References

[1] Miwa, K.; Guo, Y.; Hata, M.; Yamamoto, N.; Hoshino, T. *Chem. Pharm. Bull.*, 2023, **71**, 360-367.

[2] For note; The library has been collected at Chiba University since 1960, and all compounds have been purified. The library contains both new and known compounds. The structural information of the library is in preparation for public release.

[3] Lu, H.; Komukai, Y.; Usami, K.; Guo, Y.; Qiao, X.; Nukaga, M.; Hoshino, T. *J. Chem. Inform. Model.* 2022, **62**, 6762–6774.

[4] Fuji, H.; Qi, F.; Qu, L.; Takaesu, Y.; Hoshino, T. *Chem. Pharm. Bull.*, 2017, **65**, 461-468.

[5] Budihas, S.R.; Gorshkova, I.; Gaidamakov, S.; Wamiru, A.; Bona, M. K.; Parniak, M. A.; Crouch, R. J.; McMahon, J. B.; Beutler, J. A.; Le Grice, S. F. *Nucleic Acids Res.* 2005, **33**, 1249-1256.

[6] Martini, R.; Esposito, F.; Corona, A.; Ferrarese, R.; Ceresola, E. R.; Visconti, L.; Tintori, C.; Barbieri, A.; Calcaterra, A.; Iovine, V.; Canducci, F.; Tramontano, E.; Botta, M. *Chembiochem.* **2017**, *18*, 374–377.

* hoshino@chiba-u.jp