Molecular Docking and Molecular Dynamics Simulation to Predict Inhibitors Against HIV Envelope 1 Protein

[1,2]*Mustafa Arif*, [2]*Gaurav Sharma**
[1]*Lyons Township High School, La Grange, IL;* [2]*Eigen Sciences, Apex, NC*
Corresponding author email: research@eigensciences.com

Abstract

HIV (human immunodeficiency virus) is a virus that attacks the immune system, which serves as the body's defense mechanism. If untreated, HIV can lead to AIDS (acquired immunodeficiency syndrome), which raises risk for a host of problems, especially infections and cancers. According to data from 2022, there are 39 million people worldwide afflicted with HIV. HIV remains a leading cause of death globally, although AIDS-related deaths have declined due to ART (antiretroviral therapy). There are two main types of HIV viruses: HIV-1 and HIV-2. HIV-1 is more prevalent worldwide, while HIV-2 is less pathogenic and found predominantly in Africa. Although not fully understood, the key differences between HIV-1 and HIV-2 viruses, lie in the mechanism of retroviral pathogenesis. HIV-1 is comprised of a protein called the envelope protein, which is required for entry of the virus into the host cell. In our work, we have used computational strategies like molecular docking and molecular dynamics simulations to predict inhibitors that can bind to the envelope protein, thus inhibiting viral entry into the host cell. Based on our studies, we have proposed five chemical compounds that bind strongly to the viral envelope protein. These chemical compounds bind specifically to the Phe43 cavity on the envelope protein. The Phe43 cavity is a promising target for drug therapy. We have also introduced virtual reality (VR) technology to visualize and modify the protein-ligand complexes. Our current work will not only help in developing novel therapeutics against HIV, but also pave way for many potential new treatments, thus helping combat the global burden of this disease.

Keyword: HIV/AIDS, Envelope protein, Molecular docking, Molecular dynamics, Computer Aided Drug Design.

Introduction:

Human immunodeficiency virus (HIV) is a spherical-shaped virus belonging to the Lentivirus family. Upon infecting humans, it can lead to AIDS (acquired immunodeficiency syndrome).[1] Currently, 1.2 million people in the US are infected with HIV disease. [2] The HIV virus specifically targets CD4+ T cells (CD4 cells), which are central facilitators of the immune response. The hallmark feature of HIV infection is the depletion of CD4+ T cells, which subsequently raises risk for infections and cancers. Antiretroviral therapy (ART) aims to target the HIV virus at different stages of its life cycle. Although not curative, it reduces the patient's HIV burden, thus decreasing the risk for opportunistic infections, cancers, and other illnesses.

The surface of the HIV virus contains a protein known as the envelope protein (Figure 1), which plays an important role for viral entry into the cells. This protein exists in two conformations: open and closed. The envelope protein has two components known as gp120 and gp41. The gp120 protein contains a subprotein known as HT1 (shown in grey color in Figure 1), which binds to the CD4 receptor (shown in orange color in Figure 1) to gain entry into the cell.[3] If the interaction between the viral protein (HT1) and human cell receptor (CD4) can be prevented, viral entry can be inhibited. An example of a medication exhibiting this mechanism of action is enfuvirtide, which belongs to a new class of antiretroviral therapies known as fusion inhibitors. It mimics the viral HT1 protein and binds to CD4 receptor, thus preventing viral entry into the cell. [3-5]

Computational chemistry and biology is a field in which computers are used to simulate biomolecules. Molecular docking is a popular technique that has been used to locate receptor (proteins) and ligand interactions.[6, 7] In this technique, the 3D structure of the protein is taken and a ligand (chemical compound) is docked onto the protein. The software predicts the best binding site based on the size, shape, and interaction made by ligand with the protein.[8] Molecular docking is a popular technique in drug discovery and is used at the early stages of computer aided drug design (CADD).[9, 10] There are various software systems that are available for docking. In our work, we have use AutoDock software to perform docking.[11]

10-12 lines about molecular dynamics (Mustafa)
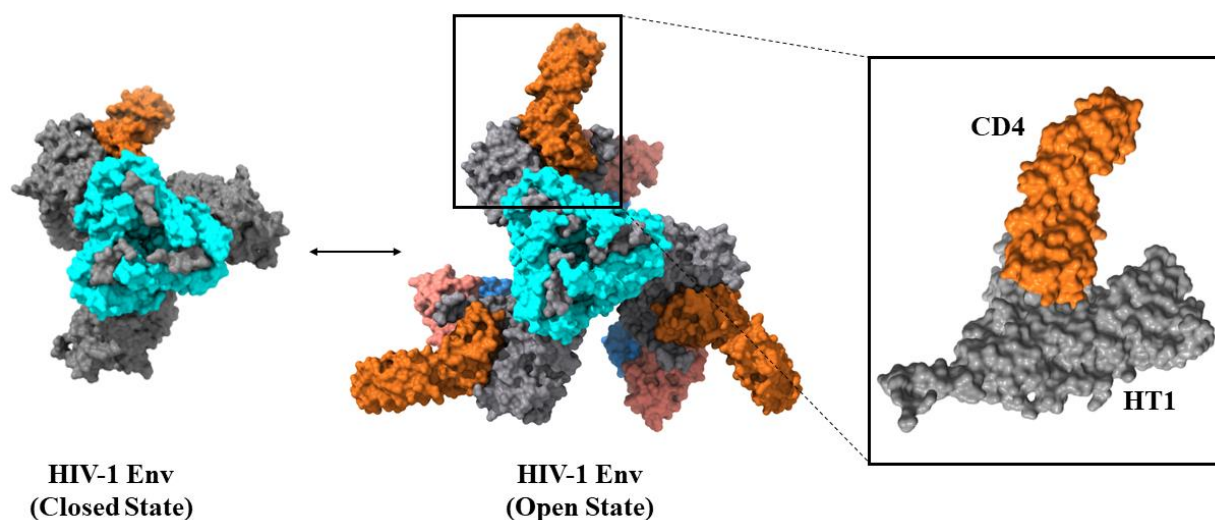
Figure 1: Closed (left) and open (middle) state of HIV-1 Envelope protein. The HT1 (viral protein) and CD4 (host cell protein) interaction is shown on right.

The management of the HIV involves the use of antiretroviral drugs that act at different stages of the HIV lifecycle.[1] Although highly effective, current therapies come with many setbacks. Eradication of the virus, hence a cure, is not possible. The treatments can cause many adverse effects and there is concern for drug resistance [1] There is an urgent need for novel

therapeutic options which can enhance the efficacy of our current treatments. In our current work, molecular docking simulations have been performed to find chemical compounds that can bind the HIV-1 Envelope protein and inhibit the viral entry in cell. Our research will help in designing novel therapeutics against the HIV virus.

Methods

The three-dimensional structure of the HT1 protein was obtained from the Protein Data Bank (PDB ID: 8FYI).[12] Ligand were downloaded from the Zinc20 Database[13], and approximately 5000 compounds were selected for virtual screening. The following are the Zinc Database Tranches settings that were used to download the ligands files: (1) only 3D models were selected; (2) in the reactivity section, "Standard" and "Exclusive" were chosen; (3) in purchasability, "In-Stock" and "Exclusive" options were selected; (4) Reference (R) pH was chosen; (5) Charge: "0"; and (6) "Lead-Like" compounds were selected. In the next step, ligands were selected based on LogP values. Since, chemical compounds with LogP=2 have good oral and intestinal absorption only these compounds were selected for virtual screening. Finally, the compounds were downloaded in pdbqt file format. The binding poses of ligands to the ZnT opening were explored using AutoDock Vina 1.5.6 software. For each protocol, 10 poses were generated for all four protein-substrate complexes.

Molecular Dynamics: The GROMACS software, employing the amber ff14sb force field, was utilized to conduct molecular dynamics simulations on both proteins and protein-ligand complexes. To maintain stability, the SETTLE algorithm was employed for water molecule bond lengths and angles, while the LINCS algorithm was applied to constrain the bond lengths of protein. Long-range electrostatic interactions were computed using the particle-mesh Ewald (PME) method. The protein-ligand complex was placed in the box of dimensions 1 Å form the

surface of the protein and was filled with TIP3P water molecules. Sodium (Na+) and chloride (Cl-) ions were introduced to uphold physiological ion levels.[14] Energy minimization comprised 5000 steps, followed by equilibration steps (nvt and npt) each lasting 1000 steps. The ensuing molecular dynamics simulations persisted for 500 ns, employing 2 fs time steps and maintaining constant particle count, pressure, and temperature (NPT ensemble). Representative structures were gleaned via cluster analysis. Binding energies between host and guest molecules were evaluated using the molecular mechanics/Poisson–Boltzmann surface area (MMGBSA) method.

Data Analysis: Autodock scorning method was used to select the top five candidates that binds strongly to the protein. The scoring criteria is the ligand should fit in the binding site and forms strong interactions with the protein. Protein Ligand Interaction Profiler (PLIP)[15] web server was used to find the protein-ligand interactions. The ChimeraX[16], PyMol[17] and VMD software were utilized to visualize and analyze the protein-ligand complexes. The pharmacological and carcinogenic properties of the compounds were assessed with the aid of SwissADME.[18]

Results and discussion:

In this study, all-atom unrestrained 500ns long MD simulations of the CD4-HT1 protein complex in an aqueous solutions were performed to explore the dynamics of the protein and also to predict the ligands that the bind and potentially inhibit the CD4-HT1 complex formation. Root-mean-square deviations (RMSDs) of the complexes were used to quantify the structural changes in the protein-ligands complex. The RMSD value remained below 1.5A. Based on our MD simulations, in the CD4-HT1 complex the CD4 binding region on the HT1 protein is partially negative as shown in electrostatic surface potential in Figure S1. To complement this binding region the CD4 binding region is mostly positive. This shows the strong binding of these two proteins. The HT1 proteins contains a unique cavity known as Phe43 cavity since the cavity

accommodates the Phe43 amino acids and is considered as a crucial in the two-protein binding and HIV infection. Therefore, the current strategy is to binding a small molecular to this biding site and hence the interaction of these two proteins can be prevented.

Based on out simulation we have proposed five ligands (**Ligand I**, **II**, **III**, **IV**, and **V**) that binds strongly to the HIV-1 Envelope protein. The 2D and 3D interactions between the protein amino acids and ligands are shown in Figure 2 and 3, respectively.
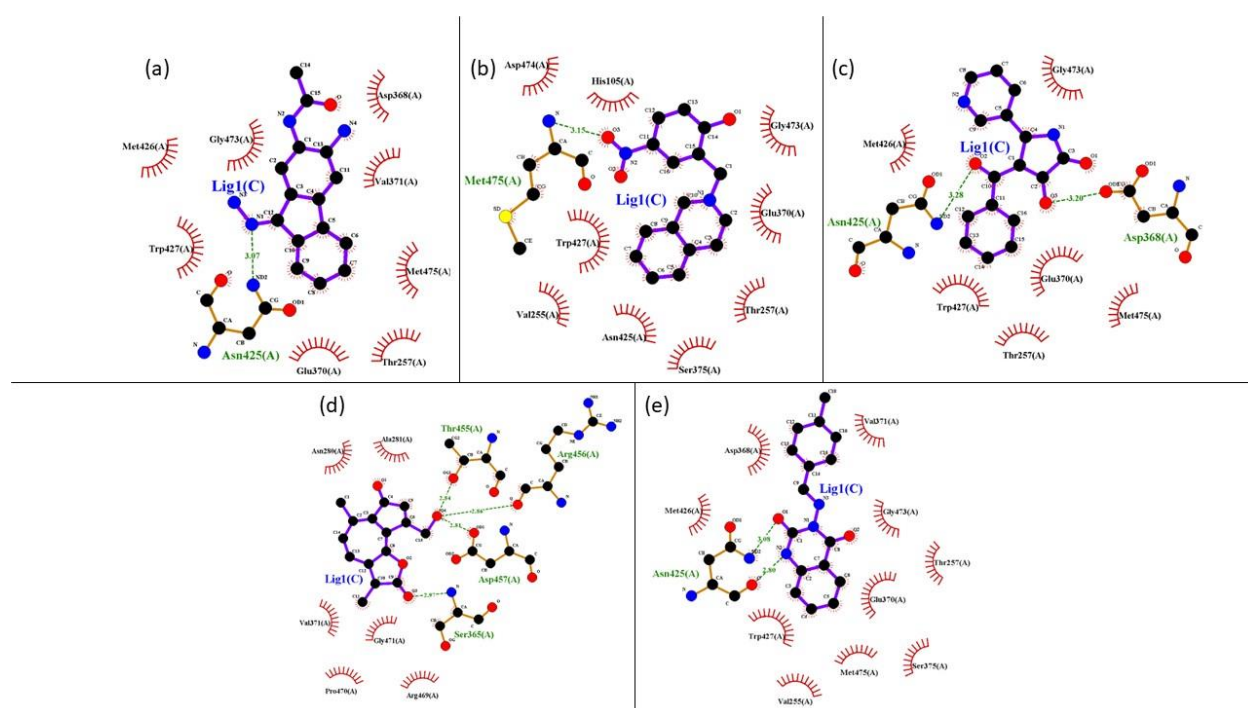
Figure 2: Protein-ligand 2D interactions. (a) **Ligand I**, (b) **Ligand II**, (c) **Ligand III**, (d) **Ligand IV**, and (e) **Ligand V.**
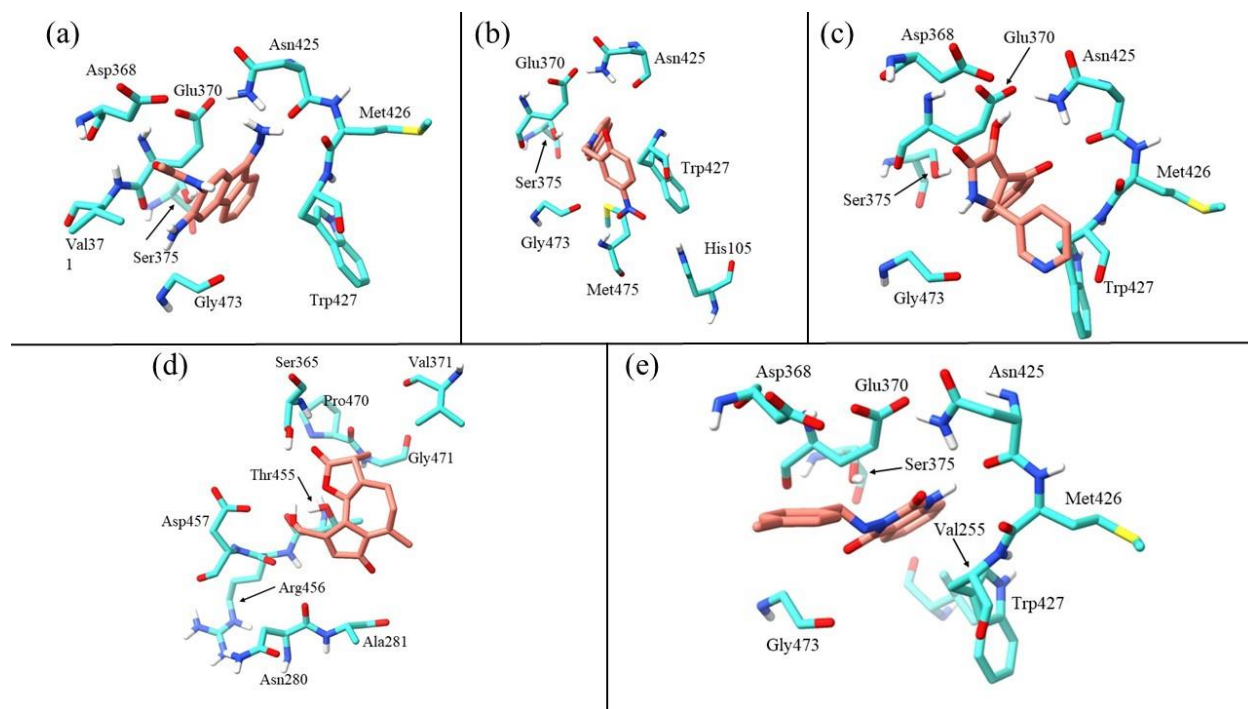
Figure 3: Protein-ligand interactions. Envelope protein amino acids are shown in cyan and ligands are shown in salmon. (a) **Ligand I**, (b) **Ligand II**, (c) **Ligand III**, (d) **Ligand IV**, and (e) **Ligand V.**

The protein-ligand interactions were calculated by Protein-Ligand Interaction Profiler (PLIP) web server as shown in Table 1. The **Ligand I** form interaction with the protein by forming hydrophobic interactions with Thr257, Glu370, Glu370, and Trp427 residues with distance of 3.81, 3.53, 3.45, and 3.40 Å respectively. In addition, it forms two hydrogen bonds with Asn425 at a distance of 2.49 and 2.17 Å, respectively. In this complex **Ligand II** interactions with the protein by forming hydrophobic interactions with Val255, Thr257, Glu370, and Glu370 at 3.67, 3.64, 3.58, and 3.30Å, respectively. Moreover, it forms one π-Stacking with Trp427 at 4.62Å. The **Ligand III** formed five hydrophobic interactions with Val255 (3.94Å), Thr257 (3.80Å), Glu370 (3.61 and 3.35Å), and Trp427 (2.64Å). In addition, it formed two hydrogen bonds with Glu370 and Asn425 at a distance of 2.64 and 2.64Å, respectively. The **Ligand IV** formed only one hydrophobic

interaction with Val371 (3.54Å) and three hydrogen bonds with Ser365, Thr455, Asp457 at a distance of 2.05, 2.06, and 1.85Å, respectively. Finally, **Ligand V** formed eight hydrophobic interactions with Residue Val255, Thr257, Asp368, Glu370, Glu370, Val371, and Trp427 at a distance of 3.54, 3.68, 3.71, 3.47, 3.85, 3.83, and 3.67Å, respectively. Moreover, it formed two hydrogen bonds with Asn425 (2.71 and 1.83Å) and one π-Stacking with Trp427 (4.53Å).

I		
Hydrophobic Interactions	Residue	Distance (Å)
	Thr257	3.81
	Glu370	3.53
	Glu370	3.45
	Trp427	3.40
Hydrogen Bonds	Asn425	2.49
	Asn425	2.17
II		
Hydrophobic Interactions	Residue	Distance (Å)
	Val255	3.67
	Thr257	3.64
	Glu370	3.58
	Glu370	3.30
π-Stacking	Trp427	4.62
III		
Hydrophobic Interactions	Residue	Distance (Å)
	Val255	3.94
	Thr257	3.80
	Glu370	3.61
	Glu370	3.35
	Trp427	3.41
Hydrogen Bonds	Glu370	2.64
	Asn425	2.65
IV		
Hydrophobic Interactions	Val371	3.54
Hydrogen Bonds	Ser365	2.05
	Thr455	2.06
	Asp457	1.85
V		
Hydrophobic Interactions	Val255	3.54
	Thr257	3.68
	Asp368	3.71
	Glu370	3.47

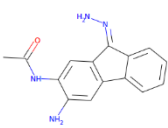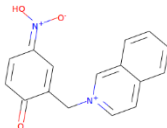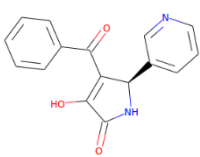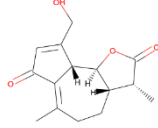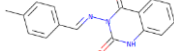|	Glu370	3.85
	Val371	3.83
	Trp427	3.67
Hydrogen Bonds	Asn425	2.71
	Asn425	1.83
π-Stacking	Trp427	4.53

Table 1: Protein-ligand interaction computed by Protein-Ligand Interaction Profiler (PLIP) web server.

Interestingly, all these ligands bind to the cavity present in the protein called as Phe43 cavity as shown in Figure 4. This cavity has been suggested as a useful target for antiviral strategies. Previous studies have revealed that the HIV entry can be inhibited by filling the Phe43 cavity.[19, 20] To evaluate the pharmaceutical properties of the proposed ligands we have performed absorption, distribution, metabolism, and excretion evaluation of the ligands by using SwissADME[18] web server. Based on this study, ligands have high gastro intestinal (GI) absorption and high blood brain barrier (BBB) permutation (Table 2). In addition, all the ligand passed Lipinski rule hence they all show high drug likeliness. Further analysis is required to find the effect of these ligands on the structure of the protein. This will be done by performing molecular dynamics simulation studies in future studies.
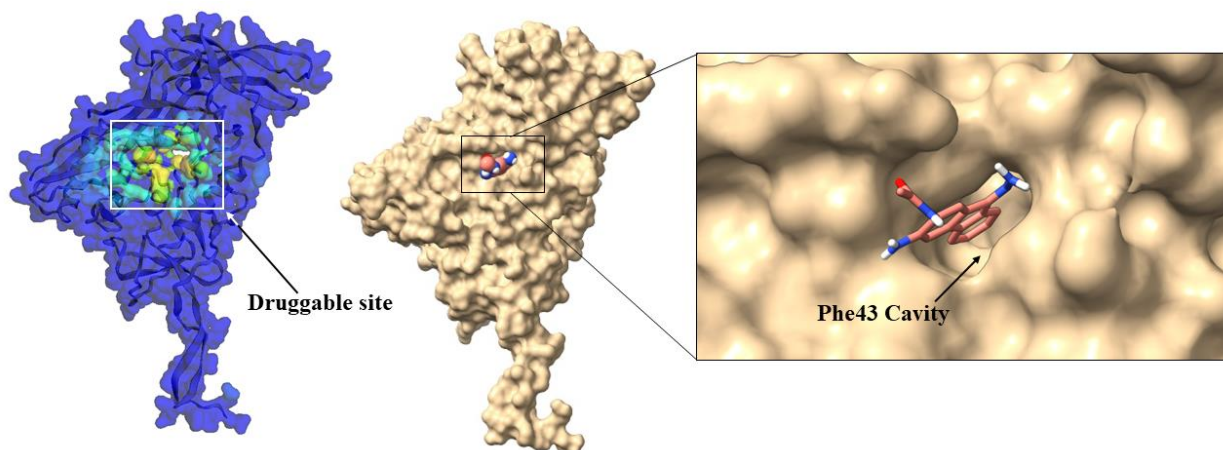
Figure 4: Phe43 cavity is the druggable site (left in yellow) as predicted by GrASP web server.

The ligand bound to the Phe43 cavity is shown in the middle and right image. The cavity is crucial in HT1-CD4 interaction.

	I	II	III	IV	V
Formula, molecular weight	$C_{15}H_{14}N_4O$, 266.30 g/mol	$C_{16}H_{13}N_2O_3^+$, 281.29 g/mol	$C_{16}H_{19}N_5O$, 297.35 g/mol	$C_{16}H_{11}N_3O_3$ 279.29 g/mol	$C_{16}H_{13}N_3O_2$ 279.29 g/mol
GI absorption	High	High	High	High	High
BBB permeation	No	Yes	No	Yes	Yes
Drug likeliness (Lipinski)	Yes	Yes	Yes	Yes	Yes
2D structure					

Table 2: Pharmaceutical properties of the five proposed ligands.

Molecular dynamics simulations were also performed starting from the molecular docking poses for the complexes of 5 high-interest compounds using the GROMACS software. In all MD simulations the ligands bind strongly to the binding site.

	I	II	III	IV	V
Binging energy	-82.1	-74.2	-81.4	-79.1	-92.1

Table 3: Binding energy of five ligands obtained from MMGBSA.

Conclusion: In the current molecular docking simulation, we have predicted inhibitors that bind to the HT1 protein of HIV-1 Envelope protein. Binding of these inhibitors to the HT1 protein will prevent the HIV-1 Envelope Protein (viral) and CD4 protein (host) interactions. This will prevent viral entry into the cell, thus inhibiting HIV infection. Based on our docking simulations we have predicted top five ligands that have shown strong interactions with the protein. These ligands

specifically bind to the Phe43 cavity, which is crucial for binding of the HIV virus to the human cell. By blocking this cavity, the drug will prevent HIV viral entry into the host human cells.

References

1. Gandhi, R.T., et al., *Antiretroviral drugs for treatment and prevention of HIV infection in adults: 2022 Recommendations of the International Antiviral Society–USA Panel.* 2023. **329**(1): p. 63-84.
2. Heath, K., J. Levi, and A.J.A. Hill, *The Joint United Nations Programme on HIV/AIDS 95–95–95 targets: worldwide clinical and cost benefits of generic manufacture.* 2021. **35**(1): p. S197-S203.
3. Maréchal, V.r., et al., *Human immunodeficiency virus type 1 entry into macrophages mediated by macropinocytosis.* 2001. **75**(22): p. 11166-11177.
4. Hu, Q., et al., *Mechanisms of peptide and phosphoester hydrolysis catalyzed by two promiscuous metalloenzymes (insulin degrading enzyme and glycerophosphodiesterase) and their synthetic analogues.* 2020. **10**(4): p. e1466.
5. Duay, S.S., et al., *Molecular dynamics investigation into the effect of zinc (II) on the structure and membrane interactions of the antimicrobial peptide clavanin A.* 2019. **123**(15): p. 3163-3176.
6. Meng, X.-Y., et al., *Molecular docking: a powerful approach for structure-based drug discovery.* 2011. **7**(2): p. 146-157.
7. Sharma, G., et al., *Effect of chemically distinct substrates on the mechanism and reactivity of a highly promiscuous metallohydrolase.* 2020. **10**(6): p. 3684-3696.
8. Morris, G.M. and M.J.M.m.o.p. Lim-Wilby, *Molecular docking.* 2008: p. 365-382.
9. Pagadala, N.S., K. Syed, and J.J.B.r. Tuszynski, *Software for molecular docking: a review.* 2017. **9**: p. 91-102.
10. Sharma, G., et al., *Advances in monolithic silica columns for high-performance liquid chromatography.* 2017. **8**: p. 1-11.
11. Eberhardt, J., et al., *AutoDock Vina 1.2. 0: New docking methods, expanded force field, and python bindings.* 2021. **61**(8): p. 3891-3898.
12. Ozorowski, G., et al., *Open and closed structures reveal allostery and pliability in the HIV-1 envelope spike.* 2017. **547**(7663): p. 360-363.
13. Irwin, J.J., B.K.J.J.o.c.i. Shoichet, and modeling, *ZINC– a free database of commercially available compounds for virtual screening.* 2005. **45**(1): p. 177-182.
14. Li, Z., et al., *Accurate Metal–Imidazole Interactions.* 2022. **19**(2): p. 619-625.
15. Salentin, S., et al., *PLIP: fully automated protein–ligand interaction profiler.* 2015. **43**(W1): p. W443-W447.
16. Pettersen, E.F., et al., *UCSF ChimeraX: Structure visualization for researchers, educators, and developers.* 2021. **30**(1): p. 70-82.
17. DeLano, W.L.J.C.N.P.C., *Pymol: An open-source molecular graphics tool.* 2002. **40**(1): p. 82-92.
18. Daina, A., O. Michielin, and V.J.S.r. Zoete, *SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules.* 2017. **7**(1): p. 42717.
19. Acharya, P., et al., *HIV-1 gp120 as a therapeutic target: navigating a moving labyrinth.* 2015. **19**(6): p. 765-783.
20. Sharma, G., et al., *Investigating coordination flexibility of glycerophosphodiesterase (GpdQ) through interactions with mono-, di-, and triphosphoester (NPP, BNPP, GPE, and paraoxon) substrates.* 2019. **21**(10): p. 5499-5509.

Supporting Information
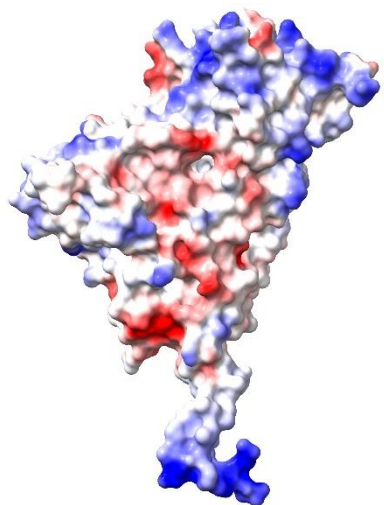
Figure S1: Electrostatic surface potential of HT1 protein.