The Drug design of HIV-1 Reverse Transcription Inhibitor

Name:鲁周语

ID:520111910142

Institute address:Shanghai Jiao Tong University

1 | Abstract

The human immunodeficiency virus-1 (HIV-1) is a retrovirus that infects T-lymphocytes in the human immune system and causes immunodeficiency. Inhibiting the reverse transcription process of HIV-1 is a key strategy in developing drugs to treat the disease. This study aimed to design drugs to inhibit the reverse transcription process of HIV-1 by utilizing a combination of 3D-QSAR models, molecular docking, and molecular dynamics simulations.The application of this method provided a more complete understanding of the interactions between inhibitors and the reverse transcriptase enzyme. This approach led to the identification of new drug candidates and a better understanding of the mechanism of drug-receptor interactions, and the optimization of the structure of existing compounds to improve their activity.

KEYWORDS

3D-QSAR, allosteric mechanism, HIV reverse transcriptase, inhibitor, MD simulation

2 | Introduction

HIV-1, or the human immunodeficiency virus, is a serious and potentially life-threatening condition that attacks the immune system. It is caused by the retrovirus HIV-1 and it leads to acquired immunodeficiency syndrome (AIDS) which is characterized by a severe impairment of the immune system, making the person susceptible to infections and cancers. HIV-1 infects T-lymphocytes, a type of white blood cell that plays a critical role in the immune response. Once the virus infects these cells, it replicates itself and begins to spread throughout the body, destroying the immune system.

The basic principle of HIV-1's pathopoiesia is that it infects T-lymphocytes, replicates itself using the reverse transcription process, and eventually leads to immunodeficiency. To treat this condition, people have developed therapy known as highly active antiretroviral therapy (HAART). HAART is

a combination of antiviral drugs that target different stages of the virus's life cycle, such as reverse transcription, protease, and integrase. By combining multiple drugs that target different stages of the virus's life cycle, HAART can reduce the amount of virus in the body, known as viral load, to undetectable levels and prevent progression to AIDS.

The Reverse transcriptase (RT) is an enzyme that plays a critical role in the replication of HIV-1, the virus that causes AIDS. RT converts the viral RNA into DNA, allowing the virus to integrate into the host cell's genome and replicate itself. This makes RT a prime target for the treatment of AIDS.

The development of reverse transcriptase inhibitors (RTIs) has been a major breakthrough in the treatment of AIDS. RTIs target the reverse transcription process of HIV-1, effectively reducing the amount of virus in the body and preventing progression to AIDS. RTIs are commonly used in combination with other antiretroviral drugs as part of highly active antiretroviral therapy (HAART) which is a combination of antiviral drugs that target different stages of the virus's life cycle.

One of the main advantages of RTIs is their ability to lower the amount of virus in the body to undetectable levels, also known as viral load suppression. This not only improves the patient's health but also reduces the risk of HIV transmission to others.

However, HIV-1 can quickly develop resistance to RTIs, which is why researchers are constantly working to develop new RTIs with improved activity and fewer side effects. The continuous development of new RTIs is crucial for the treatment of AIDS, as it allows for the development of new drug regimens that can overcome viral resistance.

Computer-aided drug design (CADD) methods can be used to develop new reverse transcriptase inhibitors (RTIs) for the treatment of AIDS. The use of CADD techniques allows for a more comprehensive understanding of the interactions between RTIs and the reverse transcriptase enzyme, leading to the identification of new drug candidates and the optimization of existing compounds. Here, we apply the combination of 3D-QSAR models, molecular docking, and molecular dynamics simulations. The 3D-QSAR model includes two methods:
comparative molecular field analysis (CoMFA) and comparative molecular similarity index analysis (CoMSIA). [1] As it may only reveal the association between substituents and biological activity, we also applied to provide insight into the mechanism of interaction between the inhibitor and the receptor, which can effectively address the limitations of the 3D-QSAR model.

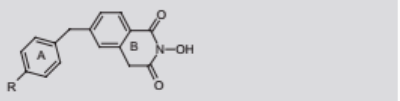
3 | Methods

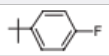
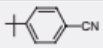
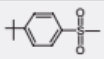
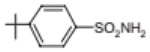
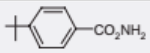
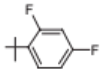
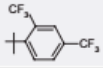
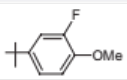
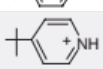
3.1 | 3D-QSAR

One of the key CADD methods used in the development of RTIs is 3D-QSAR (Quantitative Structure-Activity Relationship) modeling. This method uses regression analysis to establish the relationship between the molecular structure of a compound and its biological activity. By building a 3D-QSAR model, researchers can identify structural features that are important for activity, and use this information to design new compounds with improved activity.

All molecules were built with sybyl modeling program(sybyl-X 2.1.1 Tripos) with Gasteiger–Hückel charges and optimized with Tripos force field. The number of iterations for optimization was 10,000. [1] After optimization, all molecular was into the database “HIV_ligand”. All the molecules being included is listed in **Table 1** , the software may only be available on Win7 operation system currently.

TABLE 1 Structure and bioactivity of quinoline inhibitors



Compound	R	IC_{50} (μM)
20a*	H	1.3 ± 0.1
20b	4-CH3	1.2 ± 0.1
20c	4-F	1.3 ± 0.1
20d	4-Cl	3.9 ± 0.2
20e*	4-Br	1.4 ± 0.2
20f	4-CF3	1.5 ± 0.1
20g	2,4-F	1.6 ± 0.3
20h		0.60 ± 0.1
20i		0.80 ± 0.1
20j		2.7 ± 0.6
20k		5.4 ± 0.8
20l		0.90 ± 0.2
20m*		3.0 ± 0.4
20n		1.2 ± 0.2
20o		0.40 ± 0.1
20p*		2.4 ± 0.2
20q		1.5 ± 0.5
20r		5.1 ± 0.6
20s*		2.2 ± 0.6
20t		0.50 ± 0.05

Note:reference from [1],The compounds with “*” are consist of the test set, and the others for the training set

After all the molecules are build, we align the database with tools inside the software, here, the most active molecular 20o was selected as the template of alignment and the maximum common substructure of rings A and B was used as skeleton according to the reference [1]. The results are shown in **Fig 1**.

3.2 | CoMFA and CoMSIA models

After that, we use all the data set above to construct the CoMFA and CoMSIA models.CoMFA is commonly used to create 3D-QSAR models. The molecules are first aligned in a grid. Next, a sp3 carbon cation is selected as a probe to evaluate the steric and electrostatic interactions between the ligand and receptor. After that, regression analysis is done through the use of partial least square (PLS). CoMSIA is an extension of CoMFA, that employs similar basic processes, but instead uses Gaussian functions as a potential energy function. Fields for hydrophobic, hydrogen bond acceptor and hydrogen bond donor properties are employed to describe the structure.

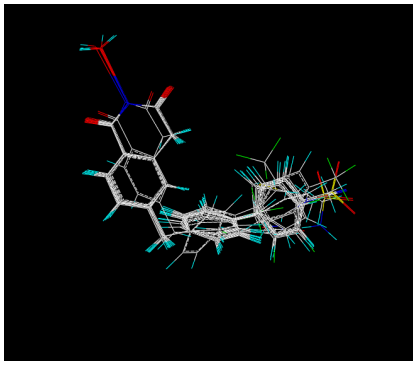


FIGURE 1. Alignment of training set based on the **maximum common substructure**

3.3 | Molecular docking

The structure of the protein-ligand complex consisting of the reverse transcriptase-associated RNase (RTRN) and LP8 was obtained from the Brookhaven Protein Databank (PDB code: 3LP1). The ligand LP8 is illustrated in **Fig 2**. Maestro software was utilized to dock the optimal conformer of inhibitors into the binding pocket of RTRN. The details of the intermolecular interactions, hydrogen bonds, van der Waals forces and rough energy information such as binding free energy were obtained from the molecular docking process.

3.4 | MD simulation

The Molecular dynamics simulations were performed using the AMBER14 simulation package and the ff14SB force field with the TIP3P water model.Due to time limitations ,we only show the coordinates of RTRN complex with the most active molecule 20o.

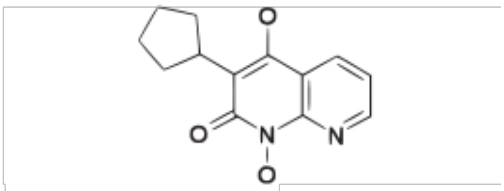


FIGURE 2 The structure of LP8

4 | Results

In order to represent our result , most of our result is showing according to the most active molecule 20o.

4.1 | The model parameter of 3D-QSAR, contour plot

The data we predict and test using our CoMFA and CoMSIA models is available in https://jbox.sjtu.edu.cn/l/y1dUPy .Here, using the model, we manage to show them in contour plots **Fig3 , Fig4.** We also showed our data of the relative figure between experimental bioactivity and prediction activity in **Fig5**. It is obvious that these data shows great correlation .(Sorry for no data proved due to time limitation)

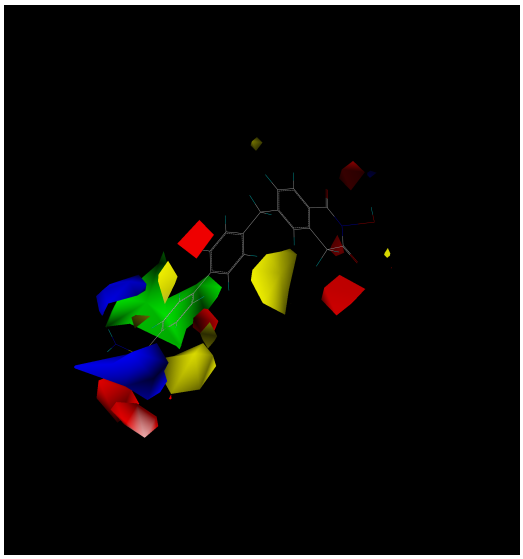


FIGURE 3 CoMFA contour plots of 20o. Green areas indicate that larger groups are beneficial for activity, yellow areas suggest that larger groups are desirable for activity, blue areas show that a positive charge is beneficial for activity, and red areas indicate that a negative charge is necessary for activity.

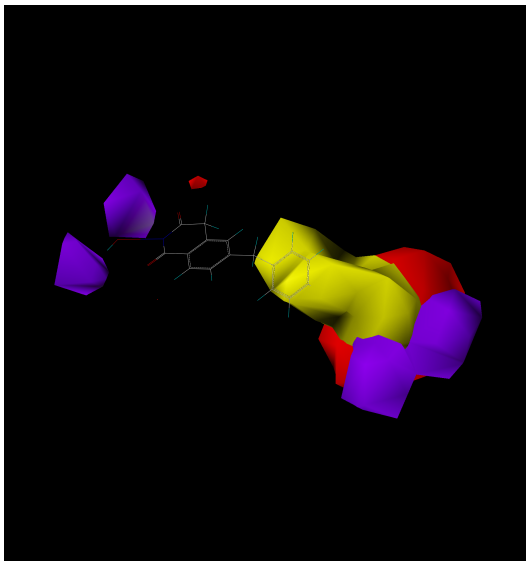


FIGURE 4 CoMSIA contour plots of 20o. The contours for steric and electrostatic interactions are similar to those of CoMFA. The magenta contours identify areas that are favorable for bioactivity when hydrophobic groups are present, while the cyan contours indicate regions where bioactivity is favored with (no exist).

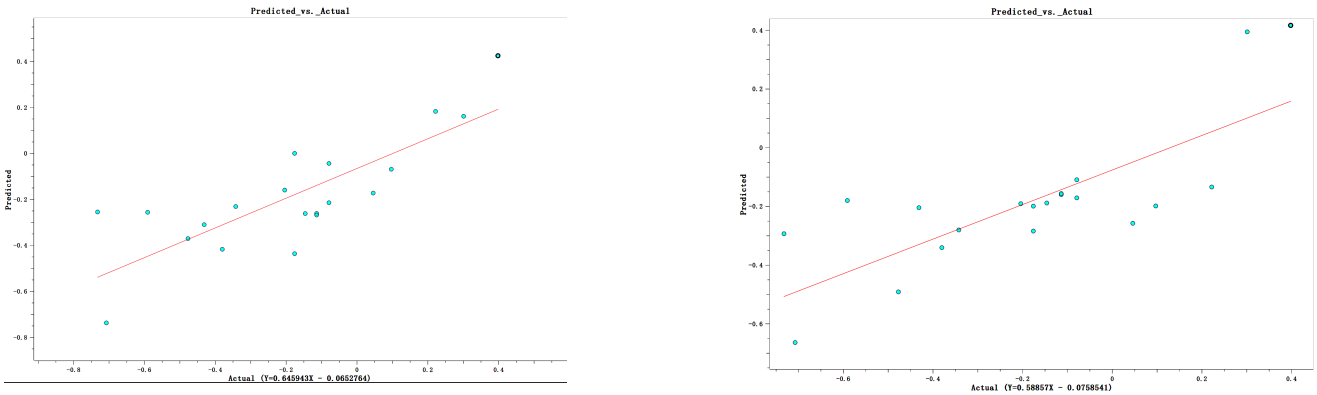


FIGURE 5 Correlation between experimental bioactivity and predicted bioactivity. (a) CoMFA model. (b) CoMSIA model .

抱歉抱歉，由于显示的时候颜色设置错，triaining和test的数据点显示为一样的了。

4.2 | Molecular docking

Here, we conduct our result using the previous method ,and succeed in showing them in **Fig 6**

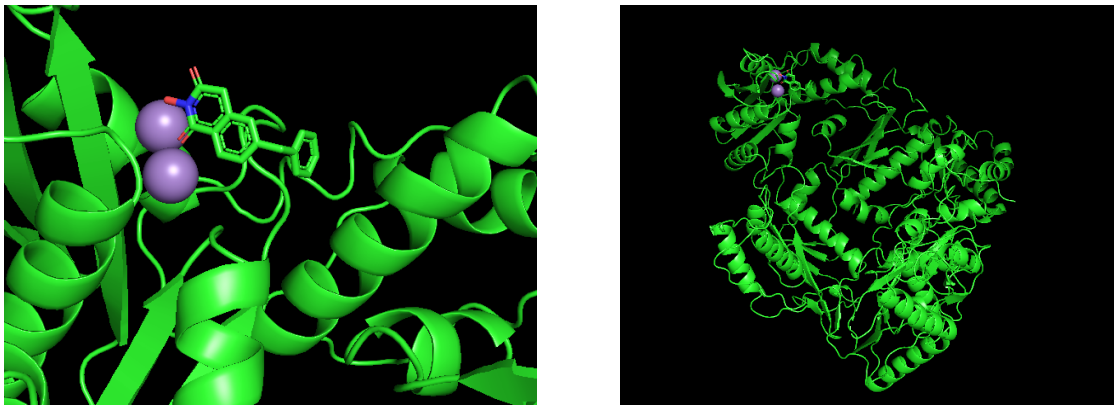


FIGURE 6 The docking complex for 20o and RTRN. The left is the more specific version of the right.

4.3 | MD simulation

For the complexes 20o-RTRN (the most active one) ,10 ns molecular dynamics simulation was performed, respectively. To examine the variations in the intramolecular conformers of RTRN, the Cα RMSD relative to initial structure is shown in **Fig7**. The results showed that the system became dynamic equilibration after 10 ns simulation

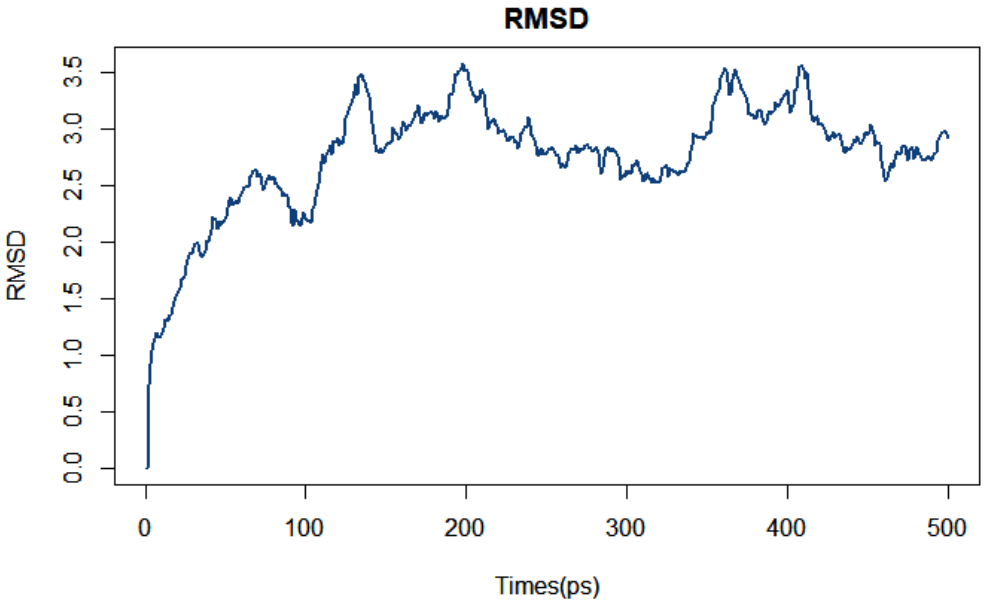


FIGURE 7 RMSD of complex 20o-RTNR.

To study the influence of 20o 's binding on the stability of RTRN, Ca variations for the complexes are illustrated in **Fig8**. In previous research[1],the Ca variation of 20o-RTRN is significantly smaller than that of 20a-RTNR and 20r-RTNR. This indicates that 20o-RTNR might become more stable than 20a-RTNR or 20r-RTNR and consistent with the bioactivity experiment.

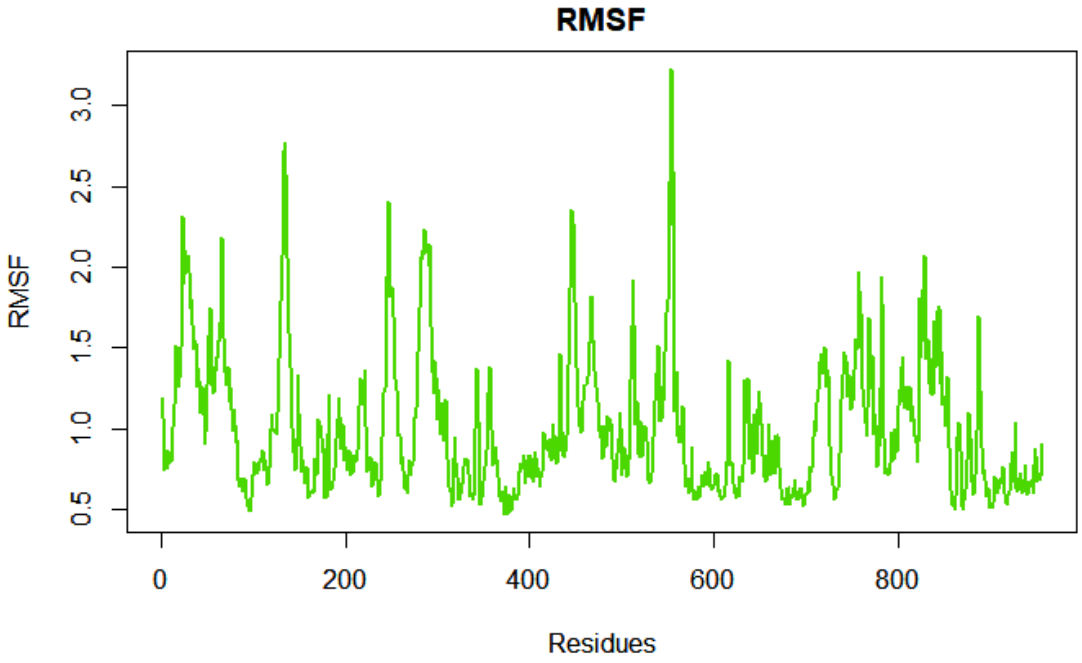


FIGURE 8 RMSF of complex 20o-RTNR.

According to the RMSF in Fig8 ,it was found that the amino acid residues in the range of

400-6000 had a large displacement.

To sum up, our result shows that the most active molecule 20o **do have a great performance and stability in docking test with our target.**

5 | Discussion

Up to now, we have experienced the whole process while it may not be able to compare with the the result in [1] due to the lack of comparisons between RTRN/20a, RTRN/20o, and RTRN/20r complexes which represents for the most active molecule 20o, the middle active 20a, and the less active 20r according to their IC50. In the future, more test may be done to verify the performance of RTRN/20a and RTRN/20r as well to make better conclusion in [1] that the compounds with higher biochemical inhibitory activity have more stable interactions than those with lower inhibitory activity, which are consistent with molecular dynamics simulation.

6 | Conclusion

Here, we do a computational test through the combination of 3D-QSAR models, molecular docking, and molecular dynamics simulations.The resuls shows the interactions between inhibitors and the reverse transcriptase enzyme though compared data may be missed due to time limitations . This approach also shows its prosper future in the identification of other drug candidates their

mechanism of drug-receptor interactions, and the optimization of the structure of existing compounds to improve their activity.

7 | Acknowledgement

This article is mostly based on the work of Professor Haifeng Chen[1]. It's my great honor to recreate his work on this meaningful topic. Without his guidance , it is hard for me to accomplish this paper. Meanwhile, I would also like to thank my friend and classmate, Haowei Tong and Ruoying Chou for their help and wisdom in the experiment. Wish everyone reading this paper a happy new year.

8 | Reference

1. Cai Y., Liu H., Chen H.F. Allosteric mechanism of quinoline inhibitors for HIV RT-associated RNase with MD simulation and dynamics fluctuation network. *Chem Biol Drug Des.* 2018, 91:805–816.
2. Cramer III, R. D.; Patterson, D. E.; Bunce, J. D., Comparative Molecular Field Analysis (CoMFA) 1. Effect of Shape on Binding of Steroids to Carrier Proteins. *J. Am. Chem. Soc.* **1988,** 110, 5959-5967.
3. Morris, G. M.; Huey, R.; Lindstrom, W.; Sanner, M. F.; Belew, R. K.; Goodsell, D. S.; Olson, A. J., AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. *J Comput Chem* **2009,** 30, (16), 2785-91.
4. D.A. Case, I. Y. B.-S., S.R. Brozell, et al, AMBER 2018. *University of California, San Francisco.* **2018**