Journal of Virology & Antiviral Research

Identification of Inhibitors Binding Site of Ebola L Polymerase Based on its Homology Model

Alfonso Trezza, Andrea Bernini and Ottavia Spiga*

Abstract

This paper describes methods, results and conclusions about our study regarding the development and the optimization of a homology model of Ebola virus RNA-dependent RNA polymerase, also called L polymerase, based on crystal structure of Vesicular stomatitis virus L polymerase. L polymerase is an essential protein for the outliving of virus; for this reason it is used as target for antiviral therapy. The aim of this work was to study the three-dimensional model of Ebola virus L polymerase and its inhibitors adopting a computational approach. We used docking simulation and multi-aligned L protein sequences in order to identify the potential binding pocket of inhibitors and their drug ability against Ebola virus L protein. Based on a dataset of ten validated molecules it was possible to understand the mechanism of action of potential inhibitors. The compounds analyzed were commercially purchasable, and experimentally tested, and the results obtained by our in silico analysis fit strongly with experimental data. This work summarizes the status quo of antiviral compounds currently used for Ebola virus disease.

Keywords

Ebola virus disease; L polymerase

Introduction

Ebola virus is a virulent pathogen which belongs to Filoviridae virus family. There are five known virus species within the genus Ebolavirus: Bundibugyo ebolavirus, Reston ebolavirus, Sudan ebolavirus, Taï Forest ebolavirus, and Zaire ebolavirus. Ebola virus causes several symptoms such as hemorrhagic fever, headache, joint muscle and abdominal pain, diarrhea and vomit; in several cases these symptoms can be fatal [1]. In 2013 Ebola virus was identified as the etiological agent of a large disease outbreak in western Africa causing almost 30,000 infections and more than 11,000 deaths and including exportation of some cases to Europe and North America [2]. The large number of cases calls for the development of medical countermeasures against EBOV disease, which has become a global priority in the field of public health. It is therefore of a core importance to look for a new drug class against EBOV. The virus has a linear, non-segmented genome with a single negative-stranded RNA. The genome codes for seven structural proteins, which include: nucleoprotein (NP), VP35, VP40, glycoprotein (GP), VP30, VP24, and L protein. The latter protein provides RNA-dependent RNA activity enclosed in an envelope [3]. Nowadays there are more therapeutic strategies against EBOV infection and each one is focused on a particular target, but previous studies have shown that the interaction between a new drug class (small molecule disruptors) and a particular active site of L polymerase, is able to inhibit processes which are essential for survival and replication [4]. The L protein of Ebola virus is the catalytic subunit of the RNA-dependent RNA polymerase complex, which, with VP35,is the key for the replication and transcription of viral genome. The sequence analysis demonstrated that L polymerase of all RNA viruses consists of six domains containing conserved functional motifs [5] so the protein may be used as a target for the design of anti-viral drugs. L polymerase has a sequence of 2212 amino acids and supplies two functional domains, the first domain is a RdRp catalytic and is localized in positions 625-809, while the second domain is a Mononegavirus-type SAM-dependent 2'-O-MTase and is localized in positions 1803-2001 [6]. Some recently published articles on L polymerase inhibitors provide us with useful information about the current knowledge on this protein [7,8]. We propose a molecular docking screening of drugs, currently in use against EBOV, on tridimensional structural models of L protein, in order to understand their mechanism of action and their potential binding pocket. The 3D structure is carried out by using homology modeling approach obtained from the 3D crystal structure of Vesicular stomatitis virus L polymerase [9], that belongs to the same order of Mononegavirales [10]. Structurally driven selection of inhibitors binding pocket seems to be essential to correctly predict the activity of drugs especially in the context of therapeutic switching [11]. Our bioinformatics approach might allow to easily selecting antiviral compounds as repurposed drugs from rich datasets.

Materials and Methods

The FASTA sequence of RNA-directed RNA polymerase L of Zaire Ebola virus (UniProtKB - Q6V1Q2) was obtained using Uniprot database [12]. The Homology model was built by hand with DeepView/Swiss-PdbViewer v. 4.1 software [13], based on the 3D structure of Vesicular stomatitis virus L protein, 5A22 entry of the Protein Data Bank [9,14], taking into account some important point of the structure like Zinc coordination systems and validated using SWISS-MODEL validation system [13]. The aminoacid sequences of all L polymerase coming from Mononegavirales order obtained by using PSI-BLAST software [15] (about 4000 whole protein sequences) were aligned using the program Clustal X v.2.1 [16]. Sequence comparison by using Scorecons Server [17] indicated that only few regions were highly conserved in all available sequences of L polymerase. Docking of ligands (as displayed in Table 1) was simulated by using flexible side chains protocol with AutoDock Vina v. 1.1 [18] that uses an iterated local search based on a succession of steps, which consisted in mutation and local optimization. Ligand structures were retrieved from ZINC Database [19] and pdbqt file were generated by using scripts in the Molecular Graphics Laboratory

***Corresponding author:** Ottavia Spiga, Department of Biotechnology, Chemistry and Pharmacy, University of Siena, Italy, via Aldo Moro 2, 53100 Siena, Italy, Tel: 39 0577 234930; E-mail: ottavia.spiga@unisi.it

Table 1: Network of interaction and binding energies.

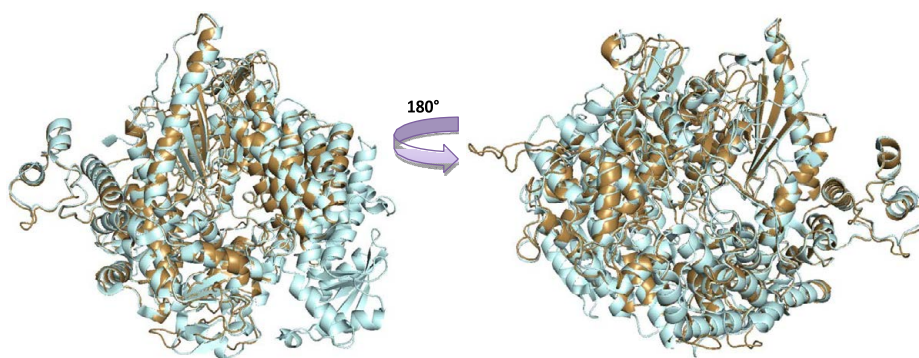| Experimental-drug	Hydrogen Bond	Hydrophobic interaction	ΔGb (Kcal/mol)
Adenosine Analogue			
Tenofovir	*Arg-571*, Lys-714**	*x*	-7.0
BCX4430	*Arg-571*, Lys-714**	*x*	-7.4
Cytidine Analogues			
Cidofovir	*x*	*x*	-7.4
Emtricitabine	*Lys-473**	*x*	-6.5
Lamivudine	*Lys-473**	*x*	-6.3
ddC	*Lys-473**	*Ala-1206, Pro-1209*	-6.5
Guanosine Analogues			
ddI	*Arg-571*, Lys-714*, Glu-578**	*Pro-1209*	-7.0
Abacavir	*Arg-571*, Lys-714*, Glu-578**	*Lys-483, Ala-579, Pro-1209*	-8.1
Thymidine Analogues			
Stavudine	*Lys473*, Arg-571*, Lys-714**	*Glu-578*, Pro-1205*	-7.1
Zidovudine	*Lys473*, Arg-571*, Lys-714**	*Ala-528, Glu-578*, Pro-1205, Ala-1206*	-7.2
*conserved residues between Ebola and Vesicular stomatitis virus.			

Figure 1: Cartoon representention of molecule overview fitting between Stomatitis (cyan) and Ebola (gold) virus of RNA-directed RNA polymerase.

(MGL) tools [20]. The generation and affinity maps, view of docking poses and analysis of virtual screening results were carried out using AutoDock plug-in of PyMOL [21]. Ligand-protein interactions were found through PLIP bioinformatics tool [22].

Results and Discussion

Protein docking procedures can predict the structure of a protein-ligand complex starting from the known structures of the individual protein components. More often, however, the structure is not known, but can be derived by homology modeling on the basis of known structures of related proteins deposited in the Protein Data Bank (PDB)[13]. We need to optimally integrate homology modeling and docking simulations with the goal of predicting the structure of a complex. In this study we report the homology modeling carried out with SWISS- MODEL [12] of the EBOV L polymerase using as template L protein of Vesicular stomatitis virus (*Rhabdoviridae*) belongs together with Ebola virus Zaire *(Filoviridae)* to the order *Mononegavirales* [9]. The molecular model obtained exhibited a medium level of similarity with the template, namely 25% identity, 30% similarity; the 3D model is showed in Figure 1. The reliability of the predicted tertiary structure is validated through the use of

SWISS-MODEL Validation process [12]. However, the models built could often differ to a significant degree from the bound conformation in the complex. In this context, we assume that multiple sequence alignments, as well as structural alignments of the templates to their corresponding subunits in the target are necessary. As a matter of fact, we have done multiple sequence alignment by using Clustal X v.2.1[15] and residues conservation was scored using the Scorecons Server [16]. The predicted structure of Ebola virus L polymerase obtained by homology modeling, together with multi-alignment sequence, gave us the necessary information on site of interaction and the mechanism of action of some nucleotide analogs as inhibitors. Furthermore, the calculated Shannon Entropy [16] providing the conservation of residues from over thousands L protein sequences coming from Mononegavirales order, shows that residues inside active sites are completely conserved or characterized by conservative mutations, thus revealing the good similarity with the template structure. In particular, we have observed how the sequence does not bear the GDD motif commonly shown by viral RNA polymerases, replaced by GDN active site motif of RdRp domain as suggested by Liang et al. [9]. Several small molecules have been tested and are currently being tested for activity against the Ebola virus, some of them are undergoing phase 3 of clinical trials [21], others have shown efficacy against the virus. In addition to that, there are ongoing in vitro and animal testing [8,22,23]. The 3D structure of Ebola virus L polymerase was subjected to protein-ligand docking simulations of ten nucleoside analogs (except from tenofovir which represents an acyclic nucleoside phosphonate) using the AutoDock Vina program [17], a program that uses a sophisticated gradient optimization method in executing local optimization in order to produce the best conformations and the
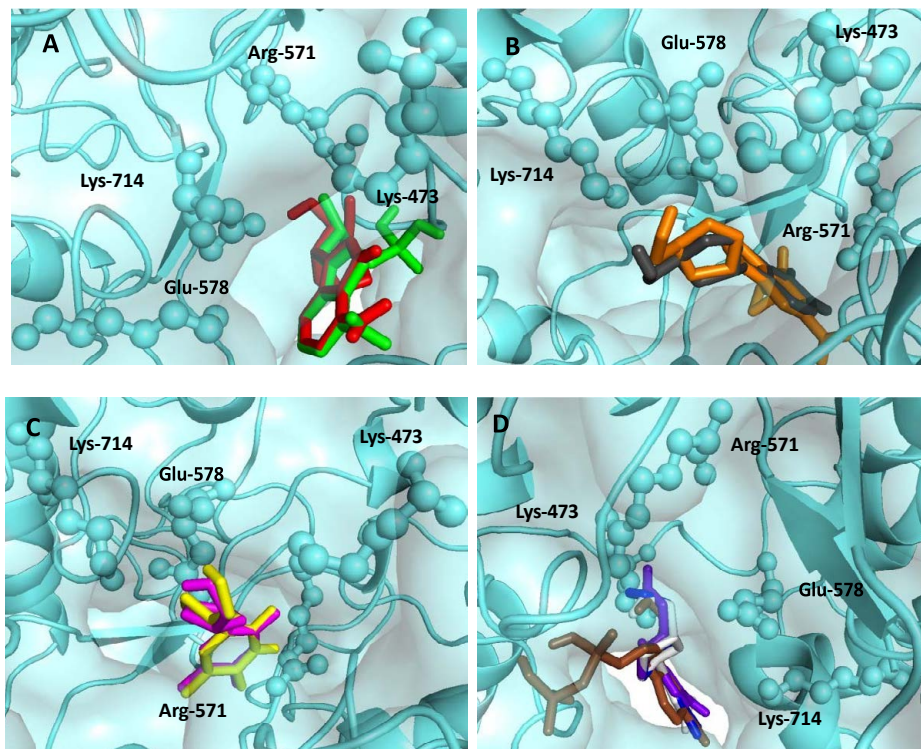
Figure 2: Molecular representation of inhibitors binding site of Ebola virus L-polymerase and inhibitors shown in colored sticks as obtained from docking simulations, while in balls and sticks are reported sensing-residue of binding phylogenetically conserved. A) Adenosine analogs binding site. In red and green are shown BCX4430 and Tenofovir respectively. B) Guanosine analogs binding site. In orange is represented Abacavir and in gray DDI. C) Thymidine analogs binding site. In yellow is displayed Zidovudine and in pink Stavudine. D) Cytidine analogs binding site. Cidofovir, Emtricitabine, Lamivudine and DDC are shown in brown, blue, purple and white respectively.

lowest possible binding energies. Nucleotide analogs were obtained from ZINC Database [18]. As shown in Table 1, ten good ΔG binding energies were generated after each run in the AutoDock Vina. Abacavir exhibits the lowest binding energy (-8.1 kcal/mol), which indicates high binding affinity later confirmed by in vitro results [8]. The P.L.I.P analysis determines the amino acid residues involved in ligand binding [24], upon binding molecules, various molecular interactions, especially hydrogen bonds and hydrophobic interactions, were formed between the L polymerase and nucleotide analogs explaining their inhibitor activity. As results of docking simulation, we observed that all molecules shared the same binding-site and, although nucleotide analogs share a similar chemical structure, modified interactions are present thus explaining their different inhibitory activities. Analyzing the network of interactions among inhibitors and ligand-sensing residues we have observed that the analogs belonging to the same nucleotide class, as shown in Figure 2, are oriented in a similar manner sharing exactly the same pose; the resulting network of interactions is described in Table 1. In conclusion, our computational analysis reveals the mechanism of interaction between different types of L polymerase inhibitors and residues of the active site as obtained by homology model. Finally, we would also like to draw your attention to the importance of knowing the interaction between ligand and an active site. As a matter of fact, this knowledge not only represents a fundamental step towards the improvement in efficacy of the inhibitors, but it will also help identify new classes of them as antiviral compounds, especially in the field of therapeutic switching.

References

1. Del Rio C, Mehta AK, Lyon GM, Guarner J (2014) Ebola hemorrhagic fever in 2014: the tale of an evolving epidemic.Ann Intern Med 161: 746-748.

2. Martins KA, Jahrling PB, Bavari S, Kuhn JH (2016) Ebola virus disease candidate vaccines under evaluation in clinical trials.Expert Rev Vaccines 15: 1101-1112.

3. Grifoni A, Lo Presti A, Giovanetti M, Montesano C, Amicosante M, et al. (2016) Genetic diversity in Ebola virus: Phylogenetic and in silico structural studies of Ebola viral proteins.Asian Pac J Trop Med 9: 337-343.

4. Ayub G, Waheed Y (2016) Sequence analysis of the L protein of the Ebola 2014 outbreak: Insight into conserved regions and mutations.Mol Med Rep 13: 4821-4826.

5. Oany AR, Sharmin T, Chowdhury AS, Jyoti TP, Hasan MA(2015) Highly conserved regions in Ebola virus RNA dependent RNA polymerase may be act as a universal novel peptide vaccine target: a computational approach. In Silico Pharmacol 3: 7-11

6. Cong Q, Pei J, Grishin NV (2015) Predictive and comparative analysis of Ebolavirus proteins.Cell Cycle 14: 2785-2797.

7. Van Hemert FJ, Zaaijer HL, Berkhout B(2015) In silico prediction of ebolavirus RNA polymerase inhibition by specific combinations of approved nucleotide analogues. J Clin Virol 73: 89-94

8. Reynard O, Nguyen XN, Alazard-Dany N, Barateau V, Cimarelli A, et al. (2015) Identification of a New Ribonucleoside Inhibitor of Ebola Virus Replication.Viruses 7: 6233-6240.

9. Liang B, Li Z, Jenni S, Rahmeh AA, Morin BM, et al. (2015) Structure of the L Protein of Vesicular Stomatitis Virus from Electron Cryomicroscopy.Cell 162: 314-327.

10. Kuhn JH, Becker S, Ebihara H, Geisbert TW, Johnson KM, et al. (2010) Proposal for a revised taxonomy of the family Filoviridae: classification, names of taxa and viruses, and virus abbreviations. Arch Virol 155:2083-2103.

11. Sleigh SH, Barton CL (2010) Repurposing Strategies for Therapeutics. Pharm Med 24: 151–159.

12. UniProt Consortium (2015) UniProt: a hub for protein information.Nucleic

Acids Res 43: D204-212.

13. Guex N, Peitsch MC (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. Electrophoresis 18: 2714-2723.

14. Berman HM, Westbrook J, Feng Z, Gilliland G, Bhat TN, et al. (2000) The Protein Data Bank. Nucleic Acids Res 28: 235-242.

15. Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, et al. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25: 3389-3402.

16. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) Clustal W and Clustal X version 2.0.Bioinformatics 23: 2947-2948.

17. Valdar WS (2002) Scoring residue conservation. Proteins 48: 227-241.

18. Trott O, Olson AJ (2010) AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization and multithreading. J Comput Chem. 2010; 31: 455- 461.

19. Irwin JJ, Shoichet BK (2005) ZINC--a free database of commercially available compounds for virtual screening. J Chem Inf Model 45: 177-182.

20. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, et al. (2009) AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility. J Comput Chem 30: 2785-2791.

21. DeLano WL (2012) The PyMOL Molecular Graphics System (2002) on World Wide Web.PyMol.

22. Salentis S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M (2015)PLIP: fully automated protein-ligand interaction profiler. Nucl Acids Res 43: 443-447.

23. Oestereich L, Lüdtke A, Wurr S, Rieger T, Muñoz-Fontela C, et al.(2014) Successful treatment of advanced Ebola virus infection with T-705 (favipiravir) in a small animal model. Antiviral Res 105: 17-21.

24. Litterman N, Lipinski C, Ekins S (2015) Small molecules with antiviral activity against the Ebola virus.F1000Res 4: 38.

Author Affiliations

Department of Biotechnology , Chemistry and Pharmacy , University of Siena, Italy