Computers in Biology and Medicine 113 (2019) 103414

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

Pharmacoinformatics-based identification of potential bioactive compounds against Ebola virus protein VP24

Samuel K. Kwofie[a,b,c,*], Emmanuel Broni[a], Joshua Teye[a], Erasmus Quansah[d], Ibrahim Issah[a], Michael D. Wilson[c,d], Whelton A. Miller III[c,e], Elvis K. Tiburu[a,b], Joseph H.K. Bonney[f]

[a] Department of Biomedical Engineering, School of Engineering Sciences, College of Basic & Applied Sciences, University of Ghana, PMB LG 77, Legon, Accra, Ghana
[b] West African Center for Cell Biology of Infectious Pathogens, Department of Biochemistry, Cell and Molecular Biology, College of Basic and Applied Sciences, University of Ghana, Accra, Ghana
[c] Department of Medicine, Loyola University Medical Center, Maywood, IL, 60153, USA
[d] Department of Parasitology, Noguchi Memorial Institute for Medical Research (NMIMR), College of Health Sciences (CHS), University of Ghana, P.O. Box LG 581, Legon, Accra, Ghana
[e] Department of Chemical and Biomolecular Engineering, School of Engineering and Applied Science, University of Pennsylvania, Philadelphia, PA, USA
[f] Department of Virology, Noguchi Memorial Institute for Medical Research (NMIMR), College of Health Sciences (CHS), University of Ghana, P.O. Box LG 581, Legon, Accra, Ghana

ARTICLE INFO

Keywords:
Ebola virus protein VP24
Virtual screening
Molecular docking
Molecular dynamics
Ebola virus inhibitors
African natural products

ABSTRACT

Background: The impact of Ebola virus disease (EVD) is devastating with concomitant high fatalities. Currently, various drugs and vaccines are at different stages of development, corroborating the need to identify new therapeutic molecules. The VP24 protein of the Ebola virus (EBOV) plays a key role in the pathology and replication of the EVD. The VP24 protein interferes with the host immune response to viral infections and promotes nucleocapsid formation, thus making it a viable drug target. This study sought to identify putative lead compounds from the African flora with potential to inhibit the activity of the EBOV VP24 protein using pharmacoinformatics and molecular docking.

Methods: An integrated library of 7675 natural products originating from Africa obtained from the AfroDB and NANPDB databases, as well as known inhibitors were screened against VP24 (PDB ID: 4M0Q) utilising AutoDock Vina after energy minimization using GROMACS. The top 19 compounds were physicochemically and pharmacologically profiled using ADMET Predictor™, SwissADME and DataWarrior. The mechanisms of binding between the molecules and EBOV VP24 were characterised using LigPlot+. The performance of the molecular docking was evaluated by generating a receiver operating characteristic (ROC) by screening known inhibitors and decoys against EBOV VP24. The prediction of activity spectra for substances (PASS) and machine learning-based Open Bayesian models were used to predict the anti-viral and anti-Ebola activity of the molecules, respectively.

Results: Four natural products, namely, ZINC000095486070, ZINC000003594643, ZINC000095486008 and sarcophine were found to be potential EBOV VP24-inhibitiory molecules. The molecular docking results showed that ZINC000095486070 had high binding affinity of −9.7 kcal/mol with EBOV VP24, which was greater than those of the known VP24-inhibitors used as standards in the study including Ouabain, Nilotinib, Clomiphene, Torimefene, Miglustat and BCX4430. The area under the curve of the generated ROC for evaluating the performance of the molecular docking was 0.77, which was considered acceptable. The predicted promising molecules were also validated using induced-fit docking with the receptor using Schrödinger and molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) calculations. The molecules had better binding mechanisms and were pharmacologically profiled to have plausible efficacies, negligible toxicity as well as suitable for designing anti-Ebola scaffolds. ZINC000095486008 and sarcophine (NANPDB135) were predicted to possess antiviral activity, while ZINC000095486070 and ZINC000003594643 to be anti-Ebola compounds.

Conclusion: The identified compounds are potential inhibitors worthy of further development as EBOV biotherapeutic agents. The scaffolds of the compounds could also serve as building blocks for designing novel Ebola inhibitors.

* Corresponding author. Department of Biomedical Engineering, School of Engineering Sciences, College of Basic & Applied Sciences, University of Ghana, PMB LG 77, Legon, Accra, Ghana.
E-mail address: skkwofie@ug.edu.gh (S.K. Kwofie).

https://doi.org/10.1016/j.compbiomed.2019.103414
Received 18 March 2019; Received in revised form 23 August 2019; Accepted 24 August 2019
Available online 28 August 2019
0010-4825/

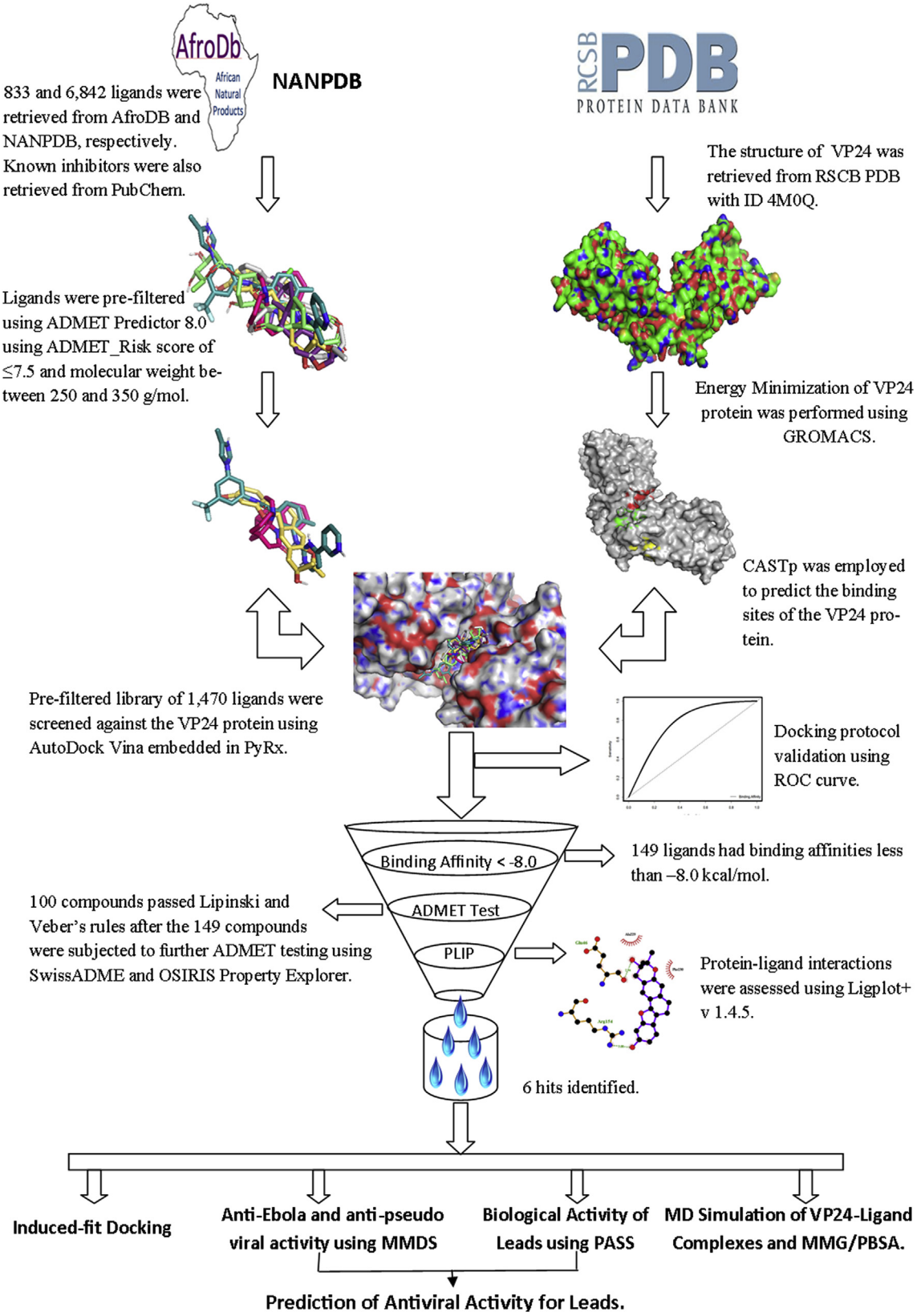


Fig. 1. Graphical representation of the schematic workflow adopted in this study. The methods included pre-filtering of ligands using toxicity, Lipinski's and Veber's rules. Molecular docking and dynamics simulations, as well as antiviral activity predictions were used to support potential lead identification. LigPlot + was used to determine the protein ligand interaction profiles (PLIP) whiles pharmacokinetics profiling was undertaken by determining absorption, distribution, metabolism, excretion and toxicity (ADMET).

1. Introduction

Ebola virus disease (EVD) is a viral haemorrhagic fever caused by Ebola viruses which affects humans and other primates [1,2]. Like other filoviruses, Ebola virus (EBOV) replicates very efficiently in many cells, giving rise to an appreciable number of the virus in cells of the mononuclear phagocyte system (MPS) and others including liver cells, fibroblasts and adrenal gland cells [3]. EBOV is reported to be transmitted to humans through direct contact with blood, mucous membranes or through skin contact. The disease is associated with a high risk of death, killing an average of about 50% of those infected with the highest death rate being up to 90% [4].

The World Health Organization (WHO) reported a total of 24 Ebola outbreaks involving 1716 cases from 1976 to 2013. Also, the largest outbreak ever recorded was the epidemic in West Africa, mostly in Liberia, Sierra Leone and Guinea from December 2013 to January 2016, claiming a little over 11,000 lives out of about 28,000 reported cases [5,6]. Even though, the end of the EVD outbreak in West Africa was declared in August 2018, it has re-emerged in the Democratic Republic of Congo. As at 1st May 2019, the total cases had reached 1510 (1444 confirmed) in the DRC with 994 deaths of which 928 were confirmed [7]. Recent studies have shown EVD reservoir in bats to be widespread with the identification of Bombali virus in the free-tailed bats (Family: Molossidae) in Sierra Leone and the discovery of a new genus of the filovirus the Mengla virus in Rousettus bats (Family: Megabats) [8,9].

The EBOV is a single-stranded RNA virus which encodes seven proteins; nucleoproteins (NP), glycoprotein (GP), polymerase (L), VP24, VP30, VP35 and VP40. The VP24 and VP35 structural proteins of EBOV are believed to play a very important role of interfering with the human immune system's response to viral infections [10]. The VP24 protein is integral in the formation of nucleocapsid (NC) and also a secondary matrix protein that plays a key role in the pathology of the EVD [11].

The VP24 disrupts signalling pathway of signal transducer and activator of transcription 1 (STAT1). It also inhibits the function of Karyopherin alpha (KPNA) by binding in a region which overlaps with the region where STAT1 binds to the KPNA due to the higher binding affinity between the VP24 - KPNA complex [12–14]. As a result, STAT1 is not able to elicit an immune response and VP24 is able to transport viral components into the nucleus of the target cell [15]. Furthermore, VP24 is responsible for forming fully functional and infectious virus-like particles (VLPs), the promotion of viral nucleocapsid formation and the regulation of replication [16,17]. Therefore, VP24 protein is pivotal in the replication and spread of the virus and enhances immune-suppression of the host, making it a viable target to combat EVD.

The search for potent EVD treatments is currently underway, but to the best of our knowledge no Food and Drug Authority (FDA) approved specific treatments exist. Currently, few drugs including Miglustat [18], Ouabain [19], Nilotinib [18], Clomiphene and Toremifene [20] among others have been documented to be VP24 inhibitors. Ouabain, Nilotinib and Miglustat are responsible for blocking viral replication, whilst Miglustat is shown to effectively inhibit EBOV particle assembly and secretion [21]. Clomiphene and Toremifene are oestrogen receptor modulators which prevent membrane fusion of EBOV [21]. Clomiphene and Toremifene were administered in mice for the treatment of EBOV, resulting in a survival rate of 50–90% [20,22]. BCX4430 (Immucillin-A), an antiviral drug, has also been developed for the treatment of deadly filovirus infections including EVD and Marburg virus disease [23,24].

The 2013–16 West African EVD outbreak accelerated the clinical development of vaccines including the recombinant VSV-ZEBOV, ChAd3-EBO-Z with or without MVA-BN-Filo, Ad26.ZEBOV with MVA-BN-Filo, Ad5-ZEBOV, and GamEvac-Combi [21,25,26]. Recent Ebola drug repurposing approach involved the screening of approved drugs to aid in the identification of novel inhibitors of EBOV [21,27]. Additionally, antiviral screening of multiple compounds using live EBOV has been reported [28].

Natural product-derived compounds are drug-like molecules which serve as rich sources of scaffolds for the discovery of novel drug leads [29]. Also, natural products have been reported to be rich in structural and chemical diversity [30]. *In silico* screening of natural products with therapeutic effects against EBOV have been reported [31]. Curcumin, curcuminoids and tetrahydrocurcumin are bioactive compounds of natural product origin which have shown potential anti-viral activity against EBOV proteins [32,33]. Previously, A total of 2020 Indonesian natural products were screened against Ebola VP24 receptor and Cycloartocarpin was suggested as an anti-Ebola lead with the least Gibbs free binding energy of −7.4847 kcal/mol [34].

Pharmacoinformatics has evolved as one of the most cost-effective and reliable techniques to discover novel leads [35–37]. *In silico* identification of potential EBOV inhibitors targeting the different viral proteins using molecular docking and quantitative structure–activity relationships (QSAR) studies have been reported [23,32,38–41]. Therefore, the identification of new bioactive compounds via *in silico* drug design is vital in unravelling of novel leads which have the potential to inhibit the activity of VP24. The African flora remains an untapped reservoir of new drug candidates for combating various kinds of diseases [42,43] and has been reported to have a diverse set of natural products with viable anti-viral activities [44]. The African continent is rich in biodiversity [45] and this can be exploited to produce novel drug candidates from its natural sources. .

The study sought to identify potential natural product-derived EBOV VP24 protein inhibitory compounds by virtual screening of an integrated African natural product library [46,47]. The predicted potential leads were evaluated using a receiver operating characteristic (ROC) curve and molecular dynamics (MD) simulations including molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) calculations and induced-fit docking (IFD). We further predicted anti-viral and anti-Ebola activity using machine learning based Open Bayesian techniques. Additionally, the study compared the binding energies of known inhibitors of EBOV VP24 protein to those in the libraries to aid in identifying compounds with potent binding affinities. Then pharmacokinetic and physicochemical profiling of compounds to identify potential novel drug-like leads were carried out. The study identified potential inhibiting compounds which warrant further development as EBOV biotherapeutic agents.

2. Methods

A graphical representation of the schematic workflow detailing the step-by-step methods implemented in this study is shown in Fig. 1. The methods involved pre-filtering of ligand libraries for downstream analysis, molecular docking, atomistic level molecular dynamics simulations and prediction of anti-viral activity.

2.1. Data sources for VP24 protein, African natural compounds and VP24-Inhibitor drugs

The 3D structure of the VP24 protein with reasonably good resolution was retrieved from the Research Collaboratory for Structural Bioinformatics Protein Data Bank (RCSB PDB) [[48]] with corresponding PDB ID: 4M0Q. An integrated library consisting of 7675 compounds was obtained from the Natural Product Library from African Medicinal Plants (AfroDB) and Northern African Natural Products Database (NANPDB). A total of 833 compounds were obtained

from AfroDB, which is a catalogue of ZINC database containing natural products originating from different geographic regions in Africa [47]. Also, 6842 compounds were obtained from NANPDB, a library of compounds isolated mainly from plants, with contributions from some endophyte, animal, fungal and bacterial sources [46]. The unique identifiers (IDs) of the NANPDB compounds used in this study were concatenated with the prefix "NANPDB". Known inhibitors of EBOV VP24 protein comprising Miglustat, Ouabain, Nilotinib, Clomiphene, Toremifene and BCX4430 were retrieved from PubChem [49].

2.2. Pre-filtering of screening library

ADMET Predictor™ (V8.0, Simulations Plus, Inc., Lancaster, CA) was used to screen the 7675 compounds to eliminate ligands with high toxicity levels and molecular weights greater than 350 g/mol and less than 250 g/mol, since the standard lead-like database of ZINC has compounds with molecular weights between 250 and 350 g/mol [50]. Also, the ADMET_Risk model embedded in ADMET Predictor™ was used to summarize the range of potential liabilities that can derail a drug candidate's development, from low solubility to unacceptably fast Cytochrome P450 (CYP) metabolism to causing phospholipidosis. A reference set of 2270 commercial drugs from the World Drug Index (WDI) were used in ADMET Predictor™ to build models. The ADMET_Risk for 90% of the reference set had scores of less than 7.5. ADMET Predictor™ has a risk value range of 0–24. The risk parameters have quantitative features where lower scores indicate the suitability of a compound to be more drug-like [51].

2.3. Preparation of protein and screening library

The NANPDB compounds obtained in 2D spatial data file (.sdf) were converted to 3D using Open Babel's "gen3d" option after removal of duplicates. The 3D structures from AfroDB were merged with NANPDB and imported into PyRx. The structures underwent energy minimization using the MMFF94 force field and Conjugate Gradient algorithm in 200 steps, and were converted to Protein Data Bank partial charge and atom type (.pdbqt) file format using Open Babel [52]. GROMACS v5.1.1 was used to perform the energy minimization of the protein using the Optimized Potentials for Liquid Simulations (OPLS)/All Atom (AA) force field [53]. The energy minimized VP24 protein which was in the GROMACS format (.gro) was then visualized in PyMOL (PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC). The water molecules were then removed from the protein and the structure was saved in the Protein Data Bank format (.pdb). PyRx was then used to convert the protein into AutoDock Vina's compatible pdbqt format using the "make macromolecule" option.

2.4. Binding site characterisation and molecular docking

Computed Atlas of Structure Surface Topography of proteins (CASTp) was used to determine the binding sites of the protein [54,55]. Previously, the binding sites of the VP24 protein were determined using Site Finder [34]. Chimera 1.12 [56] was used to analyse the binding sites predicted by CASTp [54]. The sites with no openings were eliminated including those with relatively small volumes and areas. AutoDock Vina interfaced with PyRx [57] was used to run the molecular docking and the poses were visually assessed in PyMOL (PyMOL Molecular Graphics System, Version 1.5.0.4, Schrödinger, LLC) [58,59]. The AutoDock Vina search space was maximized to cover all the four predicted binding sites using the following dimensions: center_x = 60.8092 Å, center_y = 60.8995 Å, and center_z = 60.9494 Å; and size_x = 64.5187991524 Å, size_y = 56.7348013687 Å, and size_z = 95.2103975201 Å.

2.5. Docking protocol validation

Decoys were obtained via the Directory of useful decoys and enhanced (DUD-E) [60] to generate the ROC curve using nine known VP24 specific inhibitors comprising Benzylpiperazine, Adamantane-1,3-diame, Adamantane-1,2-diamine, Ouabain, Nilotinib, Clomiphene, Torimefene, BCX4430 and Miglustat. The ROC curve evaluates the performance of the docking [61]. Forty-seven decoys were generated for each inhibitor, with the decoys possessing similar physiochemical properties but different 2-D topologies to the known inhibitors. The area under the curve (AUC) for the ROC curve generated using easyROC (Ver. 1.3) was obtained by screening a total of 423 decoys and 9 inhibitors against the VP24 receptor [62].The default parameters comprise parametrically fitted ROC curve, Standard error estimation and confidence interval of DeLong [63], and type I error of 0.05.

2.6. Characterisation of mechanism of binding and induced-fit docking (IFD)

LigPlot + v1.4.5 was used to characterise the interactions between the protein and the selected hits [64,65]. The hydrogen bonds were depicted with green dashed lines and the arcs with spokes radiating towards the ligands represented the hydrophobic interactions. Default parameters were used to generating the interaction profiles. Additionally, induced-fit docking (IFD) was used to investigate the flexibility of the ligand-receptor complexes via the Schrödinger software suite [66]. Proteins undergo conformational changes upon binding to other molecules [67–69]. The study of protein flexibility and interactions is important in drug discovery [67].

2.7. ADME screening of sub-library

Compounds with high binding affinity after molecular docking and better pose with the energy minimized VP24 protein were subjected to pharmacokinetics profiling including absorption, distribution, metabolism and excretion (ADME) testing using SwissADME [70]. An effective drug needs to reach its intended target and elicit the desired therapeutic effect with good pharmacokinetics [71].

2.8. Toxicity prediction with OSIRIS Property Explorer in DataWarrior

The toxicity profiles of the top 19 hits and 6 known inhibitors were evaluated using OSIRIS Property Explorer in DataWarrior 4.7.2 [72]. DataWarrior uses features of chemical structures to predict physiochemical properties. The compounds were processed in Simplified Molecular-Input Line-Entry System (SMILES) formats. The algorithm in the OSIRIS Property Explorer predicts relevant drug properties such as mutagenicity, tumorigenicity, irritancy and reproductive effect [72].

2.9. Prediction of anti-viral activity of compounds

The biological activity of the shortlisted compounds were predicted using the prediction of activity spectra for substances (PASS) [[73]]. PASS predicts the biological activity spectra of compounds using the SMILES files of the structures [74] based on Bayesian approach. Also, the Mobile Molecular Data Sheet (MMDS) (http://molmatinf.com/) was used as platform to calculate the Bayesian scores of the promising molecules in order to evaluate their anti-Ebola and anti-pseudo-viral activities. Open Bayesian models generated from viral Ebola replication and pseudo-type training datasets were used to identify potential anti-EBOV molecules [75]. Bayesian models embedded in MMDS compute scores which are calibrated to a probability-like scale [76]. The two models which are openly accessible (http://molsync.com/ebola/) were uploaded into MMDS to score the molecules.

2.10. Quality evaluation of shortlisted molecules

The inhibitory constant (Ki) was calculated using the binding energies of the top six selected compounds along with other metrics consisting of ligand efficiency (LE), LE scale (LE_scale), fit quality (FQ) and LE-dependent lipophilicity (LELP). The aforementioned metrics were calculated using the approach described previously [77].

2.11. Molecular dynamics simulation of protein-ligand complexes

The selected compounds and the known inhibitors were prepared with PRODRG (http://davapc1.bioch.dundee.ac.uk/prodrg/) prior to molecular dynamics (MD) simulation to generate a starting topology for the ligands with Chirality set to "yes" in order to maintain the chiral centers in the ligands. The charges were set to "full" and energy minimization set to "no". GROMOS96.1 was used to generate 43A1 parameters for the ligands. MD simulation was conducted with GROMACS 2018 under the GROMOS96 43A1 force field. The chain A of the VP24 was used for the molecular dynamics simulations. To solvate the protein-ligand complexes, the distance of complexes to box boundaries were set to 1.0 nm. Also, complex charges were neutralized with sodium and chloride ions. Each complex was minimized with 1000 steps of Steepest Descent and the final minimized structure was used as the initial structure for MD simulation. Also, equilibration protocol was used to restrain and relax protein-ligand positions. MD simulation was conducted for 100 ns with time steps of 2 fs under PME. Xmgrace [78] was used to plot the graphs generated from the MD simulations.

2.12. MM-PBSA calculations of ligand-receptor complexes

The binding free energies of the complexes were calculated using the Molecular Mechanics Poisson-Boltzmann Surface Area (MM-PBSA) method. MM-PBSA provides the combination of molecular mechanics and continuum solvent models. MM-PBSA calculations of the complexes were carried out using g_mmpbsa, which calculates binding energy components and the individual energy contributions of the residues [79,80]. R programming package [81] was then used to plot the graphs from the MM-PBSA computations.

3. Results and discussion

3.1. Binding site analysis

A search via RCSB database retrieved two EBOV VP24 protein X-ray structures with PDB IDs 4U2X and 4M0Q of resolutions 3.15 Å and 1.92 Å, respectively. 4M0Q was then selected for this study due to its higher resolution [34,82]. 4M0Q is a dimer [83] of which each chain is 37% helical comprising 10 helices and 87 residues, and 20% beta sheet composed of 10 strands and 46 residues (Fig. 3A).VP24 is a single domain, alpha/beta (α/β) structure with a shape that resembles a triangular pyramid [84], but when expressed in mammalian cells, has the tendency to oligomerize into tetramers [11].

Q-SiteFinder was used to characterise the binding site of the EBOV VP24 [34] and the residues identified were Ala99, Gln103, Leu106, Gly117, Gly120, Leu121, Ser123, Asp124, Leu127, Thr128, Thr183, Gln184, Asn185 and His186. Herein, CASTp was used to predict four major binding pockets (Table 1). Pocket 2 predicted via CASTp overlapped with the binding site predicted previously [34]. Predicted sites with very small volumes and areas such that no ligands could fit were not considered for downstream virtual screening. An *in silico* study to identify new anti-EBOV VP24 compounds screened ligands from ZINC15 as well as oleuropein and ouabain against EBOV VP24 (ID: 4U2X). The three putative leads identified had interactions with Leu127 and Thr128 [41], which were similarly predicted using CASTp in pockets 2 and 4 (Table 1). Pockets 1 and 3 predicted using CASTp are potential targets for inhibition of VP24 protein. Ser151, Ser155, Arg154 and Lys39 found in pocket 1 were also identified as the interacting residues in VP24 receptor docking studies performed with 56 compounds and their analogues [85].

3.2. Pre-filtering of library and molecular docking studies

Out of a total of 7675 natural products, 1470 ligands passed the ADMET test which was performed with ADMET Predictor™. Most of the compounds failed due to their heavy molecular weights and relatively high ADMET_Risk scores above 7.5. A sub-library composed of 1470 ligands that passed the ADMET test and six known VP24 inhibitors

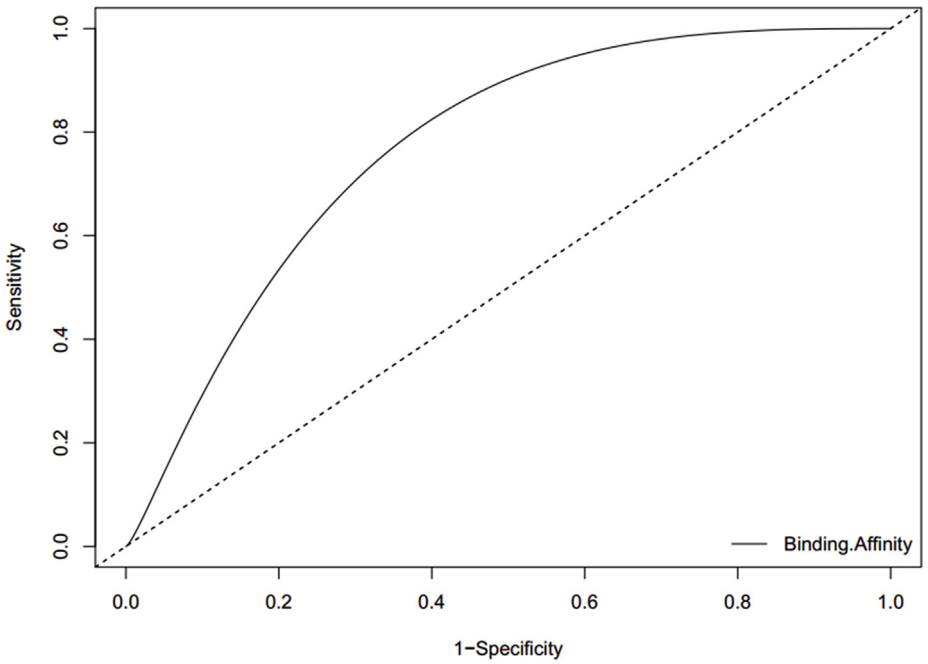


Fig. 2. Receiver Operator Characteristic (ROC) curve for evaluating the performance of the virtual screening. The curve was generated by screening active compounds and their corresponding decoys against the EBOV VP24 receptor. An AUC of 0.77 was obtained from the ROC curve.

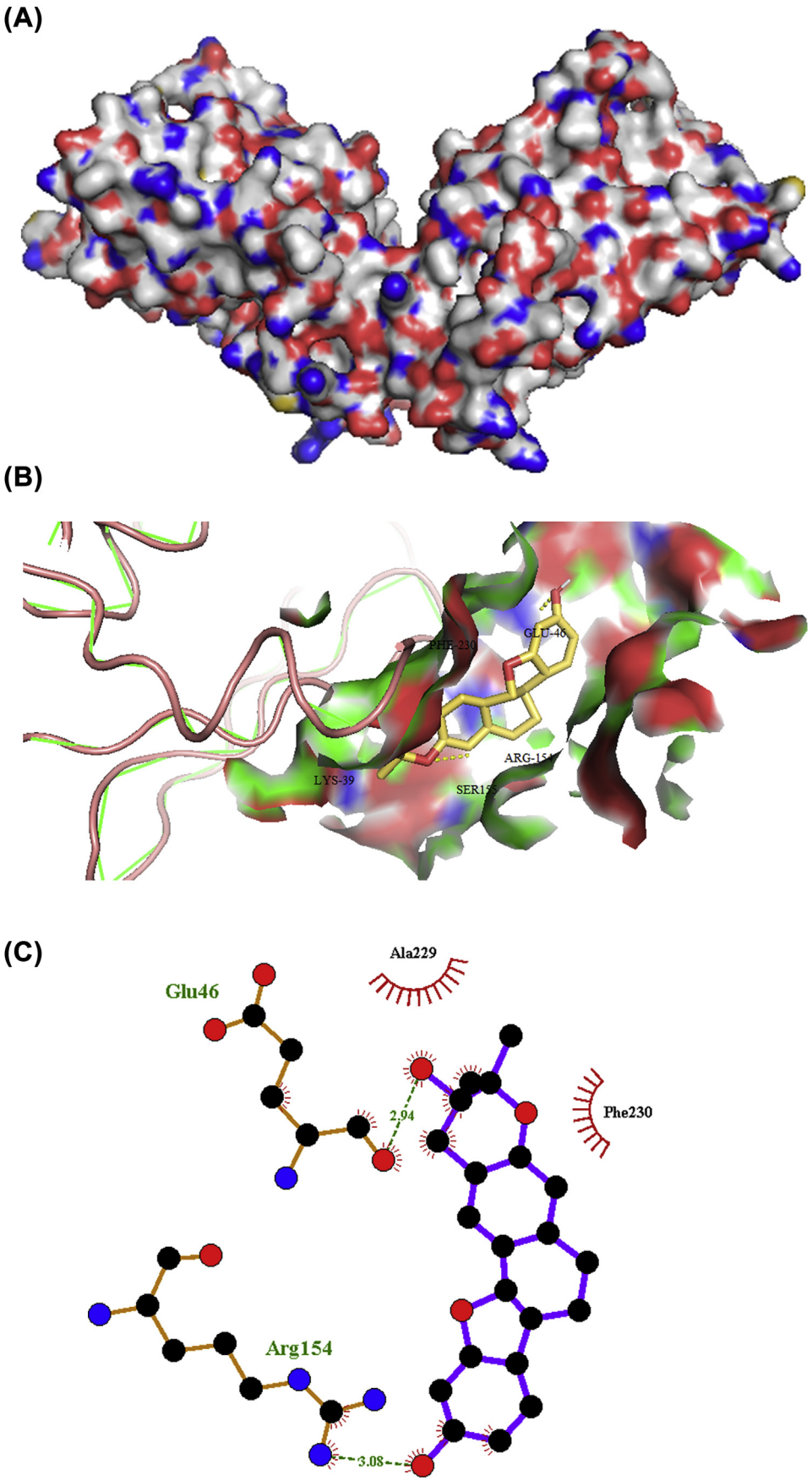


Fig. 3. The structure of EBOV VP24 protein. (A) Hydrophobicity representation of the 3D structure of VP24 (ID: 4M0Q). (B) Cartoon representation of VP24 in complex with ZINC000095486008. The binding site depicted as surface representation with ZINC000095486008 shown in sticks. (C) LigPlot + diagram of ZINC000095486008 coloured in purple showing hydrogen bonding with residues Glu46 and Arg154. Green dash lines represent the hydrogen bond interactions. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

Table 1
Selected VP24 binding sites predicted via CASTp with amino acid residues linning respective binding pockets [54,55].

Binding site	Residues lining the binding site
Pocket 1	Lys39, Gly44, Ile45, Glu46, Phel47, Asp48, Ser151, Arg154, Ser155, Ile157, Leu158, Ile161, Ser225, Thr226, Ala229, Phe230, Thr231.
Pocket 2	Val96, Ala99, Ala100, Gln103, Leu106, Glu113, Ala116, Gly117, Gly120, Ser123, Asp124, Leu127, Thr128, Ile181, Thr183, Gln184, Asn185, His186, Ile188.
Pocket 3	Gly101, Ile102, Gln105, Gln109, Ser110, Ile112, Leu115, Leu119, Gln144, Ser146, Lys148, Met149, Leu152, Ile153, Asn156, Ile157.
Pocket 4	Ile102, Gln103, Ile107, Gly120, Ser123, Asp124, Leu127, Thr128, Thr183, Gln184, Asn185, His186.

comprising Ouabain [18], Nilotinib [86], Clomiphene [87], Torimefene, Miglustat [18] and BCX4430 [24] were used for the molecular docking. Supplementary Files 1 to 8 are the PDB files generated from the molecular docking simulations.

A single grid box used for the docking was set to cover all the four predicted binding sites (Table 1). Ligands which possessed binding energies lower than −8.0 kcal/mol with the VP24 protein were selected for downstream analysis. A previous study used a threshold of −7.0 kcal/mol to pre-filter a library [88]. This threshold discriminates well between putative specific and non-specific protein-ligand bonds [88] and may be effective in filtering out weakly binding ligands. Since this threshold was defined for AutoDock users, this current study used a more stringent threshold of less than −8.0 kcal/mol with 149 ligands meeting the criteria.

3.3. Identification of hits using physicochemical profiling

The 149 pre-filtered compounds were physicochemically profiled using SwissADME and OSIRIS Property Explorer. Out of the 149 compounds, 100 complied with the Lipinski's rule of 5 [89] and Veber's rule [90] (Supplementary Table 1). Lipinski's rule evaluates drug-likeness for oral bioavailability and entails an n-octanol–water partition coefficient logP not more than 5, hydrogen bond donors not more than 5, hydrogen bond acceptors not more than 10, and molecular weight not more than 500 Da [91]. Veber's rule entails having a Total Polar Surface Area (TPSA) less than or equal to 140 Å^2 [90]. Compounds with a TPSA equal to or less than 140 Å^2 are likely to have a good oral bioavailability [90]. The TPSA values are considered as surrogate indicators of excellent human intestinal absorption (HIA) and Caco-2 permeability [39]. All the potential leads had TPSA values less than 140 Å^2.

3.4. Comparison of the binding energies of the hits with known inhibitors

The binding affinity of VP24-ligand complex is an important parameter which indicates the strength at which a ligand binds to the VP24 receptor. Specific recognition of ligands by proteins is at the core of pharmacoinformatics, providing insight into the ability of a possible drug candidate to inhibit the activity of a protein [92]. The results of the molecular docking showed that ZINC000095486070 possessed the least binding energy with VP24 compared to all the known VP24-inhibitors with a binding energy of −9.7 kcal/mol, implying ZINC000095486070 had the strongest binding affinity with VP24. Clomiphene, Torimefene, BCX4430 and Miglustat possessed binding energies of −7.5, −7.4, −6.5 and −5.2 kcal/mol with the VP24 protein, respectively (Table 2).

Ouabain was identified as an inhibitor of the EBOV VP24 protein by inhibiting viral replication in human lung cells [19]. Ouabain was used in experiments to inhibit the function of ATP1A1 in cells infected with the EBOV. The results showed a decrease in the progeny virus in treated infected cells as compared to untreated infected cells [19]. Even though, the results from the experimental data appeared encouraging, Ouabain is not yet approved by FDA for the treatment of EBOV infections and may likely exhibit toxicity in high concentrations. Furthermore, there is no available experimental record for Ouabain in living

Table 2
The binding energies and intermolecular bonds between hits and known inhibitors with EBOV VP24.

ZINC ID/Drug name	Binding energy (kcal/mol).	Number of hydrogen bonds	Hydrogen bond residues	Hydrogen bond length (Å)	Hydrophobic contacts
ZINC000095486070	−9.7	0	–	–	Gly44, Ile45, Arg154, Phe230, Leu158
Ouabain	−9.3	3	Lys39, Arg154	2.82, 3.03, 3.03	Glu46, Phe230, Leu158, Ile45, Gly44, Ser155, Arg154
NANPDB6380	−9.2	1	Arg154	3.25	Ile45, Glu46, Ala229, Phe230.
ZINC000003594643	−9.1	0	–	–	Arg154, Phe230, Asp48, Ala229, Glu46
ZINC000095486008	−9.1	2	Glu46, Arg154	2.94, 3.08	Ala229, Phe230
NANPDB102	−9.1	1	Gly44	3.08	Ile45, Glu46, Arg154, Phe230.
Nilotinib	−9.1	2	Glu46, Ser155	3.16, 2.98	Ala229, Glu46, Arg154, Leu158, Glu46, Arg154, Glu46
NANPDB135	−9.0	1	Arg154	2.98	Glu46, Leu158, Ala229, Phe230.
NANPDB2581	−9.0	0	–	–	Glu46, Arg154.
NANPDB3497	−9.0	2	Glu46, Ala229	2.70, 2.70	Ile45, Glu46, Ala229, Phe230.
NANPDB6355	−9.0	0	–	–	Ile45, Arg154, Leu158. Phe230.
ZINC000095485910	−8.8	2	Lys39, Asp48	2.91, 2.89	Glu46
NANPDB1950	−8.8	2	Arg154	2.80, 2.98	Arg154, Ser155, Leu158, Phe230.
ZINC000040393112	−8.7	0	–	–	Arg154, Glu46
NANPDB94	−8.7	0	–	–	Ile45, Glu46, Arg154, Leu158, Phe230.
NANPDB129	−8.7	0	–	–	Gly44, Glu46, Arg154, Leu158, Ala229, Phe230.
NANPDB2246	−8.7	1	Arg154	3.05	Ile45, Glu46, Arg154, Leu158, Ala229, Phe230.
NANPDB2434	−8.7	1	Arg154	3.24	Ile45, Glu46, Leu158, Phe230.
NANPDB4089	−8.7	1	Glu46	2.91	Ile45, Glu46, Arg154, Leu158, Phe230.
NANPDB4631	−8.7	1	Arg154	2.95	Ile45, Glu46, Arg154, Leu158, Ala229, Phe230.
NANPDB6786	−8.7	2	Gly44, Arg154	2.92, 3.09	Gly44, Ile45, Glu46, Arg154, Phe230.
Clomiphene	−7.5	0	–	–	Glu46, Leu158, Phe230, Glu46, Lys39, Asp48, Arg154, Glu46, Ile45
Torimefene	−7.4	0	–	–	Glu46, Arg154, Phe230, Ala229, Leu158, Ile45
BCX4430	−6.5	2	Gly44, Asp48	3.21, 3.07	Lys39, Glu46, Arg154, Phe230, Ile45, Asp48
Miglustat	−5.2	3	Arg154, Gly44	2.90, 3.13, 3.27	Arg154, Leu158, Ile45, Gly44

animals infected with EBOV [18]. Ouabain had the least binding energy of −9.3 kcal/mol among the known inhibitors considered, which implied that it had the highest binding affinity to the EBOV VP24 protein. Clomiphene and Toremifene possessed the fourth and third highest binding energies of −7.5 and −7.4 kcal/mol among the known inhibitors, respectively. Miglustat exhibited the highest binding energy of −5.2 kcal/mol among the known inhibitors, inferring least binding affinity. Miglustat is believed to block viral replication when bound to the VP24 protein [18]. Clomiphene, Toremifene and Miglustat are FDA approved drugs which have shown the potential to inhibit EBOV replication even though they have associated side effects [18,20,93]. BCX4430 possessed a binding energy of −6.5 kcal/mol and has been reported to inhibit EBOV replication but is yet to be approved by the FDA [18].

The binding energy of ZINC000095486070 was compared to the leads Bisdemethoxycurcumin [32] and Limonin [23]. Bisdemethoxycurcumin exhibited binding energy of −7.7 kcal/mol [32], whilst both ZINC000095486070 and Limonin had −9.7 kcal/mol each. Also, BCX4430 had a binding energy of −7.7 kcal/mol [23] compared to −6.5 kcal/mol (Table 2). The difference in the binding energies of BCX4430 was due to the different docking sites. However, in both studies BCX4430 exhibited the least binding affinity against the VP24 protein when compared to other screened inhibitors.

3.5. Docking protocol validation using ROC curve analysis

The ROC curve provides a representation of the sensitivity and specificity of a classifier at different thresholds. Its area under the curve (AUC) corresponds to the general performance of the docking and evaluates the ability of a docking model to distinguish docked decoys from active ligands [[62,94]]. AUC values of a ROC curve ranges between 0 and 1. AUC values closer to 1 are indicative of successful discrimination of active compounds from decoys whilst values closer to 0 indicate poor discrimination [61]. AUC value of 1 represent perfect classification of active compounds from decoys. Values less than 0.70 are indicative of moderate discrimination, whilst values ≤ 0.5 depict poor discrimination [61,95]. The AUC computed from the generated ROC curve was 0.77 (Fig. 2), which was considered as acceptable.

3.6. Molecular interactions with EBOV VP24

The molecular interactions between the ligands and the binding sites of the VP24 protein are essential when considering a ligand to be a plausible drug candidate. The VP24 protein-ligand interactions of the top five ligands with low binding energies are shown in Fig. 3 and Supplementary Figs. 1–4. ZINC000095486070 interacted with Gly44, Ile45, Arg154, Leu158 and Phe230 via hydrophobic bonds (Supplementary Fig. 1). NANPDB6380 formed a hydrogen bond with Arg154 of length 3.25 Å and hydrophobic bonds with Ile45, Glu46, Ala229 and Phe230.

ZINC000003594643 formed hydrophobic bonds with Ala229, Asp48, Glu46, Arg154 and Phe230 (Supplementary Fig. 2). ZINC000095486008 also formed hydrophobic bonds with Ala229, Glu46, Arg154 and Phe230; and hydrogen bonds with Arg154 and Glu46 of lengths 3.08 and 2.94 Å, respectively (Fig. 3). NANPDB6380 formed hydrophobic bonds with Ile45, Glu46, Ala229 and Phe230, and a hydrogen bond of length 3.25 Å with Arg154 (Supplementary Fig. 3). NANPDB102 also formed a hydrogen bond with Gly44 with bond length of 3.08 Å and formed hydrophobic bonds with Ile45, Glu46, Arg154 and Phe230 (Supplementary Fig. 4).

ZINC000095485910 also formed hydrophobic bonds with Lys39, Asp48 and Glu46. Also, ZINC000095485910 formed very strong hydrogen bonds with Lys39 and Asp48 with bond lengths of 2.91 and 2.89 Å, respectively. ZINC000040393112 also formed hydrophobic bonds with Glu46 and Arg154. Interestingly, the top 19 hits interacted with residues Lys39, Gly44, Ile45, Glu46, Phel47, Asp48, Ser151, Arg154, Ser155, Ile157, Leu158, Ile161, Ser225, Thr226, Ala229, Phe230 and Thr231, which were present in pocket 1 (Table 1 and Supplementary Table 2). Similarly, curcumin, demethoxycurcumin, bisdemethoxycurcumin and tetrahydrocurcumin formed intermolecular bonds with the active site residues Arg154, Gly44, Glu46, Gln232, Ile45, Phe230 and Leu158, which were also present in pocket 1 (Table 1) except Gln232 [32].

Limonin was reported to interact with Ser178, Tyr172, Asn171, Met209, Ile189, Leu75, Ser177 and Thr191; samarcandin interacted with Arg154, Glu46, Phe230, Ser155, Arg154, Ile45 and Leu158; and gummosin formed intermolecular bonds with Arg154, Glu46, Leu158, Arg154, Phe230 and Leu158 [23]. Additionally, these aforementioned

Table 3
ADME Prediction of Top 19 hits and 6 Known Inhibitors for Gastrointestinal (GI); Blood Brain Barrier (BBB); Estimated Solubility (ESOL) and P-glycoprotein (Pgp).

COMPOUND	ESOL Log S	ESOL Class	GI absorption	BBB permeant	Pgp substrate
ZINC000095486070	−3.24	Soluble	High	Yes	No
Ouabain	−2.13	Soluble	Low	No	No
NANPDB6380	−2.34	Soluble	High	No	No
ZINC000003594643	−3.2	Soluble	High	No	No
ZINC000095486008	−4.37	Moderately soluble	High	Yes	Yes
NANPDB102	−3.7	Soluble	High	Yes	No
Nilotinib	−6.23	Poorly soluble	Low	No	No
NANPDB135	−3.8	Soluble	High	Yes	No
NANPDB2581	−4.09	Moderately soluble	High	Yes	No
NANPDB3497	−3.16	Soluble	High	Yes	Yes
NANPDB6355	−2.66	Soluble	High	No	No
ZINC000095485910	−4.16	Moderately soluble	High	No	No
NANPDB1950	−1.79	Very soluble	High	No	No
ZINC000040393112	−3.6	Soluble	High	Yes	Yes
NANPDB94	−3.8	Soluble	High	Yes	No
NANPDB129	−2.87	Soluble	High	Yes	Yes
NANPDB2246	−3.32	Soluble	High	Yes	Yes
NANPDB2434	−2.47	Soluble	High	No	Yes
NANPDB4089	−3.53	Soluble	High	Yes	Yes
NANPDB4631	−3.99	Soluble	High	Yes	No
NANPDB6786	−1.62	Very soluble	High	Yes	Yes
Clomiphene	−6.78	Poorly soluble	Low	No	Yes
Torimefene	−6.76	Poorly soluble	Low	No	Yes
BCX4430	−0.31	Very soluble	Low	No	No
Miglustat	−0.54	Very soluble	High	No	Yes

residues also interacted with the 19 hits. BCX4430 was also reported to interact with Ser123, Gln184, Asp124, Gln103, Gly120 and Asn185 [23], which lined pocket 2 (Table 1). Arg154 has been shown to interact with all ligands that bind in pocket 1 and could be an essential residue critical for inhibition, thereby warranting further studies.

3.7. In silico pharmacokinetic properties and toxicity prediction

Most of the selected compounds were predicted to have relatively low ESOL logS values (Table 3). The value of ESOL logS provides insight into the aqueous solubility of a compound. ZINC000095486070 had a high gastrointestinal (GI) absorption which suggests a high probability of successful absorption into the intestinal tract. GI absorption is the absorption of orally administered drugs into the bloodstream. Blood Brain Barrier (BBB) permeation is the potential of a drug to cross the blood brain barrier to the brain where it binds to specific receptors for the activation of signalling pathways. Predicting the BBB permeability is crucial in the drug development pipeline since a molecule cannot demonstrate desired pharmacological activities with the brain parenchyma without permeation of the brain barrier [96]. Thirteen of the hits were found to have permeation into the BBB, which was denoted by the criterion "Yes" (Table 3).

Another parameter considered was the P-glycoprotein (P-gp), which aids in the elimination of xenobiotics from the central nervous system (CNS) by functioning as a biological barrier by removing toxins and xenobiotics from cells. It also plays an important role in the absorption and distribution of drugs [97]. Among the top 19 hits, 11 compounds which were predicted to be non-substrate of P-gp (Table 3) are more likely to possess desirable distribution in the circulatory system upon administration.

CYP450 proteins are a group of isozymes in the hemoproteins superfamily which regulate drug metabolism [98]. They are also vital in the metabolism of steroids, fatty acids, carcinogens, drugs and other xenobiotics. CYP450 inhibitors comprising 1A2, 2C19, 2C9, 2D6 and 3A4 were investigated using ADME screening. Sixteen of the hits were non-substrate for at least 3 of the CYP450 inhibitors (Supplementary Table 3). ZINC000095486070 which possessed the lowest binding energy of −9.7 kcal/mol was a substrate of CYP1A2 and CYP2D6, while NANPDB6380 and ZINC000003594643 which had the second and third lowest binding energies of −9.2 and −9.1 kcal/mol were non-substrates for all the CYP450 inhibitors, respectively. The inhibition of CYP450 proteins affects drug metabolism and leads to the rise in toxicity levels [99]. From the results obtained, 9 out of the top 19 hits were not mutagenic, tumorigenic and irritant (Supplementary Table 4). ZINC000095486070 had no mutagenicity, tumorigenicity, reproductive effect and irritation. Therefore, ZINC000095486070 may elicit negligible harmful effects when administered. Ouabain, Nilotinib, BCX4430 and Miglustat were predicted to be non-mutagenic, non-tumorigenic and non-irritant. Clomiphene was predicted to possess high mutagenicity, tumorigenicity and irritation. Also, Torimefene was predicted to have low mutagenicity and non-irritant.

3.8. Anti-viral prediction

3.8.1. Prediction of activity spectra for substances (PASS)

PASS is used to predict molecular activity by comparing the structures of the queried compounds with the structures of known active substrates present in its database [74]. A Bayesian algorithm is used to predict the probable activity (Pa) or probable inactivity (Pi) of a compound. Pa denotes the probability of a queried compound to belong to the subclass of active compounds, while Pi is the probability to belong to the subclass of inactive compounds present in the PASS database [100]. ZINC000095486008 was predicted to possess anti-Rhinovirus activity with Pa of 0.511 and Pi of 0.021; anti-Herpes activity with Pa of 0.422 and Pi of 0.026; and anti-Influenza activity with Pa of 0.431 and Pi of 0.037. NANPDB6380 was also predicted to possess anti-Rhinovirus activity with Pa of 0.305 and Pi of 0.243, whilst NANPDB135 possessed anti-Rhinovirus activity with Pa of 0.501 and Pi of 0.024. NANPDB2246 was predicted to possess activity against Rhinovirus with Pa of 0.554 and Pi of 0.011; Influenza with Pa of 0.458 and Pi of 0.030; and Herpes with Pa of 0.401 and Pi of 0.035.

When the Pa is greater than Pi for a particular compound activity, it is worth exploring the pharmacological activity [101,102]. Since the Pa obtained for each activity was greater than the corresponding Pi, ZINC000095486008, NANPDB6380, NANPDB135 and NANPDB2246 are attractive anti-viral candidates for *in vitro* experimentation [103].

3.8.2. Bayesian models for predicting anti-Ebola activity

Bayesian models were used to ascertain if a compound was likely to be active against VP24 receptor. The Bayesian models were built using training datasets composed of viral pseudo type entry and Ebola virus replication assay datasets [75]. The Bayesian models were uploaded into the MMDS to evaluate the anti-Ebola activity of the molecules [75]. Also, the Bayesian model prediction scores for the anti-Ebola activity of ZINC000095486070, NANPDB6380, ZINC000003594643, ZINC000095486008, NANPDB135 and NANPDB2246 were 0.583, 0.429, 0.551, 0.088, 0.119 and 0.009, respectively. The Bayesian model prediction scores for the pseudo viral activity of ZINC000095486070, NANPDB6380, ZINC000003594643, ZINC000095486008, NANPDB135 and NANPDB2246 were 0.577, 0.410, 0.502, 0.434, 0.234 and 0.162, respectively. Even though, ZINC000095486008 and NANPDB2246 were predicted to have the least anti-Ebola activity, they were suggested as anti-viral using PASS [73,100]. The Bayesian models were trained with total data size of 868 with 20 compounds active against Ebola and 41 compounds active against pseudo virus [75]. Even though, the Bayesian models were trained with relatively smaller datasets, they provide insight into the anti-Ebola activity of the potential leads.

3.9. Ligand efficiency-based metrics for evaluating the selected compounds

The quality of the predicted selected compounds were assessed using metrics including inhibitory constant (Ki), ligand efficiency (LE), LE scale (LE_scale), fit quality (FQ) and LE-dependent lipophilicity (LELP) (Table 4).

The predicted Ki values ranged from 0.077 to 0.419 μM (Table 4),

Table 4
The quality assessment metrics for the selected compounds. The metrics included inhibitory constant based on binding affinity (Ki), ligand efficiency (LE), LE scale (LE_scale), fit quality (FQ) and LE-dependent lipophilicity (LELP).

ID	Binding Energy	NHA	LogP	Ki	LE	LE_Scale	FQ	LELP
ZINC000095486070	−9.7	22	2.53	7.7E-08	0.441	0.429	1.028	5.738
NANPDB6380	−9.2	24	1.6	1.80E-07	0.383	0.404	0.949	4.174
ZINC000003594643	−9.1	25	2.18	2.13E-07	0.364	0.392	0.929	5.989
ZINC000095486008	−9.1	25	3.23	2.13E-07	0.364	0.392	0.929	8.874
NANPDB135	−9.0	23	3.8	2.52E-07	0.391	0.416	0.940	9.711
NANPDB2246	−8.7	24	2.72	4.19E-07	0.363	0.404	0.898	7.503

with the low Ki values indicating the inhibitory potential of the compounds [104]. LE evaluates the binding affinity in relation to the number of heavy atoms (NHA) in a molecule and should be at least 0.29 kcal/mol/HA for lead-like molecules [105,106]. The six selected compounds had very good LE values ranging from 0.36 to 0.441, similar to the average LE value (0.38) of fragment-like molecules [106]. The LE_Scale was computed to address the size-dependent limitation of LE. The LE_Scale values of the compounds ranged from 0.39 to 0.43, which were close to other active compounds with similar number of heavy atoms [107,108].

The fit quality (FQ) which is a ratio of the observed LE to the LE_Scale of a molecule provides more accuracy in determining the efficiency of a ligand. A molecule is suggested to have an ideal ligand binding if the FQ value is close to 1 [106]. The six compounds exhibited FQ values close to 1 with the lowest and highest being 0.8 and 1.028, respectively, suggesting ideal binding to the protein. The LELP addresses the issues of lipophilicity and specificity of the ligands to the receptor. The LELP values of the compounds ranged from 4.1 to 9.72 (Table 4). The lower the absolute value of LELP, the more potent a lead compound, thus a good lead will have LELP closer to 0. Compounds within the Lipinski zone have LELP of less than 16.5, though the desirable range for LELP is between 0 and 7.5 [109]. Overall, the compounds are reasonable templates for optimisation in fragment-based drug design.

3.10. Induced-fit docking

Protein flexibility is key to drug discovery since protein structures vacillate between different conformations in their native states [110–112]. For most docking tools, flexible ligand docks to rigid protein targets [113], limiting insight into the flexibility of the receptor. The IFD module in Schrödinger software suite [66] was employed in investigating the structural flexibility of ZINC000095486070, ZINC000003594643, ZINC000095486008, NANPDB6380 and

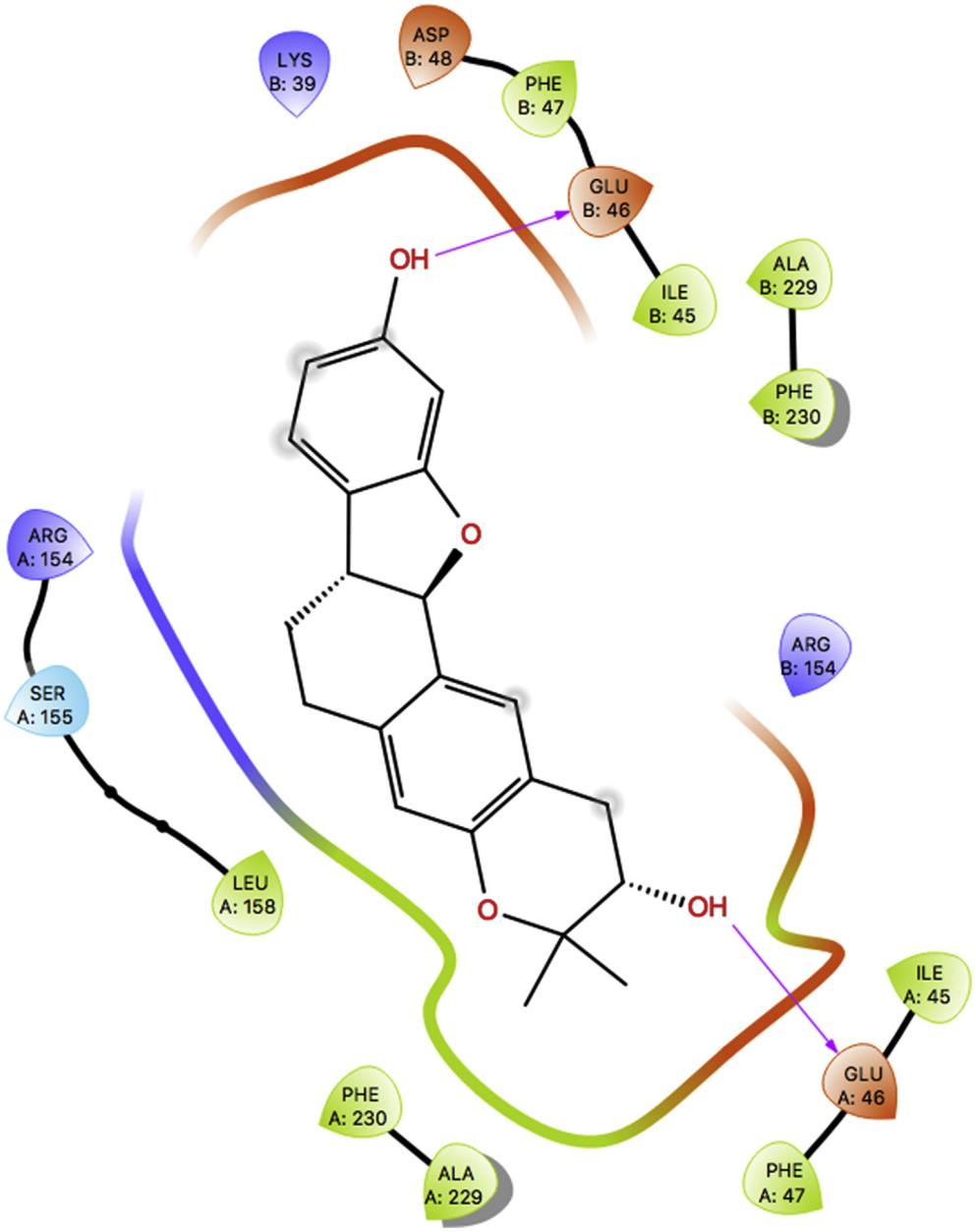


Fig. 4. Molecular interactions of ZINC000095486008 ligand complex with hydrogen bond interactions indicated as purple arrows. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

NANPDB2246 ligand complexes. The GlideScores and IFD scores of the most plausible poses were selected for downstream analysis. The GlideScore is an empirical scoring function that estimates the binding affinity of the ligand complexes, whilst the IFD score estimates the most likely conformations of the ligand complexes. Generally, lower GlideScores and IFD scores (more negative) are indicative of plausible binding between a ligand and a receptor [113–118]. The best conformers of the complexes of ZINC000095486070, NANPDB6380, ZINC000003594643, ZINC000095486008 and NANPDB2246 had GlideScores of −5.825 kcal/mol, −6.589 kcal/mol, −6.079 kcal/mol, −6.532 kcal/mol and −5.791 kcal/mol, with corresponding IFD scores of −893.707 kcal/mol, −868.98 kcal/mol, −891.056 kcal/mol, −895.226 kcal/mol and −868.13 kcal/mol, respectively. The molecular interactions of ZINC000095486008, ZINC000095486070, ZINC000003594643, NANPDB6380 and NANPDB2246 are shown in Fig. 4 and Supplementary Figs. 5, 6, 7 and 8, respectively. The IFD studies showed that the five ligands formed interactions with residues such as Glu46, Arg154, Ala229, Phe230, Leu158, Ile45, Ser155 and Gly44, which were revealed as critical interacting residues after molecular interaction analysis using LigPlot+ (Table 2 and Supplementary Table 2).

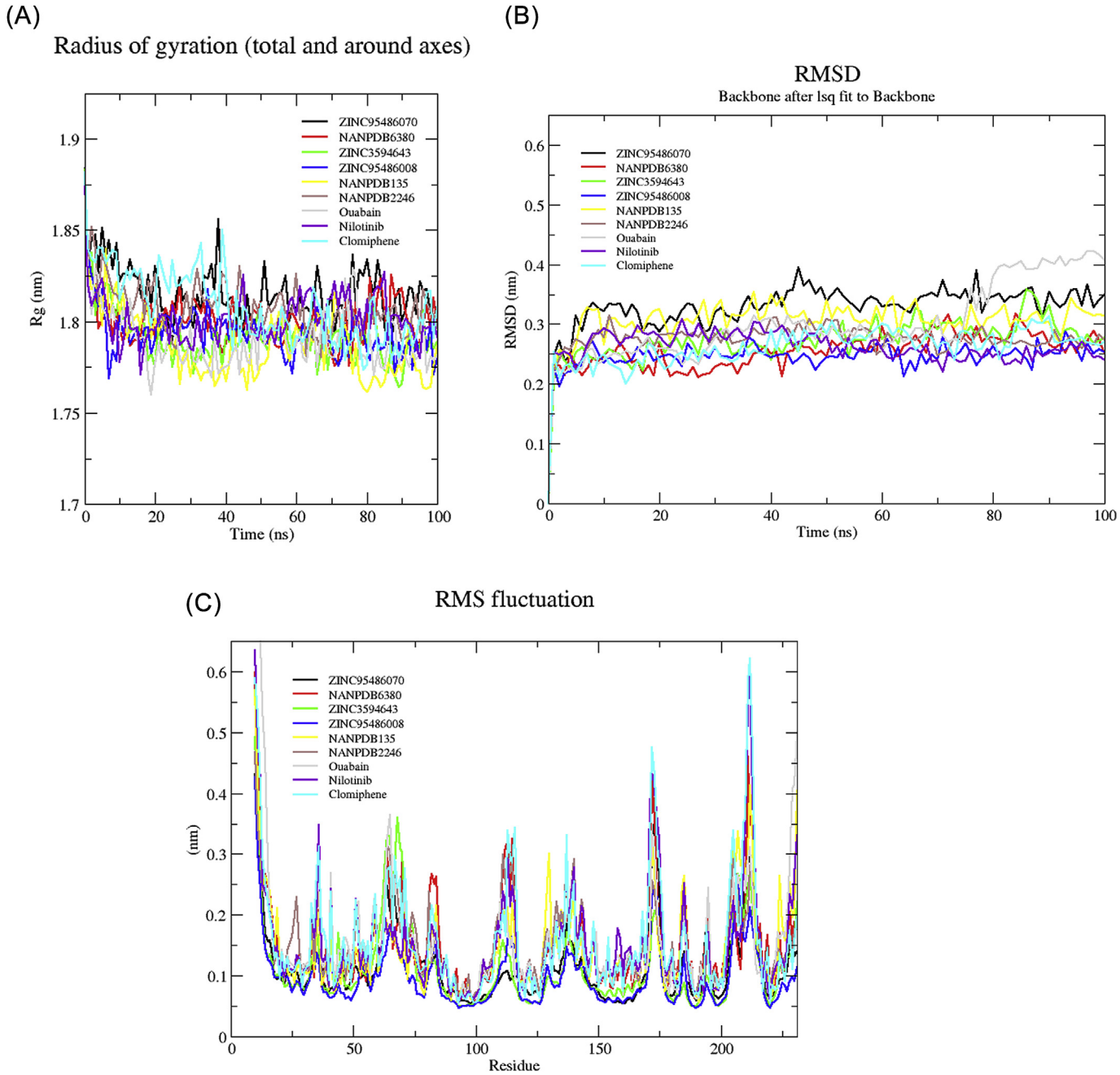


Fig. 5. Graphs of Rg, RMSD and RMSF of the EBOV VP24-ligand complexes generated over 100 ns simulation using GROMACS: (A) Radius of Gyration versus time graph of the VP24-ligand complexes. (B) RMSD versus time graph of the backbone atoms of VP24-ligand complexes over 100 ns; and (C) Analysis of RMSF trajectories of residues of VP24-ligand complexes. ZINC000095486070, NANPDB6380, ZINC0000954860008, ZINC000003594643, NANPDB135, NANPDB2246, Ouabain, Nilotinib and Clomiphene are shown in black, red, green, blue, yellow, brown, gray, purple and cyan, respectively. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

3.11. Molecular dynamics simulation of protein-ligand complexes

The stabilities of the protein-ligand complexes were analysed over 100 ns simulation period. MD was undertaken for the six selected compounds and three known inhibitors with very low binding energy to the VP24 protein (Supplementary Files 9 to 16). The compactness of the protein was determined by its radius of gyration (R_g) and a stably folded protein would likely maintain a relatively steady R_g. If a protein unfolds, its R_g would change over time. The R_g values obtained showed that the complexes remained very stable (Fig. 5A) in their compact (folded) form over the course of 100 ns. The R_g values of all nine complexes experienced a gradual fall from 0 to about 8 ns, and then remained steady till the end of the 100 ns. ZINC000003594643 and ZINC000095486008 complexes showed very little fluctuations with average R_g of about 1.78 nm each. The ZINC000095486070 complex also showed stability with an average radius of 1.82 nm with peaks recorded around 40, 50 and 80 ns. NANPDB6380, NANPDB135 and NANPDB2246 were observed to be stable over the 100 ns period with average R_g values of 1.8 nm, 1.775 nm and 1.82 nm, respectively. Ouabain had an average of 1.77 nm and was very stable throughout the 100 ns period, though slight fluctuations were observed around 19 and 74 ns. Nilotinib experienced a decrease in R_g from the beginning till about 20 ns and then stabilized with an average R_g of 1.8 nm. Clomiphene recorded the highest R_g with average of 1.825 nm till about 40 ns. Thereafter, experienced a significant fall in R_g to an average of 1.8 nm, which was maintained till the end of the 100 ns period.

To further evaluate the stability of the docked complexes, root-mean-square deviation (RMSD) plots were generated from the MD simulation of the complexes. The backbones of the six complexes were stable after a gradual rise around 10 ns (Fig. 5B). The VP24-ZINC000003594643 complex rose from 0.2 to 0.25 nm around 10 ns, and then remained stable till about 20 ns where a small rise to about 0.275 nm was observed. The VP24-ZINC000003594643 complex maintained an average RMSD of 0.275 with a deviation of 0.075 nm. The RMSD of VP24-ZINC000095486008 complex also rose from 0.2 nm to an average of 0.25 nm around 10 ns, then remained stable till the end of the 100 ns. ZINC00009548670 experienced fluctuations in RMSD till about 50 ns where it stabilized with an average RMSD of about 0.35 nm. The VP24-NANPDB6380 complex experienced a gradual rise until 20 ns where it achieved stability and rose steadily after 35 ns till about 85 ns, where it began to fall, maintaining an average RMSD of 0.25 nm. The VP24-NANPDB135 complex rose from 0.2 nm to 0.325 nm around 10 ns, then experienced fluctuations till the end of the 100 ns period, maintaining an average RMSD of 0.3 nm with a deviation of 0.5 nm. The VP24-NANPDB2246 complex showed stability after rising to 0.28 nm around 10 ns till about 40 ns, where it experienced a fall in RMSD till about 70 ns, then maintained stability till the end of the period. The VP24-NANPDB2246 complex maintained an average RMSD of 0.25 nm. VP24-ZINC000095486008 complex showed more stability as compared to the other five selected compounds (Fig. 5B). VP24-clomiphene and VP24-nilotinib complexes demonstrated stability throughout the 100 ns simulation period with an average of about 0.275 and 0.25 nm, respectively. The VP24-ouabain complex however demonstrated the least stability among all the nine complexes. It experienced a gradual rise in RMSD from 0 to about 40 ns, stabilized with an average of about 0.29 nm till about 80 ns, and then rose to about 0.4 nm.

To investigate the flexibility of key protein residues that contributed to the structural fluctuations of the complex, the root-mean-square fluctuations (RMSFs) of each residue were assessed. Nilotinib and Clomiphene complexes showed similar trends and had the highest RMSFs implying greater fluctuations (Fig. 5C). Also, ZINC000003594643, ZINC000095486070 and Ouabain complexes had similar trends. All nine compounds caused some degree of fluctuations at the same regions of the VP24 protein. Among the selected compounds, ZINC000003594643 caused the highest fluctuations to the VP24 protein. From the RMSF graph (Fig. 5C), the large fluctuated residues were shown in the regions from residue index 58 to 72, 170 to 175, 202 to 218 and 220 to 231; which were all very close to the binding pocket 1. It is probable predicted key residues Lys39, Gly44, Ile45, Glu46 and Asp48 contributed to the fluctuations observed from residues 58 to 72; while Phe147, Ser151, Arg154, Ser155, Ile157, Leu158 and Ile161 contributed to the fluctuations in the region with residue index 170 to 175. Additionally, Ser225, Thr226, Ala229, Phe230 and Thr231 could be responsible for the fluctuations from residues 202 to 218 and 220 to 231.

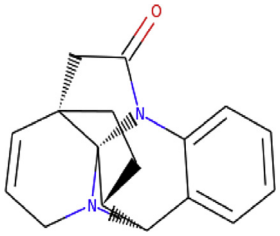
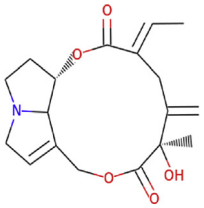
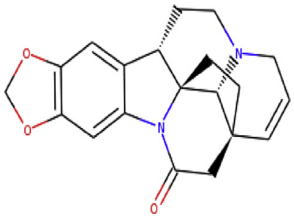
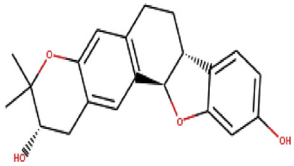
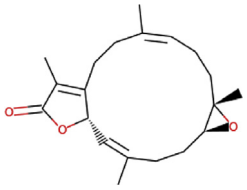
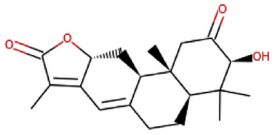
3.12. Evaluation of putative leads using MM-PBSA computations

The MM-PBSA approach was employed in determining the binding free energies of the complexes. The MM-PBSA calculations showed that Nilotinib demonstrated the minimum free binding energy of −198.039 kJ/mol (Table 5). The other 2 standard leads, Ouabain and Clomiphene had binding free energies of −74.885 and −25.428 kJ/mol, respectively. The six selected compounds NANPDB135, ZINC000003594643, ZINC000095486008, ZINC000095486070, NANPDB2246 and NANPDB6380 (Table 6) also had binding free energies of −152.832, −120.673, −114.650, −17.901, 7.575 and 99.679 kJ/mol, respectively (Table 5). Even though, ZINC000095486070 was predicted by AutoDock Vina to possess the highest binding affinity to the VP24 protein (−9.7 kcal/mol), it had a high binding free energy of −17.901 kJ/mol from the MM-PBSA computations (Table 5), thereby limiting its lead-likeness. Since clomiphene, a known inhibitor of EBOV VP24, also had high binding free energy of −25.428 kJ/mol (Table 5), it is premature to eliminate ZINC000095486070 from future experimental characterization. In some instances, compounds with high binding free energies have been reported to be active against receptors due to their very low electrostatic energies and very high polar energies [119]. Though, ZINC000095486070 showed high binding free energy, it demonstrated high polar solvation and low electrostatic energies of 337.407 and −220.892 kJ/mol, respectively. It was also observed that clomiphene

Table 5
Table showing binding energies and the contributing energy terms of the VP24-ligand complexes from MM-PBSA calculation. The values are presented in average ± standard deviations in kJ/mol.

	van der Waal energy (kJ/mol)	Electrostatic energy (kJ/mol)	Polar solvation energy (kJ/mol)	SASA energy (kJ/mol)	Binding energy (kJ/mol)
ZINC000095486070	−122.771 ± 35.866	−220.892 ± 75.060	337.407 ± 94.599	−11.645 ± 3.522	−17.901 ± 30.178
Ouabain	−108.238 ± 83.704	−27.658 ± 29.012	70.566 ± 55.666	−9.555 ± 7.801	−74.885 ± 77.527
NANPDB6380	−89.973 ± 16.053	−354.022 ± 50.858	554.724 ± 50.466	−11.050 ± 1.186	99.679 ± 29.864
ZINC000003594643	−145.636 ± 54.233	−4.704 ± 5.018	40.081 ± 21.152	−10.413 ± 4.016	−120.673 ± 53.312
ZINC000095486008	−141.705 ± 48.586	−14.432 ± 12.249	52.986 ± 27.102	−11.500 ± 4.523	−114.650 ± 46.087
Nilotinib	−226.512 ± 105.729	−4.327 ± 6.629	50.759 ± 31.034	−17.959 ± 7.615	−198.039 ± 96.467
NANPDB135	−172.127 ± 20.645	−25.583 ± 11.903	59.879 ± 14.508	−15.002 ± 1.543	−152.832 ± 21.711
NANPDB2246	−95.051 ± 19.910	−236.240 ± 53.664	348.057 ± 80.041	−9.191 ± 1.828	7.575 ± 28.358
Clomiphene	−138.665 ± 75.476	−130.299 ± 92.886	256.393 ± 134.840	−12.857 ± 7.352	−25.428 ± 50.965

Table 6
List of selected compounds with their two–dimensional structures obtained from ZINC database and IUPAC names generated using Marvin suite (http://www.chemaxon.com/).

Ligand ID	Common/IUPAC Names	Two-Dimensional Structure
ZINC000095486070	(1S,8R,16S,17R)-5,15-diazahexacyclo[13.4.2.$0^{1,16}$.$0^{5,16}$.$0^{8,17}$.$0^{9,14}$]henicosa-2,9,11,13-tetraen-21-one	
NANPDB6380	spartioidine	
ZINC000003594643	(1S,14S,18R,19R)-8,10-dioxa-4,17-diazaheptacyclo[15.4.3.$0^{1,18}$.$0^{4,19}$.$0^{5,13}$.$0^{7,11}$.$0^{14,19}$]tetracosa-5,7(11),12,22-tetraen-3-one	
ZINC000095486008	(2R,10R,18S)-17,17-dimethyl-3,16- dioxapentacyclo[11.8.0.0^{2,10}.$0^{4,9}$.$0^{15,20}$]henicosa-1(13),4,6,8,14,20-hexaene-6,18-diol	
NANPDB135	sarcophine	
NANPDB2246	helioscopinolide C	

had high binding free energy of −25.428 kJ/mol and low electrostatic energy of −130.299 kJ/mol. However, NANPDB2246 and NANPDB6380 were predicted to have unfavourable affinity to the protein with positive binding free energies of 7.575 and 99.679 kJ/mol, respectively. Therefore, NANPDB2246 and NANPDB6380 were not considered for downstream analysis, leaving NANPDB135, ZINC000003594643, ZINC000095486008 and ZINC000095486070 as potential inhibitory molecules.

The energy contribution of each residue was investigated via MM-PBSA decomposition to determine the key residues necessary for the binding of VP24. Residues that contributed energies > 5.0 or < −5.0 are worth considering as critical residues for the binding of a ligand to a protein [120]. Generally, few residues contributed energies > 5.0 or < −5.0 kJ/mol. Lys39 and Phe230 were observed to contribute energies less than −5.0 kJ/mol in most of the complexes. For the VP24-ouabain complex (Supplementary Fig. 9E), only Phe230 contributed energy less than −5.0 kJ/mol. Ile45, Phe47 and Leu158 also contributed energies beyond the ± 5.0 kJ/mol threshold. However, most of the residues contributed energies beyond the ± 5.0 kJ/mol threshold in clomiphene (Supplementary Fig. 9C), NANPDB6380 (Supplementary Fig. 9F), NANPDB135 (Supplementary Fig. 9G) and ZINC000095486070 complexes (Supplementary Fig. 9A). All the residues that contributed significantly to the binding free energy were found to be residues within pocket 1 (Table 1 and Supplementary Table 2). The MM-PBSA plot for the binding energy contribution of each residue in the VP24-ZINC000095486008 complex is shown in Fig. 6. Lys39 and Arg154 contributed positive energies of 9.9374 and 2.3993 kJ/mol, respectively. Glu46, Ser151, Ser155 and Phe230 contributed negative energies of −3.8916, −0.6065, −1.4244 and −8.2213 kJ/mol, respectively. Ile45, Phe47 and Leu158 were observed to contribute very low binding free energies of −7.9552, −6.4028 and −6.9203 kJ/mol, respectively.

4. Implications and future outlook

Even though, natural products have been shown as potent

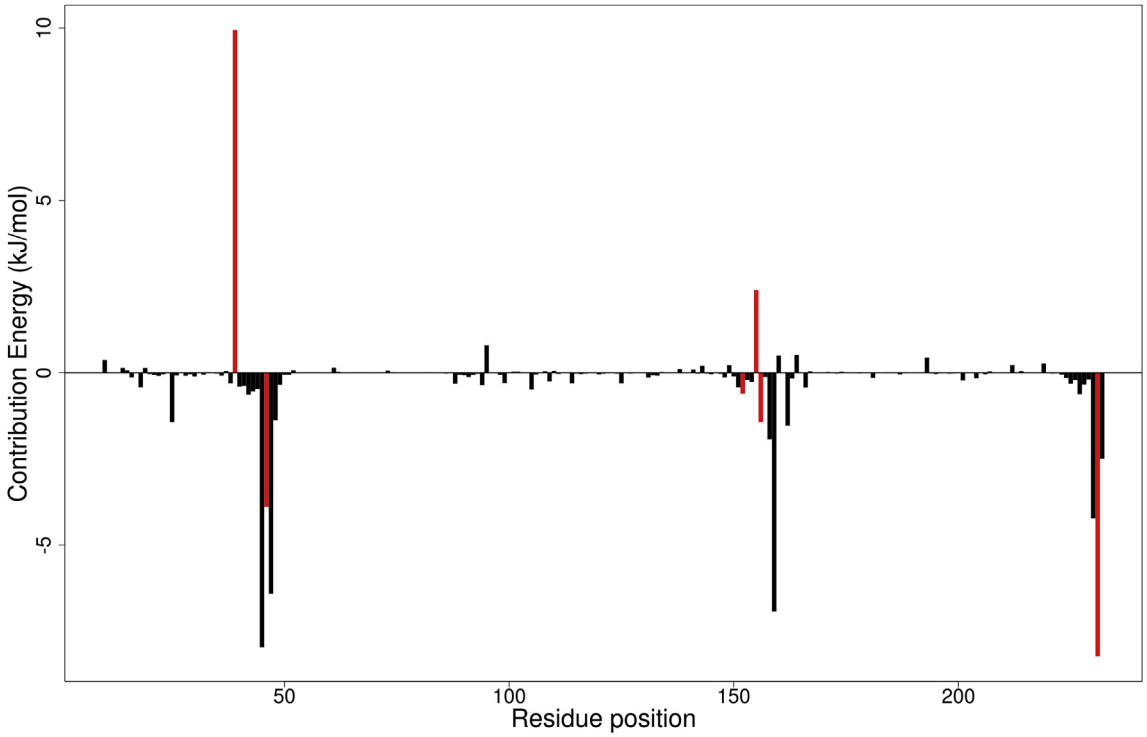


Fig. 6. Molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) plot of binding free energy contribution per residue of VP24-ZINC000095486008 complex. Fluctuations by predicted critical residues are shown in red. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

therapeutic molecules, leveraging African natural product databases to aid in unravelling novel anti-Ebola molecules is underutilized. The reported work complements current efforts geared towards identification of EBOV inhibitors [23,32,38–41]. These molecules are indispensable in the race against finding a cure as well as eradication of EBOV. The downside of the study emanates from the *in silico* methods which therefore warrant experimental validation. The study carefully predicted drug-like and lead-like small molecules, which was reinforced with machine-learning based anti-viral and anti-Ebola activity predictions. The predicted potential inhibitors could be used as baseline structures for optimisation. The fragments could form basis for de novo design of drug-like compounds for synthesis [106,121–123]. Also, these compounds could be tested both *in vitro* and *in vivo* to ascertain their anti-Ebola activity for further characterisation as potent future inhibitory molecules. The study highlights the potential of repurposing existing compounds as potential inhibitory molecules of EBOV. Therefore, making these predicted potential inhibitors available to the scientific community could invigorate the momentum towards the search for effective Ebola drugs. Also, the attrition rates of candidate drugs downstream the drug development pipeline necessitate the use of computational techniques to identify compounds with less propensity to fall off in development.

5. Conclusion

The plethora of available natural products need to be repurposed as alternative sources for identifying potent EBOV VP24 inhibitors. In summary, this study combined pharmacoinformatics with Bayesian models to identify potential EBOV inhibitors from the African flora. The study led to the identification of potential bioactive molecules comprising ZINC000095486070, ZINC000003594643, ZINC000095486008 and NANPDB135. The promising molecules exhibited high binding affinities and formed intermolecular bonding with critical residues of the EBOV VP24. ZINC000095486070 had the least binding energy of −9.7 kcal/mol, which was higher in binding affinity than all the six known EBOV VP24 inhibitors. The AUC of the ROC curve used to assess the reliability of the molecular docking was 0.77, which was considered acceptable. Also, molecular dynamics simulations including MM-PBSA calculations and IFD reinforced the potential inhibition of EBOV VP24 by the molecules. The compounds were physicochemically profiled to be druglike and ligand efficiency-based metrics indicated that they were plausible templates for fragment-based design of future EBOV inhibitors. Also, the compounds were predicted to have negligible toxicity without any potential safety concerns. Therefore, these molecules provide a valuable starting point as fragments for the design of new synthetic derivatives with enhanced activity, thus paving the way for new and exciting therapy options for EVD. The predicted compounds are potential inhibitory molecules worthy of experimental evaluation to corroborate their bioactivity.

Declarations

None.

Competing interests

There is no conflict of interest.

Funding

No funding was secured for this work.

Authors' contributions

None.

Acknowledgements

We are grateful to the West African Centre for Cell Biology of Infectious Pathogens (WACCBIP) at University of Ghana for the use of

Zuputo, a Dell EMC high performance computing cluster.

List of abbreviations

ADMET	Absorption, Distribution, Metabolism, Excretion and Toxicity
EVD	Ebola Virus Disease
EBOV	Ebola Virus
3D	Three-dimensional
ROC	Receiver Operating Characteristics
PASS	Prediction of activity spectra for substances
AfroDB	Dataset of natural products from African flora
RNA	Ribonucleic acid
NP	Nucleoproteins
GP	Glycoprotein
KPNA	Karyopherin
VLPs	Virus-Like Particles
FDA	Food and Drug Authority
CADD	Computer-aided drug design
SDF	Structure Data File
PDBQT	Protein Data Bank, Partial Charge (Q), & Atom Type (T)
RCSB	Research Collaboratory for Structural Bioinformatics
PDB	Protein Data Bank
RMSD	Root Mean Square Deviation
CYP	Cytochrome P450
TPSA	Topological Polar Surface Area
XLOGP3	an atomistic method including corrective factors and knowledge-based library
LogP	logarithm of the octan-1-ol/water partition coefficient
WHO	World Health Organization
I-TASSER	Iterative threading assembly refinement
IDs	Identification
SDF	Structure Data File
WDI	World Drug Index
OPLS/AA	Optimized Potentials for Liquid Simulations/All Atom
CASTp	Computed Atlas of Structure Surface Topography of proteins
DUD-E	Directory of useful decoys and enhanced
AUC	Area under the curve
MM-PBSA	Molecular mechanics Poisson-Boltzmann surface area
IFD	Induced fit docking
CNS	Central Nervous System
P-gp	P-glycoprotein
GI	Gastrointestinal
HIA	Human intestinal absorption
BBB	Blood Brain Barrier
Pa	Probable activity
Pi	Probable inactivity
MMDS	Mobile Molecular Data Sheet

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.compbiomed.2019.103414.

References

[1] H. Feldmann, T.W. Geisbert, Ebola haemorrhagic fever, The Lancet 377 (9768) (2011) 849–862 https://doi.org/10.1016/S0140-6736(10)60667-8.

[2] S. Gumusova, M. Sunbul, H. Leblebicioglu, Ebola virus disease and the veterinary perspective, Ann. Clin. Microbiol. Antimicrob. (2015), https://doi.org/10.1186/s12941-015-0089-x.

[3] T.W. Geisbert, H.A. Young, P.B. Jahrling, K.J. Davis, T. Larsen, E. Kagan, L.E. Hensley, Pathogenesis of ebola hemorrhagic fever in primate models, Am. J. Pathol. 163 (6) (2003) 2371–2382 https://doi.org/10.1016/S0002-9440(10)63592-4.

[4] P. Formenty, Ebola virus disease, Emerg. Infect. Dis. 328 (June) (2014) 121–134 https://doi.org/10.1016/B978-0-12-416975-3.00009-1.

[5] C. Salata, A. Calistri, G. Alvisi, M. Celestino, C. Parolin, G. Palù, Ebola virus entry: from molecular characterization to drug discovery, Viruses 11 (3) (2019), https://doi.org/10.3390/v11030274.

[6] WHO, Ebola Virus Disease, WHO, 2018 Retrieved from http://www.who.int/mediacentre/factsheets/fs103/en/.

[7] WHO, Ebola situation reports: democratic Republic of the Congo, (2019) Retrieved May 3, 2019 https://www.who.int/ebola/situation-reports/drc-2018/en/.

[8] T. Goldstein, S.J. Anthony, A. Gbakima, B.H. Bird, J. Bangura, A. Tremeau-Bravard, et al., The discovery of Bombali virus adds further support for bats as hosts of ebolaviruses, Nature Microbiology (2018), https://doi.org/10.1038/s41564-018-0227-2.

[9] X. Lou Yang, C.W. Tan, D.E. Anderson, R. Di Jiang, B. Li, W. Zhang, et al., Characterization of a filovirus (Měnglà virus) from Rousettus bats in China, Nat. Microbiol. (2019), https://doi.org/10.1038/s41564-018-0328-y.

[10] L. Falasca, C. Agrati, N. Petrosillo, A. Di Caro, M.R. Capobianchi, G. Ippolito, M. Piacentini, Molecular mechanisms of Ebola virus pathogenesis: focus on cell death, Cell Death Differ. 22 (8) (2015) 1250–1259 https://doi.org/10.1038/cdd.2015.67.

[11] Z. Han, H. Boshra, J.O. Sunyer, S.H. Zwiers, J. Paragas, R.N. Harty, Biochemical and functional characterization of the ebola virus VP24 protein: implications for a role in virus assembly and budding, J. Virol. 77 (3) (2003) 1793 https://doi.org/10.1128/JVI.77.3.1793.

[12] S.P. Reid, L.W. Leung, A.L. Hartman, O. Martinez, M.L. Shaw, C. Carbonnelle, et al., Ebola virus VP24 binds Karyopherin 1 and blocks STAT1 nuclear accumulation, J. Virol. 80 (11) (2006) 5156–5167 https://doi.org/10.1128/JVI.02349-05.

[13] S.P. Reid, C. Valmas, O. Martinez, F.M. Sanchez, C.F. Basler, Ebola virus VP24 proteins inhibit the interaction of NPI-1 subfamily Karyopherin proteins with activated STAT1, J. Virol. 81 (24) (2007) 13469–13477 https://doi.org/10.1128/JVI.01097-07.

[14] A.P.P. Zhang, Z.A. Bornholdt, T. Liu, D.M. Abelson, D.E. Lee, S. Li, et al., The ebola virus interferon antagonist VP24 directly binds STAT1 and has a novel, pyramidal fold, PLoS Pathog. 8 (2) (2012), https://doi.org/10.1371/journal.ppat.1002550.

[15] A. Watt, F. Moukambi, L. Banadyga, A. Groseth, J. Callison, A. Herwig, et al., A novel life cycle modeling system for ebola virus shows a genome length-dependent role of VP24 in virus infectivity, J. Virol. 88 (18) (2014) 10511–10524 https://doi.org/10.1128/JVI.01272-14.

[16] L. Banadyga, T. Hoenen, X. Ambroggio, E. Dunham, A. Groseth, H. Ebihara, Ebola virus VP24 interacts with NP to facilitate nucleocapsid assembly and genome packaging, Sci. Rep. 7 (1) (2017), https://doi.org/10.1038/s41598-017-08167-8.

[17] T. Noda, P. Halfmann, H. Sagara, Y. Kawaoka, Regions in ebola virus VP24 that are important for nucleocapsid formation, J. Infect. Dis. 196 (s2) (2007) S247–S250 https://doi.org/10.1086/520596.

[18] S. Yuan, Possible FDA-approved drugs to treat Ebola virus infection, Infect. Dis. Poverty 4 (1) (2015) 23 https://doi.org/10.1186/s40249-015-0055-z.

[19] I. García-Dorival, W. Wu, S. Dowall, S. Armstrong, O. Touzelet, J. Wastling, et al., Elucidation of the ebola virus VP24 cellular interactome and disruption of virus biology through targeted inhibition of host-cell protein function, J. Proteome Res. 13 (11) (2014) 5120–5135 https://doi.org/10.1021/pr500556d.

[20] L.M. Johansen, J.M. Brannan, S.E. Delos, C.J. Shoemaker, A. Stossel, C. Lear, et al., FDA-approved selective estrogen receptor modulators inhibit ebola virus infection, Sci. Transl. Med. 5 (190) (2013), https://doi.org/10.1126/scitranslmed.3005471.

[21] M. Balmith, M. Faya, M.E.S. Soliman, Ebola virus: a gap in drug design and discovery - experimental and computational perspective, Chem. Biol. Drug Des. 89 (3) (2017) 297–308 https://doi.org/10.1111/cbdd.12870.

[22] E. Picazo, F. Giordanetto, Small molecule inhibitors of ebola virus infection, Drug Discov. Today 20 (2) (2015) 277–286 https://doi.org/10.1016/j.drudis.2014.12.010.

[23] A.S. Setlur, S.Y. Naik, S. Skariyachan, Herbal lead as ideal bioactive compounds against probable drug targets of ebola virus in comparison with known chemical analogue: a computational drug discovery perspective, Interdiscip. Sci. Comput. Life Sci. 9 (2) (2017) 254–277 https://doi.org/10.1007/s12539-016-0149-8.

[24] R. Taylor, P. Kotian, T. Warren, R. Panchal, S. Bavari, J. Julander, et al., BCX4430 - a broad-spectrum antiviral adenosine nucleoside analog under development for the treatment of Ebola virus disease, Journal of Infection and Public Health 9 (3) (2016) 220–226 https://doi.org/10.1016/j.jiph.2016.04.002.

[25] A.P. Galvani, M.L. Ndeffo-Mbah, N. Wenzel, J.E. Childs, Ebola vaccination: if not now, when? Ann. Intern. Med. 161 (10) (2014) 749–750 https://doi.org/10.7326/M14-1904.

[26] D. Malvy, A.K. McElroy, H. de Clerck, S. Günther, J. van Griensven, Ebola virus disease, The Lancet (2019), https://doi.org/10.1016/S0140-6736(18)33132-5.

[27] J. Kouznetsova, W. Sun, C. Martínez-Romero, G. Tawa, P. Shinn, C.Z. Chen, et al., Identification of 53 compounds that block Ebola virus-like particle entry via a repurposing screen of approved drugs, Emerg. Microb. Infect. 3 (12) (2014) e84–e84 https://doi.org/10.1038/emi.2014.88.

[28] S.D. Dowall, K. Bewley, R.J. Watson, S.S. Vasan, C. Ghosh, M.M. Konai, et al., Antiviral screening of multiple compounds against ebola virus, Viruses 8 (11) (2016), https://doi.org/10.3390/v8110277.

[29] B. Shen, A new golden age of natural products drug discovery, Cell 163 (6) (2015) 1297–1300 https://doi.org/10.1016/j.cell.2015.11.031.

[30] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the 30 years from 1981 to 2010, J. Nat. Prod. 75 (3) (2012) 311–335 https://doi.org/10.1021/np200906s.

[31] R. Nakrumpai, Virtual screening for natural products with potential Inhibitory Effect on Ebola Virus Glycoprotein, 5 (2) (2017) 45–51.

[32] S. Baikerikar, Curcumin and natural derivatives inhibit Ebola viral proteins: an in silico approach, Pharmacogn. Res. 9 (5) (2017) 15 https://doi.org/10.4103/pr.pr_30_17.

[33] S. Nabavi, R. Thiagarajan, L. Rastrelli, M. Daglia, E. Sobarzo-Sanchez,

H. Alinezhad, S. Nabavi, Curcumin: a natural product for diabetes and its complications, Curr. Top. Med. Chem. 15 (23) (2015) 2445–2455 https://doi.org/10.2174/1568026615666150619142519.
[34] U.S.F. Tambunan, M.A.F. Nasution, Identification of novel Ebola virus (EBOV) VP24 inhibitor from Indonesian natural products through in silico drug design approach, AIP Conference Proceedings, vol. 1862, 2017 https://doi.org/10.1063/1.4991195.
[35] A. Jain, Computer aided drug design, Journal of Physics: Conference Series, vol. 884, 2017 https://doi.org/10.1088/1742-6596/884/1/012072.
[36] S.J.Y. Macalino, V. Gosu, S. Hong, S. Choi, Role of Computer-Aided Drug Design in Modern Drug Discovery, Archives of Pharmacal Research, 2015, https://doi.org/10.1007/s12272-015-0640-5.
[37] C.M. Song, S.J. Lim, J.C. Tong, Recent advances in computer-aided drug design. Briefings in Bioinformatics, https://doi.org/10.1093/bib/bbp023, (2009).
[38] M.A. Islam, T.S. Pillay, Pharmacoinformatics-based identification of chemically active molecules against Ebola virus, J. Biomol. Struct. Dyn. (2018), https://doi.org/10.1080/07391102.2018.1544509.
[39] M.U. Mirza, N. Ikram, Integrated computational approach for virtual hit identification against ebola viral proteins VP35 and VP40, Int. J. Mol. Sci. 17 (11) (2016), https://doi.org/10.3390/ijms17111748.
[40] J.X. Ren, R.T. Zhang, H. Zhang, X.S. Cao, L.K. Liu, Y. Xie, Identification of novel VP35 inhibitors: virtual screening driven new scaffolds, Biomed. Pharmacother. (2016), https://doi.org/10.1016/j.biopha.2016.09.034.
[41] Usman Sumo Friend Tambunan, Ebola viral protein 24 (VP24) inhibitor discovery BY IN silico fragment-based design, Int. J. Geomate (2018), https://doi.org/10.21660/2018.49.3534.
[42] G.M. Cragg, D.J. Newman, Natural products: a continuing source of novel drug leads, Biochim. Biophys Acta - Gen. Subj. (2013), https://doi.org/10.1016/j.bbagen.2013.02.008.
[43] D.J. Newman, G.M. Cragg, Natural products as sources of new drugs over the last 25 years, J. Nat. Prod. (2007), https://doi.org/10.1021/np068054v.
[44] P.A. Onguéné, C.V. Simoben, G.W. Fotso, K. Andrae-Marobela, S.A. Khalid, B.T. Ngadjui, et al., In silico toxicity profiling of natural product compound libraries from African flora with anti-malarial and anti-HIV properties, Comput. Biol. Chem. 72 (2018) 136–149 https://doi.org/10.1016/j.compbiolchem.2017.12.002.
[45] E.O. Farombi, African indigenous plants with chemotherapeutic potentials and biotechnological approach to the production of bioactive prophylactic agents, Afr. J. Biotechnol. 2 (2003) 662–671 (12 Cited December 31, 2003), https://doi.org/10.5897/AJB2003.000-1122.
[46] F. Ntie-Kang, K.K. Telukunta, K. Döring, C.V. Simoben, A.F.A. Moumbock, Y.I. Malange, et al., NANPDB: a resource for natural products from northern African sources, J. Nat. Prod. (2017), https://doi.org/10.1021/acs.jnatprod.7b00283.
[47] F. Ntie-Kang, D. Zofou, S.B. Babiaka, R. Meudom, M. Scharfe, L.L. Lifongo, et al., AfroDb: a select highly potent and diverse natural product library from african medicinal plants, PLoS One 8 (10) (2013) e78085https://doi.org/10.1371/journal.pone.0078085.
[48] P.W. Rose, A. Prlić, A. Altunkaya, C. Bi, A.R. Bradley, C.H. Christie, et al., The RCSB protein data bank: integrative view of protein, gene and 3D structural information, Nucleic Acids Res. 45 (D1) (2017) D271–D281 https://doi.org/10.1093/nar/gkw1000.
[49] S. Kim, P.A. Thiessen, E.E. Bolton, J. Chen, G. Fu, A. Gindulyte, et al., PubChem substance and compound databases, Nucleic Acids Res. 44 (D1) (2016) D1202–D1213 https://doi.org/10.1093/nar/gkv951.
[50] I. Doytchinova, M. Atanasova, I. Valkova, G. Stavrakov, I. Philipova, Z. Zhivkova, et al., Novel hits for acetylcholinesterase inhibition derived by docking-based screening on ZINC database, J. Enzym. Inhib. Med. Chem. (2018), https://doi.org/10.1080/14756366.2018.1458031.
[51] S. Alam, F. Khan, 3D-QSAR studies on Maslinic acid analogs for Anticancer activity against Breast Cancer cell line MCF-7, Sci. Rep. 7 (1) (2017) 1–13 https://doi.org/10.1038/s41598-017-06131-0.
[52] S. Dallakyan, A.J. Olson, Small-molecule library screening by docking with PyRx, Methods in Molecular Biology, vol. 1263, 2015, pp. 243–250 https://doi.org/10.1007/978-1-4939-2269-7_19.
[53] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A.E. Mark, H.J.C. Berendsen, GROMACS: fast, flexible, and free, J. Comput. Chem. (2005), https://doi.org/10.1002/jcc.20291.
[54] T.A. Binkowski, S. Naghibzadeh, J. Liang, CASTp: computed atlas of surface topography of proteins, Nucleic Acids Res. 31 (13) (2003) 3352–3355 https://doi.org/10.1093/nar/gkg512.
[55] J. Dundas, Z. Ouyang, J. Tseng, A. Binkowski, Y. Turpaz, J. Liang, CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues, Nucleic Acids Res. 34 (2006) W116–W118 Web Server issue https://doi.org/10.1093/nar/gkl282.
[56] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera - a visualization system for exploratory research and analysis, J. Comput. Chem. 25 (13) (2004) 1605–1612 https://doi.org/10.1002/jcc.20084.
[57] O. Trott, A.J. Olson, AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, EfficientOptimization, and multithreading, J. Comput. Chem. 31 (2) (2010) 455–461 https://doi.org/10.1002/jcc.
[58] K. Rother, Introduction to PyMOL, Methods Mol. Biol. Clifton Nj 635 (8) (2005) 0–32 https://doi.org/10.1213/ANE.0b013e3181e9c3f3.
[59] D. Seeliger, B.L. De Groot, Ligand docking and binding site analysis with PyMOL and Autodock/Vina, J. Comput. Aided Mol. Des. 24 (5) (2010) 417–422 https://doi.org/10.1007/s10822-010-9352-6.
[60] M.M. Mysinger, M. Carchia, J.J. Irwin, B.K. Shoichet, Directory of useful decoys, enhanced (DUD-E): better ligands and decoys for better benchmarking, J. Med. Chem. 55 (14) (2012) 6582–6594 https://doi.org/10.1021/jm300687e.
[61] J. Shamsara, Correlation between virtual screening performance and binding site descriptors of protein targets, Int. J. Med. Chem. 2018 (2018) 1–10 https://doi.org/10.1155/2018/3829307.
[62] D. Goksuluk, S. Korkmaz, G. Zararsiz, A.E. Karaagaoglu, easyROC: an interactive web-tool for ROC curve analysis using R language environment, The R Journal 8/2 (December) (2016) 213–230.
[63] E.R. DeLong, D.M. DeLong, D.L. Clarke-Pearson, Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach, Biometrics 44 (3) (1988) 837–845.
[64] R.A. Laskowski, M.B. Swindells, LigPlot+: multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (10) (2011) 2778–2786 https://doi.org/10.1021/ci200227u.
[65] Z. Zhao, C. Martin, R. Fan, P.E. Bourne, L. Xie, Drug repurposing to target Ebola virus replication and virulence using structural systems pharmacology, BMC Bioinf. 17 (1) (2016) 1–12 https://doi.org/10.1186/s12859-016-0941-9.
[66] W. Sherman, T. Day, M.P. Jacobson, R.A. Friesner, R. Farid, Novel procedure for modeling ligand/receptor induced fit effects, J. Med. Chem. 49 (2) (2006) 534–553 https://doi.org/10.1021/jm050540c.
[67] S.E. Dobbins, V.I. Lesk, M.J.E. Sternberg, Insights into protein flexibility: the relationship between normal modes and conformational change upon protein-protein docking, Proc. Natl. Acad. Sci. 105 (30) (2008) 10390–10395 https://doi.org/10.1073/pnas.0802496105.
[68] B.K. Ho, D.A. Agard, Probing the flexibility of large conformational changes in protein structures through local perturbations, PLoS Comput. Biol. 5 (4) (2009) e1000343https://doi.org/10.1371/journal.pcbi.1000343.
[69] E.K. Tiburu, I. Issah, M. Darko, R.E. Armah-Sekum, S.O.A. Gyampo, N.K. Amoateng, et al., Investigating the conformation of S100β protein under physiological parameters using computational modeling: a clue for rational drug design, Open Biomed. Eng. J. 12 (1) (2018) 36–50 https://doi.org/10.2174/1874120701812010036.
[70] A. Daina, O. Michielin, V. Zoete, C.L. Brooks, R. Huang, SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules, Sci. Rep. 7 (2017) 42717 https://doi.org/10.1038/srep42717.
[71] S. Farr-Jones, S. Fox, High-Throughput ADMET screening: improving the efficiency of drug discovery, Decision Resources, John Wiley & Sons, Inc, Hoboken, NJ, USA, 2003, pp. 1–18 https://doi.org/10.1002/9780470041000.cedt054.
[72] T. Sander, J. Freyss, M. Von Korff, C. Rufener, DataWarrior: an open-source program for chemistry aware data visualization and analysis, J. Chem. Inf. Model. 55 (2) (2015) 460–473 https://doi.org/10.1021/ci500588j.
[73] A. Lagunin, A. Stepanchikova, D. Filimonov, V. Poroikov, PASS: prediction of activity spectra for biologically active substances, Bioinformatics 16 (8) (2000) 747–748.
[74] S. Parasuraman, Prediction of activity spectra for substances, J. Pharmacol. Pharmacother. 2 (1) (2011) 52–53 https://doi.org/10.4103/0976-500X.77119.
[75] S. Ekins, J.S. Freundlich, A.M. Clark, M. Anantpadma, R.A. Davey, P. Madrid, Machine learning models identify molecules active against the Ebola virus in vitro, F1000Research (2015), https://doi.org/10.12688/f1000research.7217.1.
[76] S. Ekins, J.S. Freundlich, R.C. Reynolds, Fusing dual-event data sets for mycobacterium tuberculosis machine learning models and their evaluation, J. Chem. Inf. Model. 53 (11) (2013) 3054–3063 https://doi.org/10.1021/ci400480s.
[77] M.A. Islam, T.S. Pillay, Identification of promising anti-DNA gyrase antibacterial compounds using de novo design, molecular docking and molecular dynamics studies, J. Biomol. Struct. Dyn. (2019), https://doi.org/10.1080/07391102.2019.1617785.
[78] P. Turner, XMGRACE, Version 5.1. 19, Center for Coastal and Land-Margin Research, Oregon Graduate Institute of Science and Technology, Beaverton, 2005https://doi.org/10.1163/_q3_SIM_00374.
[79] N.A. Baker, D. Sept, S. Joseph, M.J. Holst, J.A. McCammon, Electrostatics of nanosystems: application to microtubules and the ribosome, Proc. Natl. Acad. Sci. (2002), https://doi.org/10.1073/pnas.181342398.
[80] R. Kumari, R. Kumar, A. Lynn, G-mmpbsa -A GROMACS tool for high-throughput MM-PBSA calculations, J. Chem. Inf. Model. (2014), https://doi.org/10.1021/ci500020m.
[81] R Development Core Team, R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, 2016, https://doi.org/10.1007/978-3-540-74686-7.
[82] M.R. Edwards, B. Johnson, C.E. Mire, W. Xu, R.S. Shabman, L.N. Speller, et al., The Marburg virus VP24 protein interacts with Keap1 to activate the cytoprotective antioxidant response pathway, Cell Rep. 6 (6) (2014) 1017–1025 https://doi.org/10.1016/j.celrep.2014.01.043.
[83] R.A. Hammou, Y. Kasmi, K. Khataby, F.E. Laasri, S. Boughribil, M.M. Ennaji, Roles of VP35, VP40 and VP24 proteins of ebola virus in pathogenic and replication mechanisms, Ebola (2016), https://doi.org/10.5772/63830.
[84] A.P.P. Zhang, D.M. Abelson, Z.A. Bornholdt, T. Liu, V.L. Woods, E.O. Saphire, The ebolavirus VP24 interferon antagonist: know your enemy, Virulence (2012), https://doi.org/10.4161/viru.21302.
[85] R. Shah, P.K. Panda, P. Patel, H. Panchal, Pharmacophore based virtual screening and molecular docking studies of inherited compounds against Ebola virus receptor proteins, World J. Pharm. Pharm. Sci. 4 (5) (2015) 1268–1282 Retrieved from http://www.wjpps.com/download/article/1430392261.pdf.
[86] M. García, A. Cooper, W. Shi, W. Bornmann, R. Carrion, D. Kalman, G.J. Nabel,

Productive replication of Ebola virus is regulated by the c-Abl1 tyrosine kinase, Sci. Transl. Med. 4 (123) (2012) 123ra24 https://doi.org/10.1126/scitranslmed.3003500.
[87] E.A. Nelson, A.B. Barnes, R.D. Wiehle, G.K. Fontenot, T. Hoenen, J.M. White, Clomiphene and its isomers block Ebola virus particle entry and infection with similar potency: potential therapeutic implications, Viruses 8 (8) (2016), https://doi.org/10.3390/v8080206.
[88] M.W. Chang, W. Lindstrom, A.J. Olson, R.K. Belew, Analysis of HIV wild-type and mutant structures via in silico docking against diverse ligand libraries, J. Chem. Inf. Model. (2007), https://doi.org/10.1021/ci700044s.
[89] C.A. Lipinski, Drug-like properties and the causes of poor solubility and poor permeability, J. Pharmacol. Toxicol. Methods 44 (1) (2000) 235–249 https://doi.org/10.1016/S1056-8719(00)00107-6.
[90] D.F. Veber, S.R. Johnson, H. Cheng, B.R. Smith, K.W. Ward, K.D. Kopple, Molecular properties that influence the oral bioavailability of drug candidates, J. Med. Chem. 45 (2002) 2615–2623 https://doi.org/10.1021/jm020017n.
[91] C.A. Lipinski, F. Lombardo, B.W. Dominy, P.J. Feeney, Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings1PII of original article: S0169-409X(96)00423-1. The article was originally published in Advanced Drug Delivery Reviews 23 (1997) 3, Adv. Drug Deliv. Rev. 46 (1–3) (2001) 3–26 https://doi.org/10.1016/S0169-409X(00)00129-0.
[92] D. Jiao, P.A. Golubkov, T.A. Darden, P. Ren, Calculation of protein-ligand binding free energy by using a polarizable potential, Proc. Natl. Acad. Sci. 105 (17) (2008) 6290–6295 https://doi.org/10.1073/pnas.0711686105.
[93] C.J. Shoemaker, K.L. Schornberg, S.E. Delos, C. Scully, H. Pajouhesh, G.G. Olinger, et al., Multiple cationic amphiphiles induce a Niemann-pick C phenotype and inhibit ebola virus entry and infection, PLoS One 8 (2) (2013) e56265https://doi.org/10.1371/journal.pone.0056265.
[94] N. Triballeau, F. Acher, I. Brabet, J.-P. Pin, H.-O. Bertrand, Virtual screening workflow development guided by the "receiver operating characteristic" curve approach. Application to high-throughput docking on metabotropic glutamate receptor subtype 4, J. Med. Chem. 48 (7) (2005) 2534–47 https://doi.org/10.1021/jm049092j.
[95] J.V. Cruz, M.F.A. Neto, L.B. Silva, R. da S Ramos, J. da S Costa, D.S.B. Brasil, et al., Identification of novel protein kinase receptor type 2 inhibitors using pharmacophore and structure-based virtual screening, Molecules 23 (2) (2018) 453 https://doi.org/10.3390/molecules23020453.
[96] C. Suenderhauf, F. Hammann, J. Huwyler, Computational prediction of blood-brain barrier permeability using decision tree induction, Molecules 17 (9) (2012) 10429–10445 https://doi.org/10.3390/molecules170910429.
[97] J.H. Lin, M. Yamazaki, Role of P-glycoprotein in pharmacokinetics: clinical implications, Clin. Pharmacokinet. 42 (1) (2003) 59–98 https://doi.org/10.2165/00003088-200342010-00003.
[98] P. Anzenbacher, E. Anzenbacherová, Cytochromes P450 and metabolism of xenobiotics, Cell. Mol. Life Sci. 58 (5) (2001) 737–747 https://doi.org/10.1007/PL00000897.
[99] T. Lynch, A. Price, The effect of cytochrome P450 metabolism on drug response, interactions, and adverse effects, Am. Fam. Physician 76 (3) (2007) 391–396 https://doi.org/10.1046/j.1365-2125.1999.00073.x.
[100] P.G. Jamkhande, S.K. Pathan, S.J. Wadher, In silico PASS analysis and determination of antimycobacterial, antifungal, and antioxidant efficacies of maslinic acid in an extract rich in pentacyclic triterpenoids, International Journal of Mycobacteriology 5 (4) (2016) 417–425 https://doi.org/10.1016/j.ijmyco.2016.06.020.
[101] P.G. Jamkhande, S.R. Barde, Evaluation of anthelmintic activity and in silico PASS assisted prediction of Cordia dichotoma (Forst.) root extract, Ancient Sci. Life 34 (1) (2014) 39–43 https://doi.org/10.4103/0257-7941.150779.
[102] S. Kwofie, B. Dankwa, E. Odame, F. Agamah, Lady Doe, J. Teye, et al., In silico screening of isocitrate lyase for novel anti-buruli ulcer natural products originating from Africa, Molecules 23 (7) (2018) 1550 https://doi.org/10.3390/molecules23071550.
[103] S.K. Kwofie, B. Dankwa, K.S. Enninful, C. Adobor, E. Broni, A. Ntiamoah, M.D. Wilson, Molecular docking and dynamics simulation studies predict munc18b as a target of mycolactone: a plausible mechanism for granule exocytosis impairment in Buruli Ulcer Pathogenesis, Toxins 11 (3) (2019), https://doi.org/10.3390/toxins11030181.
[104] C.H. Reynolds, R.C. Reynolds, Group Additivity in ligand binding affinity: an alternative approach to ligand efficiency, J. Chem. Inf. Model. (2017), https://doi.org/10.1021/acs.jcim.7b00381.
[105] A.L. Hopkins, C.R. Groom, A. Alex, Ligand efficiency: a useful metric for lead selection, Drug Discov. Today (2004), https://doi.org/10.1016/S1359-6446(04)03069-7.
[106] S. Schultes, C. De Graaf, E.E.J. Haaksma, I.J.P. De Esch, R. Leurs, O. Krämer, Ligand efficiency as a guide in fragment hit selection and optimization, Drug Discov. Today Technol. (2010), https://doi.org/10.1016/j.ddtec.2010.11.003.
[107] Y.Y. Ke, M.S. Coumar, H.Y. Shiao, W.C. Wang, C.W. Chen, J.S. Song, et al., Ligand efficiency based approach for efficient virtual screening of compound libraries, Eur. J. Med. Chem. (2014), https://doi.org/10.1016/j.ejmech.2014.06.029.
[108] C.H. Reynolds, B.A. Tounge, S.D. Bembenek, Ligand binding efficiency: trends, physical basis, and implications, J. Med. Chem. (2008), https://doi.org/10.1021/jm701255b.
[109] G.M. Keserü, G.M. Makara, The influence of lead discovery strategies on the properties of drug candidates, Nat. Rev. Drug Discov. (2009), https://doi.org/10.1038/nrd2796.
[110] H.A. Carlson, Protein flexibility and drug design: how to hit a moving target, Curr. Opin. Chem. Biol. 6 (4) (2002) 447–452 https://doi.org/10.1016/S1367-5931(02)00341-1.
[111] M. Fischer, R.G. Coleman, J.S. Fraser, B.K. Shoichet, Incorporation of protein flexibility and conformational energy penalties in docking screens to improve ligand discovery, Nat. Chem. 6 (7) (2014) 575–583 https://doi.org/10.1038/nchem.1954.
[112] S.F. Sousa, A.J.M. Ribeiro, J.T.S. Coimbra, R.P.P. Neves, S.A. Martins, N.S.H.N. Moorthy, et al., Protein-ligand docking in the new millennium – a retrospective of 10 Years in the field, Curr. Med. Chem. 20 (18) (2013) 2296–2314 https://doi.org/10.2174/0929867311320180002.
[113] H. Zhong, L.M. Tran, J.L. Stang, Induced-fit docking studies of the active and inactive states of protein tyrosine kinases, J. Mol. Graph. Model. 28 (4) (2009) 336–346 https://doi.org/10.1016/j.jmgm.2009.08.012.
[114] K.H. Deokar, H.P. Barch, J.K. Buolamwini, Homology modeling of human concentrative nucleoside transporters (hCNTs) and validation by virtual screening and experimental testing to identify novel hCNT1 inhibitors, Drug Des 6 (1) (2017), https://doi.org/10.4172/2169-0138.1000146.
[115] R.A. Friesner, J.L. Banks, R.B. Murphy, T.A. Halgren, J.J. Klicic, D.T. Mainz, et al., Glide: a new approach for rapid, accurate docking and scoring. 1. Method and assessment of docking accuracy, J. Med. Chem. 47 (7) (2004) 1739–1749 https://doi.org/10.1021/jm0306430.
[116] R.A. Friesner, R.B. Murphy, M.P. Repasky, L.L. Frye, J.R. Greenwood, T.A. Halgren, et al., Extra precision glide: docking and scoring incorporating a model of hydrophobic enclosure for Protein − Ligand complexes, J. Med. Chem. 49 (21) (2006) 6177–6196 https://doi.org/10.1021/jm051256o.
[117] H.J. Luo, J.Z. Wang, N.Y. Huang, W.Q. Deng, K. Zou, Induced-fit docking and virtual screening for 8-hydroxy-3-methoxy- 5H-pyrido [2,1-c] pyrazin-5-one derivatives as inducible nitric oxide synthase inhibitors, J. Chem. Pharm. Res. 6 (3) (2014) 1187–1194.
[118] J.L. Medina-Franco, O. Méndez-Lucio, J. Yoo, Rationalization of activity cliffs of a sulfonamide inhibitor of DNA methyltransferases with induced-fit docking, Int. J. Mol. Sci. 15 (2) (2014) 3253–3261 https://doi.org/10.3390/ijms15023253.
[119] A. Gupta, N. Chaudhary, P. Aparoy, MM-PBSA and per-residue decomposition energy studies on 7-Phenyl-imidazoquinolin-4(5H)-one derivatives: identification of crucial site points at microsomal prostaglandin E synthase-1 (mPGES-1) active site, Int. J. Biol. Macromol. (2018), https://doi.org/10.1016/j.ijbiomac.2018.07.050.
[120] S.K. Kwofie, K.S. Enninful, J.A. Yussif, L.A. Asante, M. Adjei, K. Kan-Dapaah, et al., Molecular informatics studies of the iron-dependent regulator (ideR) reveal potential novel anti-mycobacterium ulcerans natural product-derived compounds, Molecules (2019), https://doi.org/10.3390/molecules24122299.
[121] B.J. Davis, S.D. Roughley, Fragment-based lead discovery, Annual Reports in Medicinal Chemistry, 2017 https://doi.org/10.1016/bs.armc.2017.07.002.
[122] D. Joseph-McCarthy, A.J. Campbell, G. Kern, D. Moustakas, Fragment-based lead discovery and design, J. Chem. Inf. Model. (2014), https://doi.org/10.1021/ci400731w.
[123] J.P. Renaud, T. Neumann, L. Van Hijfte, Fragment-based drug discovery, Small Molecule Medicinal Chemistry: Strategies and Technologies, 2015 https://doi.org/10.1002/9781118771723.ch8.