Computer-Based Technologies for Virtual Screening and Analysis of Chemical Compounds Promising for Anti-HIV-1 Drug Design

A.M. Andrianov[1], I.A. Kashyn[1,2], and A.V. Tuzikov[2(✉)]

[1] Institute of Bioorganic Chemistry NASB, 5/2 Academician Kuprevich street, 220141 Minsk, Belarus
andrianov@iboch.bas-net.by
[2] United Institute of Informatics Problems NASB, 6, Surganov street, 220012 Minsk, Belarus
tuzikov@newman.bas-net.by

Abstract. Computer-based technologies for *in silico* drug development comprising virtual screening, high-throughput docking, molecular dynamics simulations, and binding free energy calculations are presented. The efficiency of these technologies is demonstrated by the identification of novel potential anti-HIV-1 agents able to mimic pharmacophoric properties of potent and broad neutralizing antibodies 10e8, VRC01, and 3074 that target three different functionally conserved regions of the viral envelope proteins.

Keywords: Virtual screening · Molecular docking · Molecular dynamics · Binding free energy calculations · HIV-1 entry inhibitors · Broadly neutralizing antibodies

1 Introduction

To date, over twenty five drugs have been approved by the United States Food and Drug Administration for the treatment of HIV infection (reviewed in [1, 2]). These drugs are distributed into six major classes: (1) nucleoside-analog reverse transcriptase inhibitors, (2) non-nucleoside reverse transcriptase inhibitors, (3) protease inhibitors, (4) fusion inhibitors, (5) entry inhibitors, and (6) integrase inhibitors [1, 2]. The majority of these anti-HIV drugs belong to the inhibitors of reverse transcriptase and protease [1, 2]. These inhibitors act inside a target cell and cannot block the initial steps of the HIV-1 life cycle associated with virus entry. In this context, development of novel, potent and broad-spectrum HIV-1 entry inhibitors is an area of considerable interest in the current anti-HIV drug design and discovery.

HIV-1 infection begins with virion entry into target cells through the interaction of viral envelope (Env) protein gp120 with primary receptor CD4 (reviewed in [3]). The binding of gp120 to CD4 induces the exposure of a second binding site for cellular co-receptor CCR5 or CXCR4 [3]. Following the binding, the gp41 transmembrane subunit of the Env protein undergoes a dramatic conformational change to mediate virus-cell

V.V. Krasnoproshin and S.V. Ablameyko (Eds.): PRIP 2016, CCIS 673, pp. 14–23, 2017.
DOI: 10.1007/978-3-319-54220-1_2

membrane fusion, enabling the virus capsid to enter the cell [3]. Many small molecule inhibitors that block the virus adsorption onto the host cell membrane and/or cell-mediated fusion have been developed [4]. However, the majority of these inhibitors have failed to be useful in clinical practice. Despite these disappointing results, the design of (+)-DMJ-I-228 and (+)-DMJ-II-121 inhibitors that target the Env trimer and present functional antagonists of viral entry [5] gave hope of future success in the development of novel efficient anti-HIV-1 drugs. This hope was supported by the discovery of anti-HIV-1 broadly neutralizing antibodies (bNAbs) and their specific modes of recognition on the viral Env, providing a new strategy for improved vaccine and drug design (reviewed in [6–8]). Neutralizing antibodies tend to increase in potency over time and broadly cross neutralizing responses, capable of recognizing heterologous HIV-1 variants, develop in a subset of individuals after primary infection. In some cases, the specificities of the antibodies conferring breadth have been mapped and are reactive with conserved Env regions.

In light of discovering anti-HIV-1 bNAbs [6–8], studies aimed at the identification of small molecules able to mimic pharmacophoric properties of these antibodies are of great challenge. In doing so, the latest computational technologies for structure-based drug design can be used at the first steps of solving this problem to significantly reduce drug development time and financial costs.

Recently, computer-aided methods of three-dimensional (3D) structure modeling and the study of the quantitative relationship between structure and biological activity of chemical substances (QSAR; Quantitative Structure-Activity Relationships) have occupied an important place in anti-HIV-1 drug development (reviewed in [9]). One of the more promising methods of computational drug design is molecular databases virtual screening aimed at the discovery of active structures. This method includes the following stages: (i) a choice of biological target and modeling of its high-resolution 3D structure; (ii) selection of structural databases of organic substances and the search for specific compounds which, according to the information on their structural and physicochemical properties are capable of exhibiting the biological activity to the prescribed target; (iii) molecular docking of the selected substances with the biological target using software based on criterion functions and choice of the potential ligands; (iv) post-processing of the created ligand bases using QSAR models, as a result of which a library of potential ligands for the present biological target may be obtained.

In this study, we present an integrated computational approach to *in silico* drug design involving theoretical procedures, such as virtual screening, high-throughput docking, molecular dynamics simulations (MD), and binding free energy calculations. The efficiency of this approach is demonstrated by the identification of novel potential anti-HIV-1 agents mimicking potent and broad neutralizing antibodies 10e8, VRC01 and 3074 that target three different functionally conserved regions of the HIV-1 Env trimer.

1.1 Computational Methodology for Drug Development

Computer-based approach to in *silico* drug design used in this study included the following consecutive stages: (i) generation of pharmacophore models representing 3D-arrangements of chemical functionalities that make a small molecule active towards its target; (ii) structure-based virtual screening of chemical libraries to discover new ligands on the basis of biological structures; (iii) molecular docking to predict the structure of the intermolecular complex formed between two or more constituent molecules; (iv) MD simulations of this complex followed by binding free energy calculations to evaluate its structural stability.

Brief information on the computational tools supporting the above stages is given below.

1.2 Generation of Pharmacophore Models

Pharmacophore models for virtual screening of antibody-mimetic candidates were generated in agreement with the first step of the pepMMsMIMIC web tool strategy consisting in the identification of amino-acid residues that play a key role in the protein-protein recognition process [10]. This strategy employs as input the 3D structure of a peptide bound to a protein and suggests which chemical structures are able to mimic the protein–protein recognition of this peptide by both pharmacophore and shape similarity techniques. At the same time, all possible combinations of the residues exhibiting critical structural features in 3D space may be used in generation of the templates to screen virtual compound libraries for novel ligands, which present the best similarity to the specific pharmacophore [10]. Based on this strategy, the hotspots of the antibodies of interest for their interactions with the target proteins were derived from the X-ray crystal structures of these immunoglobulins in complexes with the gp120 (VRC01 [11] and 3074 [12]) and gp41 (10e8 [13]) proteins. As a result, 3D structures of the antibody peptides including residues that greatly contribute to the binding were used as the general pharmacophore models for identification of the antibody-mimetic candidates. To identify small-molecule peptidomimetic candidates, short fragments of the general models that target different critical regions of the binding sites of the HIV-1 Env trimer were used as the additional input data for pepMMsMIMIC.

1.3 Shape and Pharmacophore-Based Virtual Screening

The pharmacophore models generated based on the antibody binding hotspots were screened against a library of 17 million conformers obtained from 3.9 million commercially available chemical structures present in the MMsINC database [14]. Screening of this virtual compound library was carried out by four scoring methods that are used in the current version of pepMMsMIMIC [10] to optimize the selection of the peptide mimetics. The tools of pepMMsMIMIC offer five search procedures including different combinations of two scoring approaches, such as ultrafast shape recognition [15] and pharmacophore fingerprints similarity [16]. All these procedures were used for search

of the antibody mimetics. The identified compounds were further screened by high-throughput docking to evaluate the efficacy of their binding to the target proteins.

1.4 Molecular Docking

The X-ray crystal structures of the antibody Fabs in complexes with the gp120 (VRC01 [11] and 3074 [12]) and gp41 (10e8 [13]) proteins were used as the rigid receptors for flexible "blind docking" with compounds from the MMsINC database by Autodock Vina [6]. These structures were prepared by adding hydrogen atoms with the AutoDockTools software [17]. For all compounds, the docked structures with the highest scores were analyzed to identify the molecules that, similarly to the antibodies, exhibit a strong attachment to the antibody-binding sites of the target proteins. As a result, the complexes of top-ranking compounds with these proteins were selected based on the values of scoring functions.

1.5 Molecular Dynamics Simulations

The MD simulations for the docked structures of the top compounds with the target proteins were performed using Amber 11 with the implementation of the Amber ff10 force field [18]. The ANTECHAMBER module was employed to use the Gasteiger atomic partial charges individually for each of the compounds, and the general AMBER force field [19] was used to prepare the force field parameters. Hydrogen atoms were added to the proteins by the xleap program of the AMBER 11 package [18]. The protonation state of histidine residues was checked visually and if necessary either Nδ or Nε protonation was chosen to ensure optimal hydrogen bonding. The systems were solvated using TIP3P water [20] as an explicit solvent and simulated in an octahedron box with periodic boundary conditions. The structures were first energy minimized by 500 steps of the steepest descent algorithm followed by 1000 steps of the conjugate gradient method. The atoms of every docked structure were then restrained by an additional harmonic potential with the force constant equal to 1.0 kcal/mol and then heated from 0 to 310 K over 1 ns using a constant volume of the unit cell. Additional equilibration was performed over 1 ns by setting the system pressure to 1.0 atm and by using a weak coupling of the system temperature to a 310 K bath with 2.0 ps characteristic time. Finally, the constraints on the complex assembly were removed and the system was equilibrated again at 310 K over 2 ns under constant volume conditions. After equilibration, the isothermal-isobaric MD simulation ($T = 310$ K, $P = 1.0$ atm) generated 30 ns trajectory using a Berendsen barostat with 2.0 ps characteristic time, a Langevin thermostat with collision frequency 2.0 ps^{-1}, a non-bonded cut-off distance of 8 Å, and a simple leapfrog integrator [18] with a 2.0 fs time step and bonds with hydrogen atoms constrained by the SHAKE algorithm [21].

1.6 Binding Free Energy Calculations

The free energy of binding was calculated in AMBER 11 by the MM/PBSA method [22]. Five hundred snapshots were selected from the last 25 ns to estimate the binding

free energy, by keeping the snapshots every 50 ps. The polar solvation energies were computed in continuum solvent using Poisson-Boltzmann and ionic strength of 0.1. The non-polar terms were estimated using solvent accessible surface areas.

Based on the MM/PBSA analyses of the MD trajectories, chemical compounds that showed negative free energies of the binding to the target proteins were selected for the final analysis.

2 Results

In the case of bNAb 10e8 that neutralizes up to 98% of diverse HIV-1 strains [13], virtual screening of the MMsINC database combined with molecular docking and MD simulations identified eight top hits that exposed the high-affinity binding to gp41 by targeting the membrane proximal external region (MPER) of this HIV-1 protein, allowing one to consider these molecules as promising peptidomimetic candidates of bNAb 10e8 [23]. Chemical structures of these compounds are shown in Fig. 1.

Figure 2 casts light on the docked structures of the identified compounds (Fig. 1) with the gp41 MPER peptide. In particular, analysis of the MMs03555010-gp41 docked structure indicates (Fig. 2a) that, similarly to bNAb 10e8, this molecule targets the central hinge region of the MPER peptide providing the conformational flexibility necessary for the Env-mediated hemifusion and fusion processes [13, 25].

The docked structures of the identified compounds with gp41 do not undergo substantial rearrangements during the MD simulations, in agreement with the low averages of free energy of their formation that are -16.1 ± 3.4 kcal/mol (MMs03555010), -12.0 ± 4.8 kcal/mol (MMs03769994), -7.3 ± 4.8 kcal/mol (MMs01288397), -7.2 ± 3.9 kcal/mol (MMs02374310), -6.6 ± 4.4 kcal/mol (MMs03064646), -5.7 ± 4.0 kcal/mol (MMs03534576), -5.4 ± 3.6 kcal/mol (MMs01100460), and -5.3 ± 4.5 kcal/mol (MMs00760407).

Analysis of the superimposed complexes of the MPER peptide with the 10e8 Fab and peptidomimetic candidates indicates (Fig. 3) that the identified compounds partially mask the region of gp41 that is targeted by bNAb 10e8. These small molecules bind to the vulnerable spots of this gp41 region and may therefore exhibit the functional mimicry of 10e8. The molecules of interest mimic segment Trp-33, Gly-52c, Pro-52b, Glu-53, Lys-97 of the 10e8 heavy chain (Fig. 3) that forms the direct intermolecular contacts with the functionally important residues of gp41.

The above computer-based technologies have been also applied to the search for potential peptidomimetics of anti-HIV-1 bNAb VRC01 [26] neutralizing over 90% of diverse HIV-1 strains by specific interactions with the CD4-binding site of the Env protein gp120 [11]. As a result, six chemical compounds from MMsINC were shown to exhibit a high affinity to this gp120 site responsible for the HIV-1 attachment to cellular receptor CD4 [26]. The docked models of these compounds with the gp120 core demonstrate intermolecular interactions involving the residues of gp120 important for the HIV-1 binding to CD4. In addition, these complexes show relative stability within the MD simulations that is validated by the values of binding free energy and corresponding standard deviations. These values were shown to be lower than that of

−9.5 ± 0.1 kcal/mol which was determined for the gp120/CD4 complex using isothermal titration calorimetry [26].

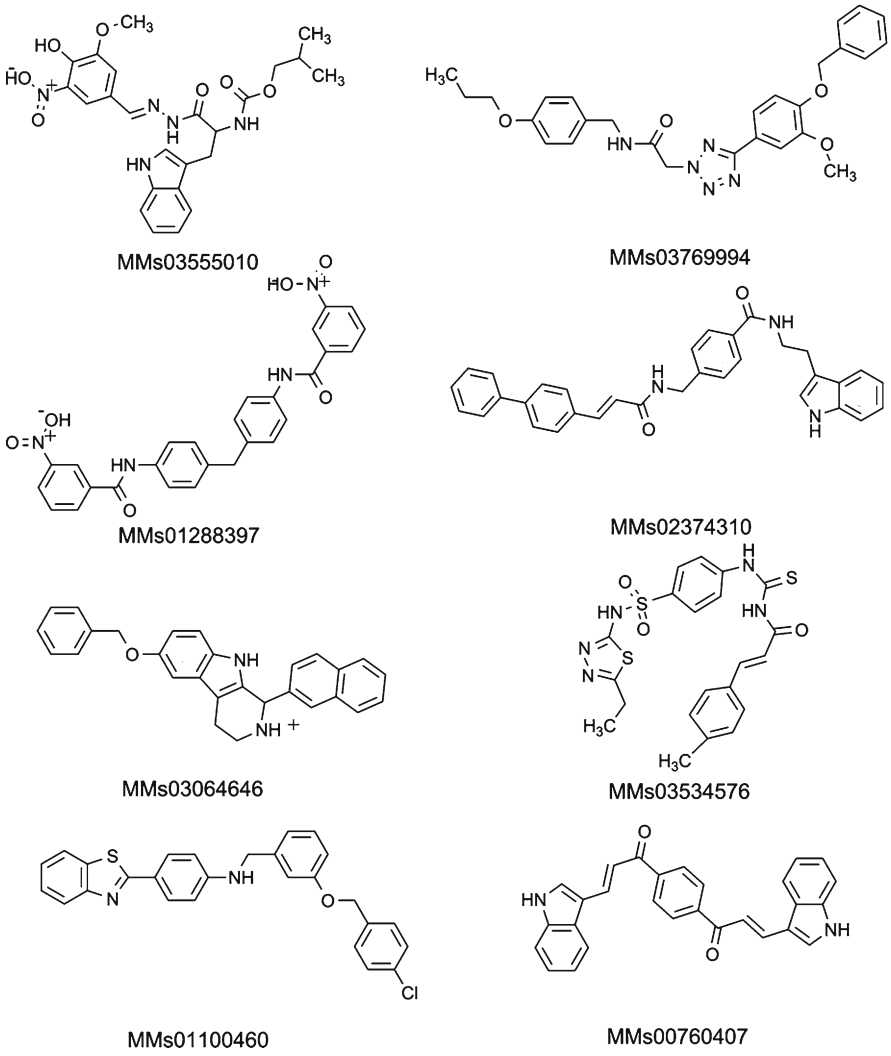

Fig. 1. Chemical structures of the most probable peptidomimetics of bNAb 10e8. The molecule codes are from the MMsINC database [14]. For these molecules, the four Lipinski's parameters - molecular weight, lipophilicity (LogP), number of H-bond donors and acceptors [24] - are respectively: 497.5 Da, 3.97, 4 and 5 (MMs03555010); 487.6 Da, 3.15, 1 and 7 (MMs03769994); 496.5 Da, 2.36, 2 and 2 (MMs01288397); 499.6 Da, 4.52, 3 and 2 (MMs02374310); 405.5 Da, 0.0, 1 and 1 (MMs03064646); 487.6 Da, 3.62, 3 and 6 (MMs03534576); 457.0 Da, 3.0, 1 and 2 (MMs01100460); 416.5 Da, 2.94, 2 and 2 (MMs00760407)

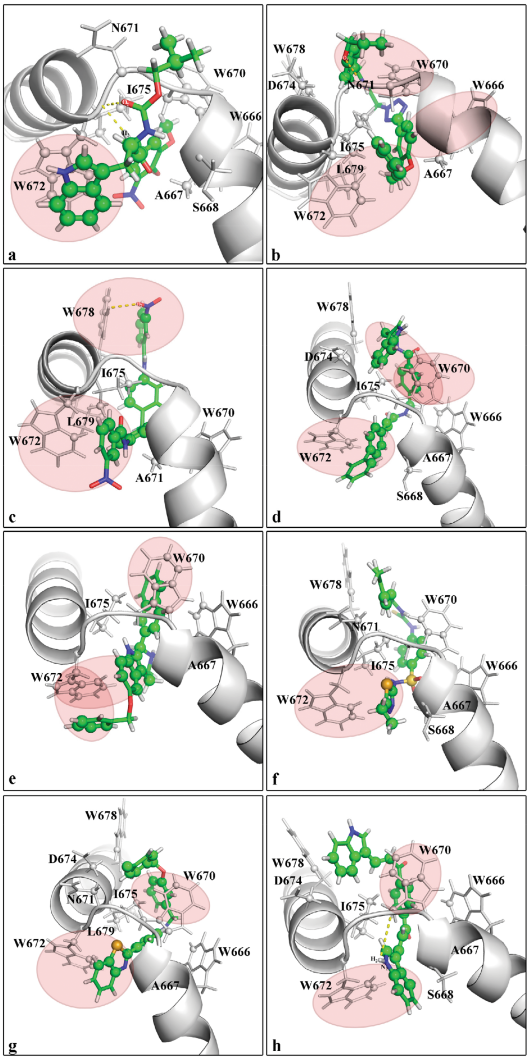

Fig. 2. The docked structures of the gp41 MPER peptide with the MMs03555010, MMs03769994, MMs01288397, MMs02374310, MMs03064646, MMs03534576, MMs01100460, and MMs00760407 compounds. Structures of these compounds are represented by a stick-ball-stick model. The residues of gp41 forming hydrogen bonds, π-π stacking and van der Waals contacts with the 10e8-mimetic candidates are indicated. Structural elements of gp41 and ligands involved in specific π-π interactions are located inside the circles. Hydrogen bonds are shown by dotted lines

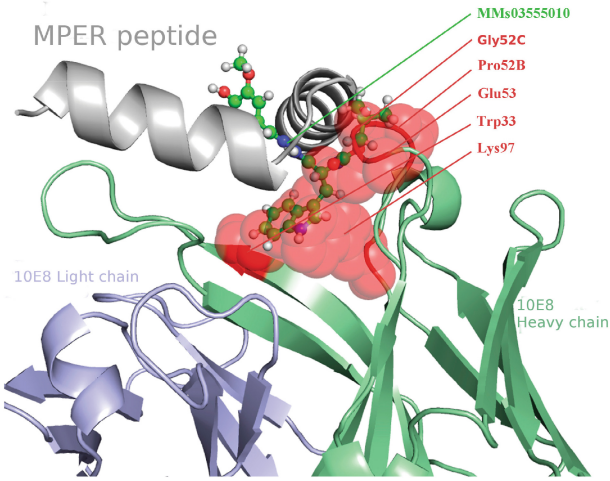

Fig. 3. Superimposed complexes of the HIV-1 MPER peptide with the 10e8 Fab and peptidomimetic candidate MMs03555010. Structure of the 10e8 peptidomimetic is shown by a stick-ball-stick model. Only MMs03555010 in complex with gp41 is shown for peptidomimetic candidates; the complexes of gp41 with the other identified compounds coincide with this supramolecular structure. A Corey-Pauling-Koltun model is used to highlight the 10e8 residues that are mimicked by the identified compounds

In contrast to VRC01 targeting CD4-binding site of gp120 [11], bNAb 3074 binds to the third variable (V3) loop of gp120 [12]. This domain of gp120 plays a central role in the biology of the HIV-1 Env trimer as a principal target for neutralizing antibodies and as a major determinant in the switch from the non-syncytium-inducing to the syncytium-inducing form of HIV-1 that is associated with accelerated disease progression [9]. Application of the above methodology to the search for the most probable mimetic candidates of 3074 resulted in the identification of four top hits demonstrating a strong attachment to the HIV-1 V3 loop [27]. Specific binding to the V3 loop was shown to be accomplished primarily by π-π interactions between the aromatic rings of the peptidomimetics and the conserved Phe-20 and/or Tyr-21 of the V3 immunogenic crown [27]. In a mechanism similar to that of bNAb 3074, these compounds were found to block the tip of the V3 loop forming its invariant structural motif that contains residues critical for the HIV-1 binding to a chemokine co-receptor, either CCR5 or CXCR4, which is required for viral entry. With these findings, the following conclusion was made: the compounds selected form promising scaffolds for the rational design of novel, potent, and broad-spectrum anti-HIV-1 therapeutics [27].

3 Conclusions

An integrated computational approach used in this study allowed one to identify eighteen top hits able to mimic anti-HIV-1 bNAbs 10e8 (eight compounds), VRC01 (six compounds), and 3074 (four compound). These small molecules target three different

vulnerable spots of the viral Env trimer responsible for the HIV-1 entry into target cells. Based on the data obtained, these compounds provide good scaffolds for the development of potent antiretroviral drugs with broad HIV-1 neutralization. In light of these data, we also suppose that a bifunctional anti-HIV-1 "cocktail" of small-molecule peptidomimetics of bNAbs 10e8, VRC01, and 3074 may suppress viral replication and reduce the plasma HIV-1 viral load.

Thus, the above findings clearly show that the computational methodology presented here is a powerful *in silico* tool for the discovery of novel small-molecule functional antagonists of viral entry based on the anti-HIV-1 bNAbs.

Acknowledgements. This study was supported by a grant from the Belarusian Foundation for Basic Research (project X15-022).

References

1. Arts, E.J., Hazuda, D.J.: HIV-1 antiretroviral drug therapy. Cold Spring Harb. Perspect. Med. **2**, a007161 (2012)
2. Kumari, G., Singh, R.K.: Highly active antiretroviral therapy for treatment of HIV/AIDS patients: current status and future prospects and the Indian scenario. HIV AIDS Rev. **11**, 5–14 (2012)
3. Wilen, C.B., Tilton, J.C., Doms, R.W.: HIV: cell binding and entry. Cold Spring Harb. Perspect. Med. **2**, a006866 (2012)
4. Acharya, P., et al.: HIV-1 gp120 as a therapeutic target: navigating a moving labyrinth. Expert Opin. Ther. Targets **19**, 1–19 (2015)
5. Courter, J.R., et al.: Structure-based design, synthesis and validation of CD4-mimetic small molecule inhibitors of HIV-1 entry: conversion of a viral entry agonist to an antagonist. Acc. Chem. Res. **47**, 1228–1237 (2014)
6. McCoy, L.E., Weiss, R.A.: Neutralizing antibodies to HIV-1 induced by immunization. J. Exp. Med. **210**, 209–223 (2013)
7. Mascola, J.R., Haynes, B.F.: HIV-1 neutralizing antibodies: understanding nature's pathways. Immunol. Rev. **254**, 225–244 (2013)
8. Haynes, B.F., McElrath, M.J.: Progress in HIV-1 vaccine development. Curr. Opin. HIV AIDS **8**, 326–332 (2013)
9. Andrianov, A.M.: Human immunodeficiency virus-1 gp120 V3 loop for anti-acquired immune deficiency syndrome drug discovery: computer-aided approaches to the problem solving. Expert Opin. Drug Discov. **6**, 419–435 (2011)
10. Floris, M.J., et al.: Swimming into peptidomimetic chemical space using pepMMsMIMIC. Nucleic Acids Res. **39**, W261–W269 (2011)
11. Zhou, T., et al.: Structural basis for broad and potent neutralization of HIV-1 by antibody VRC01. Science **329**, 811–817 (2010)
12. Jiang, X., et al.: Conserved structural elements in the V3 crown of HIV-1 gp120. Nature Struct. Mol. Biol. **17**, 955–961 (2010)
13. Huang, J., et al.: Broad and potent neutralization of HIV-1 by a gp41-specific human antibody. Nature **491**, 406–412 (2012)
14. Masciocchi, J., et al.: MMsINC: a large-scale chemoinformatics database. Nucleic Acids Res. **37**, D284–D290 (2009)

15. Ballester, P.J., Richards, W.G.: Ultrafast shape recognition to search compound databases for similar molecular shapes. J. Comput. Chem. **28**, 1711–1723 (2007)
16. Mason, J.S., et al.: New 4-point pharmacophore method for molecular similarity and diversity applications: overview of the method and applications, including a novel approach to the design of combinatorial libraries containing privileged substructures. J. Med. Chem. **42**, 3251–3264 (1999)
17. Trott, O., Olson, A.J.: Software news and update AutoDock Vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. J. Comput. Chem. **31**, 455–461 (2010)
18. Case, D.A., et al.: AMBER 11. University of California, San Francisco (2010)
19. Wang, R.M., et al.: Development and testing of a general Amber force field. J. Comput. Chem. **25**, 1157–1174 (2004)
20. Jorgensen, W.L., et al.: Comparison of simple potential functions for simulating liquid water. J. Chem. Phys. **79**, 926–935 (1983)
21. Ryckaert, J.P., Ciccotti, G., Berendsen, H.J.C.: Numerical integration of the Cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. **23**, 327–341 (1977)
22. Massova, I., Kollman, P.A.: Computational alanine scanning to probe protein-protein interactions: a novel approach to evaluate binding free energies. J. Am. Chem. Soc. **121**, 8133–8143 (1999)
23. Andrianov, A.M., Kashyn, I.A., Tuzikov, A.V.: Identification of novel HIV-1 fusion inhibitor scaffolds by virtual screening, high-throughput docking and molecular dynamics. JSM Chem. **4**(2), 1022 (2016)
24. Lipinski, C.A., et al.: Lead- and drug-like compounds: the rule-of-five revolution. Adv. Drug Deliv. Rev. **46**, 3–26 (2001)
25. Cheng, Y.: Elicitation of antibody responses against the HIV-1 gp41 Membrane Proximal External Region (MPER). Doctoral dissertation, Harvard University (2014). (http://nrs.harvard.edu/urn-3:HUL.InstRepos:12269838)
26. Andrianov, A.M., Kashyn, I.A., Tuzikov, A.V.: Computational discovery of novel HIV-1entry inhibitors based on potent and broad neutralizing antibody VRC01. J. Mol. Graph. Model. **61**, 262–271 (2015)
27. Andrianov, A.M., Kashyn, I.A., Tuzikov, A.V.: Discovery of novel Anti-HIV-1 agents based on a broadly neutralizing antibody against the envelope gp120 V3 loop: a computational study. J. Biomol. Struct. Dyn. **32**, 1993–2004 (2014)

http://www.springer.com/978-3-319-54219-5

Pattern Recognition and Information Processing
13th International Conference, PRIP 2016, Minsk,
Belarus, October 3-5, 2016, Revised Selected Papers
Krasnoproshin, V.V.; Ablameyko, S.V. (Eds.)
2017, X, 181 p. 92 illus., Softcover
ISBN: 978-3-319-54219-5