Mol Biol Rep (2017) 44:353–358
DOI 10.1007/s11033-017-4116-1

SHORT COMMUNICATION

In silico analysis of the cyanobacterial lectin scytovirin: new insights into binding properties

Andrei Santos Siqueira[1] · Alex Ranieri Jerônimo Lima[1] · Rafael Conceição de Souza[2] · Alberdan Silva Santos[2] · João Lídio da Silva Gonçalves Vianez Júnior[3] · Evonnildo Costa Gonçalves[1]

Received: 11 August 2016 / Accepted: 25 July 2017 / Published online: 29 July 2017

Abstract Scytovirin is a lectin isolated from the cyanobacterium *Scytonema varium* that has shown activity against HIV, SARS coronavirus and Zaire Ebola virus. Its 95 amino acids are divided into two structural domains (SD), the first spanning amino acids 1–48 (SD1) and the second 49–95 (SD2). Interestingly, the domains are nearly identical but differ in their affinities for carbohydrates. With the aim of enhancing understanding of the binding properties of scytovirin, we performed molecular dynamics (MD) simulations of scytovirin complexed with Man4. We set up three systems: (i) Man4 bound to both domains (SD1 + SD2) using the full-length protein; (ii) Man4 bound to an incomplete protein, containing only SD1 and (iii) Man4 bound to an incomplete protein containing only SD2. Contrary to other reports, binding free energy results suggest that Man4 can bind simultaneously to SD1 and SD2 binding regions, but SD1 individually has the best values of energy and the best affinity for Man4. Decomposition of the binding free energy showed that the residues that interact with Man4 were different in the three systems, suggesting that the binding mechanism of Man4 varies between full-length protein, SD1 and SD2. The results presented here may help to formulate strategies to use scytovirin and promote mutagenesis studies to improve the antiviral activity of scytovirin.

Keywords Molecular dynamics · Scytovirin · Lectin · Cyanobacteria · Antiviral activity

Introduction

The phylum cyanobacteria has been explored for a long time due to its high biotechnological potential, mainly because of bioactive compound production [1, 2]. In this regard, some lectins produced by algae and cyanobacteria are notable for their high antiviral activity against HIV, severe acute respiratory syndrome (SARS) coronavirus and Zaire Ebola virus [3]. In general, those proteins act by preventing the fusion of the virus with the host cell through binding to viral glycoproteins.

Among the lectins isolated from cyanobacteria, cyanovirin, one of the most well-known; it is produced by the species *Nostoc ellipsosporum* and presents 101 amino acids and a molecular weight of 10 kDa. Recombinant cyanovirin expressed in soybean seeds is shown to be biologically active in anti-HIV assays [4]. The cyanobacterial lectin microvirin, originally isolated from *Microcystis aeruginosa* PCC 7806, which shares 33% similarity with the cyanovirin, is unusual in that it potently inhibits HIV-1 with negligible toxicity compared to cyanovirin [3, 5]. A third cyanobacterial lectin with broad biotechnological potential given its biological activity is scytovirin, a lectin that was isolated from the cyanobacterium *Scytonema varium* [6]. Cyanobacterial lectin scytovirin has 95 amino acids with a molecular mass of approximately 9.7 kDa divided into two structural domains (SD), the first spanning amino acids 1–48 (SD1) and the second 49–95 (SD2). These are nearly

✉ Andrei Santos Siqueira
andrei.siqueira@icb.ufpa.br

[1] Laboratório de Tecnologia Biomolecular – Instituto de Ciências Biológicas, Universidade Federal do Pará, Belem, Pará, Brazil

[2] Laboratórios de Investigação Sistemática em Biotecnologia e Biodiversidade Molecular – Instituto de Ciências Exatas e Naturais, Universidade Federal do Pará, Belem, Pará, Brazil

[3] Centro de Inovações Tecnológicas – Instituto Evandro Chagas, Ananindeua, Pará, Brazil

identical, although they present different affinities and may bind independently to ligand carbohydrates [7]. Since scytovirin can bind to HIV glycoproteins gp160, gp120 and gp41 and to the Zaire Ebola virus mucin domain and shows antiviral activity against these viruses [8, 9], it is currently being investigated as a potential antiviral microbicide candidate.

These lectins are not HIV-specific because their target high-mannose arrays are present in all HIV subtype envelope glycoproteins and in other viruses. Because such a mannose content is uncommon in mammalian cells, these compounds are not likely to be toxic to human cells in vivo [10, 11]. However, it has been demonstrated that HIV can develop resistance to HIV by deglycosylation in site-specific residues, although multiple mutations are necessary to make this happen [12]. Therefore, it is very important to know all the structural aspects of the interaction of scytovirin with the high-mannose arrays in order to develop an effective microbicide.

Adams et al. [13] developed a study with microarray carbohydrates and determined that the minimal core structures necessary for scytovirin binding are the D3 arm of Man9, Manα(1–2) Manα(1–2) Manα(1–6)Man, also known as Man4. A binding analysis of Man4 with scytovirin by titration concluded that SD1 and SD2 act independently and not cooperatively [14]. Since some aspects of this relationship remained unclear, in our study we performed molecular dynamics (MD) simulations of scytovirin complexed with Man4 in order to better understand its binding properties.

Materials and methods

Protein structure

There are three scytovirin structures in protein data bank (PDB). For our study we chose a structure resolved by high resolution X-ray diffraction, which was 1.00 Å (PDB code: 2QSK) [15]. Its disulfide-bonding pattern differs from that described in previously published mass spectrometry and NMR studies [8].

Molecular docking

The structure of Man4 was constructed according to Adams et al. [13] from the Glycam website (http://www.glycam.org/). Gaussian09 was employed to optimize and calculate the RESP charges of the structure. To determine an initial conformation for Man4, we performed a docking simulation with Molegro Virtual Docking 5 software. MolDock scoring function with a grid resolution of 0.30 was used to rank the solutions, toggling on Internal ES, Internal HBond and Sp2-Sp2 torsions. The radius of the search area covered all key residues described as ligation sites [8]. Docking results were further refined by energy minimization and Hbonds to optimize hydrogen bonds after docking. Each docking had 10 runs, population size of 100 and max iterations of 2000. Three dockings were performed, one of them based on full-length scytovirin and the two others based on SD1 and SD2, individually.

Molecular dynamics

To determine the protonation state of the residues, we used the PDB2PQR server (http://nbcr-222.ucsd.edu/pdb2pqr_2.0.0/) considering a pH of 7.0. AMBER 12 package was used to perform all steps of preparation and MD [16]. Carbohydrate parameters were obtained from the GLYCAM06 [17] force field. For the protein and ions, FF99SB [18] and GAFF [19] force fields were used, respectively. Next, we set up three systems: (i) Man4 bound to both domains (SD1 + SD2) using the full-length protein; (ii) Man4 bound to an incomplete protein, containing only SD1, (iii) Man4 bound to an incomplete protein containing only SD2. The three systems were constructed, each one with a ligand (Man4), Cl^- ions to neutralize the charges and the TIP3P water molecules in a cubic box with 12 Å in each direction of the protein. Sander software, which is included in AMBER 12, performed the energy minimization in five stages. The first four stages used 3000 cycles of steepest descent and 5000 cycles of conjugate gradients. In the remaining stage, 5000 cycles of steepest descent and 30,000 cycles of conjugate gradients were used. The heavy atoms had their movement restrained by a harmonic potential of 1000 Kcal/mol $Å^2$, while in the final step all atoms were free. For heating and equilibration, 14 steps were used. A harmonic potential of 25 Kcal/mol $Å^2$ was employed to restrain the heavy atoms in the initial stages. In the step 13, the restraining force was set to 0. The temperature was gradually increased, until it reached 300 K. Langevin dynamics (thermostat) were employed with a collision frequency of 3.0 ps^{-1}. This method is significantly more efficient at equilibrating the system temperature than the Berendsen temperature coupling scheme that was recommended for older versions of AMBER [16]. The heating procedure lasted 650 ps until step 13 and was performed using an NVT ensemble. Afterwards, a 2 ns equilibration phase was employed in an NPT ensemble. The SHAKE algorithm was employed to restrict the vibration of the ligations of all hydrogen atoms. The Particle Mesh Ewald method was used for calculating electrostatic interactions using a cutoff value of 10.0 Å. Thus, 210 ns MD was produced in an NVT ensemble, for each system. The cpptraj module was used to compute the root mean square deviation (RMSD) of both trajectories, considering the heavy atoms of the main chain. 5000 snapshots from the last

10 ns of the MD simulations were used to calculate the free energies using MM-GBSA, MM-PBSA and SIE methods [20–22].

Results

Molecular docking

The best poses obtained in molecular docking are shown in Table 1. According to these results, the best MolDock score was obtained when Man4 binds simultaneously to SD1 and SD2, interacting with residues of both domains (Table 1). SD2 shows a higher number of residues interacting with Man4. The aromatic residues described as important to scytovirin function [14] are present in the three docking analyses performed here. These three conformations were used as an input to build the MD systems.

Molecular dynamics

We generated production runs of 210 ns for full-length scytovirin (SD1 + SD2), SD1 and SD2 complexed individually with Man4. The behavior of the protein and the ligand was different in each analysis. The most stable system was the full-length scytovirin complexed with Man4, which shows RMSD values around 1 Å for protein and 3 Å for Man4 (Fig. 1). The system based on SD1 and Man4 complex presented a higher number of conformation changes in comparison to the other systems, but the SD1 became more stable at the end of the simulation. Nonetheless, Man4 remained complexed with the protein during all MD simulations. It was also observed that during the entire simulation with the full-length scytovirin, Man4 maintained interactions with residues of SD1 and SD2 simultaneously (Fig. 2).

Binding free energy calculations

The results of MM-GBSA, MM-PBSA and SIE methods are presented in Table 2. These results corroborate the experimental approach in which SD1 individually has higher affinity for Man4 than full-length scytovirin and SD2 [7]. In the three methods used here, the SD1 showed the best values of binding free energy. When the energy decomposition was performed, it was possible to observe that different residues are involved with Man4 binding. Additionally, scytovirin presented several residues contributing along full-length protein (Fig. 3), including the regions with aromatic residues like Trp8 and Phe47. We can observe that the most important residues of SD1 binding (Glu10, Ans18 and Arg30) are conserved in SD2 as Glu58, Ans66 and Arg78, but they do not contribute to interaction when SD2 is complexed individually with Man4 (Fig. 3; Table 2).

Discussion and conclusions

This study is the first theoretical analysis to use the disulfide bonds pattern proposed by Moulaei et al. [15], which is Cys20–Cys32 and Cys26–Cys38 in SD1, Cys68–Cys80 and Cys74–Cys86 in SD2 and, Cys7–Cys55 in the inter-domain region. It is noteworthy that, based on previous reports providing contradictory results for the disulfide pairing in each individual domain, the pattern used here was proven by Moulaei et al. [23], who used N-terminal sequencing and mass spectrometry to analyze proteolytic fragments of native scytovirin obtained at an acidic pH.

Our results (Fig. 3) corroborate the McFeeters et al. [14] study wherein the key residues which interact with Man4 are essentially Cys7, Trp8, Asn9, Glu10, Arg43, Asp46, Phe47, Cys55, Asp57, Glu58, Lys60, Asp75. On the other hand, these authors propose that the two carbohydrate binding sites (one in each domain) of scytovirin function independently with no cooperativity, although, our results (Figs. 2, 3) reveal that Man4 can bind simultaneously to residues of SD1 and SD2. However, SD1 individually has the best values for binding free energy and higher affinity (Table 2). This also corroborates Xiong et al. [7], who demonstrated that scytovirin has two domains with different affinity for carbohydrates.

In this sense, we propose a way to try to explain the difference of affinity between the two nearly identical domains—since SD1 has three uncharged asparagines—where two of them are substituted by charged residues in SD2. This occurs in such a way that in SD2, charged residues Asp57 and Lys 60, replace asparagine residues

Table 1 Summary of docking scores and interactions of selected docking poses

	Moldock score	Rerank score	Hbond	Residues interactions
Scytovirin Man4	−191.928	−64.018	−17.157	C7, W8, N9, C55, W56, D57, K60, D75
SD1 Man4	−154.258	−97.238	−22.448	W8, E10, A11, P17, N18, D30, F37, C38, S42
SD2 Man4	−169.897	−11.536	−25.399	W56, D57, E58, A59, N66, D75, R78, T89, G91, H92, A93

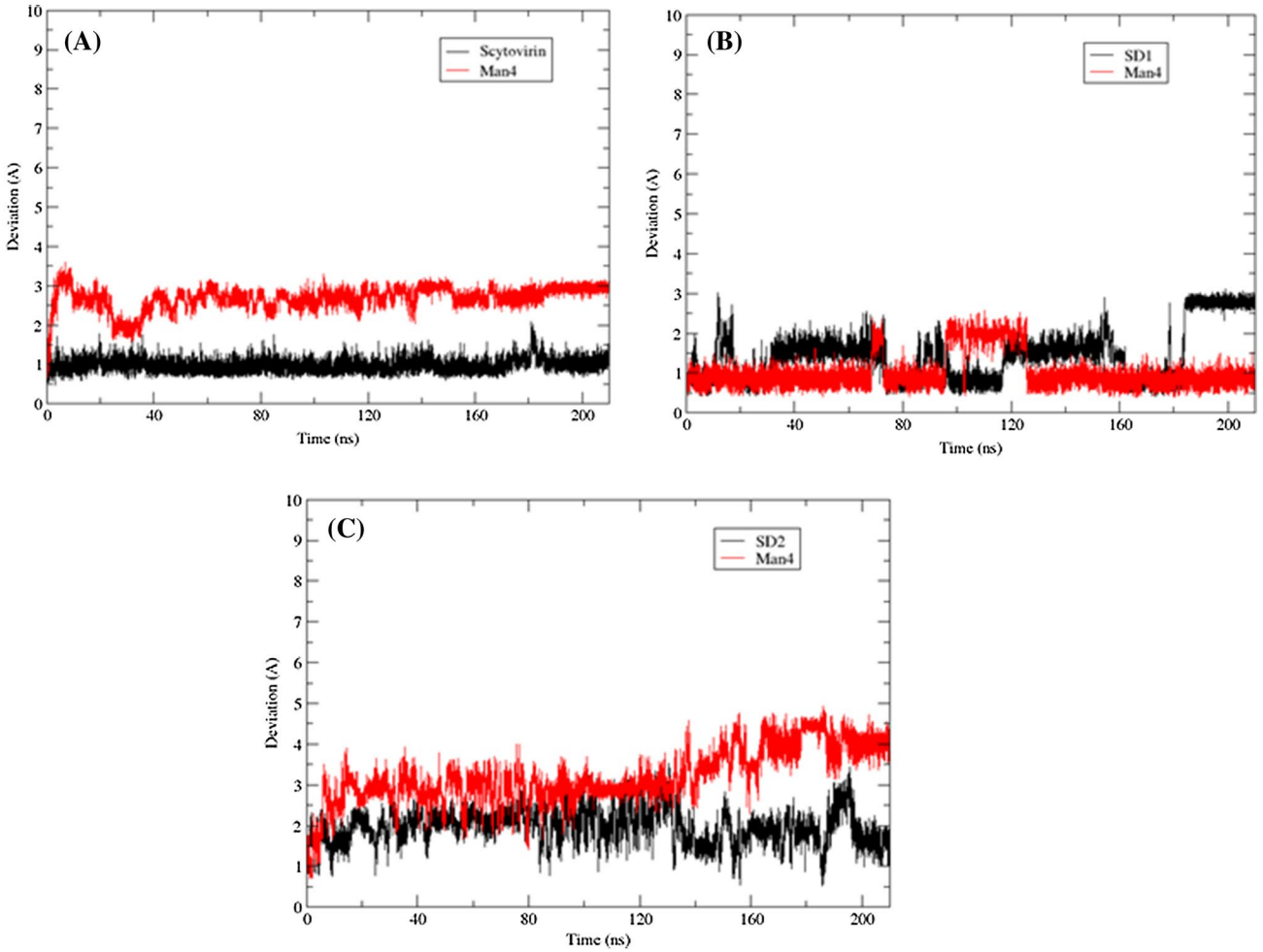

Fig. 1 RMSD graphs of MD analysis. **a** Full-length scytovirin complexed with Man4. **b** SD1 complexed with Man4. **c** SD2 complexed with Man4

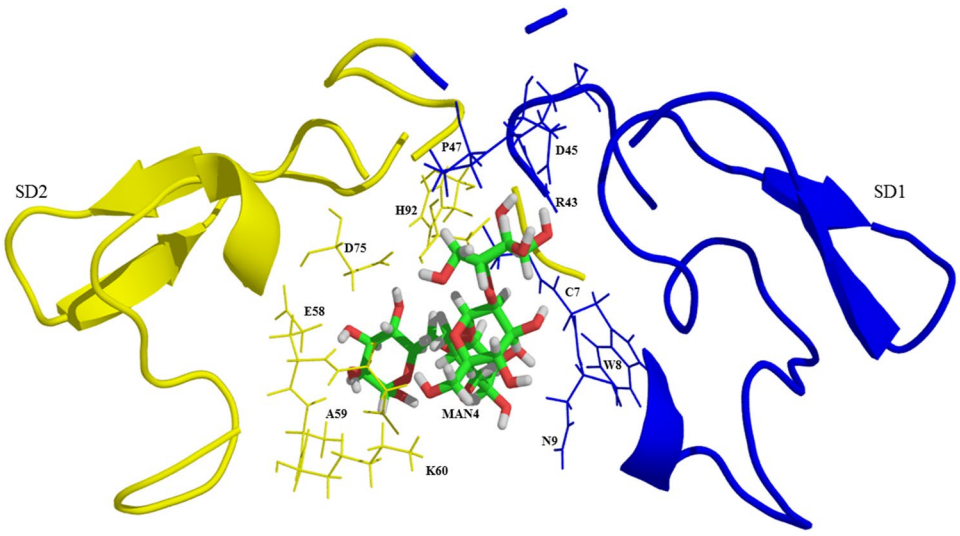

Fig. 2 Final structural conformation of full-length scytovirin complexed with Man4 (*sticks*). Residues that contributed to interaction are shown as *lines*

Table 2 Bind free energy calculations results based on last 10 ns of MD simulation

	MM-GBSA	Std. dev.	Std. error	MM-PBSA	Std. dev.	Std. error	SIE	Std. dev.	Std. error
Scytovirin Man4	−15.34	4.51	0.14	−14.03	5.49	0.17	−7.61	0.42	0.03
SD1 Man4	−32.20	3.39	0.11	−31.99	3.85	0.12	−8.09	0.02	0.35
SD2 Man4	−6.78	3.19	0.10	−14.92	5.76	0.18	−5.21	0.39	0.03

All energy values are in kcal/mol

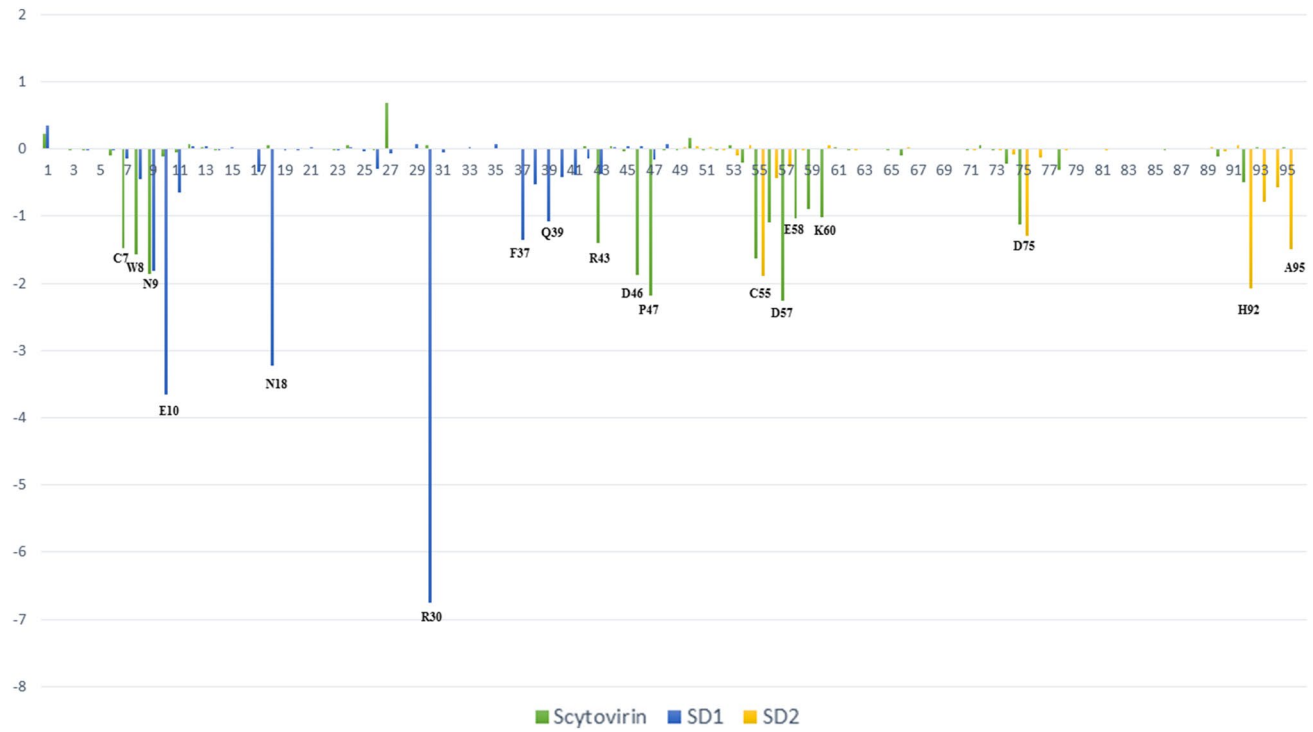

Fig. 3 Binding free energy decomposition per residue based on the MM-GBSA method. Full-length scytovirin analysis is shown in *green*, SD1 individually in *blue* and SD2 individually in *yellow*. (Color figure online)

and create an electrostatic force capable of separating Man4 from the key residues Ans66 and Arg78, which are Ans18 and Arg30 in SD1, respectively (Fig. 3). This event affects the formation of interactions in the same pattern as observed in SD1 and that affect the binding free energy of SD2.

Finally, the use of lectins in the treatment and prevention of viral diseases is a very promising application [3, 5], and it is thus important to know how theseproteins act to prevent microbial infection. In the case of scytovirin, the mechanism for binding to Man4 varies between full-length protein, SD1 and SD2. Overall, the results presented here may help to formulate strategies for using scytovirin as a potential antiviral drug as well for promoting mutagenesis studies that can improve the affinity of scytovirin for the ligand carbohydrates.

Acknowledgements We acknowledge Fundação Amazônia de Amparo a Estudos e Pesquisas do Pará (FAPESPA) for financially supporting (ICAAF 099/2014) our project. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) also supported individual authors through Grant 311686/2015-0 (ECG).

References

1. Kreitlow S, Mundt S, Lindequist U (1999) Cyanobacteria—a potential source of new biologically active substances. J Biotechnol 70:61–63. doi:10.1016/S0168-1656(99)00058-9

2. Vijayakumar S, Menakha M (2015) Pharmaceutical applications of cyanobacteria—a review. J Acute Med 5:15–23. doi:10.1016/j.jacme.2015.02.004
3. Huskens D, Schols D (2012) Algal lectins as potential HIV microbicide candidates. Marine Drugs 10:1476–1497. doi:10.3390/md10071476
4. O'Keefe BR, Murad AM, Vianna GR, Ramessar K, Saucedo CJ, Wilson J et al (2015) Engineering soya bean seeds as a scalable platform to produce cyanovirin-N, a non-ARV microbicide against HIV. Plant Biotechnol J 13:884–892. doi:10.1111/pbi.12309
5. Ziółkowska NE, Wlodawer A (2006) Structural studies of algal lectins with anti-HIV activity. Acta Biochim Pol 53:617–626
6. Bokesch HR, O'Keefe BR, McKee TC, Pannell LK, Patterson GML, Gardella RS et al (2003) A potent novel anti-HIV protein from the cultured cyanobacterium *Scytonema varium*. Biochemistry 42:2578–2584. doi:10.1021/bi0205698
7. Xiong C, O'Keefe BR, Byrd RA, McMahon JB (2006) Potent anti-HIV activity of scytovirin domain 1 peptide. Peptides 27:1668–1675. doi:10.1016/j.peptides.2006.03.018
8. McFeeters RL, Xiong C, O'Keefe BR, Bokesch HR, McMahon JB, Ratner DM et al (2007) The novel fold of scytovirin reveals a new twist for antiviral entry inhibitors. J Mol Biol 369:451–461. doi:10.1016/j.jmb.2007.03.030
9. Garrison AR, Giomarelli BG, Lear-Rooney CM, Saucedo CJ, Yellayi S, Krumpe LRH et al (2014) The cyanobacterial lectin scytovirin displays potent in vitro and in vivo activity against Zaire Ebola virus. Antiviral Res 112:1–7. doi:10.1016/j.antiviral.2014.09.012
10. Alexandre KB, Gray ES, Lambson BE et al (2010) Mannose-rich glycosylation patterns on HIV-1 subtype C gp120 and sensitivity to the lectins, Griffithsin, Cyanovirin-N and Scytovirin. Virology 402:187–196. doi:10.1016/j.virol.2010.03.021
11. Balzarini J (2005) Targeting the glycans of gp120: a novel approach aimed at the Achilles heel of HIV. Lancet Infect Dis 5:726–731. doi:10.1016/S1473-3099(05)70271-1
12. Balzarini J, Van Damme L (2007) Microbicide drug candidates to prevent HIV infection. Lancet 369:787–797. doi:10.1016/S0140-6736(07)60202-5
13. Adams EW, Ratner DM, Bokesch HR, McMahon JB, O'Keefe BR, Seeberger PH (2004) Oligosaccharide and glycoprotein microarrays as tools in HIV glycobiology; glycan-dependent gp120/protein interactions. Chem Biol 11:875–881. doi:10.1016/j.chembiol.2004.04.010
14. McFeeters H, Gilbert MJ, Wood AM (2013) Scytovirin engineering improves carbohydrate affinity and HIV-1 entry inhibition. Biochem Physiol. doi:10.4172/2168-9652.S2-003.
15. Moulaei T, Botos I, Ziółkowska NE, Bokesch HR, Krumpe LR, McKee TC et al (2007) Atomic-resolution crystal structure of the antiviral lectin scytovirin. Protein Sci 16:2756–2760. doi:10.1110/ps.073157507
16. Case D, Darden T, Cheatham T, Simmerling C, Wang J, Duke R et al (2012) AMBER 12. http://ambermd.org/. Accessed 24 July 2015
17. Kirschner KN, Yongye AB, Tschampel SM, González-Outeiriño J, Daniels CR, Foley BL et al (2008) GLYCAM06: a generalizable biomolecular force field. Carbohydrates. J Comput Chem 29:622–655. doi:10.1002/jcc.20820
18. Wickstrom L, Okur A, Simmerling C (2009) Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. Biophys J 97:853–856. doi:10.1016/j.bpj.2009.04.063
19. Wang J, Wolf RM, Caldwell JW, Kollman PA, Case DA (2004) Development and testing of a general amber force field. J Comput Chem 25:1157–1174. doi:10.1002/jcc.20035
20. Hou T, Wang J, Li Y, Wang W (2011) Assessing the performance of the MM/PBSA and MM/GBSA methods: 1. The accuracy of binding free energy calculations based on molecular dynamics simulations. J Chem Inf Model 51:69–82. doi:10.1021/ci100275a
21. Hayes JM, G. Archontis (2012) MM-GB(PB)SA calculations of protein-ligand binding free energies. Mol Dyn. doi:10.5772/37107
22. Sulea T, Cui Q, Purisima EO (2011) Solvated interaction energy (SIE) for scoring protein–ligand binding affinities. 2. Benchmark in the CSAR-2010 scoring exercise. J Chem Inf Model 51:2066–2081. doi:10.1021/ci2000242
23. Moulaei T, Stuchlik O, Reed M, Yuan W, Pohl J, Lu W et al (2010) Topology of the disulfide bonds in the antiviral lectin scytovirin. Protein Sci 19:1649–1661. doi:10.1002/pro.445