Anal Bioanal Chem (2010) 396:2559–2563
DOI 10.1007/s00216-010-3498-x

SHORT COMMUNICATION

Ligand-protein docking studies of potential HIV-1 drug compounds using the algorithm FlexX

George Patargias · Gary Ewart · Carolyn Luscombe · Wolfgang B. Fischer

Received: 31 October 2009 / Revised: 19 January 2010 / Accepted: 20 January 2010 / Published online: 18 February 2010

Abstract Four compounds are docked to a pentameric bundle representing the transmembrane part of the Vpu protein from HIV-1. Employing the docking algorithm FlexX, their free energy of binding is estimated leading to the conclusion that potential drug candidates need to form H-bonds either with neighbouring or with $n+2$ helices at the site of the serines within the bundle.

Keywords Docking approach · Vpu · HIV-1 · Ligand binding · Viral membrane proteins

Introduction

The viral genome of HIV-1 encodes a series of small regulatory proteins one of which is Vpu [1, 2]. Vpu consists of 81 amino acids and is a type I integral membrane protein [3, 4] found in subcellular units of the infected cell [5, 6].

Two functions are known [7, 8] to be associated with two topologically distinct regions of Vpu [9]: (1) the transmembrane (TM) domain is essential to facilitate the release of the newly emerging virus particles [9, 10]. This seems to be due to the homo-oligomerization of the protein [11] consequently inducing channel activity [9, 10]. The results are based on experiments with full-length Vpu [9, 10, 12] and with a synthesised peptide construct corresponding to the TM segment of Vpu [13] both reconstituted into lipid membranes. This implements the idea that the protein alters the electrochemical gradient across particular subcellular units. (2) The second role of Vpu is to initiate the downstream regulation of the CD4 receptor for which the cytoplasmic part is responsible. More recently, it has been discovered that Vpu interacts with membrane-bound host factors such as the TWIK-related acid-sensitive K^+ channel (TASK; TWIK: Tandem of P domains in a weak inward rectifying K^+ channel) which exhibits background current in cells [14] and the interferon-inducible protein bone marrow stromal cell antigen 2 (BST-2), also called tetherin [15, 16].

The structural picture of Vpu (for a review, see [17]) is based on solution [18, 19] and solid-state NMR [12, 20, 21], CD [22], and FTIR spectroscopic investigations [23]. From these investigations, a helical transmembrane domain is proposed, the cytoplasmic domain comprises of two helices. In dependence of the experimental conditions under which the solution NMR data were taken, residues from Pro-75 onward adopt either a turn [18] or a helix [19] motif. In its function as a channel, Vpu is assumed to homo-oligomerise into a pentameric assembly [24–26].

Recently, it has been suggested that a derivative of amiloride has channel blocking effects [27, 28] and putative binding sites have been suggested with docking simulations

G. Patargias · W. B. Fischer
Biomembrane Structure Unit, Department of Biochemistry,
Oxford University,
South Parks Road,
Oxford OX1 3QU, UK

G. Ewart · C. Luscombe
Biotron Limited,
Suite 1.9, 56 Delhi Rd,
North Ryde, NSW 2113, Australia

W. B. Fischer
Bionanotechnology Interdisciplinary Research Collaboration,
Clarendon Laboratory, Department of Physics, Oxford University,
Parks Road,
Oxford OX1 3PU, UK

W. B. Fischer (✉)
Institute of Biophotonics,
School of Biomedical Science and Engineering,
National Yang-Ming University,
155, Sec. 2, Li-Nong St.,
Taipei 112, Taiwan
e-mail: wfischer@ym.edu.tw

[29]. In this study, the docking results of a series of other small molecules interacting with a pentameric bundle of the transmembrane domains of Vpu are presented using the FlexX software (www.biosolveit.de). From two of the derivates with the lowest binding energy, one of them has shown excellent antiviral activity and is currently in clinical trials.

Methods

Preparation of the ligands

The structures of four given compounds (Table 1) were designed with the molecular modelling package DS ViewerPro 5.0 (Accelrys). All of the ligand models were, initially, made planar and their geometry was minimised using the package's internal intermolecular force field. The ligands' coordinates were saved in both PDB and MDL MOL file format. The latter was manually edited in order to match the specifications of the SYBYL MOL2 file format required as input by FlexX. Consequently, the ligand structures were loaded to FlexX.

For all compounds, a protonated and deprotonated form was generated. Also, two stages of "aromaticity" were assumed; one with maximum "aromaticity" (delocalised) and one in which the aromatic ring systems were separated from the guanidinium moiety (localised). In total, 12 structures were generated for which their docking energies are listed in Table 2.

First, the ligand structure was minimised and loaded into the FlexX software. Flexible ring conformations were computed by CORINA, a 3D structure generator interfaced with FlexX. The delocalisation and the protonation state of the ligands were varied so that four combinations were produced for each ligand: protonated/localised, protonated/delocalised, unprotonated/localised, and unprotonated/delocalised. All ligands had flexible torsion angles during docking.

Secondly, the protein structure was loaded. In our case, that was a pentameric alpha-helical bundle of Vpu. A radius for the size of the protein's active site was then defined. The value of the radius was chosen to be 35 Å including the whole structure of the Vpu bundle and this was assumed to be fixed throughout the docking calculations.

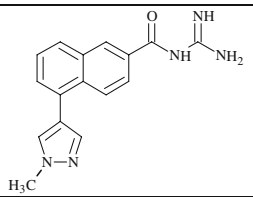
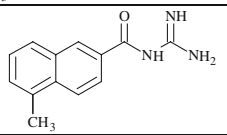
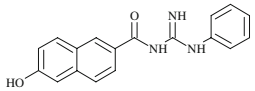
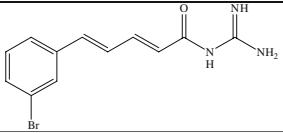
On both ligand and protein, formal charges were assigned.

Finally, the ligand was placed in the protein following an adapted version of a computer vision algorithm described in [30].

Preparation of the Vpu bundles

Pentameric helical $Vpu_{1\text{-}32}$ bundle was generated with the program X-PLOR as described elsewhere [31, 32] with the following sequence:MQPIPIVAIV^{10}ALVVAIIIAI^{20} VVWSIVIIEY^{30}RK. In brief, the bundle structure went through a simulated annealing/molecular dynamics protocol. The single helix structure was generated based on the positions of the Cα atoms, replicated and rotated in

Table 1 Chemical name and structure of the four compounds used in this study

Name	Code	Structure
N-(5-(1-methyl-1H-pyrazol-4-yl)naphthalene-2-carbonyl)guanidine	**A**	
5-methyl-2-napthoylguanidine or N-(5-Methyl-naphthalene-2-carbonyl)-guanidine	**B**	
N-(6-Hydroxy-2-napthoyl)-N'-phenylguanidine or N-(6-Hydroxy-naphthalene-2-carbonyl)-N'-phenyl-guanidine	**C**	
5-(3'-bromophenyl)penta-2,4-dienoylguanidine or N-[5-(3-Bromo-phenyl)-penta-2,4-dienoyl]-guanidine	**D**	

order to generate the bundle model of a pentameric Vpu_{1-32}. The following restraints were applied for the helices of the bundle: (1) an initial tilt of 5° was included to allow for the proper packing of the helices; (2) the interhelical distance was set initially to 9.4 Å[31, 32]; and (3) Ser-24 was oriented towards the lumen of the pore. The interhelical distance is chosen in agreement with distances found in helix packing motifs for other proteins [33–35]. The helices were built by superimposing all side chain atoms on the Cα atoms in the first stage. Gradually, the side chain atoms evolved to their proper location. At this first stage, only van der Waals interactions were taken into account and the calculations were repeated five times. In the second stage, electrostatic interactions were included and each of the five bundles underwent five times a simulated annealing protocol followed by a short molecular dynamics (MD) simulation. The averaged root mean square deviation was calculated for each of the 25 bundles and the one with the lowest value chosen for the docking experiments.

Prior to docking, a protein active site area was defined by "cutting out" all the residues that fall within a sphere of 35 Å around the minimised ligand structure. The radius was chosen so that the whole bundle was included in the search of the ligand–bundle complex. All protein atoms in this sphere are considered by FlexX as active site atoms. In our case, the active site atoms were all the protein atoms except the hydrogens.

The structures of the molecules are shown in Table 1 and denoted as A, B, C, and D for cross-reference in Table 2.

Results

Due to the common guanidinium group, each of the compounds can exist in a protonated and unprotonated state. While, under physiological conditions, the guanidinium groups are likely to be protonated (for example, the pK_a value of amiloride is 8.7 [36]), in this study, the unprotonated form is also investigated.

For compounds A (−18.6 kJ/mol) and D (−16.7 kJ/mol) in their protonated and localised states, the docking results indicate the lowest energy binding conformations to be inside the pore (Table 2). However, for compounds B and C, the lowest binding energies (B −18.7 kJ/mol; C −20.0 kJ/mol) are for conformations outside the pore. It is the 6th and the 3rd solutions in the ranking for B (17.0 kJ/mol) and C (−19.0 kJ/mol), respectively, which are suggested to be within the pore. For compound A in its delocalised stage, none of the first 20 solutions are placed within the pore. Compounds B (−12.5 kJ/mol) and D (−10.1 kJ/mol) are also found in their delocalised states to be within the pore. For compound C, the software does not generate any solution.

For all compounds in their unprotonated state, the favoured binding site is within the pore in the order of lower binding energy: D≈B>A>C for the localised state. Delocalisation leads to the same sequence with the exception that again docking of C does not lead to any solution. The sequence suggested is in accordance with experimental findings (based on IC50 vales, G. Ewart, C. Luscombe unpublished results).

The binding interaction within the pore is similar for all compounds, wherein the guanidinium group forms hydrogen bonds (1.55–2.37 Å) with at least one of the serines (Ser-24) of the Vpu peptide (for all compounds see Fig. 1). Hydrogen bonds are formed with the oxygen atoms of the side chains as well as with the carbonyl oxygens of the backbone. In addition, the aromatic ring system of the compounds allows for weak π-+ (CH_3) interactions with adjacent hydrophobic residues such as Val-21 or Ile-28 (4–5 Å).

Compound C allows for an additional hydrogen bonding pattern with another serine residue of the n+2 helix, with n

Table 2 Estimated calculated binding energies ΔG_{calc} for the compounds used in this study and found in the pore ΔG_{calc} is given in kilojoules per mole. Numbers in brackets refer to the number of solutions produced by FlexX

Protonation State	Protonated		Unprotonated	
Electron localisation Compound	Localised ΔG_{calc} k_i	De-localised ΔG_{calc} k_i	Localised ΔG_{calc} k_i	De-localised ΔG_{calc} k_i
A	−18.6 (584) 0.55	(434)[a] –	−16.0 (421) 1.6	−18.9 (447) 0.49
B	−17.0 (552) 1.0	−12.5 (400)	−15.4 (503) 2.0	−16.3 (397)
C	−19.0 (655) 0.47	–[b] –	−21.9 (516) 0.14	(0)[c] –
D	−16.7 (418) 1.2	−10.1 (33) 17	−15.1 (501) 2.3	−14.8 (432) 2.5

[a] None of the first 20 solutions is in the pore

[b] Compound is not possible to be generated by FlexX

[c] No solution is obtained with FlexX. The estimated binding constant is given in millimolar (assuming T=298 K)

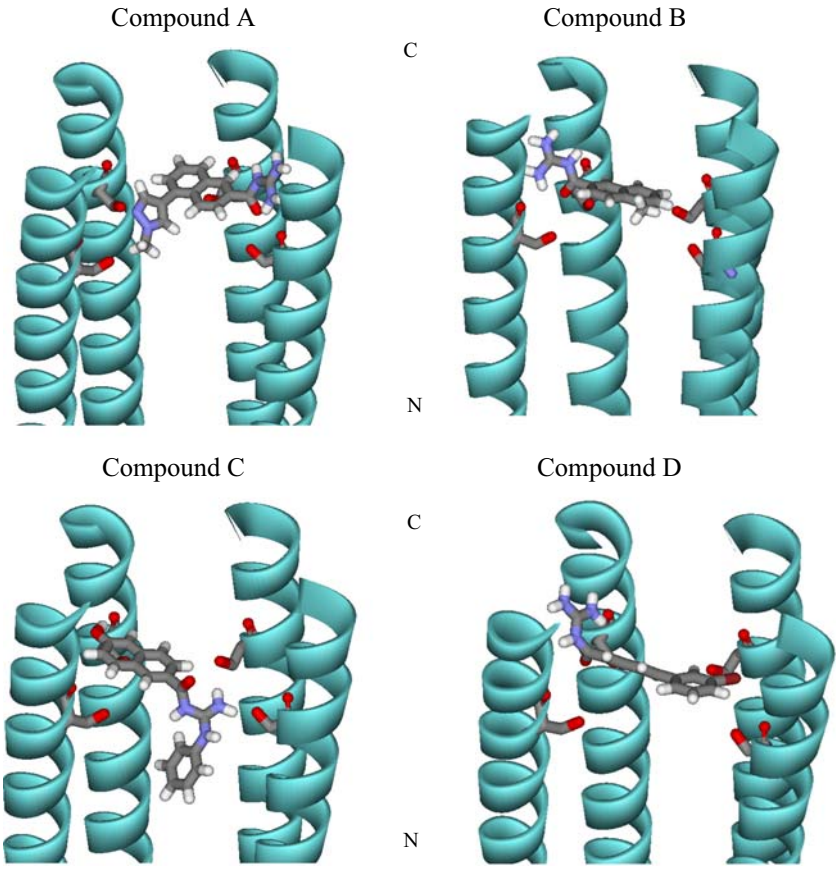


Fig. 1 Close view of *Compound A* (*upper panel*, *left*), *B* (*upper panel*, *right*), *C* (*lower panel*, *left*) and *D* (*lower panel*, *right*) in their protonated form in a close up view (colour-coding CPK). C and N denote the termini of the bundle (*light blue*). Ser-24 of each of the helices is shown. In this side view, one helix is omitted to allow a view inside the pore

the helix which forms hydrogen bonds with the guanidinium group of the compound (Fig. 1). Compound A allows for a similar pattern where the second hydrogen bond with the $n+2$ helix is formed by the N atoms of the pyrrole ring (Fig. 1).

Binding outside the pore generates structures with compounds on either side of the bundle, with a slight preference for the C terminus where Glu-29 and Trp-30 enables hydrogen bonding and π-+ interactions, respectively, with the compounds.

Discussion

The compounds in the unprotonated state and with the aromaticity localised to the guanidinium group support structural flexibility so that the compound can fit into the pore.

For docking inside the pore, it seems to be important that a maximum number of hydrogen bonds can be achieved with the serines, either on the level of the side chains or on the level of the backbone. The special separation of the hydrogen bond forming sites on the compound need to match the sites within the protein, which in this case would be sites defined by interhelical distances. The binding site can be between neighbouring helices or between the $n+2$ helix.

In the present study, the compound with the lowest docking energy (compound C) has two hydrogen bonding sites. Such sites would allow for high solubility of the compound in water. With the idea of kinked helices [37] forming the bundle [38] a vestibule may be formed at the C terminus allowing those compounds to diffuse more easily towards the binding site.

Some of the best values for binding are also found outside the pore suggesting a novel mode of inhibition.

Conclusions

The estimated binding free energies (see Table 2) rank the compounds in the following order: $Compound\ C \approx Compound\ A > Compound\ B \approx Compound\ D$ for those compounds found within the pore. In some cases, the lowest energy structures are also suggested to be outside the pore. For future drug development, potential drug candidates would need to posses two H bond donor/acceptor systems within distances so that they can form proper H-bonds with neighbouring or $n+2$ helices at the site of the serines.

Acknowledgement We thank the people from BiosolveIT GmbH and Molecular Networks GmbH for supplying the software.

References

1. Fischer W B (2005) Viral membrane proteins: structure, function and drug design. In: Atassi MZ (Ed) Protein Reviews, Vol. 1. Kluwer/Plenum
2. Fischer WB, Krüger JI (2009) Int Rev Cell Mol Biol 275:35
3. Strebel K, Klimkait T, Martin MA (1988) Science 241:1221
4. Cohen EA, Terwilliger EF, Sodroski JG, Haseltine WA (1988) Nature 334:532
5. Friborg J, Ladha A, Göttlinger H, Haseltine WA, Cohen EA (1995) J Acc Im Def Syn Hum Retr 8:10
6. Schubert U, Schneider T, Henklein P, Hoffmann K, Berthold E, Hauser H, Pauli G, Porstmann T (1992) Eur J Biochem 204:875
7. Klimkait T, Strebel K, Hoggan MD, Martin MA, Orenstein JM (1990) J Virol 64:621
8. Kimura T, Nishikawa M, Ohyama A (1994) J Biochem 115:1010
9. Schubert U, Bour S, Ferrer-Montiel AV, Montal M, Maldarelli F, Strebel K (1996) J Virol 70:809
10. Ewart GD, Sutherland T, Gage PW, Cox GB (1996) J Virol 70:7108
11. Maldarelli F, Chen MY, Willey RL, Strebel K (1993) J Virol 67:5056
12. Marassi FM, Ma C, Gratkowski H, Straus SK, Strebel K, Oblatt-Montal M, Montal M, Opella SJ (1999) Proc Natl Acad Sci USA 96:14336
13. Schubert U, Ferrer-Montiel AV, Oblatt-Montal M, Henklein P, Strebel K, Montal M (1996) FEBS Lett 398:12
14. Hsu K, Seharaseyon J, Dong P, Bour S, Marbán E (2004) Molec Cell 14:259
15. van Damme N, Goff D, Katsura C, Jorgensen RL, Mitchell R, Johnson MC, Stephens EB, Guatelli J (2008) Cell Host Microbe 3:1
16. Neil SJD, Zang T, Bieniasz PD (2008) Nature 451:425
17. Fischer WB (2003) FEBS Lett 552:39
18. Federau T, Schubert U, Floßdorf J, Henklein P, Schomburg D, Wray V (1996) Int J Pept Protein Res 47:297
19. Willbold D, Hoffmann S, Rösch P (1997) Eur J Biochem 245:581
20. Wray V, Kinder R, Federau T, Henklein P, Bechinger B, Schubert U (1999) Biochemistry 38:5272
21. Henklein P, Kinder R, Schubert U, Bechinger B (2000) FEBS Lett 482:220
22. Wray V, Federau T, Henklein P, Klabunde S, Kunert O, Schomburg D, Schubert U (1995) Int J Pept Protein Res 45:35
23. Kukol A, Arkin IT (1999) Biophys J 77:1594
24. Grice AL, Kerr ID, Sansom MSP (1997) FEBS Lett 405:299
25. Cordes FS, Tustian A, Sansom MSP, Watts A, Fischer WB (2002) Biochemistry 41:7359
26. Krüger J, Fischer WB (2009) J Chem Theory Comput 5:2503
27. Ewart GD, Mills K, Cox GB, Gage PW (2002) Eur Biophys J 31:26
28. Römer W, Lam YH, Fischer D, Watts A, Fischer WB, Göring P, Wehrspohn RB, Gösele U, Steinem C (2004) J Am Chem Soc 126:16267
29. Kim CG, Lemaitre V, Watts A, Fischer WB (2006) Anal Bioanal Chem 386:2213
30. Rarey M, Kramer B, Lengauer T, Klebe G (1996) J Mol Biol 261:470
31. Cordes F, Kukol A, Forrest LR, Arkin IT, Sansom MSP, Fischer WB (2001) Biochim Biophys Acta 1512:291
32. Lemaitre V, Ali R, Kim CG, Watts A, Fischer WB (2004) FEBS Lett 563:75
33. Bowie JUJ (1997) Mol Biol 272:780
34. Liu W, Crocker E, Constantinescu SN, Smith SO (2005) Biophys J 89:1194
35. Walters RFS, DeGrado WF (2006) Proc Natl Acad Sci USA 203:13658
36. Cuthbert AW (1976) Cell Mol Life Sci 32:1321
37. Park SH, Opella SJ (2005) J Mol Biol 350:310
38. Krüger J, Fischer WB (2008) J Comp Chem 29:2416