**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
© Springer-Verlag 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...
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 derivatives 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.

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 35Å 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 Vpu1-32 bundle was generated with the program X-PLOR as described elsewhere [31, 32] with the following sequence:MQPIPIVAIV10ALVVAIIIAI20VVWSIVIIEY30RK. 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

<table>
<thead>
<tr>
<th>Name</th>
<th>Code</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-(5-(1-methyl-1H-pyrazol-4-yl)napthalene-2-carbonyl)guanidine</td>
<td>A</td>
<td><img src="image" alt="Structure A" /></td>
</tr>
<tr>
<td>5-methyl-2-naphthoylguanidine</td>
<td>B</td>
<td><img src="image" alt="Structure B" /></td>
</tr>
<tr>
<td>N-(5-Methyl-naphthalene-2-carbonyl)-guanidine</td>
<td></td>
<td><img src="image" alt="Structure C" /></td>
</tr>
<tr>
<td>N-(6-Hydroxy-2-naphthoyl)-N'-phenylguanidine or N-(6-Hydroxy-naphthalene-2-carbonyl)-N'-phenyl-guanidine</td>
<td>C</td>
<td><img src="image" alt="Structure D" /></td>
</tr>
<tr>
<td>5-(3'-bromophenyl)penta-2,4-dienoylguanidine or N-[5-(3-Bromo-phenyl)-penta-2,4-dienoyl]-guanidine</td>
<td>D</td>
<td><img src="image" alt="Structure E" /></td>
</tr>
</tbody>
</table>
order to generate the bundle model of a pentameric Vpu1-32. 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 hydroxynium groups are likely to be protonated (for example, the pKa 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 values, 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+ (CH3) 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>
<thead>
<tr>
<th>Table 2 Estimated calculated binding energies ∆Gcalc for the compounds used in this study and found in the pore ∆Gcalc is given in kilojoules per mole. Numbers in brackets refer to the number of solutions produced by FlexX</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Protonation State</strong></td>
</tr>
<tr>
<td><strong>Electron localisation</strong></td>
</tr>
<tr>
<td><strong>Compound</strong></td>
</tr>
<tr>
<td><strong>A</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>B</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>C</strong></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td><strong>D</strong></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

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)
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 \( \pi-\pi \) 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 ≈ Compound A > Compound B ≈ 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 possess 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