J Mol Model (2009) 15:281–294
DOI 10.1007/s00894-008-0401-1

ORIGINAL PAPER

Computer-based design of novel HIV-1 entry inhibitors: neomycin conjugated to arginine peptides at two specific sites

Alexander Berchanski · Aviva Lapidot

Received: 28 July 2008 / Accepted: 3 October 2008 / Published online: 5 December 2008

Abstract Aminoglycoside–arginine conjugates (AAC and APAC) are multi-target inhibitors of human immunodeficiency virus type-1 (HIV-1). Here, we predict new conjugates of neomycin with two arginine peptide chains binding at specific sites on neomycin [poly-arginine-neomycin-poly-arginine (PA-Neo-PA)]. The rationale for the design of such compounds is to separate two short arginine peptides with neomycin, which may extend the binding region of the CXC chemokine receptor type 4 (CXCR4). We used homology models of CXCR4 and unliganded envelope glycoprotein 120 (HIV-1_{IIIB} gp120) and docked PA-Neo-PAs and APACs to these using a multistep docking procedure. The results indicate that PA-Neo-PAs spread over two negatively charged patches of CXCR4. PA-Neo-PA–CXCR4 complexes are energetically more favorable than AACs/APAC–CXCR4 complexes. Notably, our CXCR4 model and docking procedure can be applied to predict new compounds that are either inhibitors of gp120–CXCR4 binding without affecting stromal cell-derived factor 1α (SDF-1α) chemotaxis activity, or inhibitors of SDF-1α–CXCR4 binding resulting in an anti-metastasis effect. We also predict that PA-Neo-PAs and APACs can interfere with CD4–gp120 binding in unliganded conformation.

Keywords CXCR4/SDF-1α · gp120/CD4 · HIV-1 entry inhibitors · Molecular modeling and docking · Poly-arginine aminoglycoside conjugates · Predicted compounds

A. Berchanski · A. Lapidot (✉)
Department of Organic Chemistry,
The Weizmann Institute of Science,
Rehovot 76100, Israel
e-mail: aviva.lapidot@weizmann.ac.il

Introduction

Blocking human immunodeficiency virus type 1 (HIV-1) entry to its host human cells has clear therapeutic advantages over blocking subsequent stages in the life cycle of the virus. The well-coordinated multistep HIV-1 cell entry process provides multisite targeting at the entrance door through which HIV-1 enters cells. So, major efforts towards understanding the process by which HIV-1 enters its main host cells (CD4+ T-cells) continue to generate extensive interest [1, 2], and novel entry inhibitors are constantly being developed and studied (e.g. [3–5]). Potent cooperative and synergistic inhibition of HIV-1 proliferation has been observed in in vitro studies with several entry inhibitor combinations that interact with different steps of the HIV-1-cell entry cascade (e.g. [6]). Targeting a compound to several steps of viral cell entry and also to subsequent steps in the viral life cycle promises even more effective therapy, by reducing the probability that HIV-1 will develop resistance [7–9]. Thus, a single drug that can target multiple sites and/or steps in the viral life cycle has obvious advantages in clinical use.

The viral envelope protein plays a critical role in HIV-1 entry to cells. HIV-1 entry is initiated by the interaction of the viral envelope glycoprotein 120 (gp120) with the host cell receptor CD4—mainly with CXC chemokine receptor type 4 (CXCR4) and CC chemokine receptor 5 (CCR5). The CXCR4 receptor and its only natural chemokine ligand, stromal cell-derived factor 1α (SDF-1α), are crucial for embryonic development, and have been implicated in various pathological conditions, including HIV-1 infection and cancer metastasis [10]. An additional critical step in HIV-1 infection is efficacious transactivation of viral genes in the infected host cell. The selective transactivation of viral genes is facilitated by the HIV-1 transactivator protein

(Tat). Interestingly, an arginine-rich basic peptide, derived from Tat (positions 48–60), has been reported to have the ability to translocate through the cell membrane and accumulate in the nucleus. Various arginine-rich peptides with potent translocational activity very similar to Tat (48–60), including such peptides in which L-arginines were substituted with D-arginines have also been reported [11]. Several peptide-derived and other small molecule inhibitors of CXCR4- and CCR5-mediated HIV-1 infection have been reviewed by Pierson et al. [8].

CXCR4 belongs to the superfamily of G-protein-coupled receptors (GPCR). GRCRs are one of the most important drug targets and > 30% of all marketed therapeutics act on them [12, 13].

Optimal cellular and nuclear uptake was reported to be more effective for arginine polymers of 7–9 mers in length, when compared to similar lengths of lysine polymers [14]. Based on peptide models of HIV-1 Tat responsive element (TAR) RNA binding, NMR structures of TAR–ligand complexes and aminoglycoside–RNA interactions, we designed and synthesized a set of conjugates of aminoglycoside antibiotics with arginine (AACs, Fig. 1) (reviewed in [9, 15]). These AACs inhibited infectivity of a panel of HIV-1 laboratory and clinical isolates in a variety of cell lines and primary lymphocytes with EC_{50} (50% effective concentration) in the low micromolar range (0.8–4.0) [16, 17]. Of note, no antiviral activity up to 200 μM neomycin B was observed [18]. AACs inhibit binding of the monoclonal antibody (mAb) 12G5 to CXCR4 +cells with a somewhat lower potency than SDF-1α, the interaction of NeoR6 being the most potent [16, 17].

The finding that the *hexa*–arginine–neomycin conjugate (NeoR6; which contains six arginine moieties conjugated to the three pyranoside rings of neomycin B) is the most efficient anti-HIV-1 compound among of the AACs tested [9] prompted us to question whether conjugation of neomycin (or other members of this aminoglycoside group, e.g., neamine and paromomycin) with poly-arginine (6- and 9-mers), would lead to more potent HIV-1 inhibitors than a manifold of arginine conjugated via the amino groups of the aminoglycosides. Thus, a new set of poly-arginine 6-mer and 9-mer aminoglycoside conjugates (APACs) was designed and synthesized, and their cellular uptake and antiviral activities were determined [18].

To further understand how APACs might block gp120 interaction with CXCR4, and to investigate whether AACs and APACs compete with the natural ligand (SDF-1α) binding to CXCR4, we have undertaken an approach combining experimental validation and molecular modeling to propose plausible three-dimensional (3D) structural models of CXCR4 and its complexes with AACs and APACs, and common binding sites on CXCR4 of the compounds gp120, SDF-1α and mAb 12G5. Moreover, we predict new conjugates of neomycin with two arginine peptide chains binding at specific sites on rings I and IV of neomycin (poly-arginine-neomycin-poly-arginine, PA-Neo-PAs, Fig. 1). The rationale for the design of such compounds is to separate two short arginine peptides (3–5

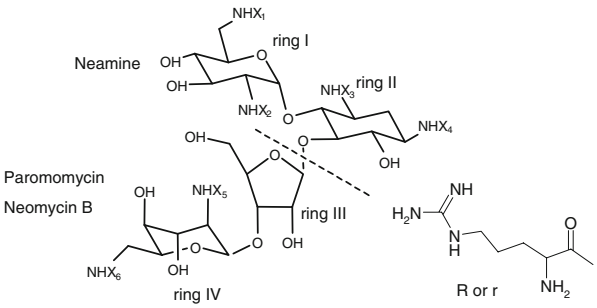

Fig. 1 Schematic representation of poly-arginine-neomycin-poly-arginine (PA-Neo-PAs), aminoglycoside–arginine conjugates (APACs), and aminoglycosides mentioned in this study. *R* L-arginine, *r* D-arginine

Compound	X_1	X_2	X_3	X_4	X_5	X_6
Neomycin B	H	H	H	H	H	H
NeoR6	R	R	R	R	R	R
Neo-r6	r6-NHAc	H	H	H	H	H
Neo-r9	r9-NHAc	H	H	H	H	H
r4-Neo-r4	r4-NHAc	H	H	H	H	r4-NHAc
r5-Neo-r5	r5-NHAc	H	H	H	H	r5-NHAc
Neamine	H	H	H	H	-	-
Neam-r9	r9-NHAc	H	H	H	-	-

amino acids) by neomycin, which may extend the binding regions of CXCR4. Indeed, our CXCR4 model [19, 20] comprises two negatively charged patches separated by neutral and positively charged residues that bind to the two predicted compounds: r4-Neo-r4 and r5-Neo-r5.

Our recently reported homology model of CXCR4 [19, 20] is in agreement with three published models of CXCR4 constructed using bovine rhodopsin as a template [21–23]. The validity of our CXCR4 model is supported by the following observations: (1) The polar residues of the extracellular domains of CXCR4 are exposed to water, and the transmembrane helices are stabilized by multiple hydrogen bonds and non-bonded interactions; (2) according to the predicted structure of the SDF-1α-CXCR4 complex [22], the 22 CXCR4 residues that participate in CXCL12 binding and signaling (Table 1) are available for direct interaction in the CXCR4 model; (3) residues identified as important for HIV-1 coreceptor activity and SDF-1α binding [24] are in line with our model; and (4) according to data collected from several studies, 32 CXCR4 residues participate in the CXCR4-gp120 complex (Table 2). Most of these residues are available for direct interactions in our CXCR4 model. Taken together, the experimental findings regarding the structure of CXCR4 and its ligand binding, found by ourselves and others [21–28], are consistent with our model, suggesting that it is appropriate for further docking investigations, such as those presented in this paper.

Previously, we identified mutations in constant regions C3 and C4 in NeoR6-resistant (NeoR6^{res}) viral isolates, and in variable region V4 of gp120, as well as two mutations in heptad repeat 2 (HR2) of gp41 [9]. Importantly, two of the three gp120 mutations (I341T and Q398K) are located in the CD4-binding cavity of gp120. These results led us to suggest that NeoR6 may also inhibit HIV-1 entry by interfering with gp120–CD4 binding [19]. Our findings indicated that the most potent APAC, nona-D-arginine-neomycin-conjugate (Neo-r9, Fig. 1) inhibits a variety of T-tropic HIV-1 isolates, including NeoR6^{res} isolates. Since the structure of Neo-r9 is relatively similar to that of NeoR6, we suggested that, like NeoR6, Neo-r9 may also inhibit HIV-1 entry by interfering with CD4–gp120 binding. Moreover, because Neo-r9 is significantly more positively charged than NeoR6, we predicted that it can also interact with the gp120 residues of NeoR6^{res} strains [19]. Docking studies of putative drugs with available HIV-1 gp120 structures representing the CD4-bound state (e.g., [29, 30]) have some limitations in light of the significant conformational changes that occur in gp120 upon CD4 binding [31]. Since we suggested that NeoR6 and Neo-r9 interact with gp120 prior to its binding with CD4, we have recently constructed a model of the unliganded HIV-1$_{IIIB}$ gp120 core using the unliganded simian immunodeficiency virus (SIV) gp120 (PDB entry 2bf1 [32]) as a template [19]. The gp120 cores of SIV and HIV-1 have 35–37% sequence identity and over 70% sequence similarity (depending on the HIV-1 isolate); alignment of these sequences is unambiguous [32]. Published data from mutagenesis and antigenic reactivity experiments [33] were used to validate our unliganded HIV-1 gp120 homology model [19].

Based on all the above, we here propose that our novel neomycin-poly-arginine conjugates (PA-Neo-PAs and APACs) competitively inhibit gp120–CXCR4 binding, and predict their binding site, which maps to the mAb 12G5/SDF-1α/gp120 binding region. We also predict that, like Neo-r9 and NeoR6, PA-Neo-PAs can interfere with CD4–gp120 binding in the unliganded conformation.

Methods

We docked PA-Neo-PAs and APACs to homology models of CXCR4 and unliganded HIV-1$_{IIIB}$ gp120 using a multistep docking procedure: geometric-electrostatic docking full scan by MolFit; flexible ligand docking by AutoDock4 and final refinement of the obtained complexes by Discover3 (Fig. 2).

Homology modeling

We have recently presented homology models of CXCR4 and unliganded HIV-1$_{IIIB}$ glycoprotein gp120 [19, 20] [using the "Homology" module of InsightII (Accelrys, San Diego, CA)], which were derived from the 3D coordinates of bovine rhodopsin (PDB entry 1l9 h [34]) and unliganded SIV glycoprotein gp120 (PDB entry 2bf1 [32]) as templates, respectively.

In brief, loops in which insertions or deletions occur were generated using the "Homology" module. The initial loops were then checked for clashes that were eliminated by selecting different rotamers for the side chains. In general, all energy minimization and molecular dynamics simulations in this study were performed for molecules immersed in a layer of 5 Å water until the maximum derivative was less than 0.001 (using the "Discover 3" module of InsightII and the CVFF force field). The modeled structures (including generated loops) were further optimized by 20 intermittent minimization and molecular dynamics stages. Two disulfide bridges, Cys28–Cys274 and Cys109–Cys186, of CXCR4 were used as distance restraints in the construction of the model. All conserved disulfide bridges of gp120 were also used as distance restraints in energy minimization. The IIIB gp120 with mutations I341T and Q398K was modeled by replacing the side-chains of these residues using the "Biopolymer" module of InsightII. The model of the mutant gp120 was further energy-minimized,

Table 1 Putative ligand interaction sites on CXC chemokine receptor type 4 (CXCR4)—structural analysis of receptor–ligand complexes

Protein/ compound	Regions																								
	Nter	Met 1	Glu 2	Gly 3	Ile 4	Ser 5	Ile 6	Tyr 7	Thr 8	Ser 9	Asp 10	Asn 11	Tyr 12	Thr 13	Glu 14	Glu 15	Met 16	Asp 20	Tyr 21	Asp 22	Ser 23	Met 24	Glu 26	Pro 27	Cys 28	Glu 32
SDF-1α[a]			+	+	+	+	+	+	+	+	+	+		+	+	+	+	+	+	+	+				+	
gp120[b]			+					+			+		+			+		+	+	+	+		+			+
Neo-r9[c]			+																	+		+				
r4-Neo-r4		+	+	+																+						
r5-Neo-r5		+	+	+																+	+	+		+		
Neam-r9			+																+	+		+				
NeoR6[d]																				+				+		
Neo-r6																			+	+						
	EL1	Phe 104	Gly 105	Phe 107																						
gp120[a]		+		+																						
NeoR6[d]		+	+																							
	EL2	Ala 175	Asn 176	Val 177	Ser 178	Glu 179	Ala 180	Asp 181	Asp 182	Arg 183	Tyr 184	Asp 187	Arg 188	Phe 189	Tyr 190	Pro 191	Asn 192	Asp 193	Leu 194	Val 196	Gln 200					
12G5			+			+	+	+	+						+											
SDF-1α[a]												+	+	+	+											
gp120[b]			+			+			+	+	+	+	+	+	+	+		+			+					
NeoR6[d]						+	+	+	+	+	+															
Neo-r9[c]			+	+	+	+		+	+	+					+	+	+	+	+							
r4-Neo-r4					+	+	+	+								+	+	+								
r5-Neo-r5		+	+		+	+	+	+	+	+					+	+	+	+	+	+						
Neo-r6			+	+	+	+		+	+		+	+														
Neam-r9			+	+		+		+	+						+		+	+	+							
	EL3	Asp 262	Ser 263	Phe 264	Ile 265	Leu 266	Leu 267	Glu 268	Gln 272	Glu 277	Asn 278	Ile 284														
SDF-1α[a]		+							+	+	+	+														
gp120[b]		+						+	+	+																
Neo-r9[c]		+	+	+	+		+	+																		
r4-Neo-r4		+				+	+	+																		
r5-Neo-r5		+				+		+	+																	
	TM2	Asp 97	TM5	Phe 201	TM7	Glu 288																				
SDF-1α[a]		+		+		+																				
gp120[b]		+				+																				

[a] [22]
[b] [21–28]
[c] [19]
[d] [19, 20]

Table 2 Final docked energies, interaction energies, and binding free energies from molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) calculations for the final ligand–CXCR4 complexes

Compound	Energy (kcal/mol)		
	Final docked energy ranges (AutoDock4)	Interaction energy (refined complex)	Binding free energies (MM-PBSA)
r5-Neo-r5	−45.09 to -37.51	−4,297	−123.9
r4-Neo-r4	−41.45 to -31.77	−3,720	−99.3
Neo-r9	−38.34 to -30.01	−3,611[a]	−97.1[a]
Neo-r6	−31.05 to -23.65	−2,503	−67.3
Neam-r9	−33.44 to -29.89	−3,335	−85.3

[a] [19]

the Cα atoms of the mutant gp120 being constrained to their initial positions.

In the present study, the amino acid numbering of IIIB gp120 is based on the sequence presented in [35], which differs from the amino acid numbering of JR-FL gp120 (PDB entry 2b4c, chain g [30]) by 30 residues. For example, Ile341 and Q398 of IIIB gp120 correspond to Ile371 (341 + 30=371) and Gln428 (398 + 30=428) of JR-FL gp120, respectively. This relation approximately holds for all the residues of these two sequences.

Full rigid-body docking of PA-Neo-PAs and APACs to CXCR4 and gp120

Models of PA-Neo-PAs and APACs were constructed by addition of arginines to neomycin B (3D coordinates taken from PDB entry 1qd3 [36]) using the "Biopolymer" module of InsightII. We used the first NMR structure (out of 17) from 1qd3. Neam-r9 was built by trimming aminoglycoside rings III and IV. All these modeled structures were solvated in water and energy minimized. PA-Neo-PAs (r5-Neo-r5 and r4-Neo-r4) and APACs (Neo-r6 and Neam-r9) were docked to gp120 and CXCR4 using the program MolFit (kindly provided by M. Eisenstein [37, 38]). Prior to the rigid body docking step, hydrogens and water molecules were omitted. MolFit treats the molecules as rigid bodies. Molecules are represented by 3D grids in which each grid point carries information concerning its position with respect to the surface/interior of the molecule. The surface grid points carry chemical information such as the electrostatic potential or the hydrophobicity of the surface. MolFit performs an exhaustive scan of the relative rotations and translations of the molecules and produces a list of models. Since gp120 and CXCR4 consist of many negatively charged residues, while PA-Neo-PAs and APACs are highly positively charged compounds, we performed full geometric-electrostatic scans [39] of these ligands against CXCR4 and gp120. The electrostatic potential around each of the docked molecules is calculated by solving the linearized Poisson-Boltzman equation, using the finite-differences method as implemented in the program Delphi [40, 41]). For this purpose, we used the program Delphi implemented in the InsightII package with default settings. The calculations are performed on a fine grid (less than 1 Å), producing accurate estimates of the potential. The calculation of the potential is separate from the docking procedure (implemented in a computer program named MolFit), which reads the potential files for the two molecules, together with the necessary data regarding the grid interval and the origin of the potential grid. The program Delphi uses a grid to calculate the potential, and this grid does not necessarily correspond to the MolFit grid points. This is a result of the different requirements of Delphi and MolFit regarding the grid interval. The potential

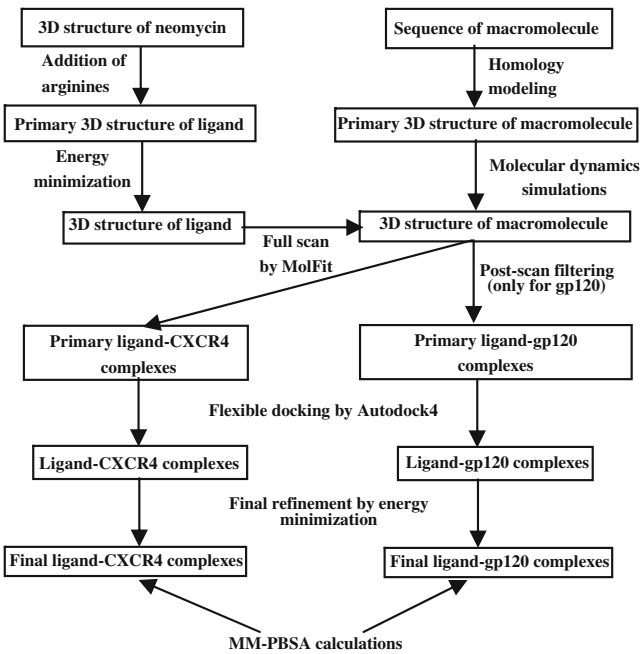

Fig. 2 Schematic representation of the molecular modeling and the multistep docking procedure followed in this study

grid of less than 1 Å in our case is too fine for the geometric representation of molecules in MolFit (1.2 Å). The different grid intervals make it impossible to map the potential grid directly onto the MolFit grid. Therefore, the potential grid was translated by MolFit into potential spheres with a radius $r_{pot} = \left(\sqrt{3/2}\right)*h*f$, where h is the potential grid interval, and f is an adjustable parameter (additional details regarding the electrostatic version of MolFit are described in Heifetz et al. [39].

Potential grid points with an electrostatic potential of more than −3 and less than 3 kT/e (i.e., relatively neutral) were disregarded. The MolFit default electrostatic descriptor (0.35) was used in all scans. Each full docking scan tested 8,760 relative orientations (translation and rotation grid intervals of 1.2 Å and 12°, respectively, were used), and for each orientation the best scoring solution was saved. These solutions were further sorted by their geometric-electrostatic complementarity score. After docking scans, the carbohydrates were manually added to the appropriate glycosylation sites on $gp120_{IIIB}$ [35]. The results of each scan were filtered in such a manner that the MolFit solutions that have contacts with glycosyls were omitted from the list.

Of note, all our tested ligands (PA-Neo-PAs and APACs) were in the unbound state only at the step of their construction and associated energy minimizations in water. In the rigid body docking (MolFit) step, they kept their unbound conformation, but in the following docking step they were flexible.

Flexible docking of PA-Neo-PAs and APACs to CXCR4 and gp120

The geometric-electrostatic MolFit scan solutions were used as input for the flexible docking program AutoDock4 [42]. In this step of our multistep docking procedure (as in the rigid body MolFit step), ligands and macromolecules were not solvated in water. We tested several possible binding sites of PA-Neo-PAs and APACs to CXCR4 and gp120 formed by the largest MolFit clusters. We used a grid spacing of 0.2 Å and 128×128×128 points. Other AutoDock4 parameters were used at their default settings. The aminoglycoside rings were defined as fixed roots. Since the maximum number of active torsions allowed was 32, 3–4 active torsions per side-chain of each arginine moiety were permitted. The aminoglycoside rings were defined as fixed roots. We permitted flexibility for two CXCR4 residues: Glu2 and Glu179.

The final refinement of top-scored docking conformations obtained from AutoDock4 was performed by energy minimization using Discover3 with default settings. Before this step, ligand and macromolecule were separated from the obtained AutoDock4 complex (without change of conformation) and again independently immersed in a 5 Å layer of water, as in the step of construction of unbound structures. This was performed to provide a water layer between ligand and macromolecule (Discover3 permits overlapping of two water layers). Ligands and macromolecules were then re-assembled and energy minimization was performed. The intracellular domains and transmembrane helices of CXCR4 were fixed, except for several residues that are close to the N-terminus, and the extracellular loops. The energy of the finally refined complexes was measured by the "Docking" module of InsightII (command "Evaluation of Intermolecular energy") according to maximal lengths of ligands (25–31 Å).

The binding free energies were calculated by the molecular mechanics Poisson-Boltzmann surface area (MM-PBSA) method implemented in the Amber9 program package [43, 44].

Results and discussion

Computational methodologies have become a crucial component of many drug discovery programs; from hit identification to lead optimization, approaches such as ligand- or structure-based virtual screening techniques are used widely in many discovery efforts [45]. When the structure of a target is available, docking is used primarily as a hit-identification tool. Docking is often also used later during lead optimization, when modifications to known active structures can be tested quickly in computer models prior to compound synthesis [46]. In the following sections we describe our analyses based on computational methodologies of the interaction of PA-Neo-PAs and APACs with two main targets: cellular CXCR4, the main coreceptor used by T-tropic HIV-1 viral isolates, and gp120, the viral glycoprotein that interacts with both cellular receptors, CD4 and CXCR4.

Evaluation of electrostatic potentials

We calculated the electrostatic potentials of our CXCR4 model, gp120 (PDB entry 2b4c, chain g [30]), SDF-1α (PDB entry 2j7z), and the aminoglycoside-arginine conjugates Neo-r9, r5-Neo-r5, and r4-Neo-r4, by Delphi as implemented in the InsightII package (Fig. 3). The extracellular loops and N-terminus of CXCR4 possess strong negative potential, while gp120, especially the V3 loop and adjacent regions (which interact with CXCR4), have positive potential. SDF-1α is a natural ligand of CXCR4 and has predominantly positive potential, which permits its binding with the highly negatively charged extracellular domains of CXCR4. Since the V3 loop and

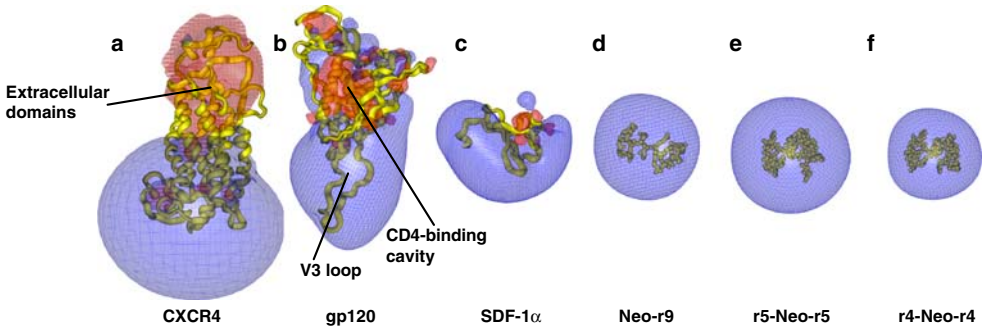

Fig. 3a–f Electrostatic potential maps calculated by Delphi as implemented in the InsightII package. **a** CXCR4 model, **b** gp120 [crystal structure of gp120 (PDB entry 2b4c, chain g)], **c** SDF-1α (PDB entry 2j7z), **d** Neo-r9, **e** r5-Neo-r5, **f** r4-Neo-r4; *red* −3 kT/e, *blue* + 3 kT/e. Neo-r9, r5-Neo-r5, and r4-Neo-r4 possess strong positive potential and could closely mimic SDF-1α and gp120 V3 loop and compete with the latter for binding with CXCR4

adjacent regions of gp120 also comprise strong positive potential, this glycoprotein binds easily to extracellular domains of CXCR4. The APACs and PA-Neo-PAs possess strong positive potential that roughly mimics SDF-1α and/or the gp120 V3 loop and thus could compete with them for binding with CXCR4. The sizes of APACs and PA-Neo-PAs are roughly comparable to those of SDF-1α and the gp120 V3 loop (Fig. 3). Since negative potential dominates at the gp120 CD4-binding cavity, we predict that r5-Neo-r5, and r4-Neo-r4 could also bind the CD4-binding cavity of gp120, as was suggested for Neo-r9 and NeoR6.

Docking of PA-Neo-PAs and APACs to CXCR4

Full geometric-electrostatic docking of PA-Neo-PAs and APACs to CXCR4

We performed full geometric and geometric-electrostatic docking scans of PA-Neo-PAs and APACs against CXCR4 using the program MolFit. The geometric docking of r5-Neo-r5 revealed its strong preference for binding CXCR4 at the mAb 12G5 binding site and adjacent regions, including multiple residues of the gp120 binding site. Geometric-electrostatic docking further restricted the possible interactions of r5-Neo-r5 to these regions only. The top ranking solution of the geometric-electrostatic scan is located exactly at the gp120 binding site. Furthermore, the 35 top ranking solutions formed a cluster in this region. The results of geometric-electrostatic docking for r4-Neo-r4, Neam-r9, and Neo-r6 were similar to those for r5-Neo-r5. The possibility that using another, quite different, conformer of neomycin B [root mean square deviation (RMSD) = 2.5 Å] for ligand construction might influence docking results was examined for Neo-r9. The cluster distributions performed by MolFit of rigid docking of this Neo-r9 "second conformer" to CXCR4 were similar to those obtained for the first-used conformer [19].

Flexible docking of PA-Neo-PAs and APACs to CXCR4

Top ranking ligand–CXCR4 complexes, obtained from the geometric-electrostatic scans, were used for flexible docking using AutoDock4. Two CXCR4 residues, Glu2 and Glu179, were flexible in these calculations. The final docked energy ranges for ligand–CXCR4 complexes obtained from AutoDock4 scans are shown in Table 2. The best ligand–CXCR4 complexes obtained from AutoDock4 scans were finally refined by energy minimizations performed for molecules immersed in a layer of 5 Å water using the "Discover 3" module of InsightII and the CVFF force field (results shown in Table 2). The values presented in Table 2 for AutoDock4 are related to "final docked energy" which includes: final intermolecular energy [van der Waals, hydrogen bonds, desolvation, electrostatic energy (Poisson-Boltzmann)] and final internal energy of the ligand. This definition of energy was used in AutoDock3 but not in the newer version AutoDock4, where sorting of conformers is executed by their "estimated free energy of binding", which comprises: final intermolecular energy + final total internal energy + torsional free energy − unbound system energy. However, we used values of final docked energy, since it was very interesting to compare results of APACs/PA-Neo-PAs docking with previously calculated results for Neo-r9 and AACs performed by AutoDock3 [19, 20]. Of note, in this step of our multistep docking procedure, the ligand and macromolecule are not immersed in water, so the docked energy comprises only the energy of the ligand and macromolecule. In contrast, the "interaction energy" of finally refined complexes by Discover3 was measured with ligand and macromolecule immersed in water. The interaction energy was calculated by the "Docking" module of InsightII, according to the maximal lengths of the ligands. Thus, this measurement related to the energy of the whole system (ligand/macromolecule/water) within

the restricted region (25–31 Å). This interaction energy is total intermolecular energy, which comprises van der Waals energy (repulsion and dispersion) and electrostatic (Coulomb) energy. The final total energy of the whole system (complex comprises ligand, macromolecule, and water) was 30,000±3,000 (for ligand/CXCR4 complex) and 24,000±3,000 kcal/mol (for ligand/gp120 complex) depending on the specific ligand.

We re-calculated the AutoDock step for Neo-r9 to compare its results with those of other APACs and PA-Neo-PAs presented in this report, since, in our previous study, AutoDock3 (in this version the macromolecule is non-flexible) was used for docking of this compound to CXCR4 and gp120 [19]. Although an improvement in final docked energy ranges for Neo-r9 was obtained, the overall shape of the best AutoDock4 Neo-r9-CXCR4 complex differs only slightly from that of the previous AutoDock3 Neo-r9-CXCR4 complex [19]. Thus, we did not refine it with Discover3 and used the results of the subsequent docking steps for this compound from our previous study [19] for comparison with those of APACs and PA-Neo-PAs presented in this study (Table 2).

The final r5-Neo-r5-CXCR4 complex with the lowest interaction energy (−4,297 kcal/mol) is shown in Fig. 4b, and Tables 1 and 2. All ten arginine residues of r5-Neo-r5 have multiple interactions (mostly electrostatic) with CXCR4 residues in the N-terminus, EL2 and EL3, including eight negatively charged residues. According to our structural model, nine residues of CXCR4 are common for CXCR4 binding with r5-Neo-r5 and gp120 (Table 1). The same docking procedure was performed for r4-Neo-r4 and APACs. In general, the final refined r4-Neo-r4-CXCR4 complex with the lowest interaction energy (Fig. 4c, Tables 1, 2) is similar to the r5-Neo-r5-CXCR4 complex. r4-Neo-r4 interacts with seven negatively charged CXCR4 residues, in comparison to eight electrostatic interactions of r5-Neo-r5; r4-Neo-r4 also has multiple common residues with gp120, suggesting strong (but less potent than r5-Neo-r5) competition with gp120-CXCR4 binding. Although r4-Neo-r4 has only eight arginines with CXCR4 in contrast to nine arginines of Neo-r9 (Table 1), the final docked energy of the r4-Neo-r4-CXCR4 complex (−3,720 kcal/mol) is even better than that of Neo-r9 (−3,611 kcal/mol) (Table 2). The finding that the interaction energy values of r5-Neo-r5-CXCR4 and r4-Neo-r4-CXCR4 complexes are lower than those of the Neo-r9-CXCR4 complex can be explained by the existence of two negatively charged patches separated by neutral and positively charged residues on the surface of CXCR4. The first patch consists of residues Asp22, Glu179, Asp181, and Asp182 (Fig. 4) and the second contains residues Glu2, Asp193, Asp262, and Asp268. Neo-r9 interacts with residues Asp262 and Asp268 only via the free amino groups of the neomycin core and not via the arginine moieties. In the r5-Neo-r5-CXCR4 and r4-Neo-r4-CXCR4 complexes, the neomycin core of aminoglycoside conjugates is situated approximately between these two patches. r5-Neo-r5 and r4-Neo-r4 form electrostatic interactions via the guanidino groups of the poly-arginine chains, resulting in more energetically favorable complexes with CXCR4. Thus, we predict that r5-Neo-r5 and r4-Neo-r4 may be more potent coreceptor (CXCR4) HIV-1 inhibitors than Neo-r9. This should be investigated experimentally.

The final CXCR4 interaction energy with Neam-r9 (−3,335 kcal/mol, Table 2), in which the neamine residue

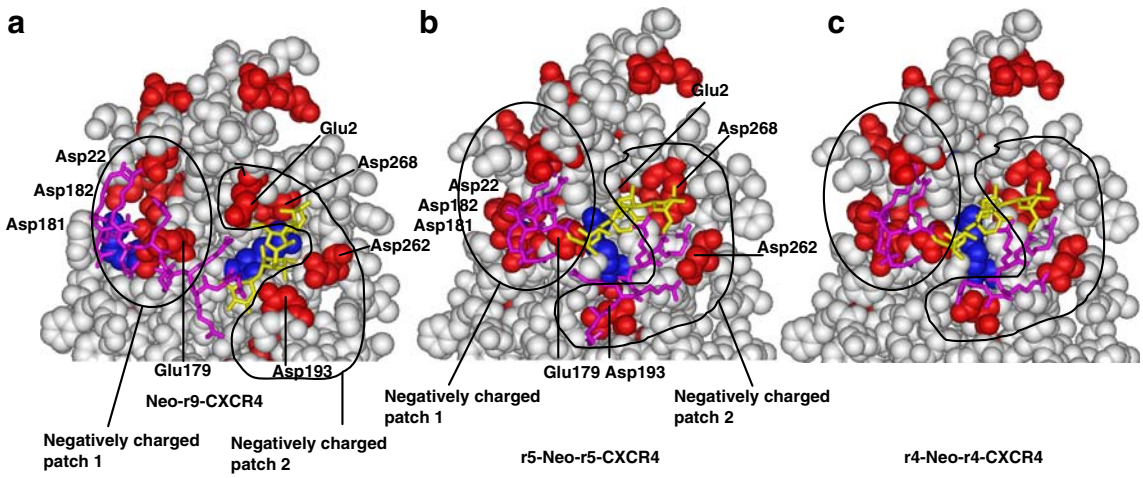

Fig. 4 **a** Neo-r9-CXCR4 complex. CXCR4 is shown in CPK representation; *red* negatively charged residues, *blue* positively charged residues. The Neo-r9 is shown in stick representation, the neomycin core is in *yellow* and arginine moieties are *magenta*. Two negatively charged patches separated by neutral and positively charged residues are visible. **b**, **c** r5-Neo-r5-CXCR4 (**b**) and r4-Neo-r4-CXCR4 (**c**) complexes. Representations and colors as in **a**

consists of only two rings (Fig. 1), is higher than with r5-Neo-r5, r4-Neo-r4, and Neo-r9. This may be due to the neamine two-ring core, which does not form interactions with Asp262 and Asp268 as is the case with the neomycin core. This complex also shows multiple common residues with gp120 binding. The final docked energy of Neo-r6-CXCR4 complexes obtained from AutoDock4 scans is higher than with other tested compounds. The final Neo-r6-CXCR4 complex (Tables 1, 2), which also has significantly higher interaction energy in comparison to r5-Neo-r5 due to its shorter arginine chain, reveals lower electrostatic interactions with CXCR4. However, Neo-r6 also interacts with six CXCR4 residues that are important for gp120 binding.

The model structures of PA-Neo-PAs and APACs-CXCR4 complexes presented in this study are consistent with our experimental findings: (1) APACs compete with anti-CXCR4 mAb 12G5 [18]; (2) APACs inhibit infection of cells with different HIV-1 clinical isolates and laboratory strains with EC_{50} in the range 0.9–4.8 μM [18]. PA-Neo-PAs and APACs structurally mimic the highly positively charged gp120 V3 loop. Of note, mutations D193K and D193A in CXCR4 significantly reduce CXCR4's ability to bind ALX40-4C (Nα-acetyl-nona-D-arginine amide) [47], which is structurally similar to the poly-arginine chains of PA-Neo-PAs and APACs. According to our docking results, Neo-r9, r5-Neo-r5, r4-Neo-r4 and Neam-r9 interact with D193 (Table 1). Several CXCR4 residues (Glu2, Asp22, Asn176, Glu179, Asp182, Tyr190, Asp193, Asp262, E268) are critical for binding more than one ligand (SDF-1α, gp120 or mAb 12G5, Table 1); they are also important in the binding of PA-Neo-PAs and APACs with CXCR4.

Previously, it was thought that T-20 (a well-known fusion and entry inhibitor) binds gp41 HR1 and interferes with the gp41 conformational changes required for membrane fusion. However, it was also found that T-20 blocks the interaction between the gp120-CD4 complex and the CXCR4 receptor by direct binding to the base of the V3 loop and β19 strand of the CXCR4-bound gp120 [48]. In contrast to T-20, PA-Neo-PAs and APACs cannot bind the V3 loop or β19 strand due to their highly positive charges, but they can interact with negatively charged residues of CXCR4, preventing gp120–CXCR4 binding. Thus, PA-Neo-PAs/APACs and T-20 can both interfere with the gp120–CXCR4 interaction, acting on opposite binding sites of this complex.

Can the interactions of PA-Neo-PAs and APACs with CXCR4 suggested by our model explain HIV-1 inhibition? Multiple CXCR4 residues are in common for binding with PA-Neo-PAs/APACs and gp120 (Table 1), thus supporting our experimental findings for APACs [18], i.e., that one of the mechanisms of the anti-HIV-1 activity of PA-Neo-PAs and APACs is interference with the gp120-CXCR4 interaction. For example, r5-Neo-r5 has contacts with 25 critical CXCR4 residues, including 9 common residues for gp120 binding to CXCR4 (Table 1). The model/docking results also show that PA-Neo-PAs and APACs interact with several CXCR4 residues participating in SDF-1α binding (i.e., Glu2, Asp22, Tyr190, Asp262) (Table 1). Moreover, PA-Neo-PAs and APACs bind to residues (Tyr190, Pro191 and Asp262) common to binding sites of well known CXCR4 antagonists, such as AMD3100 and T140 [23]. This is consistent with our experimental findings that APACs inhibit SDF-1α-induced cell migration [18]. In contrast to PA-Neo-PAs, APACs, T140 and AMD3100; AACs (e.g., NeoR6 and ParomoR5) interact with CXCR4 residues of the mAb 12G5 binding site, without overlapping with residues of the SDF-1α-CXCR4 binding site. Indeed, NeoR6 does not block SDF-1α-induced cell migration [20]. The sphere-like NeoR6-CXCR4 binding conformer [19] reveals a completely different structure in comparison to the extended structure of PA-Neo-PAs and APACs in complex with CXCR4 (Fig. 5).

The binding free energies, calculated by the MM-PBSA method implemented in the Amber9 program package [43, 44] for final complexes of CXCR4 with r5-Neo-r5, r4-Neo-r4, Neam-r9 and Neo-r6 in comparison with the previously described complex with Neo-r9 [19] (Table 2), revealed a binding free energy of the r5-Neo-r5-CXCR4 complex lower (−123.9 kcal/mol) than that of the r4-Neo-r4-CXCR4 (−99.3 kcal/mol), Neo-r9-CXCR4 (−97.1 kcal/mol), and Neam-r9-CXCR4 (−85.3 kcal/mol) complexes, and significantly lower than that of the Neo-r6-CXCR4 complex (−67.3 kcal/mol). These findings support our notion that the poly-arginine chains of PA-Neo-PAs and APACs contribute significantly more to the binding with CXCR4 than the aminoglycoside rings, thus the interaction energies of PA-Neo-PA and APACs-CXCR4 complexes are dependent mostly on the poly-arginine chain length rather than on the polar residues of the aminoglycoside rings. However, r4-Neo-r4, which has eight arginines, revealed a lower binding free energy in complex with CXCR4 than Neo-r9, which possesses nine arginines, suggesting that the overall shape of r4-Neo-r4 fits better for interaction with CXCR4. The values of binding free energy presented in Table 2 are related to the final estimated binding free energy calculated from the following terms: electrostatic energy calculated by molecular mechanics (MM) force field; van der Waals contribution from MM; internal energy arising from bond, angle and dihedral terms in the MM force field; nonpolar contribution to the solvation free energy calculated by an empirical model; electrostatic contribution to the solvation free energy, calculated by Poisson-Boltzmann.

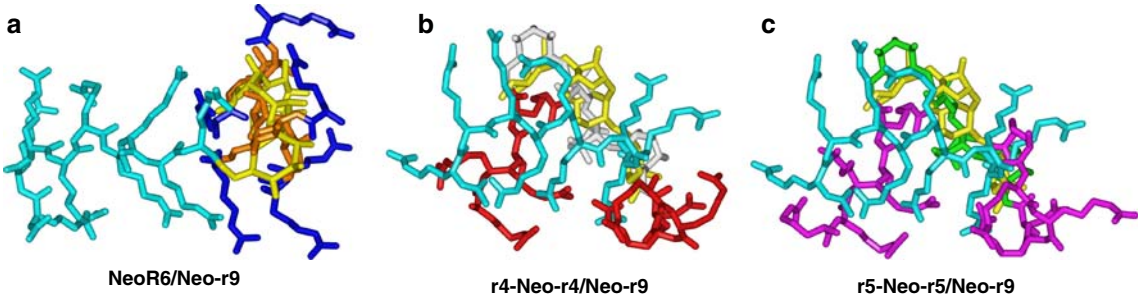

Fig. 5a–c Structural differences between aminoglycoside-arginine conjugates. Neomycin cores of NeoR6 (**a**), r4-Neo-r4 (**b**), and r5-Neo-r5 (**c**) are superimposed on the neomycin core of Neo-r9. Neomycin cores: *yellow* Neo-r9, *orange* NeoR6, *white* r4-Neo-r4, *green* r5-Neo-r5. Arginine moieties: *cyan* Neo-r9, *blue* NeoR6, *red* r4-Neo-r4, *magenta* r5-Neo-r5

Docking of PA-Neo-PAs to unliganded HIV-1_{IIIB} gp120

Full geometric-electrostatic docking of PA-Neo-PAs to unliganded HIV-1_{IIIB} gp120

Previously we found mutations in gp120 at the regions C3, C4, and V4 in $NeoR6^{res}$ isolates [49]. Our experimental findings also proved that NeoR6 does not inhibit mAb 2D7 binding to CCR5 [16], and does not compete with the CCR5 chemokine RANTES [17], suggesting that NeoR6 is not a CCR5 antagonist. However, NeoR6 inhibits HIV-1 JR-FL and Ba-L CCR5-using isolates with EC_{50} values in the range of 0.8–5 μM, similar to that of CXCR4-using isolates [16]. These results led us to suggest that NeoR6 may also inhibit HIV-1 entry by interfering with gp120–CD4 binding [16, 19]. The newly designed compounds 4r-Neo-4r and 5r-Neo-5r structurally resemble Neo-r9. A variety of T-tropic HIV-1 isolates, including $NeoR6^{res}$ strains are sensitive to Neo-r9, thus we predicted that 4r-Neo-4r and 5r-Neo-5r may also inhibit gp120-CD4 interaction.

Based on all the above findings, we present here the molecular docking of the predicted compounds against HIV-1_{IIIB} gp120. We performed a full geometric-electrostatic docking scan of PA-Neo-PAs against modeled HIV-1_{IIIB} gp120 using the program MolFit (see Methods and Fig. 2). Post-scan filtering was performed with the purpose of deleting docking solutions that overlap with glycosyls. This docking scan of r5-Neo-r5 against modeled HIV-1_{IIIB} gp120 revealed three clusters of MolFit solutions similar to the docking results of Neo-r9 [19]. According to our previously constructed model of unliganded trimer [19], cluster 1 is located at the region interacting with the V3 loop of the neighboring protomer, making this site inaccessible for ligand binding. Two remaining clusters (2 and 3) are located at the CD4-binding region of gp120, which includes the site of $NeoR6^{res}$ mutations I341T and Q398K. The top-ranking solution after post-scan filtering is located at cluster 3. The results of geometric-electrostatic docking for r4-Neo-r4 were similar to those of r5-Neo-r5.

Flexible docking of PA-Neo-PAs to unliganded HIV-1_{IIIB} gp120

The best-scoring representatives of these two clusters were used for flexible docking withy AutoDock4. The final

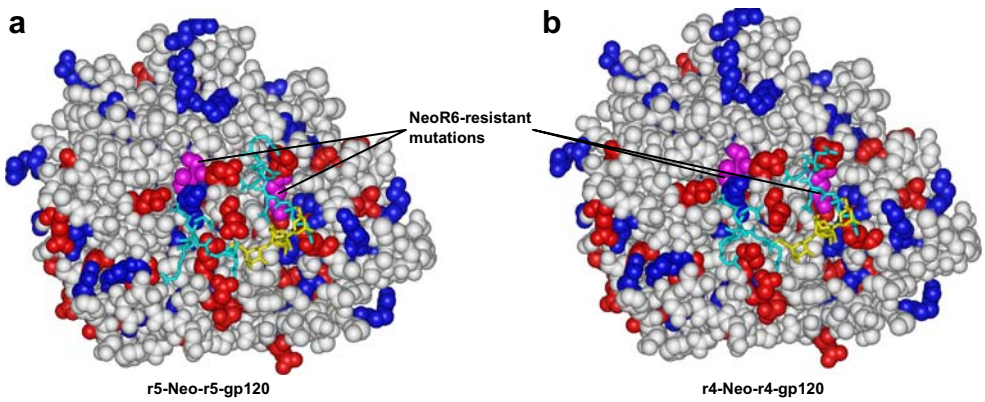

Fig. 6 **a** Complex between r5-Neo-r5 and modeled unliganded HIV-1 gp120 (IIIB isolate). Gp120 is shown in CPK representation; *red* negatively charged gp120 residues, *blue* positively charged residues, *magenta* NeoR6-resistant mutations, *yellow* neomycin core of r5-Neo-r5, *cyan* arginine moieties. Competition between r5-Neo-r5 and the CD4 receptor at this site is suggested. **b** The complex between r4-Neo-r4 and unliganded HIV-1 gp120 (IIIB isolate). Colors and representations as in **a**

Table 3 Binding sites on gp120 of CD4, Neo-r9, r4-Neo-r4, and r5-Neo-r5

gp120 region	Putative binding sites			
	CD4-gp120[a] (JR-FL isolate, PDB entry 2b4c)	Neo-r9[b] (IIIB isolate)	r4-Neo-r4 (IIIB isolate)	r5-Neo-r5 (IIIB isolate)
C1 V1/V2	Pro124, Cys126, Ala129		Phe63, Asn64, Met65, Trp66 Glu123	Asn64, Met65, Trp66 Glu123
C2	Asn279, Asn280, Ala281, Thr283			
C3	Ser365, **Gly366**, **Gly367**, **Asp368**, **Glu370**, **Ile371**	Ile329, Ile330, Phe331, Gln333, **Ser335**, **Gly336**, **Gly337**, **Asp338**, **Glu340**, **Ile341**, Val342	Lys327, Ile329, Ile330, Phe331, Lys332, Gln333, **Gly336**, **Gly337**, **Asp338**, **Pro339**, **Glu340**, **Ile341**	Lys327, Thr328, Ile330, Phe331, Lys332, Gln333, **Gly337**, **Asp338**, **Pro339**, **Glu340**, **Ile341**
V4		Glu373	Glu373	Glu373
C4	**Asn425**, **Met426**, **Trp427**, **Gln428**, **Glu429**, **Val430**, **Gly431**, Thr455, Arg456, Asp457, Gly459	Phe393, **Asn395**, **Met396**, **Trp397**, **Gln398**, **Glu399**, **Val400**, **Gly401**, **Asp427**, **Gly429**	**Trp397**, **Gln398**, **Glu399**, **Asp427**, **Gly429**	**Trp397**, **Gln398**, **Glu399**, **Asp427**, **Gly429**
V5	**Ile460**, Arg469	**Asn430**, Asn431, Glu434, Glu436	**Asn430**, Asn433, Glu434, Ser435, Glu436, Ile437	**Asn430**, Asn433, Glu434, Ser435, Glu436, Ile437
C5	Gly472, **Gly473**, **Asp474**, Asp477			

The gp120 residues that are mutated in the NeoR6^{res} isolates are *underlined*. HIV-1 gp120 residues common to CD4 and ligand binding are marked in *bold*

[a] [30]

[b] [19]

docked energy for r5-Neo-r5 ranged from −23.56 to −7.11 kcal/mol at cluster 2, and from −20.79 to −3.34 at cluster 3. r5-Neo-r5 at both these clusters has contacts with I341 and Q398, which are located at the CD4 binding site. The top-ranking AutoDock4 solutions for each putative site were refined by energy minimizations. The r5-Neo-r5-gp120 complex with the lowest energy (−2,598 kcal/mol) was obtained at the CD4 binding site (Fig. 6a, Table 3). via their guanidino groups, the eight arginine residues of r5-Neo-r5 have electrostatic interactions with gp120 residues Glu123, Asp338, Glu340, Glu373, Glu399, Asp427, Glu434, and Glu436; the neomycin core binds to gp120 residues via van der Waals interactions. Thus, the docking results reveal that r5-Neo-r5 binds gp120 preferentially at the CD4-binding site. The interaction energy (−2,577 kcal/mol) found for the best complex between r5-Neo-r5 and gp120 with NeoR6^{res} mutations I341T and Q398K, is similar to results for r5-Neo-r5 complexed with wild-type gp120$_{IIIB}$ (−2,598 kcal/mol), indicating that, in spite of the influence of these mutations, r5-Neo-r5 can successfully bind to NeoR6^{res} gp120.

Since the cluster distribution of MolFit solutions for r4-Neo-r4 is similar to that of r5-Neo-r5, we also performed the flexible docking in only the two same regions as for r5-Neo-r5. The final docked energy obtained by AutoDock4 for r4-Neo-r4 ranged from −22.43 to −5.63 at the CD4 binding site. The best finally refined r4-Neo-r4-gp120 complex at the CD4 binding site has an energy of −2,562 kcal/mol (Fig. 6b, Tables 3, 4). In general, the final r4-Neo-r4-gp120 complex is only slightly different from the r5-Neo-r5-CXCR4 complex. It also has eight electrostatic interactions with the same gp120 residues as r5-Neo-r5 (Tables 3, 4).

As in the case of CXCR4, we re-calculated the AutoDock step for Neo-r9 docking to gp120 only and used

Table 4 Final docked energies, interaction energies, and binding free energies from MM-PBSA calculations for final ligand-gp120 complexes

Compound	Energy (kcal/mol)		
	Final docked energy ranges (AutoDock4)	Interaction energy (refined complex)	Binding free energies (MM-PBSA)
Neo-r9	−20.47 to -3.83	−2,556[a]	−83.5[a]
R4-Neo-r4	−22.43 to -5.63	−2,562	−83.2
R5-Neo-r5	−23.56 to -7.11	−2,598	−87.3

[a] [19]

the results of subsequent docking steps for this compound from our previous study [19] for comparison with those of APACs and PA-Neo-PAs.

The binding free energies of PA-Neo-PA complexes with CXCR4 were consistently better than those of PA-Neo-PA complexes with wt-gp120$_{IIIB}$ (Tables 2, 4). However, the absolute values of binding free energies of PA-Neo-PAs-wt-gp120$_{IIIB}$ complexes are also sufficient. This preference of PA-Neo-PAs to bind CXCR4 rather than gp120 could be explained by the existence of two negatively charged patches separated by neutral and positively charged residues on the surface of CXCR4, and the use of two poly-arginine chains separated by a neomycin core resulting in more energetically favorable complexes with CXCR4. The negatively charged patch on gp120 at the CD4 binding site is more widely distributed, and a clear separation for two or more patches is not observed (Fig. 6). Since the docking results of PA-Neo-PAs to gp120 are also satisfactory, we predict that these compounds may also act as attachment inhibitors; however, their most important anti-HIV-1 effect is probably explained by their interference with gp120–CXCR4 binding.

The entropic contribution to binding was not estimated since we assumed that, in our specific cases, this contribution is not significant. All our tested ligands (PA-Neo-PAs and APACs) are highly positively charged compounds, and in the majority of docking steps they are solvated in water and flexible so the loss of entropy is not significantly large. In addition, CXCR4 and gp120 are relatively large proteins with negatively charged electrostatic patches, and they are also in an aqueous environment. In the docking steps containing energy calculations (e.g., Discover3) most of the protein surface (including the large hydrophobic transmembrane portions of CXCR4) was fixed, i.e., did not participate in binding, and its entropy was not changed in comparison to the previously unbound energy minimized state in water. There is some loss of entropy caused by docking but only at the ligand–macromolecule binding site. However, the loss of entropy for every conformer of a given compound is approximately similar. Thus, the free energies of a given ligand can be compared easily in the different systems (binding with CXCR4 and gp120), e.g., the r5-Neo-r5-CXCR4 complex is more favorable than the r5-Neo-r5-gp120 complex. More difficult to compare is entropy between different ligands, since entropy depends on the number of rotatable bonds. For example, binding of r5-Neo-r5, which harbors ten arginines, assumes a greater loss of entropy than binding of Neo-r9 (nine arginines) or r4-Neo-r4 (eight arginines). However, the difference in free binding energy between the r5-Neo-r5-CXCR4 complex (−123.9 kcal/mol) and the Neo-r9-CXCR4 complex (−97.1 kcal/mol) is so large that estimation of entropy loss could not significantly reduce this difference. In the case of the r4-Neo-r4-CXCR4 complex (−99.3 kcal/mol) and Neo-r9-CXCR4 complex (−97.1 kcal/mol), the estimation of entropy loss could have the opposite effect—the difference between the energies of these complexes could slightly increase, since Neo-r9 has more rotatable bonds than r4-Neo-r4.

Conclusion

To predict new compounds that could inhibit gp120–CXCR4 and CD4–gp120 binding we used our CXCR4 model, unliganded HIV-1$_{IIIB}$ gp120 and applied a multistep docking procedure including rigid body geometric-electrostatic full scans by MolFit [37–39], consequent flexible ligand docking by AutoDock4 [42] of top-scoring representatives of clusters found at the possible binding sites, and final refinement by energy minimizations of the obtained complexes. We conclude that the most favorable docked compound is r5-Neo-r5. Our results suggest that the gp120-binding site on CXCR4 is a more likely target for poly-arginine-neomycin-poly-arginines (PA-Neo-PAs) than the CD4-binding site on gp120. The structural models of unliganded HIV-1$_{IIIB}$ gp120, CXCR4 and interaction with HIV-1 inhibitors (PA-Neo-PAs and APACs) provide insight into their activity, and could be useful in structure-based drug design studies for selective targeting of HIV-1 entry.

Acknowledgments We thank Dr. G. Borkow and Dr. M. Eisenstein for critical remarks and discussions; the Amber Development Team for providing the Amber9 program package; the AutoDock Development Team for providing the AutoDock4 program package; Dr. M. Eisenstein for providing the docking program MolFit (downloaded from: http://www.weizmann.ac.il/Chemical_Research_Support/molfit).

References

1. Adamson CS, Freed EO (2008) Recent progress in antiretrovirals—lessons from resistance. Drug Discov Today 13:424–432. doi:10.1016/j.drudis.2008.02.003
2. van Lunzen J (2007) How will CCR5 antagonists influence the recommendations for the antiretroviral treatment of HIV-1 infection. Eur J Med Res 12:435–440
3. Jones R, Nelson M (2007) The role of receptors in the HIV-1 entry process. Eur J Med Res 12:391–396
4. Ho HT, Fan L, Nowicka-Sans B, McAuliffe B, Li CB, Yamanaka G, Zhou N, Fang H, Dicker I, Dalterio R, Gong YF, Wang T, Yin Z, Ueda Y, Matiskella J, Kadow J, Clapham P, Robinson J, Colonno R, Lin PF (2006) Envelope conformational changes induced by human immunodeficiency virus type 1 attachment inhibitors prevent CD4 binding and downstream entry events. J Virol 80:4017–4025. doi:10.1128/JVI.80.8.4017-4025.2006
5. Kadow J, Wang HG, Lin PF (2006) Small-molecule HIV-1 gp120 inhibitors to prevent HIV-1 entry: an emerging opportunity for drug development. Curr Opin Investig Drugs 7:721–726
6. Nagashima KA, Thompson DA, Rosenfield SI, Maddon PJ, Dragic T, Olson WC (2001) Human immunodeficiency virus type

1 entry inhibitors PRO 542 and T-20 are potently synergistic in blocking virus-cell and cell-cell fusion. J Infect Dis 183:1121–1125. doi:10.1086/319284
7. Cooley LA, Lewin SR (2003) HIV-1 cell entry and advances in viral entry inhibitor therapy. J Clin Virol 26:121–132. doi:10.1016/S1386-6532(02)00111-7
8. Pierson TC, Doms RW, Pohlmann S (2004) Prospects of HIV-1 entry inhibitors as novel therapeutics. Rev Med Virol 14:255–270. doi:10.1002/rmv.435
9. Borkow G, Lapidot A (2005) Multi-targeting the entrance door to block HIV-1. Curr Drug Targets Infect Disord 5:3–15. doi:10.2174/1568005053174645
10. Tachibana K, Hirota S, Iizasa H, Yoshida H, Kawabata K, Kataoka Y, Kitamura Y, Matsushima K, Yoshida N, Nishikawa S, Kishimoto T, Nagasawa T (1998) The chemokine receptor CXCR4 is essential for vascularization of the gastrointestinal tract. Nature 393:591–594. doi:10.1038/31261
11. Futaki S, Suzuki T, Ohashi W, Yagami T, Tanaka S, Ueda K, Sugiura Y (2001) Arginine-rich peptides. An abundant source of membrane-permeable peptides having potential as carriers for intracellular protein delivery. J Biol Chem 276:5836–5840. doi:10.1074/jbc.M007540200
12. Hopkins AL, Groom CR (2002) The druggable genome. Nat Rev Drug Discov 1:727–730. doi:10.1038/nrd892
13. Schlyer S, Horuk R (2006) I want a new drug: G-protein-coupled receptors in drug development. Drug Discov Today 11:481–493. doi:10.1016/j.drudis.2006.04.008
14. Mitchell DJ, Kim DT, Steinman L, Fathman CG, Rothbard JB (2000) Polyarginine enters cells more efficiently than other polycationic homopolymers. J Pept Res 56:318–325. doi:10.1034/j.1399-3011.2000.00723.x
15. Lapidot A, Berchanski A, Borkow G (2008) Insight into the mechanisms of aminoglycoside derivatives interaction with HIV-1 entry steps and viral gene transcription. FEBS J 275:5236–5257
16. Litovchick A, Lapidot A, Eisenstein M, Kalinkovich A, Borkow G (2001) Neomycin B-arginine conjugate, a novel HIV-1 Tat antagonist: synthesis and anti-HIV activities. Biochemistry 40:15612–15623. doi:10.1021/bi0108655
17. Borkow G, Vijayabaskar V, Lara HH, Kalinkovich A, Lapidot A (2003) Structure-activity relationship of neomycin, paromomycin, and neamine-arginine conjugates, targeting HIV-1 gp120-CXCR4 binding step. Antiviral Res 60:181–192. doi:10.1016/S0166-3542(03)00156-6
18. Hegde R, Borkow G, Berchanski A, Lapidot A (2007) Structure-function relationship of novel X4 HIV-1 entry inhibitors—L- and D-arginine peptide-aminoglycoside conjugates. FEBS J 274:6523–6536
19. Berchanski A, Lapidot A (2007) Prediction of HIV-1 entry inhibitors neomycin-arginine conjugates interaction with the CD4-gp120 binding site by molecular modeling and multistep docking procedure. Biochim Biophys Acta 1768:2107–2119. doi:10.1016/j.bbamem.2007.04.017
20. Lapidot A, Peled A, Berchanski A, Pal B, Kollet O, Lapidot T, Borkow G (2008) NeoR6 inhibits HIV-1-CXCR4 interaction without affecting CXCL12 chemotaxis activity. Biochim Biophys Acta 1780:914–920
21. Zhou N, Luo Z, Luo J, Liu D, Hall JW, Pomerantz RJ, Huang Z (2001) Structural and functional characterization of human CXCR4 as a chemokine receptor and HIV-1 co-receptor by mutagenesis and molecular modeling studies. J Biol Chem 276:42826–42833. doi:10.1074/jbc.M106582200
22. Huang X, Shen J, Cui M, Shen L, Luo X, Ling K, Pei G, Jiang H, Chen K (2003) Molecular dynamics simulations on SDF-1alpha: binding with CXCR4 receptor. Biophys J 84:171–184
23. Trent JO, Wang ZX, Murray JL, Shao W, Tamamura H, Fujii N, Peiper SC (2003) Lipid bilayer simulations of CXCR4 with inverse agonists and weak partial agonists. J Biol Chem 278:47136–47144. doi:10.1074/jbc.M307850200
24. Brelot A, Heveker N, Montes M, Alizon M (2000) Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. J Biol Chem 275:23736–23744. doi:10.1074/jbc.M000776200
25. Dragic T (2001) An overview of the determinants of CCR5 and CXCR4 co-receptor function. J Gen Virol 82:1807–1814
26. Thompson DA, Cormier EG, Dragic T (2002) CCR5 and CXCR4 usage by non-clade B human immunodeficiency virus type 1 primary isolates. J Virol 76:3059–3064. doi:10.1128/JVI.76.6.3059-3064.2002
27. Lin G, Baribaud F, Romano J, Doms RW, Hoxie JA (2003) Identification of gp120 binding sites on CXCR4 by using CD4-independent human immunodeficiency virus type 2 Env proteins. J Virol 77:931–942. doi:10.1128/JVI.77.2.931-942.2003
28. Tian S, Choi WT, Liu D, Pesavento J, Wang Y, An J, Sodroski JG, Huang Z (2005) Distinct functional sites for human immunodeficiency virus type 1 and stromal cell-derived factor 1alpha on CXCR4 transmembrane helical domains. J Virol 79:12667–12673. doi:10.1128/JVI.79.20.12667-12673.2005
29. Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA (1998) Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. Nature 393:648–659. doi:10.1038/31405
30. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD (2005) Structure of a V3-containing HIV-1 gp120 core. Science 310:1025–1028. doi:10.1126/science.1118398
31. Madani N, Perdigoto AL, Srinivasan K, Cox JM, Chruma JJ, LaLonde J, Head M, Smith AB III, Sodroski JG (2004) Localized changes in the gp120 envelope glycoprotein confer resistance to human immunodeficiency virus entry inhibitors BMS-806 and #155. J Virol 78:3742–3752. doi:10.1128/JVI.78.7.3742-3752.2004
32. Chen B, Vogan EM, Gong H, Skehel JJ, Wiley DC, Harrison SC (2005) Structure of an unliganded simian immunodeficiency virus gp120 core. Nature 433:834–841. doi:10.1038/nature03327
33. Clark S, Mellors JW, Calef CE (2005) Los Alamos National Laboratory by the HIV sequence database project. In: Korber BTM, Brander C, Haynes BF, Koup R, Moore JP, Walker BD, Watkins DI (eds) HIV molecular immunology. Los Alamos National Laboratory, Los Alamos, NM
34. Okada T, Fujiyoshi Y, Silow M, Navarro J, Landau EM, Shichida Y (2002) Functional role of internal water molecules in rhodopsin revealed by X-ray crystallography. Proc Natl Acad Sci USA 99:5982–5987. doi:10.1073/pnas.082666399
35. Capon DJ, Ward RHR (1991) The CD4-gpl20 Interaction and AIDS pathogenesis. Annu Rev Immunol 9:649–678
36. Faber C, Sticht H, Schweimer K, Rosch P (2000) Structural rearrangements of HIV-1 Tat-responsive RNA upon binding of neomycin B. J Biol Chem 275:20660–20666. doi:10.1074/jbc.M000920200
37. Katchalski-Katzir E, Shariv I, Eisenstein M, Friesem AA, Aflalo C, Vakser IA (1992) Molecular surface recognition: determination of geometric fit between proteins and their ligands by correlation techniques. Proc Natl Acad Sci USA 89:2195–2199. doi:10.1073/pnas.89.6.2195
38. Berchanski A, Shapira B, Eisenstein M (2004) Hydrophobic complementarity in protein-protein docking. Proteins 56:130–142. doi:10.1002/prot.20145
39. Heifetz A, Katchalski-Katzir E, Eisenstein M (2002) Electrostatics in protein–protein docking. Protein Sci 11:571–587. doi:10.1110/ps.26002
40. Klapper I, Hagstrom R, Fine R, Sharp K, Honig B (1986) Focusing of electric fields in the active site of Cu-Zn superoxide

dismutase: effects of ionic strength and amino-acid modification. Proteins 1:47–59. doi:10.1002/prot.340010109
41. Honig B, Nicholls A (1995) Classical electrostatics in biology and chemistry. Science 268:1144–1149. doi:10.1126/science.7761829
42. Morris GM, Goodsell DS, Halliday RS, Huey R, Hart WE, Belew RK, Olson AJ (1998) Automated docking using a Lamarcian genetic algorithm and empirical binding free energy function. J Comput Chem 19:1639–1662. doi:10.1002/(SICI)1096-987X(19981115)19:14<1639::AID-JCC10>3.0.CO;2-B
43. Gohlke H, Case DA (2003) Converging free energy estimates: MM-Pb(GB)SA studies on the protein-protein complex Ras-Raf. J Comput Chem 25:238–250. doi:10.1002/jcc.10379
44. Case DA, Cheatham TE, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods R (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688. doi:10.1002/jcc.20290
45. Kitchen DB, Decornez H, Bajorath J (2004) Docking and scoring in virtual screening for drug discovery: methods and applications. Nat Rev 3:935–949
46. Leach AR, Shoichet BK, Peishoff CE (2006) Docking and scoring. J Med Chem 49:5851–5855. doi:10.1021/jm060999m
47. Doranz BJ, Filion LG, Diaz-Mitoma F, Sitar DS, Sahai J, Baribaud F, Orsini MJ, Benovic JL, Cameron W, Doms RW (2001) Safe use of the CXCR4 inhibitor ALX40-4C in humans. AIDS Res Hum Retroviruses 17:475–486. doi:10.1089/08892220151126508
48. Yuan W, Craig S, Si Z, Farzan M, Sodroski J (2004) CD4-induced T-20 binding to human immunodeficiency virus type 1 gp120 blocks interaction with the CXCR4 coreceptor. J Virol 78:5448–5457. doi:10.1128/JVI.78.10.5448-5457.2004
49. Borkow G, Lara HH, Lapidot A (2003) Mutations in gp41 and gp120 of HIV-1 isolates resistant to hexa-arginine neomycin B conjugate. Biochem Biophys Res Commun 312:1047–1052. doi:10.1016/j.bbrc.2003.11.011