



RESEARCH

Open Access

Modelling binding between CCR5 and CXCR4 receptors and their ligands suggests the surface electrostatic potential of the co-receptor to be a key player in the HIV-1 tropism

Olga V Kalinina*, Nico Pfeifer and Thomas Lengauer

Abstract

Background: CCR5 and CXCR4 are the two membrane-standing proteins that, along with CD4, facilitate entry of HIV particles into the host cell. HIV strains differ in their ability to utilize either CCR5 or CXCR4, and this specificity, also known as viral tropism, is largely determined by the sequence of the V3 loop of the viral envelope protein gp120.

Results: With statistical and docking approaches we have computationally analyzed binding preferences of CCR5 and CXCR4 to both V3 loop sequences of virus strains of different tropism and endogenous ligands.

Conclusions: We conclude that the tropism cannot be satisfactorily explained by amino-acid interactions alone, and suggest a two-step mechanism, by which initial coreceptor selection and approach of the ligand to the binding pocket is dominated by charge and glycosylation pattern of the viral envelope.

Keywords: HIV-1 tropism, Structural model, Coreceptor usage, Entry inhibitor, Tropism test

Background

To enter a host cell, the HIV envelope protein gp120 binds to the cellular CD4 receptor and to one of the two cellular co-receptors, CCR5 or CXCR4. We call viruses (or their genome sequences) that can bind exclusively to CCR5 R5-tropic, those that can bind exclusively to CXCR4 X4-tropic and those that can bind to either coreceptor dual-tropic viruses. X4 and dual-tropic viruses together form the set of X4-capable viruses. Both CCR5 and CXCR4 belong to a prominent family of G-protein coupled receptors. They bind to chemokines: CCR5 binds to a number of inflammatory CC-chemokines including CCL3, CCL4 and CCL5 [1,2], CXCR4 is a receptor for SDF-1 [3] and extracellular ubiquitin [4]. Viral usage of one or the other coreceptor can vary, and is of pivotal importance for the correct choice of antiviral therapy with a class of drugs called entry inhibitors. The first CCR5 inhibitor, maraviroc, has entered clinical

practice in the year 2007 [5], but no inhibitor of CXCR4 is on the market yet. Since an entry inhibitor targeting CCR5 or CXCR4, respectively, can only be effective against viruses with the corresponding tropism, administering entry inhibitors, such as maraviroc, requires an advance test of viral tropism. There are two classes of such tests: phenotypic and genotypic. In phenotypic tests, viral tropism is determined in a laboratory assay. In genotypic tests the viral tropism is inferred from the relevant regions of the viral genome by computational means. Several computational prediction models based on the sequences of a part of the gp120 sequence called the V3 loop have been proposed [6-9]. Even though recently a genotypic test has become available that points to regions in the V3 loop determining tropism [10], neither class of tests provides a mechanistic understanding of the coreceptor choice. In this study, various methods of computational structural biology have been applied to elucidate this mechanism.

To date, only an experimentally resolved structure of nearly full-length gp120 with a short N-terminal peptide

* Correspondence: kalinina@mpi-inf.mpg.de
Department for Computational Biology and Applied Algorithmics, Max Planck Institute for Informatics, Campus E1 4, Saarbrücken 66123, Germany

of CCR5 [11] is available. This structure has been used recently in a modeling attempt to understand the mechanistic basis of viral tropism [12]. In contrast, we have studied interactions between V3 loop and the full-length co-receptor.

Results

Structural analysis of V3 loops corresponding to different virus subtypes does not reveal any tropism preferences

In our study, we used two sets of V3 loop sequences: one set of 7 R5-tropic and 13 X4-capable sequences with experimentally verified tropism [10] and another set of 47 R5-tropic and 49 X4-capable sequences with tropism predicted using *geno2pheno-C_NGS-Sanger* [8] (see Methods for more details).

The net charge of V3 loop sequences from the set of experimentally verified cases is significantly higher for X4-capable sequences than for R5-tropic sequences (Figure 1a). However, this may be a consequence of a bias during selection of sequences for experimental testing. The set of sequences with predicted tropism is more diverse and evenly distributed in sequence space than the sequences with experimentally verified tropism: the median and 1st and 3rd quartiles of the pairwise sequence identity are 0.76, 0.71, 0.79 for the predicted R5-tropic; 0.88, 0.88, 0.92 for the experimentally verified R5-tropic; 0.60, 0.68, 0.74 for the predicted X4-capable; and 0.65, 0.71, 0.76 for the experimentally verified X4-capable sequences. The difference of the sequence charge between the two tropism types for the predicted sequences is not very pronounced, although significant, and the median charge is equal for R5-tropic and X4-capable sequences. There is no significant correlation between the net charge of the V3 loop and the probability of exhibiting the X4-capable phenotype, as calculated by *geno2pheno-C_NGS-Sanger* [8] (p-values -0.17 and 0.25 for R5-tropic and X4-capable sequences, respectively, Figure 1b).

PepSite is a computational tool that scans the surface of a given protein for patches that are likely to bind individual amino acid residues or peptides up to ten amino acids [13,14], providing a score that reflects the propensity of the peptide to bind to the protein. The PepSite score is expressed in relative units and the higher scores mean better binding. We apply PepSite in a sliding window of 10 residues to assess the binding of the experimentally verified X4-capable and R5-tropic sequences to CXCR4, and acquire higher average scores for the X4-capable sequences (Figure 2A, low statistical significance here is due to the limited size of the sample). The same procedure for the sequences of the predicted set does not reveal such a trend (Figure 2B). Still, both experimentally verified and predicted X4-capable sequences have significantly higher propensity to bind to CCR5

than to CXCR4 (p-value 0.064 for experimentally tested and 6.156×10^{-08} for predicted sequences based on a two-sided Wilcoxon test). None of the X4-capable sequences has a score for CXCR4 binding that would exceed the score for CCR5 binding by more than 1.5 percentage points. This may indicate that these viruses are equally capable of using both CCR5 and CXCR4 for entry. For R5-tropic sequences, on the contrary, we find a statistically significant higher propensity to bind to CCR5 rather than to CXCR4 in both experimentally verified and predicted sequence sets, in agreement with the assigned phenotype.

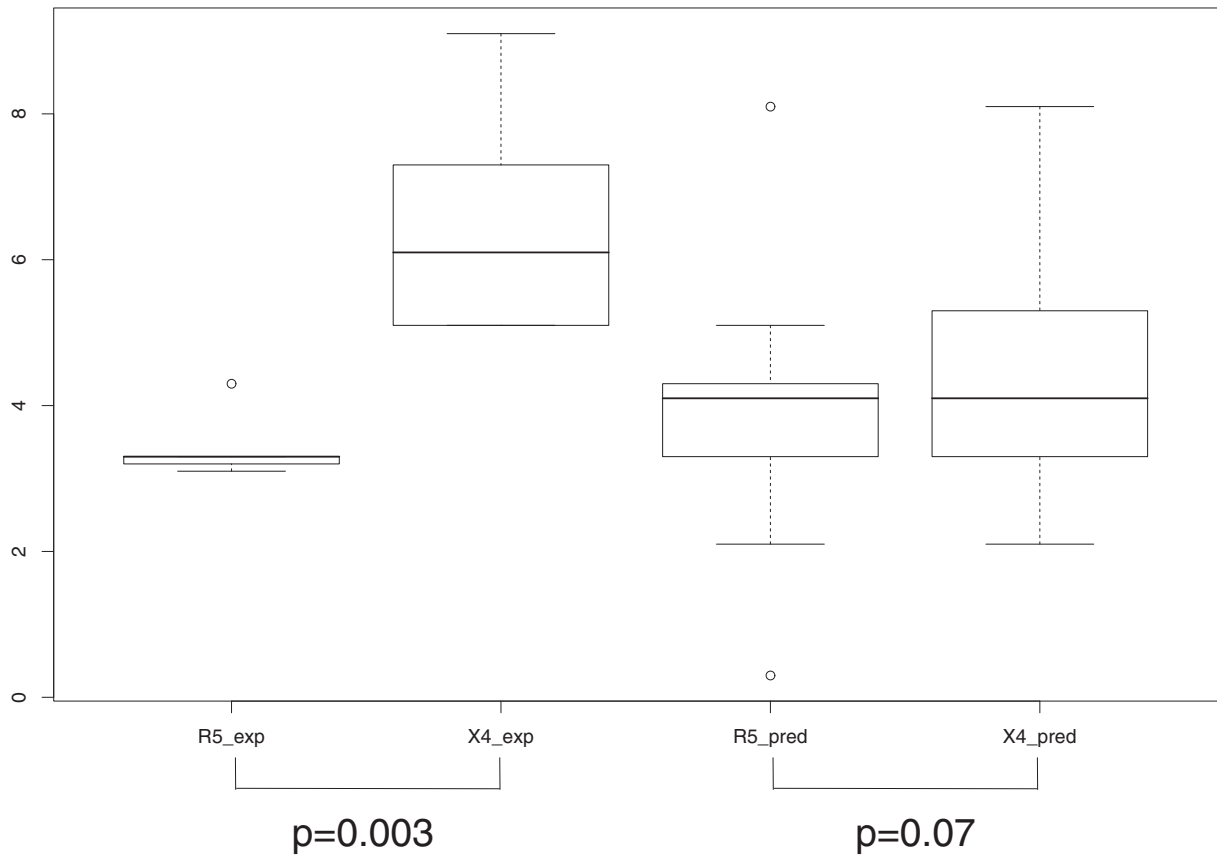
Both CCR5 and CXCR4 belong to GPCR family of transmembrane proteins and have seven transmembrane helices (Figure 3a). In the crystal structure, CXCR4 is bound to an inhibitor peptide (magenta in Figure 3a) that penetrates deeply into the extracellular pocket of the receptor. The hotspots for individual amino acids are predominantly places inside this pocket, as well. Proline residues are most favorably placed deeply inside the pocket (magenta in Figure 3c), which suggests that the β -hairpin structure of the V3 loop penetrates deeply into the binding pocket of the coreceptor [15]. Aromatic residues are more highly preferred along the channel, and charged residues are situated near the mouth of the pocket (Figure 3c). These preferences are not as pronounced in the CCR5 model as in the structure of CXCR4 (data not shown).

The docking experiment was unable to differentiate between X4-capable and R5-tropic sequences when they were placed into the pocket of CXCR4. The docking energies (FlexPepDock [16] energy scores between -25 and -40 for interface energy and -120 and -280 for whole-complex energy) and the docking poses of the whole loop and its parts do not significantly differ between the sequences of the two tropisms neither in the set of experimentally verified, not in the set of sequences with predicted tropism. Without more experimental information on the precise positioning of the V3 loop in the binding pocket of the coreceptor, it is probably impossible to model this interaction to anything approaching atomic detail. The structural bioinformatics approaches fail to further explain the process of coreceptor binding.

Structural comparison of the CCR5 and CXCR4 chemokine receptors

CCR5 and CXCR4 are chemokine receptors of the G-protein coupled receptor family, with a global sequence identity of 33%. CXCR4 was recently crystallized as a dimer [15], and it is reported to dimerize in cells independent of ligand binding [17]. The structure forms a large pocket on the extracellular side where the ligand binds. The same pocket is presumably used by the V3

a.



b.

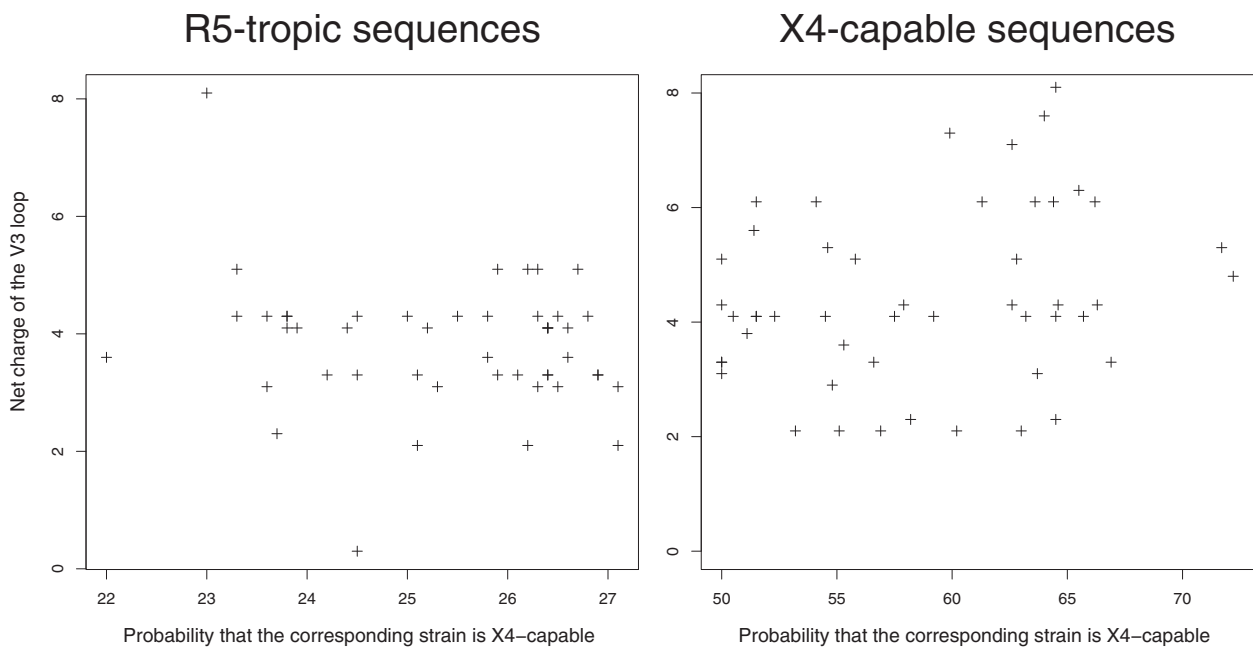


Figure 1 (See legend on next page.)

(See figure on previous page.)

Figure 1 Charge of V3 loop sequences. **a.** Charges of the V3 loops of experimentally tested sequences (exp) and sequences predicted to be prototypic for their tropism (pred). The box plots show median, first and third quartile ± 1.5 interquartile range. P-values displayed are based on a two-sided Wilcoxon test. **b.** Dot-plots showing the relationship between the net charge of the V3 loop and the probability of the corresponding strain to show the X4-capable phenotype, as calculated by geno2pheno-C_NGS-Sanger [8], for the set of predicted sequences. P-values based on Pearson's product-moment correlation test: -0.17 for R5-tropic and 0.25 for X4-capable sequences.

loop of gp120 [15]. The structure of CCR5 has not been experimentally resolved, so we have modeled it using a CXCR4 structure (PDB entry 3oe0) as a template. The model checks out as acceptable when tested with PROCHECK [18].

Despite their high sequence identity and hence structural similarity, the surface electrostatic potentials of the two receptors differ substantially (Figure 4): the binding pocket of CXCR4 and the area near its mouth are negatively charged, while CCR5 has a negatively charged pocket and a neutral to positive entrance to that pocket. The N-terminus that bears a significant negative charge due to sulfation [19,20] is disordered and absent from both structures.

Comparison of endogenous ligands

Both CCR5 and CXCR4 are chemokine receptors. CCR5 binds to a variety of CC-chemokines including CCL5, as well as CCL3 and CCL4 [1,2]. The ligands of CXCR4 are SDF-1 [3] and extracellular ubiquitin [4]. When structurally compared using DaliLite [21], CCL5, CCL4 and CCL3 exhibit significant structural similarity but little, while still detectable, resemblance to the CXCR4 ligand SDF-1 (Figure 5). CXCR4 ligands exhibit no structural similarity among each other. There is no structural similarity between endogenous ligands of CCR5 or CXCR4 and the structure of the V3 loop as it is crystallized in PDB entries 2b4c or 2qad.

The electrostatic potentials of the ligands are largely compatible with the electrostatic potential of the receptors: CCL3 and CCL4 have negatively charged surfaces, CCL5, although generally positive, is glycosylated [22], hence potentially it can inherit a negative charge from the glycan. SDF-1, in contrast, has a positive potential on the surface, while ubiquitin is generally neutral (Figure 6).

Analysis of patterns of glycosylation in R5-tropic and X4-capable sequences

We have analyzed the distribution of potential glycosylation sites in R5-tropic and X4-capable sequences. For the dataset of experimentally verified sequences, all R5- and all but two X4-capable sequences contain such a glycosylation site at position 301 (predicted with N-GlycoSite [23]). Of the predicted sequences, all R5-tropic and all but three X4-capable sequences contain this site. This glycosylation site has been proven to be crucial for coreceptor tropism: its removal switches the virus to use CXCR4 [24]. Yet, this site appears to be present in many X4-capable sequences, but it is impossible to assess the extent of its occupancy by computational tools.

There are two other potential glycosylation sites just downstream of the V3 loop at positions 332 and 339. For a larger set of 199 sequences from a study on HIV neutralization by combination of monoclonal anti-bodies [25], R5-tropic strains tend to contain

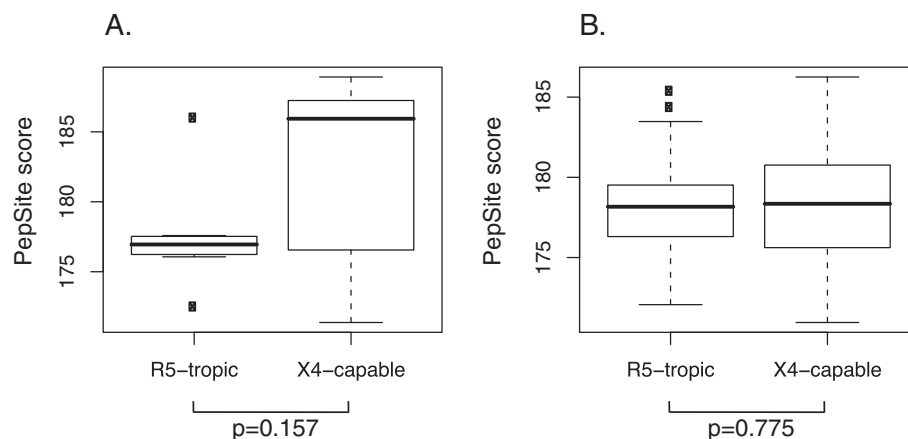
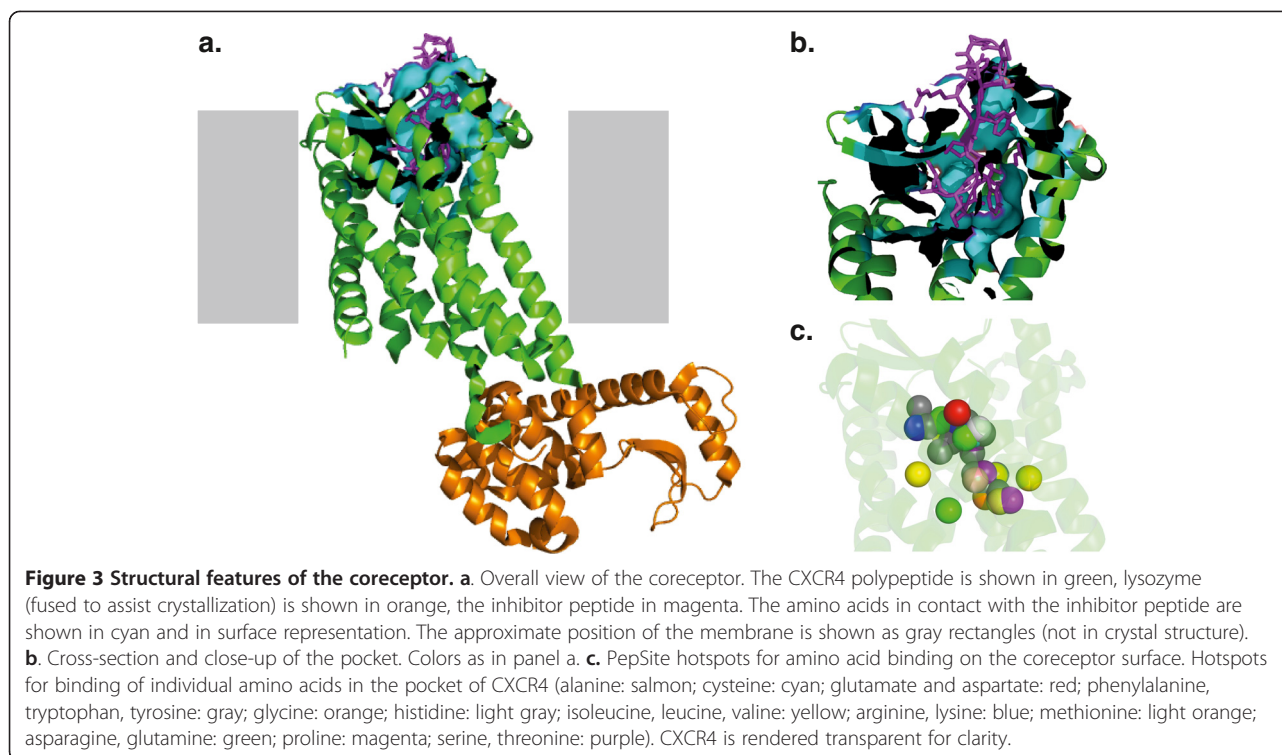


Figure 2 PepSite scores of V3 loop sequences. PepSite scores for a 10-amino acid sliding window reflect the propensities of R5-tropic and X4-capable sequences to bind to CXCR4. **A:** experimentally verified sequences; **B:** predicted sequences. P-values displayed are based on a two-sided Wilcoxon test.



more glycosylation sites at positions 301, 332 and 339 (Figure 7, tropism predicted with geno2pheno [coreceptor] [6] at 10% false positive rate cutoff). For the position 301, the association with tropism type is statistically significant by Fisher's exact test (p-value $7.64e-06$, Bonferroni corrected p-value: 0.0013). Almost all R5-tropic sequences have a glycosylation site at position 301, as opposed to about 80% of X4-capable sequences. Positions 332 and 339 contain a glycosylation site in 70% to 80% of R5- and 50 to 60% of X4-tropic sequences.

Discussion

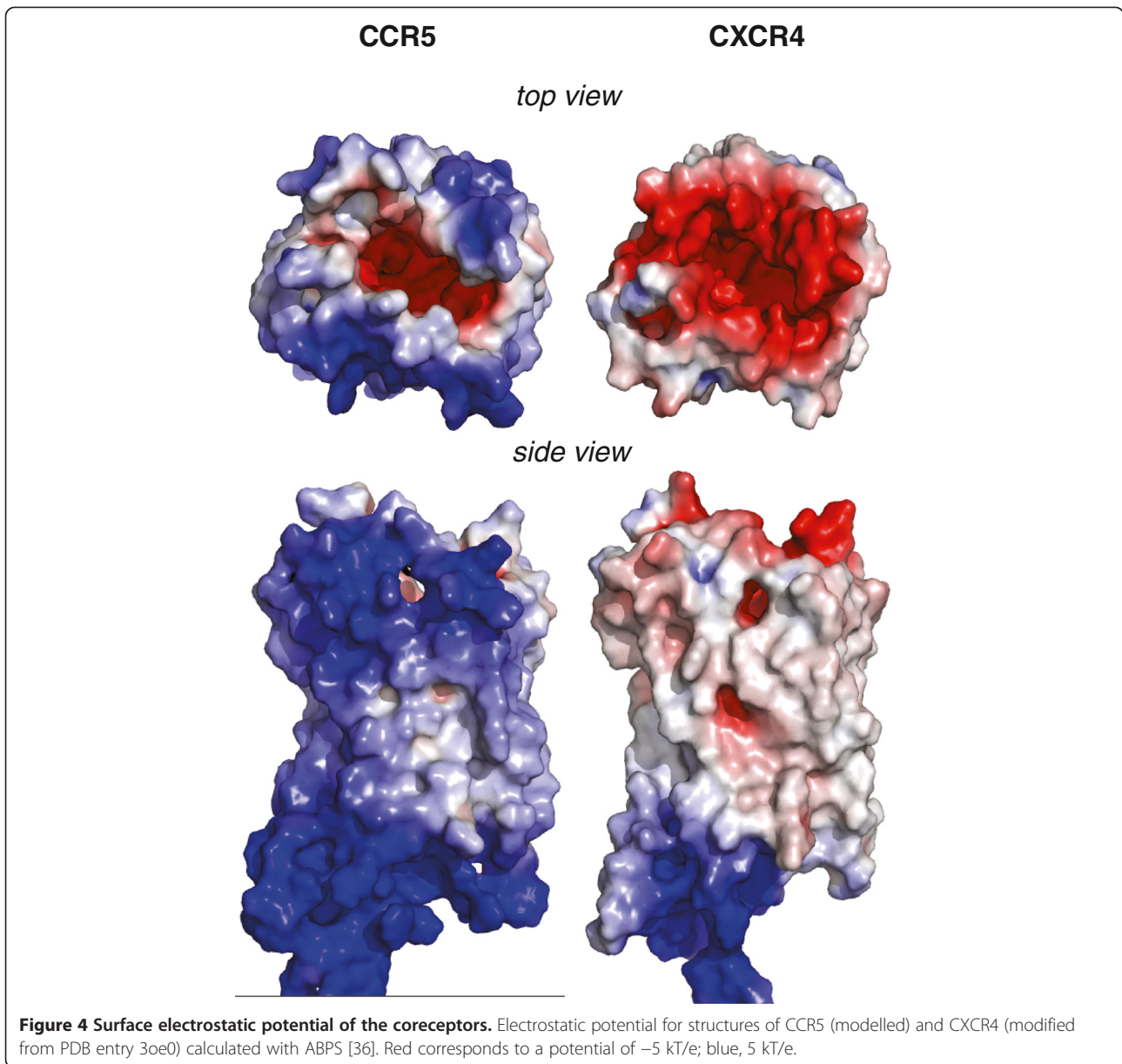
We have analyzed binding preferences of the V3 loops of HIV strains of different tropism. PepSite, pursuing a statistical approach, clearly indicates binding of the V3 loop inside the pocket of the co-receptor, but is unable to discriminate between R5- and X4-tropism satisfactorily. Docking also fails to discriminate between the two kinds of tropism. Based on our analysis, we can propose that V3 loops with both kinds of tropism, once placed inside the binding pocket of either coreceptor, would bind comparably tightly.

However, the electrostatics is strikingly different between CXCR4 and CCR5. The electrostatics of their cognate ligands suggests that electrostatic interactions play an important role in the endogenous recognition process [15]. We propose a two-step model of the binding process, in which the first step is characterized

by a long-range electrostatic interaction between a coreceptor and its ligand. A suitable potential may preclude the initial contact between the co-receptor and the virus of the non-cognate tropism. This model is consistent with the general two-step model for ligand-receptor recognition, which states that long-range electrostatic interactions govern the formation of non-specific encounter complex, and then binding partners reorient themselves to increase the complementarity of the surfaces [26].

Charge has long been recognized to differ between R5-tropic and X4-capable V3 loop sequences [27], and mutation of neutral to positively charged, or negatively charged to neutral amino acids can switch the virus to using CXCR4 [28]. This may explain the fact that the charge of the V3 loop is a reasonable predictor for virus tropism [29], but not a perfect one. Machine-learning based classifiers improve the prediction [6-9]. Indeed, although the net charge of the V3 loop sequences predicted to be prototypic for their tropism differs significantly between the tropism classes, it cannot account for the entire phenomenon of tropism (Figure 1). We can hypothesize that, although the contribution of the charge is important for the determination of tropism, other factors, such as post-translational modifications, may play a role.

Both CCR5 and CXCR4 are extensively post-translationally modified in their extracellular part; however, most of these modifications reside in a disordered N-terminal region,



which is absent from our models. The sulfated acidic N-terminus is characteristic for both proteins [19,20]. The sulfation is critical for interaction between HIV gp120 and CCR5 [30], but not so much for CXCR4 [31]. If the negative N-terminus plays some role in the binding of the more highly conserved base of the V3 loop, which seems plausible by its position on top of the pocket mouth, it can be compensated for by the more negative charge of the pocket entrance of CXCR4. Differential N-glycosylation within the N-terminus of CXCR4 also alters the ability of CXCR4 to bind R5-tropic viruses, but this could not be investigated in the context of the presented study.

As many viral envelope proteins, gp120 is heavily glycosylated, and glycosylation is known to play an important role in viral tropism. Limited evidence in the literature [23,32,33] suggests that R5-tropic viruses tend to be more heavily glycosylated, and the glycans tend to exhibit more complex branching patterns and to be sialylated more frequently, which equips them with negative charge. Removal of a glycosylation site at position 301 leads to an unambiguous switch of tropism from R5 to X4 [24]. Since the exact extent to which every single potential glycosylation site is occupied is not known, structural modeling appears to be inappropriate in this case. However, we observe a statistically

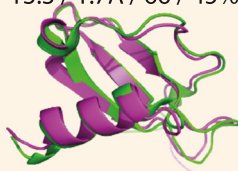
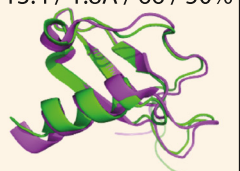

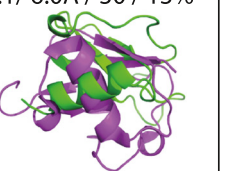


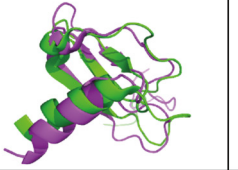
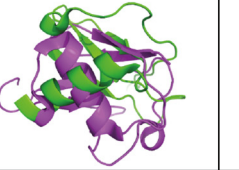
CCL3	CCL4	SDF-1	Ub	Z-score/RMSD/Nfit/seq id
13.3 / 1.7Å / 66 / 45% 	13.4 / 1.8Å / 66 / 50% 	5.9 / 3.1Å / 60 / 28% 	0.1 / 6.0Å / 30 / 13% 	CCL5
	13.8 / 0.7Å / 66 / 70% 	6.0 / 3.1Å / 62 / 22% 	No detectable similarity	CCL3
		5.9 / 3.0Å / 62 / 29% 	0.1 / 6.3Å / 31 / 16% 	CCL4
<p>Orange background: comparison of CCR5 ligands Green background: comparison of CXCR4 ligands Plain background: comparison of ligands of different receptors</p>			No detectable similarity	SDF-1

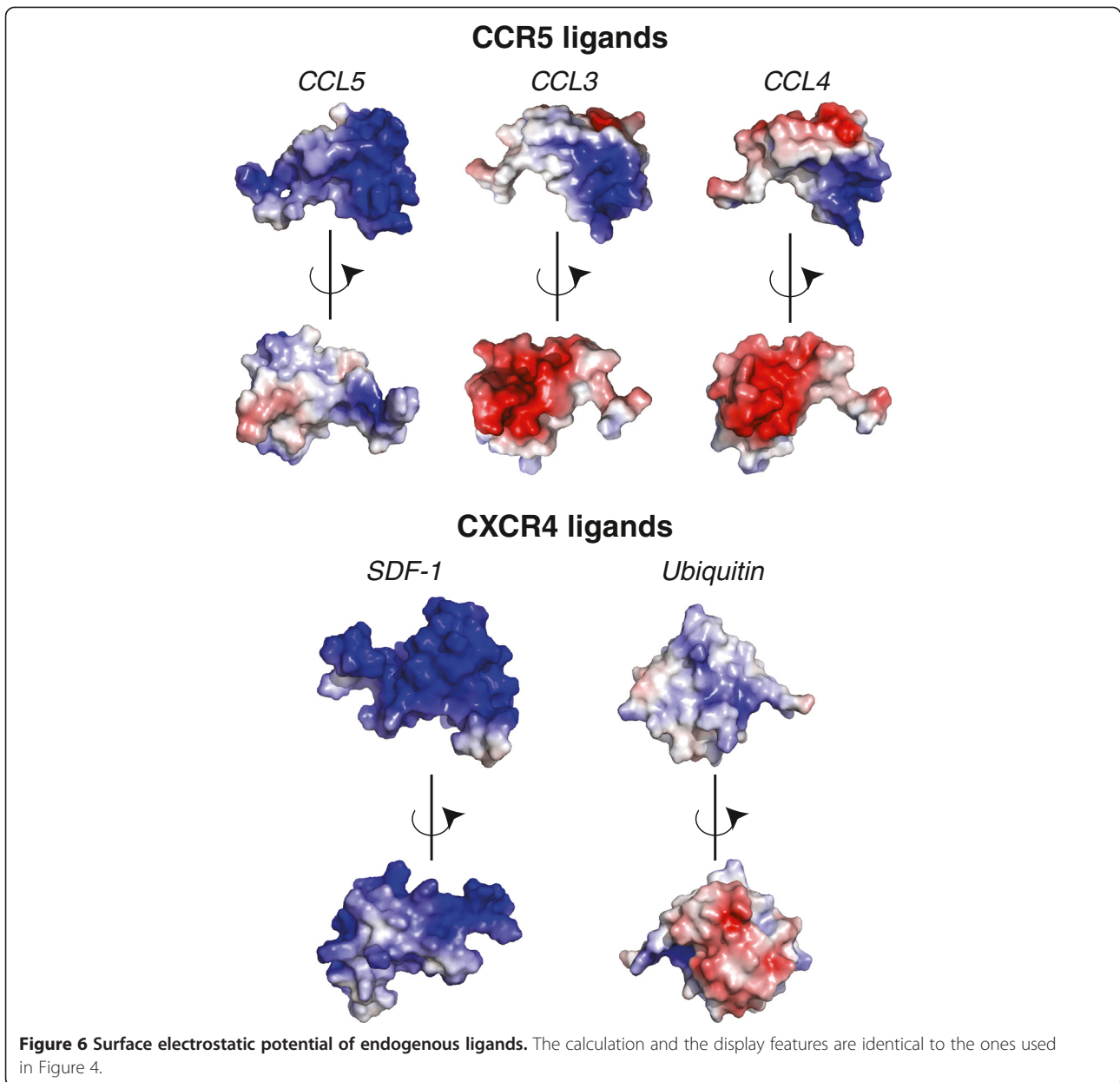
Figure 5 Structural similarity of endogenous ligands of CCR5 and CXCR4. Each cell presents a structural alignment of two ligands (one shown in green and the other in magenta; for clarity, only alignment of the first chain of the structures of each ligand is shown). The comparison was performed with DalLite [21], the obtained Z-scores, RMSD (root-mean-square deviation) of C_α atoms, number of aligned residues, and percent of the identical residues in the alignment are reported in each cell.

significant preference for more potential glycosylation sites in R5-tropic than in X4-capable sequences, position 301 being the most prominent case. Although the trend is pronounced, there remains a large fraction of X4-capable sequences with the glycosylation sites at the positions mentioned above. Not necessarily all these sites are occupied. The known consensus for N-linked glycosylation is rather degenerate (NXT where X is any amino acid except proline). There might be a more complicated, yet unknown, sequence signal that determines glycosylation of 301 N, which assures that R5-tropic sequences are glycosylated more often than X4-capable sequences, or that this glycan is more highly branched and/or more frequently sialylated (which would increase its negative charge) In summary, we hypothesize that differential glycosylation of the HIV envelope gp120 protein is

an important factor for choice of co-receptor, and this choice is primarily guided by the V3 loop charge, either defined by the sequences, or conferred by glycosylation.

Conclusions

We propose a two-step model of interaction between the V3 loop of the HIV Env protein and the CCR5 or CXCR4 coreceptor of the host cell. The choice of the coreceptor is determined at an initial stage by the charge of the V3 loop. We hypothesize that differential glycosylation of HIV envelope gp120 protein is an important factor for choice of co-receptor, as it may alter the charge. This model explains well the available data on the importance of charged residues at certain positions of the V3 loop, as well as the observed glycosylation profile of R5-tropic and X4-capable Env proteins.

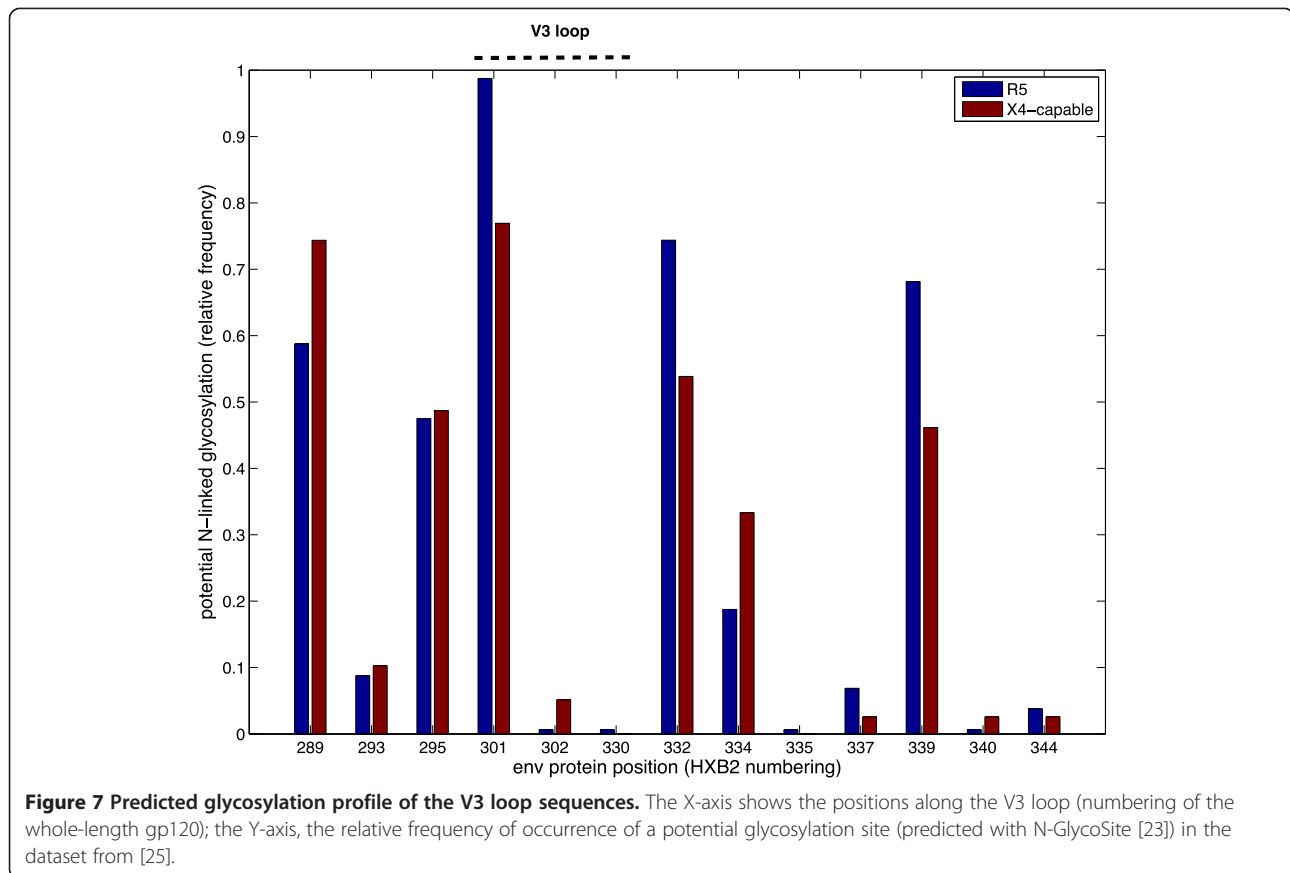


Methods

Experimentally verified V3 loop sequences were retrieved from [10]. As described in [10], viruses were derived from blood samples of HIV-1 infected patients, cultivated in a cell line, and the SI/NSI phenotype was determined by light microscopy. The dataset consists of 7 R5-tropic and 13 X4-tropic sequences.

A set of V3 loop sequences prototypic for each tropism type was obtained with the tropism prediction tool geno2pheno-C_NGS-Sanger [8]. This data set was constructed by predicting coreceptor usage using the coreceptor prediction method introduced in [8] on the Sanger sequences of data from the MOTIVATE as

well as the 1029 trial (see [8] for more details on the data set). The sequences were ordered by predicted risk of X4 emergence and the most extreme 47 (risk of X4 emergence smaller than 27.2%) and 49 (risk of X4 emergence larger than or equal to 50%) sequences were selected. The sequences of the second group are predicted to be able to use CXCR4 as the coreceptor (X4-capable), whereas the sequences of the first group are incapable of using CXCR4 and can bind only to CCR5 (R5-tropic). This does not exclude that the sequences in the group of the X4-capable viruses can use CCR5 as well. The new prediction tool geno2pheno-C_NGS-Sanger [8] does not report FPR (false positive



rate). Instead the tool outputs a probability that the corresponding virus strain is capable of using CXCR4 (cut-offs mentioned above). In an additional analysis, we predicted the FPR for these sequences using the original tool `geno2pheno[coreceptor]` and only kept sequences, for which both tools agreed on the tropism call (10% FPR as recommended by the European guidelines [34]). We achieved similar results with this sequence subset (data not shown).

An experimentally resolved three-dimensional structure of CXCR4 (PDB entry 3oe0, [15]) was edited to remove the portion corresponding to T4 lysozyme that was fused to facilitate crystallization. This structure was also used as a template for building a homology model of the CCR5 structure with MODELLER [35] (sequences of CXCR4 and CCR5 have 33% identical and 55% similar amino acids). Preferential peptide localization was calculated using PepSite [14]. Electrostatic potentials were calculated using ABPS [36]. The structural comparison of the endogenous receptor ligands was performed using the DaliLite web server [21].

All sequences of the V3 loop under consideration were modeled onto the structure of the V3 loop extracted from the PDB entry 2b4c by introducing amino acid mutations using FoldX [37]. Only sequences of the

same length as in the template structure were used for modeling, hence 17 experimentally verified sequences (7 R5- and 10 X4-tropic) and 72 sequences with predicted tropism (28 R5- and 44 X4-tropic) were modeled. Docking was performed using FlexPepDock protocol [16], a peptide-protein docking protocol with a fully flexible peptide, of the Rosetta modelling suite [38]. The 25 best-scoring modelled structures were considered for each complex. The charge of V3 loops was calculated using Protein Calculator v.3.3 [39]. This tool is not applicable to folded proteins, but due to its high structural flexibility, the V3 loop can be regarded as a flexible peptide.

Competing interests

The authors declare no competing interests.

Authors' contributions

OVK designed the study and performed the experiments. NP designed the study, performed the experiments and the statistical analysis. TL contributed ideas to the design and interpretation of the study. All authors wrote and approved the manuscript.

Acknowledgements

We thank Nicole Doria-Rose and John Mascola for providing Genbank accession numbers for the *env* sequence panel used in the HIV antibody neutralization study [25].

Note added in proof

While this manuscript was under peer review, an X-ray structure of human CCR5 in complex with maraviroc was resolved at 2.7 Å [40] (PDB id 4MBS). It supports the model proposed here in several respects:

- (i) The structure of the protein is very similar to our model (1.15 Å all atom RMSD over 278 amino acids, which covers the whole CCR5 protein except the inserted rubredoxin).
- (ii) The electrostatic surface potential is also observed to differ between CXCR4 and CCR5 [40], thus supporting our two-step binding model.
- (iii) The position of maraviroc in the extracellular pocket coincides with the position of the tip of the V3 loop proposed here.

Received: 17 July 2013 Accepted: 4 November 2013

Published: 11 November 2013

References

1. Combadiere C, Ahuja SK, Tiffany HL, Murphy PM: **Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1(alpha), MIP-1(beta), and RANTES.** *J Leukoc Biol* 1996, **60**(1):147–152.
2. Raport CJ, Gosling J, Schweickart VL, Gray PW, Charo IF: **Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha.** *J Biol Chem* 1996, **271**(29):17161–17166.
3. Bleul CC, Farzan M, Choe H, Parolin C, Clark-Lewis I, Sodroski J, Springer TA: **The lymphocyte chemoattractant SDF-1 is a ligand for LESTR/fusin and blocks HIV-1 entry.** *Nature* 1996, **382**(6594):829–833.
4. Saini V, Marchese A, Majetschak M: **CXC chemokine receptor 4 is a cell surface receptor for extracellular ubiquitin.** *J Biol Chem* 2010, **285**(20):15566–15576.
5. MacArthur RD, Novak RM: **Reviews of anti-infective agents: maraviroc: the first of a new class of antiretroviral agents.** *Clin Infect Dis* 2008, **47**:236–241.
6. Lengauer T, Sander O, Sierra S, Thielen A, Kaiser R: **Bioinformatics prediction of HIV coreceptor usage.** *Nat Biotechnol* 2007, **25**(12):1407–1410.
7. Swenson LC, Mo T, Dong WW, Zhong X, Woods CK, Jensen MA, Thielen A, Chapman D, Lewis M, James I, Heera J, Valdez H, Harrigan PR: **Deep sequencing to infer HIV-1 co-receptor usage: application to three clinical trials of maraviroc in treatment-experienced patients.** *J Infect Dis* 2011, **203**(2):237–245.
8. Pfeifer N, Lengauer T: **Improving HIV coreceptor usage prediction in the clinic using hints from next-generation sequencing data.** *Bioinformatics* 2012, **28**(18):i589–i595.
9. Jensen MA, Coetzer M, van 't Wout AB, Morris L, Mullins JI: **A reliable phenotype predictor for human immunodeficiency virus type 1 subtype c based on envelope v3 sequences.** *J Virol* 2006, **80**(10):4698–4704.
10. Bozek K, Eckhardt M, Sierra S, Anders M, Kaiser R, Kräusslich HG, Müller B, Lengauer T: **An expanded model of HIV cell entry phenotype based on multi-parameter single-cell data.** *Retrovirology* 2012, **9**:60.
11. Huang CC, Tang M, Zhang MY, Majeed S, Montabana E, Stanfield RL, Dimitrov DS, Korber B, Sodroski J, Wilson IA, Wyatt R, Kwong PD: **Structure of a V3-containing HIV-1 gp120 core.** *Science* 2005, **310**(5750):1025–1028.
12. Chen M, Svicher V, Artese A, Costa G, Alteri C, Ortuso F, Parrotta L, Liu Y, Liu C, Perno CF, Alcaro S, Zhang J: **Detecting and understanding genetic and structural features in HIV-1 B subtype V3 underlying HIV-1 co-receptor usage.** *Bioinformatics* 2013, **29**(4):451–460.
13. Petsalaki E, Stark A, García-Urdiales E, Russell RB: **Accurate prediction of peptide binding sites on protein surfaces.** *PLoS Comput Biol* 2009, **5**(3):e1000335.
14. Trabuco LG, Lise S, Petsalaki E, Russell RB: **PepSite: prediction of peptide-binding sites from protein surfaces.** *Nucleic Acids Res* 2012, **40**:W423–W427.
15. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC: **Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists.** *Science* 2010, **330**(6007):1066–1071.
16. Raveh B, London N, Schueler-Furman O: **Sub-angstrom modeling of complexes between flexible peptides and globular proteins.** *Proteins* 2010, **78**:2029–2040.
17. Babcock GJ, Farzan M, Sodroski J: **Ligand-independent dimerization of CXCR4, a principal HIV-1 coreceptor.** *J Biol Chem* 2003, **278**(5):3378–3385.
18. Laskowski RA, MacArthur MW, Moss DS, Thornton JM: **PROCHECK - a program to check the stereochemical quality of protein structures.** *J App Cryst* 1993, **26**:283–291.
19. Farzan M, Mirzabekov T, Kolchinsky P, Wyatt R, Cayabyab M, Gerard NP, Gerard C, Sodroski J, Choe H: **Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry.** *Cell* 1999, **96**(5):667–676.
20. Seibert C, Veldkamp CT, Peterson FC, Chait BT, Volkman BF, Sakmar TP: **Sequential tyrosine sulfation of CXCR4 by tyrosylprotein sulfotransferases.** *Biochemistry* 2008, **47**(43):11251–11262.
21. *DaliLite*. [http://www.ebi.ac.uk/Tools/services/web/toolform.ebi?tool=dalilite]
22. Kameyoshi Y, Dörschner A, Mallet AI, Christophers E, Schröder JM: **Cytokine RANTES released by thrombin-stimulated platelets is a potent attractant for human eosinophils.** *J Exp Med* 1992, **176**(2):587–592.
23. Zhang M, Gaschen B, Blay W, Foley B, Haigwood N, Kuiken C, Korber B: **Tracking global patterns of N-linked glycosylation site variation in highly variable viral glycoproteins: HIV, SIV, and HCV envelopes and influenza hemagglutinin.** *Glycobiology* 2004, **14**(12):1229–1246.
24. Pollakis G, Kang S, Kliphuis A, Chalaby MI, Goudsmit J, Paxton WA: **N-linked glycosylation of the HIV type-1 gp120 envelope glycoprotein as a major determinant of CCR5 and CXCR4 coreceptor utilization.** *J Biol Chem* 2001, **276**(16):13433–13441.
25. Doria-Rose NA, Louder MK, Yang Z, O'Dell S, Nason M, Schmidt SD, McKee K, Seaman MS, Bailer RT, Mascola JR: **HIV-1 neutralization coverage is improved by combining monoclonal antibodies that target independent epitopes.** *J Virol* 2012, **86**(6):3393–3397.
26. McCammon JA, Northrup SH, Allison SA: **Diffusional dynamics of ligand receptor association.** *J Phys Chem* 1986, **90**:3901–3905.
27. Fouchier RAM, Groenink M, Kootstra NA, Tersmette M, Huisman HG, Miedema F, Schuitemaker H: **Phenotype-associated sequence variation in the third variable domain of the human immunodeficiency virus type 1 gp120 molecule.** *J Virol* 1992, **66**(5):3183.
28. De Long J-J, De Ronde A, Keulen W, Tersmette M, Goudsmit J: **Minimal requirements for the human immunodeficiency virus type 1 V3 domain to support the syncytium-inducing phenotype: analysis by single amino acid substitution.** *J Virol* 1992, **66**(11):6777.
29. Sing T, Low AJ, Beerenwinkel N, Sander O, Cheung PK, Domingues FS, Büch J, Däumer M, Kaiser R, Lengauer T, Harrigan PR: **Predicting HIV coreceptor usage on the basis of genetic and clinical covariates.** *Antivir Ther* 2007, **12**(7):1097–1106.
30. Schnur E, Noah E, Ayzenshtat I, Sargsyan H, Inui T, Ding FX, Arshava B, Sagi Y, Kessler N, Levy R, Scherf T, Naider F, Anglister J: **The conformation and orientation of a 27-residue CCR5 peptide in a ternary complex with HIV-1 gp120 and a CD4-mimic peptide.** *J Mol Biol* 2011, **410**(5):778–797.
31. Farzan M, Babcock GJ, Vasilieva N, Wright PL, Kiprilov E, Mirzabekov T, Choe H: **The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 alpha association and HIV-1 entry.** *J Biol Chem* 2002, **277**(33):29484–29489.
32. Go EP, Hewawasam G, Liao HX, Chen H, Ping LH, Anderson JA, Hua DC, Haynes BF, Desaire H: **Characterization of glycosylation profiles of HIV-1 transmitted/founder envelopes by mass spectrometry.** *J Virol* 2011, **85**(16):8270–8284.
33. Wilen CB, Parrish NF, Pfaff JM, Decker JM, Henning EA, Haim H, Petersen JE, Wojcechowskyj JA, Sodroski J, Haynes BF, Montefiori DC, Tilton JC, Shaw GM, Hahn BH, Doms RW: **Phenotypic and immunologic comparison of clade B transmitted/founder and chronic HIV-1 envelope glycoproteins.** *J Virol* 2011, **85**(17):8514–8527.
34. Vandekerckhove LP, Wensing AM, Kaiser R, Brun-Vézinet F, Clotet B, De Luca A, Dressler S, Garcia F, Geretti AM, Klimkait T, Korn K, Masquelier B, Perno CF, Schapiro JM, Soriano V, Sönnnerborg A, Vandamme AM, Verhofstede C, Walter H, Zazzi M, Boucher CA, European Consensus Group on clinical management of tropism testing: **European guidelines on the clinical management of HIV-1 tropism testing.** *Lancet Infect Dis* 2011, **11**:394–407.
35. Eswar N, Webb B, Marti-Renom MA, Madhusudan MS, Eramian D, Shen MY, Pieper U, Sali A: **Comparative protein structure modeling using Modeller.** *Curr Protoc Bioinformatics*. Hoboken, NJ: John Wiley & Sons, Inc; 2006. Supplement 15, 5.6.1–5.6.30.
36. Baker NA, Sept D, Joseph S, Holst MJ, McCammon JA: **Electrostatics of nanosystems: application to microtubules and the ribosome.** *Proc Natl Acad Sci U S A* 2001, **98**:10037–10041.
37. Schymkowitz J, Borg J, Stricher F, Nys R, Rousseau F, Serrano L: **The FoldX web server: an online force field.** *Nucleic Acids Res* 2005, **33**(Web Server issue):W382–W388.

38. *RosettaCommons*. <https://www.rosettacommons.org/>.
39. *Protein Calculator v.3.3*. <http://www.scripps.edu/~cdputnam/protcalc.html>.
40. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Li J, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q, Wu B: **Structure of the CCR5 Chemokine receptor–HIV entry inhibitor maraviroc complex**. *Science* 2013, **341**:1387–1390.

doi:10.1186/1742-4690-10-130

Cite this article as: Kalinina et al.: Modelling binding between CCR5 and CXCR4 receptors and their ligands suggests the surface electrostatic potential of the co-receptor to be a key player in the HIV-1 tropism. *Retrovirology* 2013 **10**:130.

**Submit your next manuscript to BioMed Central
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at
www.biomedcentral.com/submit

