SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
PROTEIN FUNCTION PREDICTIONS
PROTEIN STRUCTURE PREDICTIONS

Received
14 April 2014

Accepted
5 June 2014

Published
26 June 2014

Correspondence and requests for materials should be addressed to C.A.F. (floudas@titan.princeton.edu)

Elucidating a Key Anti-HIV-1 and Cancer-Associated Axis: The Structure of CCL5 (Rantes) in Complex with CCR5

Phanourios Tamamis & Christodoulos A. Floudas

Department of Chemical and Biological Engineering, Princeton University, NJ, USA.

CCL5 (RANTES) is an inflammatory chemokine which binds to chemokine receptor CCR5 and induces signaling. The CCL5:CCR5 associated chemotactic signaling is of critical biological importance and is a potential HIV-1 therapeutic axis. Several studies provided growing evidence for the expression of CCL5 and CCR5 in non-hematological malignancies. Therefore, the delineation of the CCL5:CCR5 complex structure can pave the way for novel CCR5-targeted drugs. We employed a computational protocol which is primarily based on free energy calculations and molecular dynamics simulations, and report, what is to our knowledge, the first computationally derived CCL5:CCR5 complex structure which is in excellent agreement with experimental findings and clarifies the functional role of CCL5 and CCR5 residues which are associated with binding and signaling. A wealth of polar and non-polar interactions contributes to the tight CCL5:CCR5 binding. The structure of an HIV-1 gp120 V3 loop in complex with CCR5 has recently been derived through a similar computational protocol. A comparison between the CCL5 : CCR5 and the HIV-1 gp120 V3 loop : CCR5 complex structures depicts that both the chemokine and the virus primarily interact with the same CCR5 residues. The present work provides insights into the blocking mechanism of HIV-1 by CCL5.

Chemokines and their corresponding chemokine receptors constitute key regulators of immune activities. Chemokines are divided to two major families, homeostatic and inflammatory[1]. Homeostatic chemokines are mainly expressed in lymphoid organs and mediate leukocyte trafficking to these sites during immune homeostasis, while inflammatory chemokines are inducibly expressed at infected/damaged tissues, and thereby recruit leukocytes to sites that have been exposed to an inflammatory insult[1,2]. CCL5 (RANTES) is an inflammatory chemokine which acts as a key regulator of T-cell migration to inflammatory sites, directing the migration of T cells to damaged or infected sites. In addition, CCL5 regulates T-cell differentiation, and this, is supported by evidence depicting that CCR5 is expressed in Th1 cells[3,4]. Chemokine receptor CCR5 is one of the three corresponding high-affinity receptors of CCL5, along with CCR1 and CCR3[2,5].

The CCL5:CCR5 axis acquires a beneficial biological role, as it provides antiapoptotic signals for macrophage survival during infection, through the protection of tissue macrophages from virus-inducible cell death[6]. Recent experimental findings also suggest that CCL5, owing to the CCL5:CCR5 chemokine-mediated signaling, may be important as a general B cell coactivator[7], and that the CCL5:CCR5 interaction is a major regulator of endothelial progenitor cells homing during wound healing[8]. In addition, as the gp120 protein of HIV-1 binds to chemokine receptors CCR5[9], or CXCR4[10], a primary step of the HIV-1 entry to the host cell, the binding of CCL5, as well as of CCL5 derivatives, to CCR5 is considered a potential HIV-1 therapeutic axis[11–15].

A series of studies have provided growing evidence of the expression of CCL5 and CCR5 in predominantly non-hematological malignancies[1,3]. Several studies identified correlations between high levels of intratumoral CCL5 expression and advanced stages of breast cancer[1,16–18]. Furthermore, CCL5 possesses an important role in promoting pro-cancerous activities in tumor cells, as it acts directly on the cancer cells, leading to increased proliferation in breast, colorectal, gastric as well as prostate cancers[1,19,20]. Moreover, CCL5 was identified as a potent inducer of tumor cell migration and invention in tumor cells involved in breast, colorectal, osteosarcoma and prostate cancers[1,19–21]. Overall, the experimental evidence supports that the CCL5:CCR5 signaling leads to pro-cancerous consequences[1], and therefore, it constitutes a potential therapeutic target against cancer.

The key role of the CCL5:CCR5 pathway in the primary and advanced stages of different types of tumors suggests that the delineation of the CCL5:CCR5 complex structure can pave the way for discovering novel CCR5-targeted drugs. No high-accuracy computational or complete experimental structure exists for the CCL5:CCR5

complex. Two previous attempts[22,23] to model the CCL5:CCR5 complex structure have not reported a high-degree of agreement with previous experimental findings[24–36], and did not succeed - among others - to meet key experimental evidence depicting that the N-terminus of CCL5 (i) interacts with the transmembrane helical bundle of CCR5[31] and (ii) is crucial for activation[25]. A recent study by Schnur *et al.*[37] provided useful insights on the key CCL5 interacting residues in complex with a recombinant CCR5 construct "Nt-ECL1-ECL2"containing only the N-terminal segment and residues 89–101 and 168–194 of extracellular loops 1 and 2. Despite the valuable results derived in[37], the CCR5 transmembrane domain is crucial (i) for CCL5 binding and signaling[14,25,27,30–33], and (ii) for CCL5 to adopt an appropriate orientation to optimally interact with the entire CCR5. Thus, there is no explicit evidence which supports that the identified CCL5 residues which are affected by CCR5 binding in complex with CCR5 domains, reported in the NMR study[37], would also correspond to the important CCL5 residues in complex with the entire chemokine receptor; see *Discussion* in the *Supplementary Information*.

The presence of NMR[38,39] structures for CCL5, as well as the recently reported X-ray structures of CCR5[40], and its homologous CXCR4[41], provide the basis for the computational derivation of the CCL5:CCR5 complex structure. In this work, we exploit these structures, as well as our recent computationally derived structure of an HIV-1 gp120 V3 loop in complex with CCR5[9], to derive the first complex CCL5 : CCR5 structure which is in excellent agreement with experiments (see CCL5 and CCR5 residues marked in bold face in Table 1). A similar computational protocol to the one applied in the present study has already demonstrated its power in elucidating the complex structures of a specific dual tropic HIV-1 gp120 V3 loop in complex with CXCR4[10] and CCR5[9], as well as the structure of chemokine CXCL12 (SDF-1α) in complex with CXCR4[42].

Results

We present the complete complex structure of CCL5 in complex with CCR5. The structure corresponds to an ensemble of snapshots derived from the simulation "Complex 14A" (see *Methods*). The conformations extracted at 2 ns intervals are provided as Supplementary Coordinates in PDB format. We analyzed 1000 simulation snapshots of "Complex 14A" which were extracted at 20-ps intervals: (i) we calculated the average intermolecular interaction free energies between CCL5 : CCR5 residue pairs (Supplementary Figure 1), (ii) we summed up the total intermolecular interaction free energies of every CCR5 residue, so as to provide insights into the role of each interacting CCR5 residue in complex with CCL5, and in addition, (iii) we evaluated the intermolecular hydrogen bond occupancies (Supplementary Table 1). Based on these analyses, we summarize the basic intermolecular CCL5 : CCR5 interactions in Table 1.

According to the computationally derived structure, the 1–15 residue moiety of CCL5 is inserted into the CCR5 binding pocket; the 1–6 N-terminal domain of CCL5 is buried within the transmembrane region of CCR5, and the 7–15 residue moiety of CCL5 is predominantly encompassed by the N-terminal domain and extracellular loops of CCR5. CCL5 residues Ala16, Arg17 and additional residues of the 24–50 residue moiety interact with the upper N-terminal domain and extracellular loop interface of CCR5 (Figure 1).

Conformational Analysis of CCL5 and CCR5 within the Simulation. According to STRIDE[43], the CCL5 conformation acquires the following secondary structure features: (i) β-turns in residue domains 11–14, 20–23, 31–37, 44–47, 52–55, and to a smaller extent 6–10 (the 20–23 β-turn is frequently interchanged with a 3_{10} helix formed by residues 21–23); (ii) β-extended antiparallel sheets which involve residue domains 24–29, 39–43 and 47–51; and (iii) α-helical conformation within the 56–65 residue moiety. We aligned the complex structures of the simulation based on the receptor's backbone transmembrane region, and calculated the average backbone root mean square deviation (RMSD) of CCL5, with respect to the first simulation snapshot after equilibration. The calculated RMSD values for the total CCL5, the 1–11, 1–15 and 16–68 CCL5 residue moieties are 1.7 ± 0.5 Å, 0.9 ± 0.3 Å, 1.0 ± 0.3 Å and 1.6 ± 0.5 Å, respectively. The results show that the overall CCL5 conformation is well maintained throughout the simulation, and that the more CCL5 is buried into CCR5, the more rigid it is. Thus, similarly to the CXCL12 binding to CXCR4[42], the bound properties of CCL5 differ from its unbound properties, as the 1–11 residue moiety of CCL5 is the most flexible region in the unbound conformation[39]. The CCR5 conformation is very well maintained within the simulation, and this is reflected by the relatively low average backbone RMSD values of the transmembrane region and N-terminal domain (0.8 ± 0.1 Å and 1.2 ± 0.3 Å, respectively).

Interactions of CCL5 Residues 1:5 with CCR5. CCL5 residue Ser1 is the most buried CCL5 residue within CCR5 and participates in a highly interacting salt bridge with CCR5 residue Glu283 through its positively charged N-terminal (Figure 2). The charged N-terminal is also (i) hydrogen bonded to CCR5 residue Gln280, (ii) is polarly attracted to CCR5 residue Asp276, and (iii) is in proximity to the aromatic ring of CCR5 residue Tyr37. The side chain hydroxyl group of Ser1 forms a hydrogen bond with CCR5 Tyr251 OH and less frequently Met279 SD. In addition, throughout the simulation, the charged N-terminal of Ser1 participates in a cation-π interaction with the aromatic ring of Trp86 (Figure 2), and in general, Ser1 forms contacts with a cluster of CCR5 aromatic residues: Trp86, Tyr89, Tyr108 and Tyr251. CCL5 residue Pro2 is buried within a pocket comprised of CCR5 residues Trp86, Thr105, Tyr108, Phe109, Thr167, Cys178, Ser179, Tyr251 and Glu283. CCL5 residue Tyr3 forms non-polar contacts with CCR5 residues Thr105, Phe109, Gly163, Thr167, Cys178, Ser179, Ser180, Lys191, Thr195, Ile198, Tyr251, Leu255, Met279; its aromatic side chain position is predominantly stabilized due to aromatic interactions with CCR5 residue Phe109, and a hydrogen bond interaction of its side chain hydroxyl group with CCR5 Thr167 OG1. CCL5 residue Ser4 intercalates among the side chain moieties of CCR5 residues Lys22, Glu172, Ser179, Lys191, Asn258, Asp276 and Met279; its main chain carbonyl group is frequently hydrogen bonded to CCR5 Lys191 NZ, and its side chain hydroxyl group is also frequently hydrogen bonded to CCR5 Lys22 and Asp276 OD1/2 (Figure 2). CCL5 residue Ser5 participates in a series of hydrogen bond interactions with CCR5 residues, Ser5 O : Lys22 NZ, Ser5 N : Asn258 OD1/ND2, Ser5 OG : Lys22 NZ (Figure 2), Ser5 OG : Asn258 OD1 and Ser5 OG : Ser272 O/OG; Ser5 of CCL5 is also polarly attracted to CCR5 residue Glu262, and it participates in non-polar contacts with CCR5 residues Gln261, Leu275 and Asp276.

Interactions of CCL5 Residues 6:17 with CCR5. CCL5 residue Asp6 is predominantly polarly attracted to CCR5 residues Met1, Asp2, Lys22, Ser179, Lys191 and Glu262. These polar interactions comprise a highly interacting salt bridge with CCR5 residue Lys191 (Figure 2), and less frequent hydrogen bond interactions among atom pairs Asp6 O : Met 1 N, Asp6 N : Glu262 OE1/2 and Asp 6 OD1 : Ser179 OG; also, the negatively charged side chain group of CCR5 Asp2 is proximal to the backbone amide group of CCL5 Asp6. CCL5 residues Thr7 and Thr8 intercalate among CCR5 residues Asp2, Ser17, Cys20, Gln21, Lys22, Gln261, Glu262, Phe264, Cys269, Ser272, and Met1, Asp2, Gln21, Lys22, Asn24, Glu172, respectively, and form non-polar and polar interactions with the aforementioned CCR5 residues; the polar interactions include hydrogen bonds between atom pairs: Thr7 N : Asp2 OD1/2, Thr7 O : Lys22 N, Thr7 OG1 : Ser272 OG, Thr8 O : Met1 N, Thr8 OG1 : Gln21 NE2 and Thr8 OG1 : As24 OD1/ND2. CCL5 residue Pro9 forms non-polar contacts with CCR5 residues Met1, Cys20 and Gln21, and infrequently, its backbone carbonyl forms a hydrogen

Table 1 | Important CCL5:CCR5 Intermolecular Interactions: Residue pair-wise interaction free energies, salt bridges and hydrogen bonds

CCL5[§]	CCR5[¶]	Salt Bridges[†] and Hydrogen Bonds[‡]
Ser1	□**Tyr37**(0.0,-0.5), **Trp86** (-6.8, -2.4), **Tyr89** (1.0, -0.8), **Tyr108** (4.5, -1.0), **Tyr251** (-2.5, -0.3), **Asp276** (-0.6, -0.1), Met279 (-1.1, -1.5), Gln280 (-16.2, 0.8), **Glu283** (-70.7, 4.5)	‡**Ser1** OG : **Tyr251** OH, ‡**Ser1** OG : Met279 SD, ‡**Ser1** N : Gln280 OE1, ‡**Ser1** N : Gln280 NE2, †**Ser1**:**Glu283**
Pro2	**Trp86** (0.4, -1.9), Thr105 (-0.8, -1.1), **Tyr108** (-0.3, -1.9), **Phe109** (0.1, -1.0), Thr167 (-0.2, -0.9), **Cys178** (-0.5, -1.5), Ser179 (-0.5, -0.8), **Tyr251** (-0.4, -0.6), **Glu283** (-1.0, -0.3)	
Tyr3	Thr105 (0.2, -0.7), **Phe109** (-0.3, -1.2), **Gly163** (-0.7, -0.5), Thr167 (-4.9, -0.5), **Cys178** (-0.2, -0.6), Ser179 (0.0, -1.6), **Ser180** (0.0, -1.8), **Lys191** (0.5, -0.7), **Thr195** (-0.1, -1.0), **Ile198** (-0.1, -0.8), **Tyr251** (-0.4, -0.6), Leu255 (0.0, -1.0), Met279 (0.0, -1.2)	‡**Tyr3** OH : Thr167 OG1
Ser4	**Lys22** (2.3, -0.9), **Glu172** (-0.7, -0.3), Ser179 (-0.1, -0.6), **Lys191** (-2.9, -0.4), Asn258 (-0.3, -1.6), **Asp276** (-10.0, 0.2), Met279 (0.0, -1.5)	‡**Ser4** OG : **Lys22** NZ, ‡**Ser4** O : **Lys191** NZ, ‡**Ser4** OG : **Asp276** OD*
Ser5	**Lys22** (-12.3, -0.9), Asn258 (-2.1, -2.0), Gln261 (0.9, -1.4), Glu262 (-6.7, -2.3), Ser272 (-3.9, -0.8), Leu275 (0.0, -0.5), **Asp276** (0.6, -0.6)	‡**Ser5** O : **Lys22** NZ, ‡**Ser5** OG : **Lys22** NZ, ‡**Ser5** N : Asn258 OD1, ‡**Ser5** N : Asn258 ND2, ‡**Ser5** OG : Asn258 OD1, ‡**Ser5** OG : Ser272 O/OG
Asp6	Met1 (-1.3, -1.2), Asp2 (-2.3, -1.9), **Lys22** (-3.5, -1.6), Ser179 (-0.4, -0.3), **Lys191** (-11.2, -0.1), Glu262 (-5.9, -1.2)	‡**Asp6** O : Met1 N, ‡**Asp6** OD1 : Ser179 OG, †**Asp6**:**Lys191**, **Asp6** N : Glu262 OE*
Thr7	Asp2 (-9.0, -0.9), Ser17 (-0.1, -0.9), **Cys20** (-0.1, -0.6), Gln21 (-2.1, -1.2), **Lys22** (-4.2, -2.2), Gln261 (-0.6, -1.3), Glu262 (0.5, -0.7), Phe264 (-0.1, -1.3), Cys269 (0.0, -1.6), Ser272 (-3.1, -1.0)	‡Thr7 N :Asp2 OD*, ‡Thr7 OG1 : Asp2 OD2, ‡Thr7 O : **Lys22** N, ‡Thr7 OG1 : Ser272 OG
Thr8	Met1 (-6.7, -1.8), Asp2 (0.5, -0.6), Gln21 (-1.0, -2.5), **Lys22** (0.0, -2.7), Asn24 (-0.6, -0.4), **Glu172** (-0.4, -1.1)	‡Thr8 O : Met1 N, ‡Thr8 OG1 : Gln21 NE2, ‡Thr8 OG1 : Asn24 OD1, ‡Thr8 OG1 : Asn24 ND2
Pro9	Met1 (-0.3, -3.10), **Cys20** (0.0, -0.5), Gln21 (0.2, -2.2)	‡Pro9 O : Met1 N
Cys10	Met1 (-7.4, -1.4), **Tys3** (1.7, -0.6), Gln21 (0.0, -0.7), Asn24 (0.0, -0.7), **Glu172** (0.0, -0.6), Gly173 (0.0, -1.7)	‡Cys10 O : Met1 N
Cys11	Met1 (1.3, -1.1), **Tys3** (-7.5, -2.7), **Lys171** (-0.4, -1.0)	
Phe12	Met1 (0.9, -1.8), **Tys3** (-7.1, -3.6), Gln170 (-1.0, -3.2), **Lys171** (-0.4, -3.7), **Glu172** (-0.1, -2.9), Ser179 (0.0, -0.6), **Tyr184** (-1.1, -2.3)	‡**Phe12** N : **Tys3** OS2/OS3, ‡**Phe12** O : Gln170 NE2, ‡Phe12 O : **Lys171** N
Ala13	**Tys3** (-8.7, -1.8), Gln170 (-0.2, -1.5), **His181** (0.0, -0.5), **Tyr184** (-0.4, -2.7)	‡Ala13 N : **Tys3** OS3
Tyr14	**Tys3** (-9.3, -2.5), **Tyr184** (-0.1, -1.3)	‡**Tyr14** N : **Tys3** OS3
Ile15	**Tys3** (0.1, -3.1), **Tyr184** (-0.7, -2.0), Tyr187 (0.0, -0.5)	
Ala16	**His181** (-0.1,-0.4), **Tyr184** (-1.9, -2.3), Ser185 (0.0, -1.7)	‡Ala16 N : **Tyr184** O
Arg17	**Tyr184** (-0.3, -0.4),Ser185 (0.1, -0.7), Tyr187 (-3.5, -1.5)	‡**Arg17** NH1 : **Tyr184** O
Lys25	Tys14 (-8.1, -1.1)	†Lys25:Tys14
Ser31	Gln21 (-1.0, -1.1)	‡Ser31 OG : Gln21 OE1, ‡Ser31 OG: Gln21 NE2
Lys33	Gln21 (-0.9, -1.7), Asn24 (-1.0, -2.7), Leu174 (0.2, -1.4)	‡Lys33 NZ : Gln21 OE1, ‡Lys33 O : Asn24 ND2, ‡Lys33 NZ : Asn24 OD1
Cys34	Gly173 (-1.3, -1.4)	
Ser35	**Lys171** (-5.3, -1.2), Gly173 (-3.5, -0.7), Leu174 (0.0, -2.8)	‡Ser35 OG : **Lys171** NZ, ‡Ser35 N : Gly173 O, Ser35 OG : Ser173 O
Asn36	**Lys171** (-1.7, -1.4)	‡Asn36 OD1 : **Lys171** NZ
Val40	Met1 (0.2, -1.4)	
Val42	Tys14 (-0.7, -2.0)	
Arg44	Val5 (0.3, -2.0), Tyr187 (-5.7, -2.1)	‡**Arg44** NH* : Tyr187 OH
Lys45	Val5 (-0.1, -0.6), Ser6 (-0.1, -2.5), **Asp11** (-11.4, -0.5), Tys14 (-1.5, -0.6)	†**Lys45**:**Asp11**
Asn46	Gln4 (-0.2, -0.8), **Asp11** (-0.7, -2.7), Tys14 (-8.0, -7.2), Tyr15 (-1.6, -1.5)	‡Asn46 N :Tys14 OS*, ‡Asn46 ND2 : Tys14 OS*, ‡Asn46 ND2 : Tyr15 OH
Arg47	**Tys3** (1.1, -2.4), Gln4 (-11.6, -1.6), Val5 (-0.5, -3.1), Tys14 (-0.1, -3.6), Glu18 (-1.2, -0.5), Tyr187 (0.6, -1.7)	‡**Arg47** NE : Gln4 O, ‡**Arg47** NH2 : Gln4 O, ‡**Arg47** N : Tys14 OS*
Gln48	Met1 (-0.3, -3.8), **Tys3** (-0.9, -0.6), Tys14 (-1.7, -1.8), Glu18 (-1.0, -2.8)	‡Gln48 NE2 : Tys14 OS*, ‡Gln48 NE2 : Tys14 OE*
Val49	Met1 (1.3,-2.6), **Tys3** (-4.6, -4.3)	
Cys50	Met1 (-0.1, -2.7), **Tys3** (-6.2, -2.4)	Cys50 N : **Tys3** OS4

CCL5 and CCR5 residues which are marked in **boldface** are experimentally associated with binding-signaling (see *Discussion*). The results correspond to analysis of 1000 snapshots of the Complex 14A simulation. □Principal interacting CCL5[§]- CCR5[¶] residue pairs. □For each pair, the average polar and nonpolar average interaction free energies (polar, non-polar), are provided in parentheses next to each CCR5 residue; all energies are in kcal/mol. †Salt bridges between CCL5 and CCR5 residue pairs. ‡Hydrogen bonds between CCL5 and CCR5 atom pairs. The asterisk (*) symbol used after any CCL5/CCR5 atom in the hydrogen bonding pair denotes that any of the atoms in the charged, carboxyl or amide, side-chain group may participate in the hydrogen-bond formation. Residue ''Tys'' corresponds to a sulfated tyrosine.

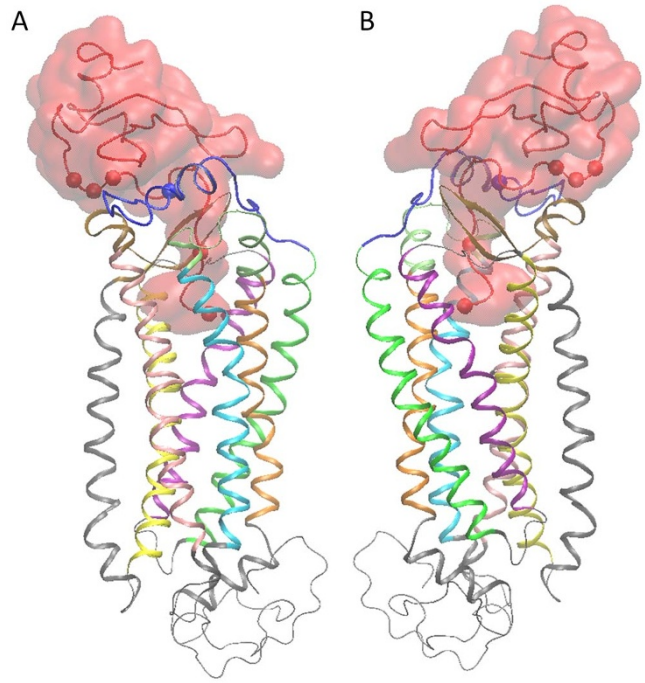

Figure 1 | Entire Simulation System of the CCL5 : CCR5 Complex Structure: Molecular graphics images of the entire simulation system corresponding to simulation 14A. Panel (B) is rotated around the z-vertical axis with regard to panel (A). CCL5 is shown in tube and transparent surface representation in red color. CCR5 is shown in cartoon representation, and the coloring used for different protein domains is as follows: (i) N-terminal domain is colored in blue, (ii) Transmembrane helix 1 (TH1) is colored in green; (iii) Intracellular loop 1 (ICL1) is colored in light gray; (iv) TH2 is colored in purple, (v) Extracellular loop 1 (ECL1) is colored in light gray; (vi) TH3 is colored in yellow; (vii) ICL2 is colored in light gray; (viii) TH4 is colored in gray; (ix) ECL2 is colored in ochre; (x) TH5 is colored in pink; (xi) ICL3 is colored in light gray; (xii) TH6 is colored in cyan; (xiii) ECL3 is colored in lime; (xiv) TH7 is colored in orange; (xv) C-terminal domain is colored in light gray. The N-terminal Cα atom of CCR5 is shown in a small van der Waals sphere. The Cα atoms of CCL5 residues 1, 5, 15, 16, 17, from bottom to top, are shown in small van der Waals spheres.

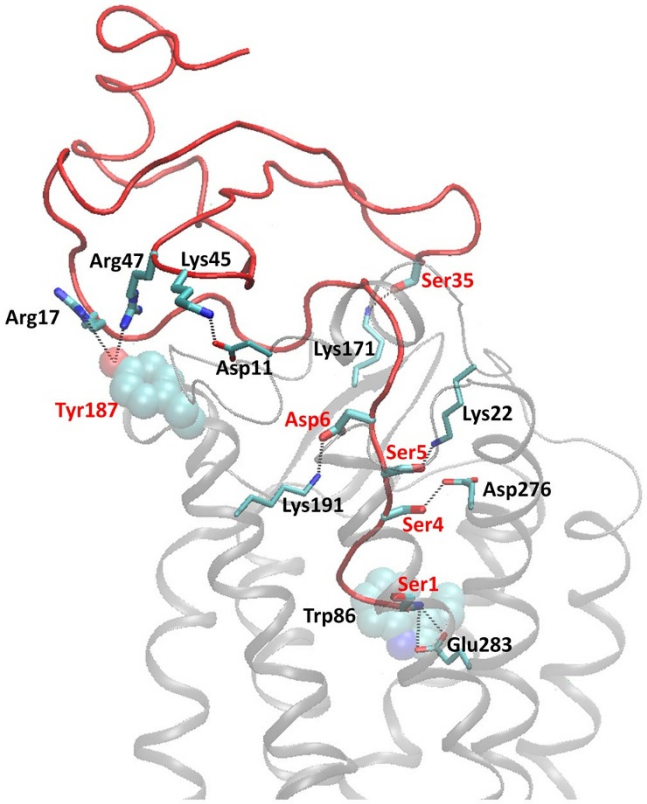

Figure 2 | Salt Bridges, Important Hydrogen Bonds and Important Cation-π Interactions Formed between CCL5 and CCR5 residues: Molecular graphics image depicting the salt bridges, important hydrogen bonds and important cation-π interactions between CCL5 and CCR5 residues in the simulation. CCL5 is shown in red tube representation and CCR5 is shown in light gray transparent tube representation. The hydrogen bonds are denoted in dashed lines and the participating CCL5 and CCR5 residue moieties are shown in licorice. CCR5 residues participating in cation-π interactions are shown in transparent van der Waals representation. CCL5 and CCR5 residues are annotated in red and black, color, respectively. Hydrogen atoms are omitted for clarity.

bond with the charged N-terminal of CCR5 residue Met1. CCL5 residues Cys10 and Cys11 are enveloped between, and interact with, the non-polar side chain moieties of CCR5 residues Met1, Tys3 (sulfated tyrosine 3), Gln21, Asn24, Lys171, Glu172 and Gly173; apart from the non-polar interactions, the backbone carbonyl of Cys10 is hydrogen bonded to the charged N-terminal of CCR5 residue Met1, and the sulfur atom of Cys11 is polarly attracted to the charged side chain group of CCR5 residue Tys3. CCL5 residue Phe12 participates in strong hydrophobic interactions with the non-polar side chain moieties of CCR5 residues Met1, Glu170, Lys171, Glu172, Ser179 and the aromatic groups of CCR5 residues Tys3 and Tyr184; also, the backbone moiety of CCL5 residue Phe12 forms a high-occupancy hydrogen bond with the charged side chain group of CCR5 residue Tys3, and less frequently with CCR5 atoms Gln170 NE2 and Lys171 N. The group of CCL5 residues Ala13, Tyr14, Ile15, Ala16 and Arg17 participates in strong hydrophobic interactions with a cluster comprising the non-polar side chain moieties of CCR5 residues Tys3, Gln170, His181, Tyr184, Ser185 and Tyr187; these interactions are facilitated in the presence of high occupancy polar hydrogen bond interactions between atom pairs: Ala13 N : Tys3 OS3, Tyr14 N : Tys3 OS3, Ala16 N : Tyr184 O, and Arg17 NH1 : Tyr187 OH (Figure 2).

Interactions of CCL5 Residues 25:50 with CCR5. CCL5 residue Lys25 forms a high occupancy salt bridge with CCR5 residue Tys14. The hydroxyl side chain group of CCL5 residue Ser31 is throughout the simulation attracted to the polar side chain group of CCR5 residue Gln21, owing to frequent hydrogen bond interactions among atom pairs Ser31 OG : Gln21 OE1/NE2. CCL5 residue Lys33 is in proximity to CCR5 residues Gln21, Asn24 and Leu174, and is involved in the following low occupancy hydrogen bonds: Lys33 O : Asn24 ND2, Lys33 NZ : Gln21 OE1, and Lys3 NZ : Asn24 OD1. CCL5 residue Cys34 is in the vicinity of CCR5 residue Gly173. CCL5 residue Ser35 forms non-polar contacts with CCR5 residue Leu174, and is polarly attracted to CCR5 residues Lys171 and Gly173 due to the occurrence of hydrogen bonds among atom pairs: Ser35 N : Gly173 O, Ser35 OG : Lys171 NZ (Figure 2), Ser35 OG: Gly173 O. CCL5 residue Asn36 is in the vicinity of CCR5 residue Lys171, and consequently, Asn36 OD1 is infrequently hydrogen bonded to Lyz171 NZ. CCL5 residues Val40 and Val42 are adjacent to CCR5 residues Met1 and Tys14, respectively. The non-polar moiety of CCL5 residue Arg44 is attracted to the hydrophobic CCR5 residue Val4, and the charged amide of Arg44 is hydrogen bonded to the hydroxyl group of Tyr187. CCL5 residue Lys45 forms a highly interacting salt bridge with CCR5 residue Asp11 (Figure 2), is polarly attracted to oppositely charged CCR5 residue Tys14, and its non-polar side chain moiety is non-polarly attracted to CCR5 residues Val5 and Ser6. CCL5 residue Asn46 is enveloped among

the side chain groups of CCR5 residues Gln4, Asp11, Tys14 and Tys15, and is position is stabilized owing to the hydrogen bonding interaction between the backbone and side chain amide groups of Asn46 with (i) the negatively charged side chain group of Tys14, as well as (ii) the hydroxyl group of Tyr15. CCL5 residue Arg47 participates in a wealth of polar and non-polar interactions with CCR5 residues which include (i) cation-π interactions formed between Arg47 and CCR5 residues Tys3 and Tyr187 (Figure 2), (ii) a hydrogen bond interaction formed between the backbone amide of Arg47 and the negatively charged side chain group of CCR5 residue Tys14, (iii) a hydrogen bond interaction formed between the positively charged side chain amide of Arg47 and the backbone carbonyl of CCR5 residue Gln4, (iv) non-polar and –mainly– polar interactions with CCR5 residues Val5 and Glu18, respectively. CCL5 residue Gln48 intercalates among the side chains of CCR5 residues Met1, Tys3, Tys14 and Glu18, and its stable positioning is enhanced due to hydrogen bond interactions which are formed between Gln48 NE2 and the charged side chain groups of CCR5 residues Tys14 and Glu18. CCL5 residues Val49 and Cys50 are proximal to CCR5 residues Met1 and Tys3, and the backbone amide of Cys50 is hydrogen bonded to the charged side chain group of Tys3.

Comparison between the CCL5 : CCR5 and CXCL12 : CXCR4 Binding Modes. We used TM-align[44] to superimpose the CCL5 : CCR5 and CXCL12 : CXCR4[42] complex structures (see Supplementary Figure 2A). Owing to the approximately 30% sequence homology between the two chemokine receptors, and the fact that the CCR5/CXCR4 domains (e.g. N-terminal domains, extracellular loops) have different conformations and sequence lengths, an RMSD comparison cannot enlighten and fully describe the similarities/differences between the CCL5 : CCR5 and CXCL12 : CXCR4 binding modes: the backbone RMSD between the CCR5 and CXCL12 1–48 residue moieties is approximately 9 Å; this result indicates the two chemokines do not possess the same structural arrangement in complex with the chemokine receptors, but it does not necessarily imply a dissimilarity between the CCL5 : CCR5 and CXCL12 : CXCR4 binding modes. Thus, to compare between the two binding modes, we performed a sequence alignment of the two chemokine receptors to identify the homologous receptor residues, and subsequently, we compared the intermolecular Cα : Cα distances of CCL5/CXCL12 : (homologous) CCR5/CXCR4 residue pairs which are located on the protein binding interface in the final simulation snapshots (see Supplementary Figure 2B). The advantage of this analysis is that it does not depend on the structural superposition of the two complex structures, and depends only on relative intermolecular distances which are indicative of the binding mode properties. The results show that the main dissimilarities between the intermolecular Cα : Cα distances of CCL5/CXCL12 : (homologous) CCR5/CXCR4 residue pairs involve contacts of chemokine residues with the N-terminal domains of the two chemokine receptors, and this can be attributed to the different conformations and sequence lengths of the CCR5/CXCR4 N-terminal domains. In addition, some dissimilarities of intermolecular Cα : Cα distance pairs involve contacts between the chemokine residue moieties 32–37, 46–53 with the second extracellular loop of the chemokine receptors. The overall analysis shows that the similar neighboring chemokine : chemokine receptor residue pairs are approximately twice compared to the dissimilar ones, for all chemokine receptor domains excluding the N-terminal regions. Therefore, the CCL5/CXCL12 : homologous CCR5/CXCR4 neighboring residue pairs, outside the N-terminal regions of the receptors, are mostly similar, and consequently, our analysis suggests that the binding mode of CCL5 in complex with CCR5 is similar to the binding mode of CXCL12 in complex with CXCR4[42] (see Supplementary Figure 2A).

Molecular recognition of CCR5 by CCL5 versus the HIV-1 gp120 V3 loop. The residue pair-wise interaction free energies (Supplementary Figure 1) are summed up for every CCR5 residue in complex with CCL5 and are presented in Figure 3. Figure 3 also presents the residue pair-wise interaction free energies summed up for every CCR5 residue in complex with the dual tropic HIV-1 gp120 V3 loop of[9]. Thirty CCR5 residues (i.e., Glu283, Tys3, Tys14, Asp11, Met1, Lys22, Asp276, Glu18, Asp2, Gln4, Ser179, Glu262, Tyr184, Gln280, Gln21, Tyr15, Lys191, Glu172, Trp86, Cys178, Met279, Ser272, Tyr251, Asn24, Gln170, Val5, Asn258, Cys20, Gln261, Phe264), interact significantly with both the chemokine and the virus[9] and are included in panel A (see *Supplementary Information*). Panel B includes five CCR5 residues (i.e., Lys171, Tyr187, Gly173, Thr167 and Leu174) which interact with the chemokine and to a smaller (or considerably smaller) extent with the virus[9], and in addition, eleven CCR5 residues (i.e., His289, Ser180, Pro183, Ile12, Thr177, Gly286, Ser17, Tyr89, Gln188, Tyr108, and Tys10) which interact with the virus[9] and to a smaller (or considerably smaller) extent with the chemokine. Panel A comprises approximately twice as many residues than panel B. Thus, the chemokine and the virus[9] primarily interact mostly with the same CCR5 residues and share to a large extent the same CCR5 binding pocket. Figure 4 presents the superimposed CCR5 structures in complex with the HIV-1 gp120 V3 loop[9] and CCL5 and depicts in fat tube representation the virus and chemokine residues which exhibit significant overlap with regard to the conformational space they occupy in complex with CCR5. Specifically, the HIV-1 gp120 V3 loop residue domain 5–33 (residue domain 300–329 in the complete gp120 sequence[45]) and CCL5 residue domains 1–8, 12, 13, 15, 43–50, occupy the same conformational space.

Discussion

A series of experiments depicted that certain CCR5 residues are involved in the CCL5 mediated binding-signaling. In 1998, a study by Rabut *et al.*[29] showed that CCR5 mutants Y3A and C20A caused a 25% and 70%, respectively, impairment in the inhibition of HIV-1 coreceptor function by CCL5. Also, Dragic *et al.*[35] in 1998 showed that CCR5 mutants D11A, K22A, R31A, H181A, Y184A, K171A/E172A, K191A/N192A, R274A/D276A were relatively insensitive to inhibiting the HIV-1 coreceptor function by CCL5. Moreover, Blanpain *et al.*[32] suggested that residues 10–13 of CCR5 are important for CCL5 binding. In 2001, a study by Govaerts *et al.*[26] indicated that CCL5 displays an unaffected affinity for the CCR5 mutants T82S, T82C, T82A, and T82V. In a subsequent study Govaerts *et al.*[27] in 2002, showed that CCR5 mutants F85L, F85L/L104F, F85L/Y89S do not significantly influence binding, the mutant Y89S/L104F reduces binding significantly, and mutants F85L, L104F, F85L/L104F, Y89S, F85L/Y89S, Y89S/L104F, Y108A, F112Y and L104F/F109H/F112Y result in reduced signaling; among the aforementioned mutants, Y89S/L104F results in the most significant loss in signaling. In addition, Blanpain *et al.*[31] in 2002 concluded that the core domain of CCL5 binds distinct residues in CCR5 extracellular domains, whereas the N terminal domain of chemokines mediates receptor activation by interacting with the transmembrane helix bundle. In the same study the authors showed that (i) specific binding to K191A-expressing cells was below the limit of detection, even using higher concentrations of [^{125}I]CCL5, making it impossible to determine chemokine binding affinities for this mutant, and (ii) specific CCR5 mutants L104F, F85L/L104F, which involve aromatic residues in transmembrane domains, exhibited reduced functional response compared to the native CCL5. Maeda *et al.*[33] in 2006 investigated the effect of CCR5 amino acid substitutions on the inhibition of HIV-1 coreceptor function by CCL5 and depicted that CCR5 mutants D11A, Y37A, F112Y, C178A, K191A, Y251A and M287E experience a significant loss in anti HIV-1 function, mutants F112L and E283A experience a moderate loss in anti HIV-1 function, and

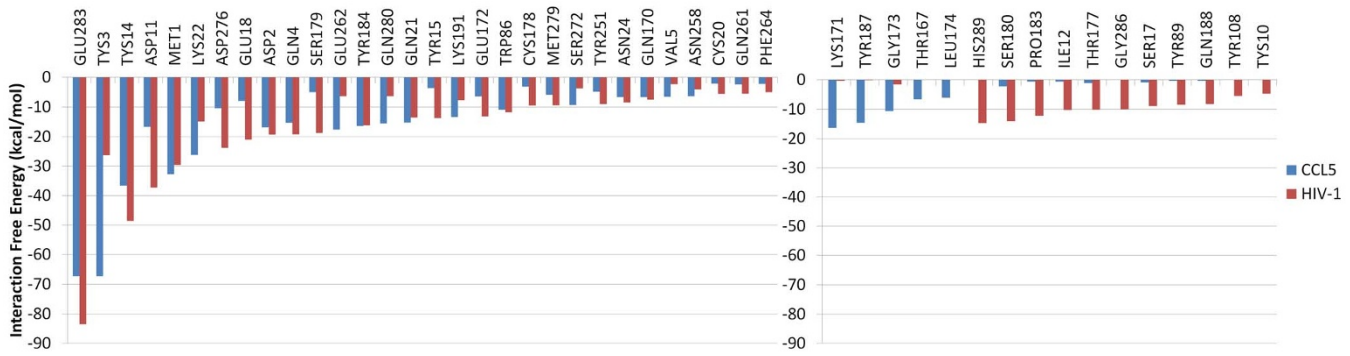

Figure 3 | Interaction Free Energies of CCR5 Residues in Complex with CCL5/HIV- gp120 V3 Loop: The residue pair-wise interaction free energies were summed up (y axis) for every CCR5 residue (x axis), in complex with (i) CCL5 (first column per CCR5 residue) and (ii) the dual tropic HIV-1 gp120 V3 loop of[9] (second column per CCR5 residue). The figure includes only CCR5 residues which possess at least −4.5 kcal/mol total interaction free energy in at least one of the two complexes (i) or (ii), and is partitioned in panels (A) and (B). If a CCR5 residue interacts strongly and approximately equally with both CCL5 and the HIV-1 gp120 V3 loop, it is listed in panel A, whereas, if a CCR5 residue interacts strongly with CCL5 and weakly with HIV-1 gp120 V3 loop, or vice versa, it is listed in panel B (see *Supplementary Information*). Residues in panels A and B are presented in descending order of magnitude of interaction free energy (averaged for the two complexes); residues in panel B are first categorized according to their property to interact strongly either with CCL5 (first five residues) or the HIV-1 gp120 V3 loop (last eleven residues). Residue name "Tys" corresponds to a sulfated tyrosine.

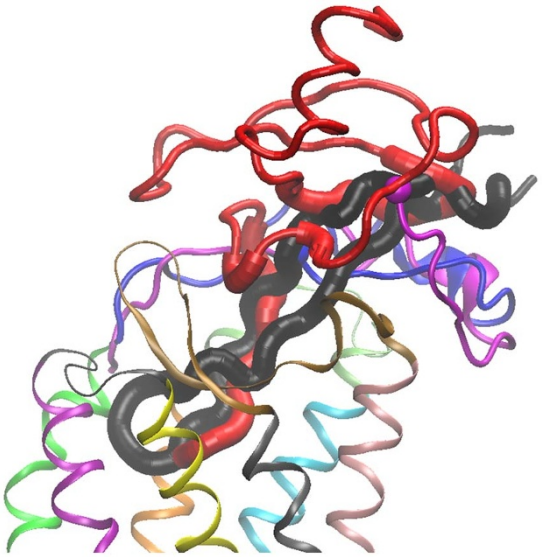

Figure 4 | Superposition of the HIV-1 gp120 : CCR5 and CCL5 : CCR5 Complex Structures: Superimposed molecular graphics images of the HIV-1 gp120 V3 loop[9] and CCL5 in complex with CCR5. The structures are aligned (superimposed) with regard to the CCR5 transmembrane backbone region and correspond to the intermediate (10 ns) snapshot of the simulations in the HIV-1 gp120 V3 loop[9] and CCL5 in complex with CCR5 (in this work). The molecular graphics representation of CCR5 is described in the legend of Figure 1. The CCR5 N-terminal domain in complex with the HIV-1 gp120 V3 loop is colored in magenta. The HIV-1 gp120 V3 loop[9] and CCL5 are shown in tube representation in black and red colors, respectively. Owing to the significant overlap between the conformational space occupied by the HIV-1 gp120 V3 loop residue moiety 5-33 (which corresponds to 300–329 in the entire gp120 sequence[45]) and CCL5 residue moieties 1–8, 12, 13, 15, 43–50, the aforementioned regions are shown in fat tube representation. According to backbone RMSD calculations of the two structures presented, the RMSD of the transmembrane helical domains is 2.22 Å, the RMSD of the N-terminal domains is 6.31 Å, the RMSD of extracellular loops 1, 2, 3 is 1.61, 3.08 and 3.17 Å, respectively. Despite the relatively small RMSD values obtained for the transmembrane region and extracellular loop domains, the N-terminal domain of CCR5 is "locked" in different relative orientations in complex with the HIV-1 gp120 V3 loop[9] and CCL5 (in this work).

mutants Y108A, G163R, K171A/E172A, and S180T experience a small loss in anti HIV-1 function, whereas no loss was reported for mutants I198A and M287A. Furthermore, Kondru *et al.*[30] in 2007 depicted that CCR5 mutants T195A, I198A, W248A, Y251A, and M287A, did not significantly change the binding affinity of CCL5 to CCR5; however, mutations W86A, F109A, and E283A reduced the binding affinity of RANTES to CCR5 by 5-, 10-, and 20-fold, respectively. The recent study by Thiele *et al.*[28] in 2011 showed that CCR5 mutants D276A and E283A result in a decrease in signaling.

In addition, a series of experiments provided evidence for the involvement of certain CCL5 residues in the CCR5 mediated binding-signaling. Proudfoot *et al.*[25] in 1996 showed that extension of recombinant human CCL5 by the retention of the initiating methionine produces a potent antagonist; based on these findings, the authors state that the integrity of the amino terminus of CCL5 is crucial to receptor binding and cellular activation. Pakianathan *et al.*[24] in 1997 demonstrated that alanine mutations at specific CCR5 residues of the N-terminal (residues 1:9) and N-loop (residues 12–20) domains of CCL5 result in reduced signaling. Specifically, P2A and F12A mutants almost abrogate signaling; S1A, Y3A, D6A, I15A mutants cause at least a 50% reduction in signaling; mutants S4A, S5A, Y14A, R17A cause a reduction in signaling within the range of 25–50%. In addition, Martin *et al.*[36] in 2001 introduced acetylations at positively charged residues of CCL5, and suggested that (i) the cluster of basic residues Arg44, Lys45 and Arg47 plays a moderate role with regard to the CCR5 binding and functional response, and (ii) an acetylation of CCR5 residue Lys33 impairs the ability of CCL5 to stimulate CCR5; nevertheless, this result is most probably a consequence of the high tendency of the acetylated CCL5 Lys33 derivative to form aggregates at high concentration, as suggested by gel filtration experiments. Also, according to Choi *et al.*[14] in 2011, (i) CCR5 mutations distinguish N-terminal modifications of CCL5 with agonist versus antagonist activity, and (ii) the transmembrane region of CCR5, which is a key interaction site for inhibitors, is a sensitive molecular switch modulating receptor activity.

The complex structure derived in our study sheds light into, and provides compelling evidence for the role of all aforementioned CCR5 and CCL5 residues and their involvement in binding and/or signaling. Firstly, the computationally derived structure is in exceptional accordance with the fact that the core domain of CCL5, i.e. the so called "N-loop" (which comprises CCL5 residues 10–17), three antiparallel β-strands and the C-terminal α-helix, binds distinct residues in CCR5 extracellular domains, whereas the N-terminal domain

of CCL5 mediates receptor activation by interacting with the transmembrane helix bundle[31]. In our simulation, the charged N-terminal of Ser1 forms a highly interacting salt bridge with CCR5 residue Glu283 and is also polarly attracted to Asp276; as experiments underlined the key role of both the N-terminal end of CCL5 and CCR5 residues Glu283 and Asp276 in signaling, the aforementioned interactions should be also considered most critical for signaling. The significantly strong polar interaction between the charged N-terminal of CCL5 and CCR5 residue Glu283 suggests the critical role of Ser1 in guiding CCL5 to obtain a proper position and orientation with regard to CCR5 for optimal binding. Experiments show that CCR5 mutant M287A does not affect the binding of CCL5 to CCR5, however the M287E mutant results in a significant loss in anti HIV-1 function; owing to the proximity of residue position 287 to the charged N-terminal, our results suggest that the M287E mutant could falsely orient the charged N-terminal of CCL5 toward M287E, and thus disallow the formation of the most highly interacting salt bridge formed between the charged N-terminal and Glu283. In addition, a juxtaposition of the knowledge derived by previous experimental studies to the findings of our study, suggests that the identified strong non-polar interactions formed between CCL5 residues Ser1, Pro2 and aromatic CCR5 residues Tyr37, Trp86, Tyr89, Tyr108, Phe109, should be of utmost importance for signaling. It is worth noting that CCR5 residues Arg31, Phe85, Leu104 and Phe112 also contribute to the formation and stabilization of the aforementioned interactions, as (i) the Arg31 forms a cation-π interaction with CCR5 residue Tyr89, and (ii) aromatic residues Phe85, Leu104, Phe112 are part of an aromatic cluster which additionally comprises CCR5 residues Tyr37, Trp86, Tyr89, Tyr108, Phe109 and Tyr251. Also, the important for signaling CCL5 residues Ser2, Tyr3, Ser4 and Ser5 are involved in polar and non-polar interactions, which include experimentally key CCR5 residues Lys22, Gly163, Glu172, Cys178, Ser180 and Asp276; additionally, the low intensity interactions formed by Tyr3 of CCL5 and Thr195, Ile198 of CCR5 also justify the experimentally-defined minor role of the latter residues. Furthermore, the high occupancy salt bridge formed between CCL5 residue Asp6 and CCR5 residue Lys191, as well as the strong polar interactions formed between CCL5 residue Asp6 and CCR5 residue Lys22, justify the critical role of CCL5 residue Asp6 in signaling, and CCR5 residues Lys22, Lys191 in binding; our study suggests that Lys22 and/or Lys191 can also be important for signaling. CCL5 residue Phe12 is, according to experiments, most important for signaling, and according to our study, this can be attributed to its numerous interactions with CCR5; these interactions are associated with CCR5 residues Tys3, Lys171, Glu172, Tyr284, which are also deemed important for binding, and thus, it is possible that all, or a portion of the aforementioned CCR5 residues, are/is important for signaling, as well. While experiments did not investigate the role of CCL5 residues Cys10, Cys11, Ala13 and Ala16, our computationally derived structure predicts their involvement in binding as they are involved in interactions with several CCR5 residues, which include the key CCR5 experimentally defined residues: Tys3, Lys171, Glu172, Tyr184 and His181. CCL5 residues Tyr14, Ile15 and Arg47 are experimentally important for signaling and this can be attributed to their interactions with CCR5 residues, which include CCR5 residues Tyr3 and Tyr184. These residues are important according to experiments and this finding additionally suggests that Tyr3 and/or Tyr184 can be significant for signaling, as well. Moreover, the role of CCL5 residues Arg44, Lys45 and Arg47 with regard to binding and functional response can be interpreted by our findings, as they are involved in an abundance of interactions with CCR5. These interactions comprise a high occupancy formed by CCL5 residue Lys45 and CCR5 residue Asp11; owing to this salt bridge, our study can also provide evidence for the key role of CCR5 residue Asp11 in blocking HIV-1 function which occurs owing to the CCL5 binding to CCR5. Last but not least, the important role of CCR5 residues Tys3 and Asp11 could additionally be attributed to their interactions with the CCL5 residue moiety 47–50 and residue Asn46, respectively.

Our study presents the first structural (Figure 4) and interaction free energy (Figure 3) based comparison between the HIV-1 : CCR5 and CCL5 complexes, and thus, it sheds light into the blocking mechanism of HIV-1 by CCL5. Interestingly, while (i) the bound structures of the HIV-1 gp120 V3 loop[9] and CCL5 (Figure 1) differ significantly, and (ii) the unbound HIV-1 gp120 V3 loop[10,46] and the N-terminal domain of CCL5[39] are highly flexible, the present work shows that both the chemokine and the virus primarily interact mostly with the same CCR5 residues, and share the same chemokine receptor binding pocket (Figures 3 and 4). Both the virus[9] and the chemokine interact most strongly with CCR5 residue Glu283; CCR5 residue Glu283 forms Coulombic interactions with (i) the "central-tip" Arg18 (Arg315 in the complete HIV-1 gp120 sequence[9], (ii) and the charged N-terminal of CCL5.

The binding mode of CCL5 : CCR5 is similar to the binding mode of CXCL12 : CXCR4[42] (see Supplementary Figure 2). According to both complex structures, the bound properties of CCL5/CXCL12 vary from their unbound properties: while the 1–11 domain is the most flexible region of the CCL5/CXCL12 solution conformations[39,42], the same domain turns out to be the most rigid domain upon binding to CCR5/CXCR4[42]. In addition, in both complex structures, the charged N-termini of CCL5/CXCL12 interact with Glu283/Glu288 of CCR5/CXCR4, respectively; interestingly, the aforementioned chemokine receptor residues are most critical for the HIV-1 gp120 binding, as well[9,10].

The design of novel CCR5 antagonists which can therapeutically block (i) the HIV-1 entry to CCR5[9], or (ii) the CCL5:CCR5 axis which is associated with cancer[1,3,16–21] is of significant medical importance. As a proof of concept for the first direction, the binding of CCL5 and CCL5 derivatives, to CCR5 is – experimentally considered –a potential HIV-1 therapeutic axis[11–15]. Thus, we propose that the computational derived CCL5 : CCR5 structure which is in excellent agreement with previous experimental findings can be used to design[47–53] modified-novel CCL5-based peptides which can potentially constitute therapeutic agents for HIV-1 and cancer.

Methods

We employed a computational protocol, which consists of six major steps and is primarily based on free energy calculations and molecular dynamics simulations to derive the CCL5 : CCR5 complex structure. 1) We generated and selected structural templates for CCL5 and CCR5; 2) We docked selected CCL5 conformations on selected CCR5 conformations; 3) We used experimental constraints to discard falsely-docked complex structures which are not associated with signaling; 4) We employed the membrane GBSA[54] approximation, as in[9,10,42], to minimize the energy and calculate the binding free energy for the selected docked complexes of step 3; 5) We selected the docked complexes which possess the lowest binding free energy according to step 4 and performed MD simulations of the selected complexes; 6) We calculated the binding free energy for the simulated complex structures using the membrane MM GBSA[54] approximation, as in[9,10,42], and identified the complex structure with the lowest average binding free energy. CHARMM[55] was used to perform the free energy calculations and the MD simulations. Steps 1–6 are analytically described below.

1) Sixty six CCL5 structures were selected as flexible ligand templates for docking. Thirteen structures correspond to NMR conformations which are deposited in the PDB entry 1HRJ[38], and two structures correspond to NMR conformations which are deposited in the PDB entry 2L9H[39]. The rest of the conformations were produced by considering the 7th and 13th NMR structures of[38] as initial coordinates in two independent 0.28 μs replica exchange MD (REMD) simulations with the FACTS19 implicit solvent model[56]. The FACTS implicit solvent model in conjunction with replica exchange simulations possesses the capacity to reproduce results of explicit solvent simulations[57–59]. We introduced light shape "bestfit" harmonic constraints on the backbone of the 12–68 residue moiety. The choice to use 7th and 13th conformations of[38] as initial coordinates in the REMD simulations was based on the fact that they encompass the key secondary structural elements of the CCL5 12–68 residue moiety. REMD simulations were conducted so as to achieve enhanced conformational sampling of the flexible N-terminal 1–11 domain of CCL5; experiments have confirmed that the N-terminal domain of CCL5 is extremely flexible[39]. Upon the completion of the REMD simulations,

the CCL5 conformations were clustered based on the Cα coordinates, using a 1.5 Å clustering radius, in WORDOM[60]. Prior to docking, the sixty six CCL5 structures were subjected to two hundred steps of steepest minimization in implicit solvation[61]; the backbone heavy atoms were constrained during the minimization. As far as CCR5 is concerned, the fifteen clustered CCR5 conformations produced in[9] and the two CCR5 X-ray structures[40] were considered as flexible receptor templates for docking. The construction of any missing N-terminal domain and intracellular loop residues of the X-ray structures was achieved using the refined conformations of the corresponding regions produced in[9].

2) We docked sixty six CCL5 structures to the seventeen CCR5 structures using the parallel linux version of Zdock v.3.0.2[62]. For each docking run, 2000 docked structures were produced, and consequently, the total number of complex structures produced was 2,244,000; see reference[42] for additional information regarding the methodology used.

3) Experimental evidence depicts that the N-terminus of CCL5 (i) interacts with the transmembrane helical bundle of CCR5[31] and (ii) is crucial for activation[25]. We utilized this information to exclude all docked conformations at which the N-terminal end of CCL5 was not penetrating into the CCR5 transmembrane region. As a result, we collected 17,113 conformations according to the following criteria: (i) the z coordinate value of atom CCL5 Ser1 Cα was less than 15.5 Å, and (ii) the N-terminal domain was not contacting the exterior surface of the CCR5 transmembrane helical bundle.

4) The selected 17,113 complex conformations were subjected to three hundred steps of steepest minimization in a heterogeneous water-membrane-water environment, implicitly modeled by GBSW[54] as in[9,10,42]. Subsequently, as in[9,10,42], we used the expression $\Delta G = E_{PL} - E_P - E_L$, where E_X corresponds to the total (free) energy of molecule X (complex PL: CCR5 : CCL5, free protein (P): chemokine receptor CCR5, or free ligand (L): chemokine CCL5), to calculate the binding free energy. In the calculations, we assumed that the conformations of the chemokine and the chemokine receptor are identical in their complex and unbound states, as in[9,10,42,49,63,64], and thus, bonded-energy contributions are canceled. The heterogeneous water-membrane-water environment[54] was employed to calculate the solvation free energy components of the complex and unbound chemokine receptor, while a homogeneous aqueous environment[61] was employed to calculate the solvation free energy of the unbound chemokine. As a result of this procedure, we spotted the complex structure acquiring the lowest binding free energy −170.4 kcal/mol, and subsequently, we selected the twenty five complex structures acquiring the lowest binding free energy, within the range of (−170.4 kcal/mol : −146.5 kcal/mol), to be included in the analysis of the following steps (see Supplementary Table 2).

5) We used the twenty five complex structures as starting conformations to conduct twenty five independent MD simulations. Prior to the production MD simulation runs, a 400-ps heating stage and a subsequent 700-ps equilibration stage were conducted; during the equilibration stage, the harmonic constraints were gradually removed from the chemokine receptor and the chemokine. During the 30-ns production MD simulation runs, the complexes were not constrained. We considered the first 10-ns of each trajectory as additional equilibration time, and thus, only the last 20-ns of the simulation trajectories were analyzed; see references[9,10,42] for additional information regarding the methodology and the force field parametrization used in the simulations.

6) We calculated the average binding free energy for 1,000 complex structures per simulation trajectory, using the MM GBSA method[9,10,42] (see Supplementary Table 2). The 1,000 snapshots were extracted every 20-ps intervals, from the last 20-ns of each of the twenty five trajectories. As in step 4, the heterogeneous water-membrane-water environment[54] was employed to calculate the solvation free energy components of the complex and unbound chemokine receptor, while a homogeneous aqueous environment[61] was employed to calculate the solvation free energy of the unbound chemokine. According to the results which are included in Supplementary Table 2, the binding free energy range of the simulated complexes is expanded (−346.1 kcal/mol : −263.3 kcal/mol) compared to step 4, and in addition, the average binding free energy values of the structures produced in the simulations are significantly decreased with respect to the binding free energy values of the corresponding starting structures produced in step 4 (see Supplementary Table 2). The values of the binding free energies are significantly reduced in the MD simulations compared to the starting conformations, and this underlines the beneficial role of MD simulations with regard to providing structural relaxation, refinement and optimization of interactions within the simulated complexes. As in[42], the simulations enable the occurrence of salt bridges between adjacent oppositely charged residues and hydrogen bonds between adjacent hydrogen donor-acceptor pairs, and in addition, the simulations contribute to stabilization of non-polar contacts. Based on Supplementary Table 2, Complex 14 possesses the lowest binding free energy among all complexes. As the average binding free energy of Complex 14 falls within a standard deviation (≈16 kcal/mol) of the second most promising complex, Complex 17, we aimed at validating that Complex 14 is indeed energetically favored compared to Complex 17. To address this, we performed an additional 20-ns simulation for each of the two complexes, using the lowest binding free energy snapshot from the previously simulated Complexes, 14 and 17, as a starting structure for a new additional round of MD simulations; the additional simulations were conducted with the MD protocol of step 5, and are referred as Complex 14A and Complex17A, respectively. Subsequently, we extracted 1000 snapshots from Complex 14A and Complex17A and evaluated the average binding free energy by employing the MM GBSA approximation. This additional analysis verified that Complex 14A is indeed more energetically favored that Complex 17A (by (≈13 kcal/mol). Also, it is worth noting that, while the interactions formed within Complex 14A (as well as Complex 14) are in excellent agreement with previous experimental findings, and thus Complex 14A is capable of shedding light into the role of key CCL5 and CCR5 residues for binding and signaling, the interactions formed within Complexes 17 or 17A do not possess the capacity to sufficiently interpret previous experimental findings. The 1000 simulation snapshots extracted from the new simulation, Complex 14A, are considered for subsequent analysis. Supplementary Table 2 presents the average binding free energies of the 25 complexes, as well as for the additional simulations of Complexes 14A and 17A.

Analysis of Complex 14A. We calculated the average intermolecular CCL5 : CCR5 residue pair-wise interaction free energies for 1,000 complex structures which correspond to Complex 14A using Eq. (1):

$$\Delta G^{\text{inte}}_{RR'} = \underbrace{\sum_{i\in R}\sum_{j\in R'}\left(E^{\text{Coul}}_{ij} + E^{\text{GB}}_{ij}\right)}_{\Delta G^{\text{polar}}_{RR'}} + \underbrace{\sum_{i\in R}\sum_{j\in R'} E^{\text{vdW}}_{ij} + \sigma\sum_{i\in R,R'}\Delta S_i}_{\Delta G^{\text{nonpolar}}_{RR'}} \quad (1)$$

The first term of Eq. (1) describes the polar interactions between residues R and R', and the second term of Eq. (1) describes the non-polar interactions between residues R and R'. The results of this analysis are presented in Table 1 and Supplementary Figure 1. The analysis of intermolecular interaction free energies between two interacting groups R and R' provides invaluable insights into the key intermolecular interactions[9,10,42] and key interacting residues[49,50,63,64] in an MD simulation trajectory. This methodology was also applied by Tamamis and Floudas to analyse (i) the key interactions between an HIV-1 dual tropic V3 loop in complex with CCR5[9] and CXCR4[10], and (ii) the key interactions between chemokine CXCL12 and CXCR4[42]; see reference[42] for additional information regarding the methodology used.

1. Weitzenfeld, P. & Ben-Baruch, A. The chemokine system, and its CCR5 and CXCR4 receptors, as potential targets for personalized therapy in cancer. *Cancer Lett.* In Press, doi:10.1016/j.canlet.2013.10.006. (2013).
2. Bachelerie, F. *et al.* International Union of Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. *Pharmacol. Rev.* **66**, 1–79 (2013).
3. Soria, G. & Ben-Baruch, A. [The CCL5/CCR5 Axis in Cancer]. *Cancer Drug Discovery and Development: Chemokine Receptors in Cancer.* [Fulton, A. (ed.)] [109–130] (Humana Press, New York, 2009).
4. Luther, S. A. & Cyster, J. G. Chemokines as regulators of T cell differentiation. *Nat. Immunol.* **2**, 102–107 (2001).
5. Murphy, P. M. *et al.* International union of pharmacology. XXII. Nomenclature for chemokine receptors. *Pharmacol. Rev.* **52**, 145–176 (2000).
6. Tyner, J. W. *et al.* CCL5-CCR5 interaction provides antiapoptotic signals for macrophage survival during viral infection. *Nat Med.* **11**, 1180–1187 (2005).
7. Sullivan, N. L. *et al.* Importance of the CCR5-CCL5 axis for mucosal Trypanosoma cruzi protection and B cell activation. *J. Immunol.* **187**, 1358–1368 (2011).
8. Ishida, Y. *et al.* Pivotal role of the CCL5/CCR5 interaction for recruitment of endothelial progenitor cells in mouse wound healing. *J. Clin. Invest.* **122**, 711–721 (2012).
9. Tamamis, P. & Floudas, C. A. Molecular recognition of CCR5 by an HIV-1 gp120 V3 loop. *PLoS ONE* **9**, e95767 (2014).
10. Tamamis, P. & Floudas, C. A. Molecular recognition of CXCR4 by a dual tropic HIV-1 gp120 V3 loop. *Biophys J.* **105**, 1502–1514 (2013).
11. Nardese, V. *et al.* Structural determinants of CCR5 recognition and HIV-1 blockade in RANTES. *Nat Struct Biol.* **8**, 611–615 (2001).
12. Coffey, M. J., Woffendin, C., Phare, S. M., Strieter, R. M. & Markovitz, D. M. RANTES inhibits HIV-1 replication in human peripheral blood monocytes and alveolar macrophages. *Am. J. Physiol.* **272**, L1025-9 (1997).
13. Lusso, P. *et al.* Molecular engineering of RANTES peptide mimetics with potent anti-HIV-1 activity. *FASEB J.* **25**, 1230–1243 (2011).
14. Choi, W. T. *et al.* CCR5 mutations distinguish N-terminal modifications of RANTES (CCL5) with agonist versus antagonist activity. *J. Virol.* **86**, 10218–10220 (2012).
15. Gaertner, H. *et al.* Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide. *Proc. Natl. Acad. Sci. U S A.* **105**, 17706–17711 (2008).
16. Luboshits, G. *et al.* Elevated expression of the CC chemokine regulated on activation, normal T cell expressed and secreted (RANTES) in advanced breast carcinoma. *Cancer Res.* **59**, 4681–4687 (1999).

17. Hartmann, M. C. *et al.* Relationship between CCL5 and transforming growth factor-β1 (TGFβ1) in breast cancer. *Eur. J. Cancer.* **47**, 1669–1675 (2011).
18. Niwa, Y. *et al.* Correlation of tissue and plasma RANTES levels with disease course in patients with breast or cervical cancer. *Clin. Cancer Res.* **7**, 285–289 (2001).
19. Cambien, B. *et al.* CCL5 neutralization restricts cancer growth and potentiates the targeting of PDGFRβ in colorectal carcinoma. *PLoS ONE.* **6**, e28842 (2011).
20. Vaday, G. G., Peehl, D. M., Kadam, P. A. & Lawrence, D. M. Expression of CCL5 (RANTES) and CCR5 in prostate cancer. *Prostate.* **66**, 124–134 (2006).
21. Wang, S. W. *et al.* CCL5 and CCR5 interaction promotes cell motility in human osteosarcoma. *PLoS ONE.* **7**, e35101 (2012).
22. Zhou, N. *et al.* Molecular modeling and site-directed mutagenesis of CCR5 reveal residues critical for chemokine binding and signal transduction. *Eur. J. Immunol.* **30**, 164–173 (2000).
23. Fano, A., Ritchie, D. W. & Carrieri, A. Modeling the structural basis of human CCR5 chemokine receptor function: from homology model building and molecular dynamics validation to agonist and antagonist docking. *J. Chem. Inf. Model.* **46**, 1223–1235 (2006).
24. Pakianathan, D. R., Kuta, E. G., Artis, D. R., Skelton, N. J. & Hébert, C. A. Distinct but overlapping epitopes for the interaction of a CC-chemokine with CCR1, CCR3 and CCR5. *Biochemistry.* **36**, 9642–9648 (1997).
25. Proudfoot, A. E. *et al.* Extension of recombinant human RANTES by the retention of the initiating methionine produces a potent antagonist. *J. Biol. Chem.* **271**, 2599–2603 (1996).
26. Govaerts, C. *et al.* The TXP motif in the second transmembrane helix of CCR5. A structural determinant of chemokine-induced activation. *J. Biol. Chem.* **276**, 13217–13225 (2001).
27. Govaerts, C. *et al.* Activation of CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3. *J. Biol. Chem.* **278**, 1892–1903 (2003).
28. Thiele, S. *et al.* Allosteric and orthosteric sites in CC chemokine receptor (CCR5), a chimeric receptor approach. *J. Biol Chem.* **286**, 37543–37554 (2011).
29. Rabut, G. E., Konner, J. A., Kajumo, F., Moore, J. P. & Dragic, T. Alanine substitutions of polar and nonpolar residues in the amino-terminal domain of CCR5 differently impair entry of macrophage- and dualtropic isolates of human immunodeficiency virus type 1. *J. Virol.* **72**, 3464–3468 (1998).
30. Kondru, R. *et al.* Molecular interactions of CCR5 with major classes of small-molecule anti-HIV CCR5 antagonists. *Mol. Pharmacol.* **73**, 789–800 (2008).
31. Blanpain, C. *et al.* The core domain of chemokines binds CCR5 extracellular domains while their amino terminus interacts with the transmembrane helix bundle. *J. Biol. Chem.* **278**, 5179–5187 (2003).
32. Blanpain, C. *et al.* Multiple charged and aromatic residues in CCR5 amino-terminal domain are involved in high affinity binding of both chemokines and HIV-1 Env protein. *J. Biol. Chem.* **274**, 34719–34727 (1999).
33. Maeda, K. *et al.* Structural and molecular interactions of CCR5 inhibitors with CCR5. *J. Biol. Chem.* **281**, 12688–12698 (2006).
34. Nardese, V. *et al.* Structural determinants of CCR5 recognition and HIV-1 blockade in RANTES. *Nat. Struct. Biol.* **8**, 611–615 (2001).
35. Dragic, T. *et al.* Amino-terminal substitutions in the CCR5 coreceptor impair gp120 binding and human immunodeficiency virus type 1 entry. *J. Virol.* **72**, 279–285 (1998).
36. Martin, L. *et al.* Structural and functional analysis of the RANTES-glycosaminoglycans interactions. *Biochemistry.* **40**, 6303–6318 (2001).
37. Schnur, E. *et al.* NMR mapping of RANTES surfaces interacting with CCR5 using linked extracellular domains. *FEBS J.* **280**, 2068–2084 (2013).
38. Chung, C. W., Cooke, R. M., Proudfoot, A. E. & Wells, T. N. The three-dimensional solution structure of RANTES. *Biochemistry.* **34**, 9307–9314 (1995).
39. Wang, X., Watson, C., Sharp, J. S., Handel, T. M. & Prestegard, J. H. Oligomeric structure of the chemokine CCL5/RANTES from NMR, MS, and SAXS data. *Structure.* **19**, 1138–1148 (2011).
40. Tan, Q. *et al.* Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science.* **341**, 1387–1390 (2013).
41. Wu, B. *et al.* Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science.* **330**, 1066–1071 (2010).
42. Tamamis, P. & Floudas, C. A. Elucidating a Key Component of Cancer Metastasis: CXCL12 (SDF-1α) Binding to CXCR4. *J. Chem. Inf. Model.* **54**, 1174–1188 (2014).
43. Frishman, D. & Argos, P. Knowledge-based protein secondary structure assignment. *Proteins.* **23**, 566–579 (1995).
44. Zhang, Y. & Skolnick, J. TM-align: A protein structure alignment algorithm based on TM-score. *Nucleic Acids Res.* **33**, 2302–2309 (2005).
45. Huang, C. C. *et al.* Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science.* **317**, 1930–1934 (2007).
46. López de Victoria, A., Tamamis, P., Kieslich, C. A. & Morikis, D. Insights into the Structure, Correlated Motions, and Electrostatic Properties of Two HIV-1 gp120 V3 Loops. *PLoS ONE.* **7**, e49925 (2012).
47. Bellows-Peterson, M. L. *et al.* De novo peptide design with C3a receptor agonist and antagonist activities: theoretical predictions and experimental validation. *J. Med. Chem.* **55**, 4159–4168 (2012).
48. Bellows, M. L. *et al.* A. Discovery of entry inhibitors for HIV-1 via a new de novo protein design framework. *Biophys. J.* **99**, 3445–3453 (2010).
49. Tamamis, P. *et al.* Molecular dynamics in drug design: new generations of compstatin analogs. *Chem. Biol. Drug. Des.* **79**, 703–718 (2012).
50. Gorham, R. D. Jr. *et al* Novel compstatin family peptides inhibit complement activation by drusen-like deposits in human retinal pigmented epithelial cell cultures. *Exp. Eye. Res.* **116**, 96–108 (2013).
51. Smadbeck, J., Peterson, M. B., Khoury, G. A., Taylor, M. S. & Floudas, C. A. Protein WISDOM: a workbench for in silico de novo design of biomolecules. *J. Vis. Exp.* **77**, e50476 (2013).
52. Klepeis, J. L. *et al.* Integrated Computational and Experimenal Approach for Lead Optimization and Design of Compstatin Variants with Improved Activity. *J. Am. Chem. Soc.* **125**, 8422–8423 (2003).
53. Klepeis, J. L., Floudas, C. A., Morikis, D. & Lambris, J. D. Design of Peptide Analogs with Improved Activity using a Novel de novo Protein Design Approach. *Ind. Eng. Chem. Res.* **43**, 3817–3826 (2004).
54. Im, W., Feig, M. & Brooks, C. L. III. An implicit membrane generalized born theory for the study of structure, stability and interactions of membrane proteins. *Biophys. J.* **85**, 2900–2918 (2003).
55. Brooks, B. R. *et al.* CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **30**, 1545–1614 (2009).
56. Haberthür, U. & Caflisch, A. J. FACTS: Fast analytical continuum treatment of solvation. *J. Comput. Chem.* **29**, 701–715 (2008).
57. Tamamis, P., Kasotakis, E., Mitraki, A. & Archontis, G. Amyloid-like self-assembly of peptide sequences from the adenovirus fiber shaft: insights from molecular dynamics simulations. *J. Phys. Chem. B.* **113**, 15639–15647 (2009).
58. Tamamis, P. & Archontis, G. Amyloid-like self-assembly of a dodecapeptide sequence from the adenovirus fiber shaft: Perspectives from molecular dynamics simulations. *J. Non-Cryst. Solids.* **357**, 717–722 (2011).
59. Tamamis, P. *et al.* Self-assembly of an aspartate-rich sequence from the adenovirus fiber shaft: insights from molecular dynamics simulations and experiments. *J. Phys. Chem. B.* **118**, 1765–1774 (2014).
60. Seeber, M., Cecchini, M., Rao, F., Settanni, G. & Caflisch, A. Wordom: a program for efficient analysis of molecular dynamics simulations. *Bioinformatics.* **23**, 2625–2627 (2007).
61. Im, W., Lee, M. S. & Brooks, C. L. III. Generalized Born Model with a Simple Smoothing Function. *J. Comput. Chem.* **24**, 1691–1702 (2003).
62. Pierce, B. G., Hourai, Y. & Weng, Z. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One.* **6**, e24657 (2011).
63. Tamamis, P., Morikis, D., Floudas, C. A. & Archontis, G. Species specificity of the complement inhibitor compstatin investigated by all-atom molecular dynamics simulations. *Proteins.* **78**, 2655–2667 (2010).
64. Tamamis, P. *et al.* Design of a modified mouse protein with ligand binding properties of its human analog by molecular dynamics simulations: the case of C3 inhibition by compstatin. *Proteins.* **79**, 3166–3179 (2011).

Acknowledgments

C.A.F. acknowledges funding from the National Institutes of Health (R01 GM052032). The authors are pleased to acknowledge that the work reported in this paper was performed at the TIGRESS high performance computer center at Princeton University which is jointly supported by the Princeton Institute for Computational Science and Engineering and the Princeton University Office of Information Technology's Research Computing department.

Author contributions

P.T. and C.A.F. conceived the project. The computational protocol was developed and executed by P.T. and supervised by C.A.F. P.T. and C.A.F. analyzed the data and wrote the manuscript.

Additional information

Supplementary information accompanies this paper at http://www.nature.com/scientificreports

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Tamamis, P. & Floudas, C.A. Elucidating a Key Anti-HIV-1 and Cancer-Associated Axis: The Structure of CCL5 (Rantes) in Complex with CCR5. *Sci. Rep.* **4**, 5447; DOI:10.1038/srep05447 (2014).