JOURNAL OF CHEMICAL INFORMATION AND MODELING

Article

pubs.acs.org/jcim

Open Access on 03/24/2015

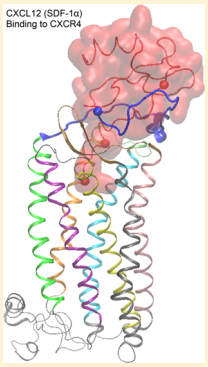
Elucidating a Key Component of Cancer Metastasis: CXCL12 (SDF-1α) Binding to CXCR4

Phanourios Tamamis and Christodoulos A. Floudas*

Department of Chemical and Biological Engineering, Princeton University, New Jersey 08544, United States

Supporting Information

ABSTRACT: The chemotactic signaling induced by the binding of chemokine CXCL12 (SDF-1α) to chemokine receptor CXCR4 is of significant biological importance and is a potential therapeutic axis against HIV-1. However, as CXCR4 is overexpressed in certain cancer cells, the CXCL12:CXCR4 signaling is involved in tumor metastasis, progression, angiogenesis, and survival. Motivated by the pivotal role of the CXCL12:CXCR4 axis in cancer, we employed a comprehensive set of computational tools, predominantly based on free energy calculations and molecular dynamics simulations, to obtain insights into the molecular recognition of CXCR4 by CXCL12. We report, what is to our knowledge, the first computationally derived CXCL12:CXCR4 complex structure which is in remarkable agreement with experimental findings and sheds light into the functional role of CXCL12 and CXCR4 residues which are associated with binding and signaling. Our results reveal that the CXCL12 N-terminal domain is firmly bound within the CXCR4 transmembrane domain, and the central 24−50 residue domain of CXCL12 interacts with the upper N-terminal domain of CXCR4. The stability of the CXCL12:CXCR4 complex structure is attributed to an abundance of nonpolar and polar intermolecular interactions, including salt bridges formed between positively charged CXCL12 residues and negatively charged CXCR4 residues. The success of the computational protocol can mainly be attributed to the nearly exhaustive docking conformational search, as well as the heterogeneous dielectric implicit water-membrane-water model used to simulate and select the optimum conformations. We also recently utilized this protocol to elucidate the binding of an HIV-1 gp120 V3 loop in complex with CXCR4, and a comparison between the molecular recognition of CXCR4 by CXCL12 and the HIV-1 gp120 V3 loop shows that both CXCL12 and the HIV-1 gp120 V3 loop share the same CXCR4 binding pocket, as they mostly interact with the same CXCR4 residues.



INTRODUCTION

Chemokine protein CXCL12, also known as stromal cell-derived factor 1 alpha (SDF-1α), binds to chemokine receptor CXCR4 and initializes chemotactic signaling.[1−6] The signaling related to the CXCL12:CXCR4 pathway is of significant biological importance, as the chemotactic responsiveness of hematopoietic stem cells (HSCs) is restricted to CXCL12,[7,8] and this unique selectivity is critical for retention and maintenance of HSCs in the hematopoietic microenvironment and the marrow-specific homing of circulating HSCs.[7,9−11] Chemokine CXCL12 is widely expressed in the central nervous system and is essential for the proper functioning of neural progenitor cells.[12] Furthermore, an important function of the CXCL12:CXCR4 axis is associated with the tissue repair and regeneration.[7,13] Also, as HIV-1 gp120 binds to CXCR4[14] (or CCR5), in one of the first and most critical steps of the HIV-1 entry to the host cell, the binding of CXCL12 to CXCR4 is a potential therapeutic axis against HIV-1.[15]

Despite the significant biological and the potential anti HIV-1 therapeutic perspective of the CXCL12:CXCR4 pathway, recent studies have provided growing evidence that CXCR4 is overexpressed in certain cancer cells,[4,16−22] and, as a result, the CXCL12:CXCR4 axis is involved in tumor progression, angiogenesis, metastasis, and survival.[23] Therefore, the CXCL12-mediated signaling is a potential mechanism of tumor resistance to both conventional therapies and biological agents through the following mode of actions: (i) by directly promoting cancer cell survival, invasion, and the cancer stem and/or tumor-initiating cell phenotype, (ii) by recruiting "distal stroma" (i.e., myeloid bone marrow-derived cells) to indirectly facilitate tumor recurrence and metastasis, and (iii) by promoting angiogenesis directly or in a paracrine manner.[3] The CXCL12:CXCR4 pathway is encountered in the trafficking of hematopoietic malignancies including chronic lymphocytic leukemia,[24−26] multiple myeloma,[27−29] other B-cell lymphomas,[30,31] and in acute leukemias.[32−35] The CXCL12:CXCR4 pathway is also involved in nonhematopoietic malignancies including breast cancer[5,36−38] and lung cancer.[39−41] Specifically, CXCR4 mediates breast cancer invasion in breast cancer metastasis.[38] Furthermore, the CXCL12:CXCR4 pathway induces migration and/or survival of the neoplastic cells, including tumor cells from brain neoplasm,[42,43] neuroblastoma cells,[44] colorectal cancer,[45] prostate cancer,[46] melanoma,[47] renal cell cancer,[48] ovarian cancer,[48] and others; CXCR4 expression of primary tumor cells correlates with recurrence, metastasis, and survival in patients with colorectal cancer[49] and melanoma.[47]

Owing to the pivotal role of the CXCL12:CXCR4 pathway in the spread and progression of a series of different types of

Received: February 3, 2014
Published: March 24, 2014

   dx.doi.org/10.1021/ci500069y | J. Chem. Inf. Model. 2014, 54, 1174−1188

tumors, the elucidation of the CXCL12:CXCR4 complex structure is of utmost biological and medical importance. No high-accuracy computational or complete experimental structure has been reported for the CXCL12:CXCR4 complex. The report of a nuclear magnetic resonance (NMR) structure of a constitutively dimeric CXCL12 in complex with a CXCR4 N-terminal fragment[50] has provided knowledge on the molecular recognition of a peptide fragment of the N-terminal of CXCR4 by CXCL12. The structure derived for the CXCR4 N-terminal 1−27 residue moiety is an important finding of the NMR study,[50] especially owing to its absence from the CXCR4 crystal structure.[51] However, there is no specific evidence that any of the interactions between CXCL12 and only the N-terminal fragment of CXCR4, reported in the NMR study, would correspond to the binding of CXCL12 to the entire CXCR4. As a proof of concept for this, in the corresponding problem of HIV-1 interaction with CCR5, Cormier et al. showed that the binding of HIV-1 gp120 to CCR5 is different when gp120 binds to (i) an N-terminal sulfopeptide of CCR5 and (ii) the entire CCR5.[52] Four computational studies aimed at investigating the molecular interactions of the entire CXCL12:CXCR4 complex structure.[53−56] Although all of the studies provide valued information, none of them reported a structure in a remarkable agreement with experiments, so as to accurately interpret the critical role of CXCL12 and CXCR4 residues for binding and signaling. Despite the fact that the last and very recently published computational study[56] utilized the CXCR4 X-ray structure,[51] and the complex structure derived was sufficiently stable, the derived models possessed specific key weaknesses which question the validity of CXCL12:CXCR4 interactions which are critical for signaling. The key weaknesses of previous computational studies (i) are mainly associated with the position and interactions formed by the critical for signaling N-terminal domain of CXCL12 (see Discussion) and (ii) can predominantly be attributed to the difficulty in sufficiently sampling and correctly modeling the bound conformation of both the highly flexible CXCL12 N-terminal domain, as well as the complete CXCR4 structure with an appropriate conformation and orientation of the N-terminal domain, as residues 1−26 are missing from the X-ray structure of CXCR4.[51]

In the present study, we exploit (i) the CXCR4 crystallographic structure,[51] (ii) our recent computationally derived CXCR4 structure in complex with a dual tropic HIV-1 gp120 V3 loop,[14] as well as (iii) the twenty CXCL12 NMR PDB structures corresponding to entry 2KEE[57] and the two CXCL12 crystallographic PDB structures corresponding to entry 2J7Z,[58] to computationally derive the first complete CXCL12:CXCR4 complex structure. We show subsequently that the derived structure is in remarkable agreement with experiments (see CXCL12 and CXCR4 residues marked in boldface in Table 1).

METHODS

Derivation of the CXCL12:CXCR4 Complex Structure. The methodology used in the present study to derive the CXCL12:CXCR4 complex structure consists of the following major steps: 1) Selection of the initial CXCL12 and CXCR4 structural templates; 2) Docking of selected CXCL12 conformations on selected CXCR4 conformations; 3) First round of energy minimization and binding free energy calculations of the docked complexes using the membrane GBSA[14,59] approximation; 4) Second round of energy minimization and binding free energy calculations of the docked complexes using the membrane GBSA[14,59] approximation; 5) MD simulations of the docked complexes acquiring the lowest binding free energy of the previous step; 6) Binding free energy calculations, using the membrane MM GBSA[14,59] approximation, of the complex structures produced in the MD simulations to identify the complex structures with the lowest average binding free energy; 7) Combination of the two structures with the lowest binding free energy to construct the final complex; 8) Validation of the final constructed complex through the calculation of binding and interaction free energies using the membrane MM GBSA[14,59] approximation. All free energy calculations and MD simulations were performed in CHARMM.[60] In what follows, we provide a detailed description of the aforementioned steps.

1) We used twenty-two CXCL12 structures as flexible ligand templates for docking. The first twenty structures are NMR-derived and correspond to PDB entry 2KEE,[57] and the last two structures correspond to the two crystal structures of PDB entry 2J7Z.[58] Prior to docking, the structures were relaxed using 200 steps of steepest descent minimization, using the GBSW implicit solvent model,[61] with the backbone heavy atoms constrained using a force constant of 5 kcal/(mol*Å^2). As for CXCR4, the seventeen clustered CXCR4 conformations of ref 14 were considered as flexible receptor templates for docking as they acquire an optimum opening of the binding pocket for molecular recognition.

2) We used the parallel linux version of Zdock v.3.0.2[62] to dock the twenty-two CXCL12 structures to the seventeen CXCR4 structures. For each run of Zdock, 2000 docked structures were produced with a dense rotational sampling and a masking applied on the region with protein coordinates $z < 0$ Å, so as to exclude the nonpotential binding region from the docking calculations. As a result, 748,000 complex structures were produced from docking.

3) All 748,000 complexes were subjected to a 100 steps of steepest descent minimization to alleviate bad contacts, and the binding free energy was evaluated subsequently for all complexes using the GB (Generalized Born) SA approximation in a heterogeneous water-membrane-water environment, modeled by GBSW,[59] as in ref 14. The binding free energy is evaluated via the expression, $\Delta G = E_{PL} - E_P - E_L$, where E_X is the total (free) energy of molecule X (complex PL:CXCR4:CXCL12, free protein P:CXCR4, or free ligand L:CXCL12), as in refs 14 and 63−65. The protein and ligand conformations were assumed identical in the complex and in their free (unbound) states as in refs 14 and 63−65. The solvation free energy components of the complex and unbound protein were computed in the heterogeneous water-membrane-water environment, while the solvation free energy component of the unbound ligand was computed in a homogeneous aqueous environment, also modeled by the implicit GBSW model.[61] With this assumption, any bonded-energy contributions to ΔG are canceled in the equation. Contributions due to the protein and ligand entropy changes on association were not included in the binding free-energy calculations. These terms are associated with large errors (see ref 66 and references therein, for a detailed discussion). The significant correlation ($R^2 = 0.84$) between the binding free energy, calculated using the membrane MM GBSA approximation, and the degree of agreement with experimental findings in this study (see Discussion) confirm the validity of the method in scoring different complexes.

Table 1. Important Intermolecular Polar and Nonpolar Interaction Free Energies, Hydrogen Bonds and Salt Bridges, between CXCL12 and CXCR4 Residue Pairs within the Final Simulation (see Methods)[a]

CXCL12[§]	CXCR4[¶] residues (polar, nonpolar interaction free energies)	salt bridges[c] and hydrogen bonds[d]
Lys1	Tyr45 (2.1, −0.6)[b], Trp94 (−2.6, −2.4), **Asp97** (−0.4, 0.0), His113 (−6.4, −0.9), Tyr116 (−4.1, −1.7), **Asp171** (−69.9, 2.6), Arg188 (15.8, −2.9), Gln200 (−17.2, 0.7), His203 (−14.6, 0.4), **Tyr255 (−0.4, −1.1)**, Asp262 (−0.4, −0.1), Ile284 (−1.2, −0.7), Ser285 (−1.6, −0.3), **Glu288** (−79.8, 5.4)	[d]**Lys1** N: Tyr45 OH, [d]**Lys1** N: **Tyr255** OH, [c]**Lys1**:**Glu288**, [d] **Lys1** NZ: Tyr116 OH, [c]**Lys1**: **Asp171**, [d]**Lys1** NZ: Gln200 OE1, [d]**Lys1** NZ: His203 NE2
Pro2	Trp94 (−0.1, −1.8), **Asp97** (0.1, −0.8), His113 (−0.3, −0.7), Cys186 (0.1, −1.2), **Asp187** (−0.3, −1.9), Arg188 (−12.4, −1.8)	[d]**Pro2** O: Arg188 NE/NH2
Val3	Arg188 (−0.1, −1.5), His281 (−0.2, −2.3), Ile284 (−0.2, −2.3)	
Ser4	Arg30 (−0.1, −1.5), His281 (−0.6, −2.0)	[d]**Ser4** OG: Arg30 NH2
Leu5	Lys25 (−2.6 −1.3), Asp262 (0.4, −2.2), Ile265 (0.0, −2.1), **Glu268** (0.0, −2.2), Glu277 (−0.5, −3.6), Val280 (0.2, −1.7), His281 (−0.5, −3.0)	[d]**Leu5** O: Lys25 NZ, [d]**Leu5** N: His281 NE2
Ser6	Cys28 (0.0, −1.7), Arg30 (−6.3, −2.9), Glu277 (−1.0, −2.6), Asn278 (−1.6, −2.0)	[d]**Ser6** OG: Arg30 NH2, [d]**Ser6** OG: Asn278 OD1
Tyr7	Pro27 (−0.1, −1.2), Cys28 (−0.2, −1.0), Arg30 (0.2, −3.4)	
Arg8	Glu2 (−25.7, −0.8), Met16 (−0.1, −0.7), **Asp187** (−18.1, 1.3), Arg188 (−1.3, −1.7), Phe189 (−0.7, −2.7), **Tyr190** (−1.1, −2.7), Leu266 (−0.1, −0.6)	[c]**Arg8**:Glu2, [c]**Arg8**:**Asp187**, [d]**Arg8** NH2: Arg188 O
Cys9	Glu2 (−5.3, −1.7)	[d]Cys9 N: Glu2 OE*
Pro10	**Glu14** (0.0, −1.1), Met16 (0.0, −1.5), Met24 (−0.2,-3.0), Lys25 (0.1, −2.2), Pro27 (0.0, −0.4)	
Cys11	Ile6 (0.1, −0.8), **Tys7** (0.2, −0.8), **Glu14**(−4.4, −2.2), Met24 (−0.2, −2.1)	
Arg12	Glu2 (−7.4, −0.1), Ile6 (−1.3, −1.9), **Tys7** (−6.8, −4.4), Tys12 (−2.4, −1.3), **Glu14** (−12.3 −1.7), Met16 (−0.6, −1.1), **Tyr190** (−3.9, −2.8), Pro191 (−0.3, −2,1), Asn192 (0.9, −1.5), Asp193 (−21.8, 0.3)	[d]**Arg12** O:**Tys7** N, [d]**Arg12** N: **Glu14** OE2, [c]**Arg12**:Glu2, [c]**Arg12**:**Tys7**, [d]**Arg12** NH2: Met16 SD, [d]**Arg12** NH1: **Tyr190** O, [c]**Arg12**:Asp193
Phe13	**Tys7** (−1.8, −5.4), Thr8 (−0.4, −1.5), Asp10 (0.4, −2.1), Tys12 (0.1, −4.3), **Glu14** (−7.9, −1.1)	[d]**Phe13** N: **Glu14** OE2
Phe14	Ile6 (0.1, −2.4), **Tys7** (−2.6, −2.0), Thr8 (−1.7 −2.5), Ser9 (−1.5, −1.2)	[d]**Phe14** N: **Tys7** O, [d]**Phe14** N: Thr8 O, [d]**Phe14** O: Ser9 N, [d]**Phe14** O: Asp10 N, [d]**Phe14** O: Asn11 N
Glu15	Ser9 (−6.7, −3.2), Asp10 (−9.2, −0.8), Asn11 (−10.2, −1.2), Tys12 (−3.3, −2.0)	[d]**Glu15** OE*: Asp10 N, [d]**Glu15** OE*: Asn11 N, [d]**Glu15** OE1: Tys12 N
Ser16	Ser9 (−3.0, −1.2)	[d]**Ser16** N/OG: Ser9 OG
His17	Asn11 (0.0, −2.3)	[d]**His17** ND1: Asn11 ND2
Lys24	**Asp20** (−0.5, 0.0), **Tys21** (−10.5, −1.2)	[c]Lys24:**Tys21**
His25	**Tys21** (0.0, −0.9)	
Lys27	**Asp20** (−0.6, 0.0), **Tys21** (−0.9, −0.3), **Glu26** (−5.5, −1.6), Pro27 (−0.6, −2.2)	[c]**Lys27**:**Glu26**
Leu29	Pro27 (0.2, −2.6), Cys28 (−0.1, −2.0), Phe29 (−0.1, −2.9)	
Asn30	Cys28 (0.2, −0.4), Phe29 (−0.3, −1.4), Arg30 (−0.1, −0.4)	
Thr31	Pro27 (0.2, −1.6), Cys28 (0.0, −0.6), Arg30 (0.4, −0.4)	
Pro32	Arg30 (0.0, −1.9), Glu32 (0.1, −0.4), Asp181 (0.0, −0.5)	
Asn33	Met1 (0.1, −2.4), Glu2 (−0.7, −1.8), Glu189 (0.0, −0.5), Ala180 (−0.9, −2.2), Asp181 (−4.3, −1.5)	[d]**Asn33** OD1/ND2: Asp181 N, [d]**Asn33** ND2: Asp181 OD*
Cys34	Met1 (0.6, −3.4), Glu2 (−0.1, −2.8)	[d]Cys34 N: Met1 SD
Ala35	Met1 (−0.2, −2.0), Ile6 (0.0, −0.9)	
Val39	Met24 (−0.1, −1.7), Pro27 (0.1, −1.3)	
Arg41	**Asp20** (−0.6,-0.1), **Tys21** (−25.1, −5.2), Asp22 (−1.5, −1.2), Ser23 (−0.9, −1.2)	[d]**Arg41** NH*: **Tys21** O, [c]**Arg41**:**Tys21**, [d]**Arg41** NH1: Ser23 OG
Leu42	Thr13 (0.0, −0.4)	
Lys43	**Tys21** (−0.4, −0.2)	
Asn46	**Tys21** (−3.4, −2.5)	[d]Asn46 ND2: **Tys21** OS*
Arg47	Thr13 (−13.0, −2.5), **Glu14** (−0.8, −0.6), **Glu15** (−15.8, −0.2), **Asp20** (−0.4, 0.0), Asp22 (−7.9, −0.8), Ser23 (−0.6, −4.9)	[d]**Arg47** NH2: Thr13 O, [d]**Arg47** NE/NH2: Thr13 OG1, [c]**Arg47**:**Glu15**, [c]**Arg47**:Asp22, [d]**Arg47** NH1: Ser23 N
Gln48	Thr13 (−0.8, −1.0), Ser23 (−5.7, −0.8), Met24 (−0.8, −5.3), Lys25 (−0.6, −1.0), **Glu26** (−0.1, −1.4)	[d]**Gln48** N: Ser23 OG, [d]**Gln48** O: Met24 N,[d]**Gln48** NE2: Met24 O,[d]**Gln48** NE2: Lys25 O
Val49	Thr13 (0.1, −1.9), **Glu14** (−0.7, −0.5), Met24 (−0.3, −2.5)	
Cys50	**Glu14** (1.0, −0.8), Met24 (0.1, −2.3)	

dx.doi.org/10.1021/ci500069y | J. Chem. Inf. Model. 2014, 54, 1174−1188

Table 1. continued

CXCL12[§]	CXCR4[¶] residues (polar, nonpolar interaction free energies)	salt bridges[c] and hydrogen bonds[d]
Asp52	Thr8 (−0.2, −0.4), Ser9 (0.0, −0.4)	

[a]CXCL12 and CXCR4 residues marked in **boldface** are experimentally associated with binding-signaling (see Discussion − Agreement with Experiments). The results presented correspond to analysis of 1000 snapshots of the final simulation. [b]Principal interacting CXCL12[§]−CXCR4[¶] residue pairs. For each pair, the average polar and nonpolar average interaction free energies (polar, nonpolar) are provided in parentheses next to each CXCR4 residue; all energies are in kcal/mol. [c]Salt bridges between CXCL12 and CXCR4 residue pairs. [d]Hydrogen bonds between CXCL12 and CXCR4 atom pairs. The asterisk (*) symbol used after any CXCL12/CXCR4 atom in the hydrogen bonding pair denotes that any of the atoms in the charged, carboxyl or amide, side-chain group can participate in the hydrogen-bond formation.

4) Out of the 748,000 complexes, we selected the 6,000 CXCL12:CXCR4 complexes with the lowest GBSA binding free energy, and, subsequently, we performed an additional round of 200 steps of steepest descent minimization and recalculated the binding free energy using the GBSA approximation and the same setup used in ref 14. At the end of this procedure, we excluded all complexes at which, after the second round of minimization, the z-coordinate of the Lys1 Cα atom of CXCL12 was greater than 14 Å, so as to eliminate binding conformations which were not in line with experimental evidence required for signaling.[67,68] Subsequently, we identified the complex structure with the lowest binding free energy −174.2 kcal/mol, and, additionally, we selected the twenty complex structures with the lowest binding free energy within the range of (−174.2 kcal/mol: −146.3 kcal/mol) for subsequent investigation. Supporting Table 1 presents the binding free energies of the 20 different complex structures produced.

5) We performed 20 independent MD simulations for each of the complexes identified in the previous step. The MD simulations comprised a 400-ps heating procedure and an additional 700-ps equilibration procedure at which the harmonic restraints were gradually removed from the protein and the peptide. No restraints were imposed during the production run at 300 K, the duration of which was equal to 20-ns for every individual complex. The simulation methodology and force field[69,70] parametrization used is identical to ref 14. All molecular dynamics simulations were performed using CHARMM.[60]

6) We extracted 1000 snapshots, corresponding to 20-ps intervals, from the twenty MD 20-ns simulations and re-evaluated the binding free energy by employing the MM GBSA approximation, as in ref 14. The results included in Supporting Table 1 underline the beneficial role of the MD simulations in decreasing and expanding the binding free energy range (−393: −293 kcal/mol), compared to the binding free energy range after docking and minimization (−174: −147 kcal/mol). The decrease of the binding free energies is associated with the fact that the MD simulations, during the equilibration and production runs, allowed the relaxation and refinement of the complex structures. The relaxation and refinement was accompanied by the configuration of optimum interactions between CXCL12 and CXCR4 such as (i) the formation of salt bridges between neighboring oppositely charged residues, (ii) the formation hydrogen bonds between neighboring hydrogen donor−acceptor pairs, and (iii) the formation of nonpolar interactions between neighboring nonpolar moieties.

The lowest binding free energies corresponded to complexes 2 and 3, which possessed comparable average binding free energies (Supporting Table 1). The RMSD comparison between the conformations in the production runs with respect to the minimized docked conformations of Complexes 2 and 3 shows that (i) the N-terminal domain, and to a lesser extent the extracellular loops of CXCR4, as well as, (ii) the CXCL12 residue moiety 16:68, and to a lesser extent CXCL12 residue moiety 1:15, experienced a conformational variability so as to optimize their relative interactions during the simulations (Supporting Table 2). The RMSD comparison between the conformations in the production runs with respect to the starting conformations in the production run, in Complexes 2 and 3, depict a conformational variability in the 16−68 residue moiety of CXCL12 and the N-terminal domain of CXCR4 (Supporting Table 2). A visual inspection of the simulated

system shows that the optimization of interactions during the simulation between the aforementioned CXCL12:CXCR4 domains acts as a driving force for the conformational variability to occur. Despite the conformational variability between the aforementioned CXCL12:CXCR4 interacting segments, the interacting domains preserved sufficiently their individual conformations. The CXCR4 transmembrane domain experienced almost no variability during the entire simulations (Supporting Table 2).

7) Complexes 2 and 3 differ with regard to the relative orientation of CXCL12 segments 1−12 and 12−68. The two CXCL12 binding modes in both complexes coincide at residue 12, while the conformation of the CXCR4 binding site is quite similar in Complexes 2 and 3. To obtain additional insights into Complexes 2 and 3, we calculated (i) the sum of interaction energies between CXCL12 residues in the 1−11 moiety and all CXCR4 residues, as well as (ii) the sum of interaction energies between CXCL12 residues in the 12−68 moiety and all CXCR4 residues, for Complexes 2 and 3, individually. The analysis was performed using eq 1 of step 8; see step 8 for additional details regarding the equation. Interestingly, the calculations showed that Complex 3 possessed improved interactions between the CXCL12 1−11 residues and CXCR4, whereas Complex 2 possessed improved interactions between the CXCL12 12−68 residues of CXCL12 and CXCR4. Therefore, we extracted the intermediate −10-ns − structures from both Complexes 2 and 3 and constructed a new complex which combined the interaction properties of Complex 3 for CXCL12 residues 1−11 and Complex 3 for CXCL12 residues 12−68, using the CXCR4 structure of Complex 3. Owing to the remodeling procedure, the conformation of the final complex was subjected to an initial 20-ns simulation, to allow structural refinement and improvement of intermolecular interactions, and the lowest binding free energy structure was selected as a starting point for a subsequent − concluding −20-ns simulation of the final complex. We performed a thorough analysis on the latter simulation, which is referred to as "final simulation" in the text. In the validation and analysis we used 1000 snapshots, extracted every 20-ps intervals, from the 20-ns "final simulation".

8) To validate that the derived structures in the final simulation are energetically favored compared to the simulations of all twenty complexes referred to in step 5, we calculated the average binding free energy of the final simulation (Supporting Table 1) and showed that the binding free energy of the structures in the final simulation is approximately ≈50 kcal/mol less compared to the top Complexes (2 and 3) of step 6 (i.e., −438 kcal/mol compared to −393 kcal/mol and −388 kcal/mol). This result validates that the newly constructed complex in the final simulation is energetically more favored than Complexes 2 and 3. The use of the already refined structures from Complexes 2 and 3 so as to produce the complex structure for the final simulation, as well as the additional simulation conducted prior to the final production run in the final simulation, provide justification for why the complex structure in the final simulation is very well preserved (see Results and Supporting Table 2).

Analysis of the Final Simulation. We analyzed the interaction free energies between CXCL12 and CXCR4 residue pairs of the final simulation using eq 1:

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

The first and second group of terms on the right-hand side of eq 1 describe, respectively, polar and nonpolar interactions between R and R'. A similar methodology for the analysis of interacting residues has been used for the elucidation of the molecular recognition of CXCR4 by a dual tropic V3 loop,[14] the delineation of problems related to species specificity of proteins,[63] the design of modified-"transgenic" proteins,[64] and in problems related to drug design.[65,71,72] In the calculations, R corresponds to a CXCL12 residue and R' to a CXCR4 residue. To compute the GB term in eq 1, we included all atoms and set the charges of atoms outside the two − under investigation − residues R and R' to zero. The last term contains the difference in solvent accessible surface areas of residues R and R' in the complex and unbound states. The generalized-Born energies and the atomic accessible-surface areas (ΔS_i) entering in eq 1 depend on the location of R and R' in the complex. The polar component contains a Coulombic term and a GB contribution, modeling the interaction between residue R and the solvent polarization potential induced by R' (or vice versa). Similarly, the nonpolar component contains a van der Waals interaction between R, R' and a surface term, expressing cavity contributions and nonpolar interactions with the surrounding solvent. The nonpolar and polar solvation terms were calculated using the heterogeneous water-membrane-water GBSW[59] using the same parameters as in ref 14. The sum of the two components, polar and nonpolar, reflects the total direct interaction between R and R' in the solvated complex. Subsequently, we decomposed the polar and nonpolar interaction free energy contributions and present the results of the average intermolecular interaction free energies of the lowest binding free energy complex in two-dimensional density maps in Supporting Figure 1. In addition, we summed up the total intermolecular interaction free energies for every CXCR4 residue, so as to provide insights into the role of each interacting CXCR4 residue with CXCL12 (first column per CXCR4 residue in Figure 3), and the results are presented in Figure 3. Also, in Figure 3, we provide a comparison to the sum of intermolecular interaction free energies summed up for every CXCR4 residue with regard to the HIV-1 gp120 V3 loop binding (second column per CXCR4 residue in Figure 3) using data from Tamamis and Floudas.[14]

RESULTS

We applied the computational protocol which is analytically explained in Methods and derived the complete complex structure of CXCL12:CXCR4, which corresponds to the ensemble of snapshots contained in the "final simulation"; snapshots extracted at 2 ns intervals are provided as Supporting Coordinates. In what follows, we focus our analysis on the final simulation. For all subsequent analyses, we extracted 1000 snapshots from the final simulation, corresponding to 20-ps intervals, and calculated the intermolecular CXCL12:CXCR4 residue pairwise interaction free energies (Supporting Figure 1), as well as the hydrogen bond occupancies (Supporting Table 3). The intermolecular interactions are summarized in Table 1.

Our results reveal that the 1−15 domain of CXCL12 penetrates into the CXCR4 binding pocket. Specifically, the 1−

5 N-terminal domain of CXCL12 is buried within the transmembrane region of CXCR4, and the 6−15 residue moiety of CXCL12 is predominantly embraced by the N-terminal domain and extracellular loops of CXCR4; CXCL12 residues 16, 17 and additional residues of the 24−50 residue moiety interact with the upper extracellular interface of CXCR4 which is mainly composed by the N-terminal domain of CXCR4 (see Figure 1).

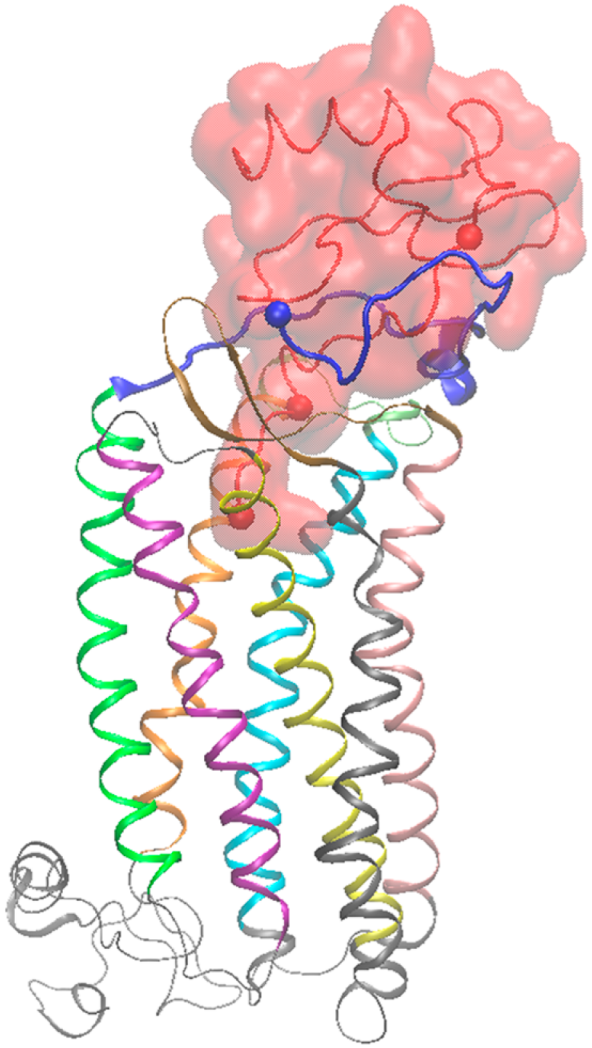


Figure 1. Entire simulation system of the CXCL12:CXCR4 complex structure: Molecular graphics image of the entire simulation system corresponding to the final simulation. CXCL12 is shown in tube and transparent surface representation in red color. CXCR4 is shown in cartoon representation, and the coloring 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 CXCR4 is shown in a small van der Waals sphere. The Cα atoms of residues 1, 5, and 15 of CXCL12, from bottom to top, are shown in small van der Waals spheres.

Conformational Analysis of CXCL12 and CXCR4 within the Final Simulation. The bound CXCL12 conformation within the final simulation combines the 1−11 CXCL12 conformation of Complex 3 and the 12−68 CXCL12 conformation of Complex 2 (see Methods). Interestingly, in the docking procedure, Complex 3 was initially produced from the 13th NMR structure of PDB entry 2KEE,[57] while Complex 2 was initially produced from the second X-ray structure of PDB entry 2J7Z.[58] Thus, the complex structure in the final simulation acquires a hybrid-origin CXCL12 conformation produced by a combination of an NMR (region 1−11) and an X-ray (12−68) structure. According to secondary structure definitions of STRIDE,[73] the CXCL12 conformation within the final simulation possesses the following: (i) β-turns which occur mainly in residue moieties 4−8, 16−22, 34−37, 43−46, and 52−55 and to a smaller extent within residue moieties 11−14 and 31−34 (the 19−22 β-turn is frequently interchanged to a 3_{10} helix formed by residues 20−22); (ii) β-extended antiparallel conformations which are formed between residue moieties 23−37, 38−42, and 47−50; (iii) a parallel β-bridge between residues Glu15 and Ile51, which is facilitated by the Cys11-Cys50 disulfide bridge; (iv) an α-helical conformation extending from residues 56−66. The average backbone root-mean-square deviation (RMSD) of CXCL12 within the simulation, with regard to the first simulation snapshot after equilibration, is 1.7 ± 0.5 Å, without superposition. The corresponding backbone RMSD values for the 1−11, 1−15, and 16−68 residue moieties are 1.1 ± 0.2 Å, 1.3 ± 0.3 Å, and 1.8 ± 0.6 Å. These results indicate that the overall CXCL12 conformation is firmly preserved throughout the simulation, and that the more CXCL12 is buried within CXCR4, the less flexibility it experiences. Thus, the bound properties of CXCL12 differ from its unbound properties in that the 1−11 domain of CXCL12 is the most flexible region in the unbound conformation.[67]

The CXCR4 conformation is also very well preserved within the final simulation. The average backbone RMSD of the transmembrane helical region and N-terminal domain (comprising residues 1−37) within the simulation is 0.8 ± 0.1 Å and 1.6 ± 0.4 Å, respectively. Despite the larger flexibility of the CXCR4 N-terminal domain compared to the inner transmembrane helical region of CXCR4, the N-terminal domain is maintained in a sufficiently stable conformation as, with regard to nonpolar interactions, it is the most highly interacting CXCR4 domain with CXCL12.

Interactions of CXCL12 Residues 1:5 with CXCR4. CXCL12 residue Lys1 forms two concurrent highly interacting salt bridges with CXCR4 residues Asp171 and Glu288 of CXCR4 (see Figure 2); Asp171 interacts with the charged side chain amide of Lys1, and Glu288 interacts with the charged N-terminal end of Lys1. In addition, the charged side chain amide of Lys1 is hydrogen bonded to CXCR4 Tyr116 OH, Gln200 OE1, and His203 NE2, and the backbone charged amide of Lys1 is occasionally hydrogen bonded to the hydroxyl groups of CXCR4 residues Tyr45, Tyr116, and Tyr255 and is also in the vicinity of CXCR4 residue Asp97. The nonpolar moiety of CXCL12 Lys1 is buried in a pocket comprising the nonpolar moieties of CXCR4 residues Trp94, His113, Tyr116, Arg188, Tyr255, Ile259, Ile284, and Ser285. Residue Pro2 of CXCL12 is also buried in a pocket comprising the nonpolar moieties of CXCR4 residues Trp94, Asp97, His113, Cys186, Asp187, and Arg188, and, in addition, Pro2 O is strongly hydrogen bonded to the charged amide of CXCR4 residue Arg188. Residue Val3

dx.doi.org/10.1021/ci500069y | J. Chem. Inf. Model. 2014, 54, 1174−1188

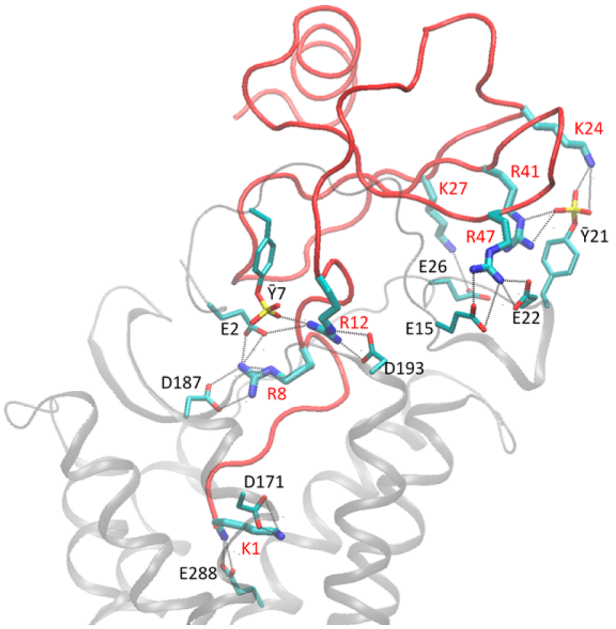


Figure 2. Salt bridges between CXCL12 and CXCR4 residues: Molecular graphics images of the salt bridges between CXCL12 and CXCR4 residues in the final simulation. CXCL12 is shown in red tube representation, and CXCR4 is shown in light gray transparent tube representation. The hydrogen bonds leading to the salt bridge formation are denoted in dashed lines, and the participating CXCL12 and CXCR4 residue moieties are shown in licorice; CXCL12 and CXCR4 residues are annotated in red and black, color, respectively. Residue Ȳ corresponds to a sulfated tyrosine. Hydrogen atoms are omitted for clarity.

of CXCL12 is buried in a pocket composed of the nonpolar side chain moiety of CXCR4 Arg188 on one site and the side chain moieties of His281 and Ile284 on the opposite site. The backbone of CXCL12 residue Ser4 is proximal to the side chain of residue CXCR4 His281, and the side chain of Ser4 is proximal to the charged amide of CXCR4 Arg30; as a result, a low interacting hydrogen bond is formed between the hydroxyl group of Ser4 and the charged amide of Arg30. The side chain of CXCL12 residue Leu5 is buried within the nonpolar moieties of CXCR4 residues Lys25, Asp262, Ile265, Glu268, Glu277, Val280, and His281. Due to these interactions, the backbone amide and carboxyl groups of Leu5 are occasionally hydrogen bonded to the side chain amide groups of CXCR4 Lys25 and His281, respectively.

Interactions of CXCL12 Residues 6:17 with CXCR4. The side chain of CXCL12 residue Ser6 is enveloped within the nonpolar side chain moieties of CXCR4 residues Cys28, Arg30, Glu277, and Asn278, and its position is additionally stabilized by two high occupancy hydrogen bonds formed by its hydroxyl group with (i) the charged amide of CXCR4 Arg30 and (ii) the side chain carboxyl group of CXCR4 Asn278. Furthermore, the side chain of CXCL12 residue Tyr7 forms hydrophobic contacts with CXCR4 residues Pro27, Cys28, and the nonpolar side chain moiety of Arg30. Arg8 of CXCL12 forms two simultaneous highly interacting salt bridges with CXCR4 residues Glu2 and Asp187 (see Figure 2). The charged amide of CXCL12 Arg8 is additionally hydrogen bonded to the backbone carbonyl of CXCR4 Arg188 and forms a cation-π interaction with CXCR4 residue Tyr190; the nonpolar moiety of CXCL12 Arg8 forms nonpolar contacts with CXCR4 residue Phe189 and Tyr190, and is in the vicinity of CXCR4 residues Met16 and Leu266. Residue Cys9 of CXCL12 is solely interacting with CXCR4 residue Glu2, as the backbone amide of the former is during the last 2/3 of the simulation hydrogen bonded to the charged carboxyl group of the latter. Residue Pro10 of CXCL12 participates in hydrophobic contacts with the nonpolar moieties of CXCR4 residues Glu14, Met16, Met24, and Lys25 and is in the vicinity of CXCR4 residue Pro27. Residue Arg12 of CXCL12 forms polar and nonpolar interactions with a series of CXCR4 residues. As for polar interactions, its charged amide (i) forms salt bridges with the oppositely charged carboxyl groups of CXCR4 Glu2, Tys7 (sulfated tyrosine 7), and Asp193 (see Figure 2), (ii) is in the vicinity of the oppositely charged groups of CXCR4 Tys12, and (iii) is hydrogen bonded to CXCR4 Tyr190 O and to a smaller extent CXCR4 residue Met16 SD; also, its backbone carbonyl and amide are hydrogen bonded to CXCR4 Tys7 N and Glu14 OE2, respectively. As for nonpolar interactions, the nonpolar moiety of CXCL12 residue Arg12 is proximal to the nonpolar moieties of CXCR4 residues Ile6, Tys7, Tys12, Pro191, and Asn192. Phe13 of CXCL12 is buried within a hydrophobic pocket comprising the nonpolar moieties of CXCR4 residues Tys7, Thr8, Asp10, Tys12, and Glu14; in addition, the backbone amide of Phe13 is strongly hydrogen bonded to CXCR4 Glu14 OE2. The side chain of CXCL12 residue Phe14 is attracted to the hydrophobic side chain of CXCR4 residue Ile6, and the backbone of CXCL12 residue Phe14 is close to the backbone moieties of CXCR4 residues Tys7, Thr8, Ser9, and Asp10 owing to intermolecular backbone−backbone hydrogen bond interactions. The charged carboxyl group of CXCL12 residue Glu15 is hydrogen bonded to the CXCR4 backbone amides of Asp10, Asn11, and Tys12 and is in the vicinity of the backbone amide of CXCR4 Ser9. CXCL12 residues Ser16 and His17 are proximal to CXCR4 residues Ser9 and Asn11, respectively, owing to the formation of hydrogen bonds between the CXCL12 and CXCR4 atom pairs Ser16 N/OG:Ser9 OG and His17 ND1:Asn11 ND2.

Interactions of CXCL12 Residues 24:52 with CXCR4. Residues Lys24 and His25 of CXCL12 are adjacent to CXCR4 residue Tys21, and, as a result, Tys21 forms a salt bridge with CXCR4 Lys24 which is preserved throughout the simulation (see Figure 2). Lys27 of CXCL12 is attracted to the oppositely charged CXCR4 residues Glu26 and Tys21, and occasionally a salt bridge is formed with the former (see Figure 2); in addition, the nonpolar moiety of Lys27 forms nonpolar contacts with CXCR4 residue Pro27. Both Lys24 and Lys27 of CXCL12 are in the vicinity of the oppositely charged CXCR4 residue Asp20. CXCL12 residues Leu29, Asn30, Thr31, and Pro32 participate in nonpolar interactions with neighboring to CXCR4 residues Pro27, Cys28, Phe29, and Arg30. Residue Asn33 of CXCL12 buries its nonpolar side chain moiety within the nonpolar side chain moieties of CXCR4 residues Met1, Glu2, Glu189, and Ala180, while its polar side chain group faces toward CXCR4 Asp181; as a result, the polar group of Asn33 is hydrogen bonded to both the backbone amide and the charged carboxyl group of CXCR4 residue Asp181. CXCL12 residues Cys34 and Ala35 are adjacent to CXCR4 residues Met1, Glu2, and Met1, Ile6, respectively, predominantly owing to nonpolar contacts. Val39 of CXCL12 is inserted in a hydrophobic pocket comprising CXCR4 residues Met24 and Pro27. Residue Arg41 of CXCL12 forms a highly interacting salt bridge with CXCR4 residue Tys21 (see Figure 2). The charged amide of Arg41 is hydrogen bonded to the backbone carbonyl of Tys21 and the side chain

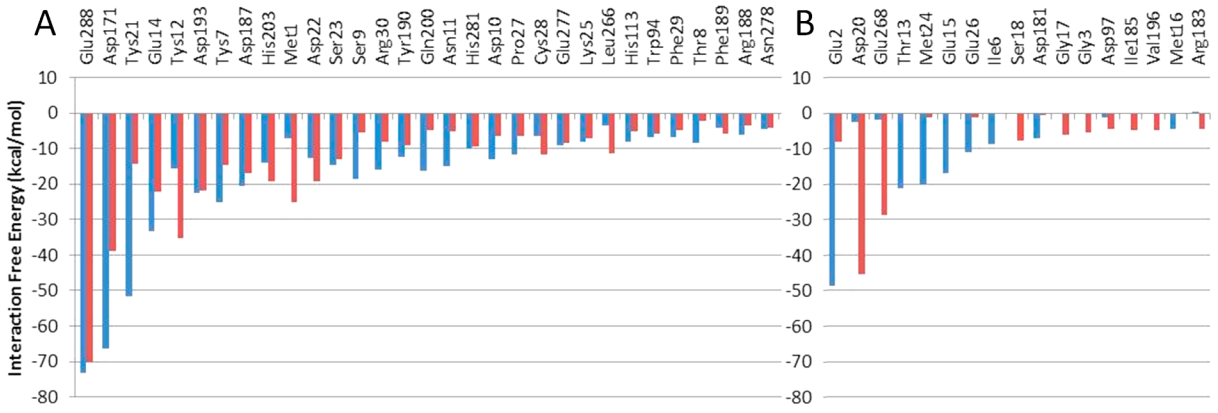


Figure 3. Interaction free energies of CXCR4 residues in complex with CXCL12/HIV- gp120 V3 loop: The residue pairwise interaction free energies were summed up (*y* axis) for every CXCR4 residue (*x* axis), in complex with (i) CXCL12 (first column per CXCR4 residue) and (ii) the dual tropic HIV-1 gp120 V3 loop of ref 14 (second column per CXCR4 residue). The figure includes only CXCR4 residues which possess at least −4.0 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 CXCR4 residue interacts strongly and approximately equally with both CXCL12 and the HIV-1 gp120 V3 loop, it is listed in panel A, whereas, if a CXCR4 residue interacts strongly with CXCL12 and weakly with HIV-1 gp120 V3 loop, or vice versa, it is listed in panel B (see Supporting Information).

hydroxyl of CXCR4 Ser23 and is also adjacent to the oppositely charged Asp20 and Asp22 of CXCR4. The side chain polar group of CXCL12 Asn46 is regularly hydrogen bonded with the negatively charged side chain group of CXCR4 residue Tys21. The positively charged group of Arg47 (i) forms two concurrent salt bridges with CXCR4 residues Glu15 and Asp22 (see Figure 2), (ii) is hydrogen bonded to the backbone carbonyl and side chain hydroxyl group of CXCR4 Thr13, and (iii) is also in the vicinity of CXCR4 residues Glu14 and Asp20; in addition, the nonpolar moiety of CXCL12 Arg47 is proximal to CXCR4 residue Ser23. Residue Gln48 of CXCL12 intercalates between the CXCR4 residues Thr13, Ser23, Met24, Lys25, and Glu26, and, as a result, Gln48 N is hydrogen bonded to Ser23 OG throughout the simulation, Gln48 O is hydrogen bonded to Met24 N during the last 2−4 ns of the trajectory, and also, Gln48 NE2 is hydrogen bonded to Met24 or Lys25 O. Residue Val49 of CXCL12 is buried within a hydrophobic pocket of CXCR4 residues Thr13 and Met24 and is also in the vicinity of CXCR4 residue Glu14. Cys50 of CXCL12 is adjacent to CXCR4 residue Met24 and to a smaller extent CXCR4 residue Glu14, while Asp52 of CXCL12 is in the vicinity of CXCR4 residues Thr8 and Ser9.

Molecular Recognition of CXCR4 by CXCL12 versus the HIV-1 gp120 V3 Loop. We summed up the residue pairwise interaction free energies (presented in the Supporting Figure 1) for every CXCR4 residue in complex with (i) CXCL12 (first column of Figure 3 per CXCR4 residue) and (ii) the dual tropic HIV-1 gp120 V3 loop of ref 14 (second column of Figure 3 per CXCR4 residue). CXCR4 residues Glu288, Asp171, Tys21, Glu14, Tys12, Asp193, Tys7, Asp187, His203, Met1, Asp22, Ser23, Ser9, Arg30, Tyr190, Gln200, Asn11, His281, Asp10, Pro27, Cys28, Glu277, Lys25, Leu266, His113, Trp94, Phe29, Thr8, Phe189, Arg188, and Asn278, presented in descending order of magnitude of interaction free energy (averaged for the two complexes), interact significantly with both CXCL12 and the HIV-1 gp120 V3 loop[14] and are listed in panel A. On the contrary, in Panel B, CXCR4 residues Glu2, Thr13, Met24, Glu15, Glu26, Ile6, Asp181, and Met16, presented in descending order of magnitude of interaction free energy, interact with CXCL12 and to a smaller (or considerably smaller) extent with the HIV-1 gp120 V3 loop,[14] while CXCR4 residues Asp20, Glu288, Ser18, Gly17, Gly3, Asp197, Ile185, Val196, and Arg183, presented in descending order of magnitude of interaction free energy, interact with HIV-1 gp120 V3 loop[14] and to a smaller (or considerably smaller) extent with CXCL12. Based on these results, approximately 2/3 of CXCR4 residues belong to panel A, depicting a significant overlap in the CXCR4 interacting residues which interact both with CXCL12 and the specific HIV-1 gp120 V3 loop, investigated by Tamamis and Floudas.[14] Despite the fact that CXCR4 residue Glu2 interacts stronger with CXCL12, and also CXCR4 residues Asp20 and Glu268 interact stronger with the HIV-1 gp120 V3 loop, the total interaction free energy of Glu268 in complex with the HIV-1 gp120 V3 loop is ≈−8 kcal/mol, and the total free energy of Asp20 and Glu268 in complex with CXCL12 is ≈−3 kcal/mol and ≈−2 kcal/mol, respectively; thus, CXCR4 residue Glu2 is also involved in the HIV-1 gp120 V3 loop binding, and, in addition, CXCR4 residues Asp20 and Glu268 are also involved in the CXCL12 binding. The interaction free energy based analysis identified that the HIV-1 gp120 V3 loop[14] and CXCL12 mostly interact with the same CXCR4 residues, and, thus, they share the same CXCR4 binding pocket (see Supporting Figure 2 which presents CXCR4 in complex with the superimposed conformations of the HIV-1 gp120 V3 loop[14] and CXCL12).

DISCUSSION

Agreement with Experiments. Since 1997, a series of experimental studies aimed at elucidating the key CXCL12 and CXCR4 interacting residues with regard to binding and signaling upon the formation of the complex. These studies mainly used site-directed mutagenesis and showed that specific CXCL12 and CXCR4 residues are critical to, are involved in, or influence the molecular recognition of CXCR4 by CXCL12, depending on the degree of impairment in binding or signaling.

Experiments on CXCR4 Residues. As for CXCR4, experimental studies suggest that specific CXCR4 residues are associated with the CXCL12 mediated binding-signaling. According to Doranz et al. in 1999,[74] the D193K has no effect on signaling, whereas the EADD(179−182)QAAN mutant is capable of binding but not signaling. According to Brelot et al. in 2000,[75] EE(14−15)AA, Y21A, D97N, and E288Q mutants cause approximately a 75% loss in CXCL12 binding, and the D187A mutant reduces binding by approximately a factor of 1/

3; D97N, D187A, and E288Q mutants abolished signaling as well. According to Zhou et al. in 2001,[53] alanine mutations in CXCR4 residues Tys7, Asp20, Tys21, Glu26, and Glu268 reduce CXCL12 binding. Among these CXCR4 residues, Tys21 is identified as most critical since an alanine mutation at position 21 causes a loss of more than 75% loss in binding; in the same study, the Y190A mutant impaired signaling. According to Choi et al. in 2005,[76] a D97A mutation does not alter the binding activity, whereas alanine mutations at residues Phe87, Asp171, and Phe292 cause a significant reduction in binding activity of CXCL12. Therefore, by comparing the effect of asparagine and alanine mutations at position 97, respectively in studies,[75,76] (i) Cβ at position 97 should be important for binding owing to nonpolar interactions, and (ii) the presence of an amide in the side chain group of position 97 should not be tolerated for neither binding nor signaling. Moreover, residues Phe87, Asp171, and Phe292 were identified as most critical for CXCL12 binding in a subsequent study by Tian et al. in 2005[77] as alanine mutations at these positions impair binding; consequently, the same mutations should lead to an impairment of the signaling, as well. Tian et al.[77] also showed that E288A or E288D mutants cause a critical loss in CXCL12 signaling and that additional alanine mutations at residues Trp252A, Tyr255A, Asp262A, Asp187, and Phe189, which are situated above the middle of the membrane, cause a reduction of binding.

Experiments on CXCL12 Residues. As for CXCL12, a series of experimental studies suggest that specific residue positions in the CXCL12 residue moiety 1−17, as well as CXCL12 residue positions 25, 27, 31, 33, 35, 39, 41, 47, 48, 49, and 55, are associated with (or influence) binding-signaling to CXCR4. Mutations K1A[58] or K1R[67] abrogate chemotactic activity, while the latter mutation only slightly impairs binding to receptor. Mutants P2G, SLS(4−6)AQA, and Y7A also abolish chemotaxis and slightly impair binding to receptor, whereas the Y7H mutant has no effect on binding or chemotaxis.[67] While the R8K mutation only slightly impairs chemotaxis and binding to CXCR4, the R8Q mutation impairs chemotaxis to a significant extent.[78] As a result, the positively charged group at position 8 should be considered critical for signaling. Furthermore, the RFFESH(12−17)AAAAAA mutation impairs chemotactic activity, and this can mainly be attributed to the R12A mutation which also significantly impairs chemotaxis.[79] Also mutants F13A, F14A, E15A, S16A, H17A,[79] FF(13−14)AA, and E15Q|H17N[78] have an effect on both chemotactic and binding activities, showing that the RFFESH motif, which corresponds to residue moiety 12−17, is involved in binding. A glycine triple mutant at positions 31, 33, and 35 and an asparagine quintuple mutant at positions 25, 27, 41, 47, and 48 decreased the binding activity by approximately 20% and 33%, respectively, showing that these residues are involved in CXCR4 binding.[78] A subsequent study by Veldkamp et al. in 2008[50] investigated the chemotaxis reduction by single mutations within the 25−49 moiety of CXCL12 and showed that (i) mutant R47E abolishes chemotactic activity, (ii) mutants V39A, K27E, and R47A impair to a critical extent the chemotactic activity, (iii) mutants K27A, V49A impair considerably the chemotactic activity, and (iv) mutants H25R and R41A decrease chemotaxis to a small extent. Also, a study by Kofuku et al. in 2009[68] used methyl-utilizing TCS experiments and showed that residues Leu29, Val39, Val49, and Leu55 are in proximity to CXCR4.

Agreement with Experiments - Interactions between CXCL12 Residues 1−17 and CXCR4. Within the final simulation, Lys1 of CXCL12 forms two highly interacting salt bridges with CXCR4 residues Asp171 and Glu288 which are critical for binding-signaling. As Lys1 of CXCL12 and both Asp171, Glu288 of CXCR4 are critical residues for signaling, the formation of concurrent salt bridges must be considered as the switching mechanism required for signaling. This is also in line with experiments showing that even a E288D substitution results in a critical loss of signaling, and this most presumably can be attributed to the inability of aspartic acid to participate in the concurrent salt bridge formation owing to its shorter size. Our results reveal that the charged side chain moiety of Lys1 is hydrogen bonded to Tyr116 OH, its charged N-terminal end is hydrogen bonded to Tyr45 OH, and its nonpolar moiety forms contacts with aromatic CXCR4 residues Trp94, Tyr116, and Tyr255; all the aforementioned aromatic residues comprise a cluster of aromatic residues which also includes CXCR4 residues Phe87, Trp252, Tyr255, and Phe292. Despite the fact that the critical CXCR4 residues Phe87, Trp252, and Phe292 are positioned toward the center of the membrane[51] and do not strictly belong to the binding site, our results depict that an alanine mutation at these positions should impair (i) the clustering of aromatic residues and, as a result, (ii) the anchoring of Lys1 inside the CXCR4 binding pocket, directly interacting with Trp94, Tyr116, and Tyr255 and indirectly with Phe87, Trp252, and Phe292. Experiments suggest that the side chain moiety of CXCL12 Pro2 and the methyl Cβ group of CXCR4 Asp97 are important in signaling. In line with this, the structure derived in the present study shows that the side chain of Pro2 forms contacts with the nonpolar moiety of Asp97. Specific modifications at positions 4−7 of CXCL12 abolish chemotaxis; in our computationally derived structure, these modifications would impair (i) the hydrogen bonding interactions between CXCR4 and the CXCL12 hydroxyl groups of Ser4 and Ser6 and (ii) the burial of CXCL12 residue Leu5 and Tyr7 within specific pockets of CXCR4, which include the nonpolar side chain moieties CXCR4 residues Glu262, Glu268. Therefore, the interactions, or a portion of the interactions, formed by CXCL12 residues 4−7 with CXCR4 (see Results) should be considered critical with regards to the chemotactic signaling. Our data provide a compelling argument for the most critical role of both (i) a positively charged residue at position 8 of CXCL12 and (ii) Asp187 and Tyr190 of CXCR4. In the simulation, the charged side chain group of Arg8 of CXCL12 forms a highly interacting salt bridge with CXCR4 Asp187 and a cation-π interaction with CXCR4 Tyr190. In addition, the nonpolar moiety of CXCL12 Arg8 is attracted to the aromatic ring of Tyr190. The critical role of CXCL12 Arg12 in chemotaxis can be attributed to its capacity to form rich nonpolar and predominantly polar interactions with CXCR4 residues within the simulation. Polar interactions include both an abundance of salt bridges and hydrogen bonds with CXCR4 negatively charged residues Glu2, Tys7, Glu14, and Asp193. According to experiments, CXCR4 residues Glu14 and Tys7 are important for CXCL12 binding, whereas a D193K mutation does not affect signaling. To understand why the lysine substitution at position 193 does not affect signaling, we performed an additional 20-ns simulation of CXCL12 in complex with a D193K CXCR4 mutant, starting from the optimized initial coordinates of the final simulation. Interestingly, we observed that the calculated average binding free energy is similar when a lysine is present at 193; this is

attributed to the fact that in the new simulation of the D193K CXCR4 mutant, Tys7 is capable of forming a highly interacting salt bridge with CXCL12 Arg12 without being repelled by its neighboring Asp193 as is the case in the natural CXCR4 protein. Furthermore, our computationally derived structure in the final simulation is in remarkable agreement with experiments, as the 12−17 residue moiety of CXCL12 is indeed involved in the binding; according to our computationally derived structure, this region forms interactions with CXCR4 N-terminal domain residues 6−14.

Agreement with Experiments - Interactions between CXCL12 Residues 24−55 and CXCR4. Within the final simulation, CXCL12 residue Lys24 forms a salt bridge with CXCR4 Tys21 which, according to experiments, is a highly critical CXCR4 residue for binding. Experiments suggested that mutations at CXCL12 residue His25 affect to a small extent the binding or signaling; in our study, His25 is part of the CXCL12 binding site and is proximal to Tys21. CXCL12 residue Lys27 is according to experiments critical for signaling, and the final simulation supports this, as Lys27 forms a salt bridge with an important CXCR4 residue Glu26 and is in the vicinity of oppositely charged CXCR4 residues Asp20 and Tys21. Residue Leu29 of CXCL12 is in proximity to CXCR4 residues Pro27, Cys28, and Phe29, in line with experiments. Our study provides evidence with regard to the involvement of CXCL12 residues Thr31, Asn33, and Ala35 in binding, as within the final simulation they participate in an abundance of intermolecular polar and nonpolar interactions with CXCR4 residues. These interactions include nonpolar contacts between CXCL12 residue Asn33 and CXCR4 residues Glu189, Ala180, and Asp181, as well as multiple hydrogen bond interactions between the polar side chain group of Asn33 with the backbone amide and/or the side chain carboxyl group of CXCR4 Asp181. The aforementioned interactions should be important as the EADD(179−182)QAAN CXCR4 mutant is capable of binding but not signaling. Our study provides evidence on the critical nonpolar role of CXCL12 Val39 in chemotaxis, as within the simulation, Val39 is involved in two hydrophobic contacts with Met24 and Pro27 of CXCR4. Moreover, our simulation provides evidence on the role of CXCL12 residue Arg41 in binding, as its charged side chain group (i) forms a salt bridge with CXCR4 Tys21, (ii) is hydrogen bonded to the hydroxyl group of CXCR4 Ser23, and (iii) is attracted to oppositely charged CXCR4 residues Asp20 and Asp22. According to experiments the positively charged side chain group of Arg47 is of utmost importance as the R47E CXCL12 mutant abolishes chemotactic activity. Our simulation data provide strong evidence for this as its charged side chain group (i) forms salt bridges with CXCR4 residues Glu15 and Asp22, (ii) is strongly hydrogen bonded to CXCR4 residue Thr13, and (iii) is attracted to oppositely charged residues Glu14 and Asp20; it is worth noting that CXCR4 residues Glu14, Glu15, and Asp20 are according to experiments involved in the CXCL12 binding to CXCR4. In addition, experiments suggest that CXCL12 residue Gln48 is involved in the binding with CXCR4, and, in line with this, our computationally derived structure depicts that it is hydrogen bonded to CXCR4 residues Ser23, Met24, and Lys25 and participates in additional nonpolar contacts with CXCR4 residues Thr13, Glu26. Our study provides evidence on the important role of CXCL12 Val49 for chemotactic activity, as it forms important hydrophobic contacts with the nonpolar moieties of CXCR4 residues Thr13, Glu14, and Met24. In addition, Leu55 CD1/2 atoms are frequently within the simulation 7−8 Å away from CXCR4 Ser9 CB; owing to this position, the side chain of Leu55 forms intramolecular nonpolar interactions with CXCL12 Ser16 which stabilize its position in the complex and facilitates its polar interactions with CXCR4 residue Ser9.

Predictions for Future Experimental Studies. MD simulations, in general, can interpret experimental data, and, in addition, they possess the capacity to provoke new experiments.[80] Our study suggests that specific residues of CXCL12, shown in italics in Table 1, (e.g., Val3, Cys9, Pro10, Cys11, Lys24, Asn30, Pro32, Cys34, Asn46, and Cys50), and CXCR4 (e.g., Met1, Thr8, Ser9, Asp10, Asn11, Tys12, Asp22, Ser23, Lys25, Pro27, Cys28, Phe29, Arg30, Trp94, His113, Arg188, Phe189, Gln200, His203, Leu266, Glu277, Asn278, and His281) take part in the binding and may be associated with signaling, as well. As, according to our knowledge, the role of the aforementioned CXCL12 and CXCR4 residues was not investigated by experiments, we suggest that future experimental studies be performed to examine the involvement of the aforementioned CXCL12 and CXCR4 residues in binding and/or signaling.

Comparison of the Final Simulation and the Top Ranked Complexes with Regard to Experiments. We evaluated the total interaction free energy for CXCL12:CXCR4 residue pairs belonging to the final simulation and the top five ranked Complexes: 2, 3, 19, 1, and 16. Subsequently, we selected only the residue pairs which meet the following two conditions: (i) at least one of the two interacting residues is, according to experiments, involved in the binding and/or signaling, and (ii) the total interaction free energy of the interacting residue pair is at least −4.5 kcal/mol. This additional analysis aimed at assessing (i) the degree of agreement with experiments of the top ranked complexes (see Supporting Figure 3A) and (ii) the correlation between the degree of agreement with experiments and the ranking according to the membrane MM GBSA approximation (see Supporting Figure 3B).

The final simulation contains the largest number of interactions (35) between CXCL12 and CXCR4 residues which are experimentally associated with binding and/or signaling. Complexes 2 and 3 which were used as building blocks for the construction of the CXCL12:CXCR4 complex in the final simulation contain the second (31) and third (29) largest number of interactions between CXCL12 and CXCR4 residues which are experimentally associated with binding and/or signaling. The key weaknesses of Complexes 2 and 3 with regard to the final simulation is that they cannot provide evidence for the key role of CXCR4 residue Asp171 and CXCL12 residue Lys27, respectively. The binding mode of Complex 19 is also similar to the binding mode of the final simulation. Despite the weakness of Complex19 in providing evidence adequate information on the important role of CXCR4 residues Glu26 and Glu268, and the charged amide of CXCL12 residue Arg12, (i) the relatively large number (23) of interactions between CXCL12 and CXCR4 residues which are experimentally associated with binding and/or signaling in Complex 19 and (ii) the similarity between the binding modes of Complex 19 and the final simulation suggest that a portion of the CXCL12:CXCR4 interacting residue pairs observed in Complex 19 (e.g., Ser4:Asp187, Arg8:Asp181, Lys27:Tys21, Arg41:Glu14, and Arg47:Glu14) may − upon binding − alternate with the interacting residue pairs encountered in the

final simulation. Also, Complexes 1 and 16 contain a large number (24 and 22, respectively) of interactions between CXCL12 and CXCR4 residues which are experimentally associated with binding and/or signaling; nevertheless, they fail − among others − to provide evidence for the key role of CXCR4 residue Asp171.

Interestingly, the aforementioned analysis reveals that, at least for the specific system investigated here, there is a significant correlation (R^2 = 0.84) between the binding free energy, which was calculated using the membrane MM GBSA approximation, and the degree of agreement with experimental findings. The results show that a decrease of the binding free energy is associated with an increase of the number of interactions between CXCL12:CXCR4 residues which, according to experiments, are related to binding and/or signaling of CXCL12 in complex with CXCR4 (see Supporting Figure 3B)

Comparison with Previous Computational Studies. It is possible that prior to the optimum and final binding, which is associated with signaling, CXCL12 may undergo rapid conformational changes in its N-terminal end to adopt an appropriate conformation for binding, as suggested in the two-site binding model of Crump et al.[67] Nevertheless, our computationally derived structure suggests that - upon binding - the ligand and the receptor form a unique-optimized complex structure; this structure encompasses all key interactions formed between the key CXCL12 and CXCR4 residues which possess a functional role with regard to binding or signaling. The most recent binding models suggested by Xu et al.[56] could constitute probable initial binding conformations of CXCL12 to CXCR4 which could occur prior to the stabilization of the optimum, unique, and final CXCL12:CXCR4 complex conformation, which is identified in our study, and leads to signaling. The CXCL12:CXCR4 complex structure of Xu et al.[56] is unable to provide evidence on the role of the most critical CXCL12 residue, Lys1, with regard to signaling,[58,67] as according to Xu et al. Lys1 of CXCL12 is not buried within the transmembrane helical regions of CXCR4.[56] Hence, Xu et al.[56] cannot provide evidence on the key binding-signaling role of CXCR4 residues Phe87, Asp171, Trp252, Tyr255, Glu288, and Phe292.[77] In contrast, our study (i) presents Lys1 to be buried in the transmembrane CXCR4 domain, as suggested experimentally,[67,68] and (ii) provides compelling evidence on the combined role of Lys1 of CXCL12 and the aforementioned CXCR4 residues in the complex structure. It is possible that the binding of Lys1 in the N-terminal domain which was reported in the recent computational study[56] had affected negatively the formation of optimized interactions between CXCL12 and CXCR4 in the models reported.[56] For example, experiments verified the critical role with regards to signaling of CXCL12 residue Arg8 and CXCR4 residue Tyr190. This is in agreement with our study since we show that these residues have a combined-common role in signaling owing to the formation of a cation-π interaction between the two in the complex structure; on the contrary, Xu et al.[56] provide no evidence for the key signaling role of CXCR4 residue Tyr190. The higher degree of agreement with experimental findings of our computationally derived structure compared to the most recent CXLC12:CXCR4 modeled structure[56] is also reflected in the binding free energy which is approximately twice lower in our final simulation than the three predicted models of.[56] Moreover, it is worth noting that in our study the binding free energy was calculated with a quite more rigorous[81] Generalized Born representation[59] which was developed specifically to take into account the heterogeneous-dielectric water-membrane-water environment.

The inability of previous computational attempts[53−56] to report a CXLC12:CXCR4 structure in high agreement with previous experimental findings could be attributed to the difficulty in sufficiently sampling and correctly modeling the bound conformation of both (i) the highly flexible CXCL12 N-terminal domain as well as (ii) the complete CXCR4 structure with an appropriate conformation and orientation of the N-terminal domain, as residues 1−26 are missing from the X-ray structure of CXCR4.[51] As for CXCR4, the modeling and the nearly exhaustive conformational search to produce multiple receptor structures in Tamamis and Floudas[14] aided significantly the production of an appropriate CXCR4 conformation in this study, too. As for CXCL12, the selection of the proper bound CXCL12 conformation required for binding and signaling is considerably important owing to the high flexibility of region 1−11 and, thus, the numerous possible orientations which may occur between domains 1−11 and 12−68 of CXCL12. According to our findings, the CXCL12 bound conformation is not included neither in the NMR ensemble of ref 57 nor in any of the two X-ray structures of ref 58; nevertheless a hybrid-origin NMR/X-ray CXCL12 conformation which was simulated in the final simulation was shown to be appropriate to optimize the intermolecular interactions with CXCR4 upon binding. This suggests that the high flexibility of the 1−11 CXCL12 domain[67] and the large number of orientations between domains 1−11 and 12−68, prior to binding, could be most important for CXCL12 to adopt an appropriate conformation for binding.[67] Nevertheless, can the high flexibility of the 1−11 CXCL12 N-terminal domain lead the CXCL12 structure of Complex 2 to conformationally adopt the exact N-terminal orientation of Complex 3 and result in the hybrid-origin CXCL12 structure of the final simulation? To answer this, we performed replica exchange MD simulations[82−84] using the FACTS implicit solvent model,[85] as in ref 86−88 (see Supporting Information). Based on the analysis, we show that by starting from the bound conformation of Complex 2, and by imposing very light "bestfit" conformational constraints on the heavy backbone atoms of region 12−68, so as to maintain its structural integrity at elevated temperatures, the unbound CXCL12 conformation at of Complex 2 indeed possesses the capacity to adopt the N-terminal conformation of Complex 3 in region 1−11 at 300 K and, consequently, to be conformationally transformed into the bound structure of the final simulation (see Supporting Information *and* Supporting Figure 4). Despite the high flexibility of the CXCL12 N-terminal domain which is needed so as to adopt the appropriate and optimum conformation in complex with CXCR4, the 1−11 domain of CXCL12 upon binding is, interestingly, maintained in a "locked" low-flexibility conformation owing to its burial into CXCR4; the burial leads to the formation of significantly low binding free energy binding conformation, containing an abundance of intermolecular interactions with CXCR4. The remarkable agreement with experiments was additionally facilitated by the nearly exhaustive docking conformational search, as well as the high accuracy heterogeneous dielectric implicit model used[59] to simulate and select the optimum conformations (see Methods); the same computational protocol was recently applied and succeeded in reporting the first computationally derived HIV-1 gp120 V3 loop:CXCR4

complete complex structure, also in remarkable agreement with experiments.[14]

Molecular Recognition of CXCR4 by CXCL12 versus the HIV-1 gp120 V3 Loop. According to our knowledge, this study presents the first structural and interaction free energy based comparison between the HIV-1 gp120 V3 loop:CXCR4[14] and CXCL12:CXCR4 complexes. Despite the key differences between the bound structures of the HIV-1 gp120 V3 loop[14] and CXCL12 (Figure 1), and the high flexibility of both the unbound HIV-1 gp120 V3 loop[14,89] and the N-terminal domain of CXCL12[67] (see Supporting Information), the comparative analysis between the CXCR4 residues which interact with (i) the specific dual tropic HIV-1 gp120 V3 loop[14] and (ii) CXCL12 provides evidence that the specific V3 loop investigated in ref 14 and CXCL12 share the same CXCR4 binding pocket, as they mostly interact with the same CXCR4 residues. Therefore, there is a significant overlap between the conformational space captured by the HIV-1 gp120 V3 loop residue moiety 4−32 (which corresponds to 299−328 in the entire gp120 sequence[90]) and CXCL12 residue moieties 1−12, 31, 39−42, and 45−49 (see Supporting Figure 2). The RMSD between the last snapshot of the simulations in the HIV-1 gp120 V3 loop[14] and CXCL12 in complex with CXCR4 (with superposition in the entire transmembrane CXCR4 region), for the highly interacting CXCR4 residues listed in Figures 3A, 3B, is equal to 8.4 Å, 2.7 Å, and 1.8 Å for the N-terminal domain, extracellular loops, and transmembrane region, respectively. This result shows that the transmembrane region and N-terminal domain of CXCR4 exhibit the highest and lowest degree of similarity, respectively, with regard to the HIV-1 gp120 V3 loop[14] and CXCL12 binding to CXCR4. Therefore, this suggests that the binding of the HIV-1 gp120 V3 loop[14] and CXCL12 leads to the "locking" of the CXCR4 N-terminal domain in different conformations in the two complexes which are correlated with optimum intermolecular interactions. It is worth noting that most of the CXCR4 residues with the highest degree of dissimilarity between the HIV-1 gp120 V3 loop[14] and CXCL12 binding are situated in the N-terminal domain of CXCR4.

In general, the most highly interacting residues of CXCR4, in both complexes, are negatively charged; this result underlines the key role of negatively charged CXCR4 residues in attracting (and interacting with) CXCL12 and the HIV-1 gp120 V3 loop.[14] Both the HIV-1 gp120 V3 loop[14] and CXCL12 form their strongest interactions with CXCR4 transmembrane residues Asp171 and Glu288, as a consequence of Coulombic interactions between oppositely charged residues. In the case of the HIV-1 gp120 V3 loop,[14] its "central-tip" residue Arg18, which corresponds to residue Arg315 in the entire gp120 sequence,[90] forms two concurrent salt bridges with CXCR4 residues Asp171 and Glu288, and, in the case of CXCL12, its charged N-terminal end forms a salt bridge with Glu288 and the side chain positively charged group of CXC12 Lys1 forms a salt bridge with Asp171.

CONCLUSIONS

The discovery of novel CXCR4 antagonists is of utmost medical importance owing to their potential therapeutic application in blocking (i) the HIV-1 entry to CXCR4[14] or (ii) the CXCL12:CXCR4 axis which is involved in tumor metastasis, as well as tumor progression, angiogenesis, and survival.[23] The discovery of novel CXCL12 based agonists is also an imperative need, as CXCR4 agonists can among others be used for the mobilization of bone marrow hepatopoietic cells.[91] The high-accuracy computational derived structure of our study can be exploited for the *de novo* design[65,71,92−97] of new generations of CXCL12-based peptides which can serve as CXCR4 agonists or antagonists.

ASSOCIATED CONTENT

Supporting Information

The MD coordinates, extracted every 2-ns in PDB format. The structures are aligned with regard to the backbone of the CXCR4 transmembrane helical region. A video demonstrating the final simulation trajectory of CXCL12:CXCR4 and depicting the gradual stabilization and preservation of the key salt bridges, which are shown in Figure 2. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: 609-258-4595. Fax: 609-258-0211. E-mail: floudas@titan.princeton.edu.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

C.A.F. acknowledges funding from the National Institutes of Health (R01 GM052032). All MD simulations and free energy calculations were performed on the Tiger computer cluster at the TIGRESS high performance computer center at Princeton University.

REFERENCES

(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.* **2013**, DOI: 10.1016/j.canlet.2013.10.006.

(2) Bachelerie, F.; Ben-Baruch, A.; Burkhardt, A. M.; Combadiere, C.; Farber, J. M.; Graham, G. J.; Horuk, R.; Sparre-Ulrich, A. H.; Locati, M.; Luster, A. D.; Mantovani, A.; Matsushima, K.; Murphy, P. M.; Nibbs, R.; Nomiyama, H.; Power, C. A.; Proudfoot, A. E; Rosenkilde, M. M.; Rot, A.; Sozzani, S.; Thelen, M.; Yoshie, O.; Zlotnik, A. International Union of Pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. *Pharmacol. Rev.* **2014**, *66* (1), 1−79.

(3) Duda, G. D.; Kozin, S. V.; Kirkpatrick, N. D.; Xu, L.; Fukumura, D.; Jain, R. K. CXCL12 (SDF1alpha)-CXCR4/CXCR7 pathway inhibition: an emerging sensitizer for anticancer therapies? *Clin. Cancer Res.* **2011**, *17* (8), 2074−2080.

(4) Sun, X.; Cheng, G.; Hao, M.; Zheng, J.; Zhou, X.; Zhang, J.; Taichman, R. S.; Pienta, K. J.; Wang, J. CXCL12/CXCR4/CXCR7 chemokine axis and cancer progression. *Cancer Metastasis Rev.* **2010**, *29* (4), 709−722.

(5) Müller, A.; Homey, B.; Soto, H.; Ge, N.; Catron, D.; Buchanan, M. E.; McClanahan, T.; Murphy, E.; Yuan, W.; Wagner, S. N.; Barrera, J. L.; Mohar, A.; Verástegui, E.; Zlotnik, A. Involvement of chemokine receptors in breast cancer metastasis. *Nature* **2001**, *410* (6824), 50−56.

(6) Kojima, Y.; Acar, A.; Eaton, E. N.; Mellody, K. T.; Scheel, C.; Ben-Porath, I.; Onder, T. T.; Wang, Z. C.; Richardson, A. L.; Weinberg, R. A.; Orimo, A. Autocrine TGF-beta and stromal cell-derived factor-1 (SDF-1) signaling drives the evolution of tumor-promoting mammary stromal myofibroblasts. *PNAS* **2010**, *107* (46), 20009−20014.

(7) Burger, J. A.; Kipps, T. J. CXCR4: a key receptor in the crosstalk between tumor cells and their microenvironment. *Blood* **2006**, *107* (5), 1761−1767.
(8) Wright, D. E.; Bowman, E. P.; Wagers, A. J.; Butcher, E. C.; Weissman, I. L. Hematopoietic stem cells are uniquely selective in their migratory response to chemokines. *J. Exp. Med.* **2002**, *195* (9), 1145−1154.
(9) Laird, D. J.; von Adrian, U. H.; Wagers, A. J. Stem cell trafficking in tissue development, growth and disease. *Cell* **2008**, *132*, 612−630.
(10) Aiuti, A.; Webb, I. J.; Bleul, C.; Springer, T.; Gutierrez-Ramos, J. C. The chemokine SDF-1 is a chemoattractant for human CD34+ hematopoietic progenitor cells and provides a new mechanism to explain the mobilization of CD34+ progenitors to peripheral blood. *J. Exp. Med.* **1997**, *185* (1), 111−120.
(11) Peled, A.; Petit, I.; Kollet, O.; Magid, M.; Ponomaryov, T.; Byk, T.; Nagler, A.; Ben-Hur, H.; Many, A.; Shultz, L.; Lider, O.; Alon, R.; Zipori, D.; Lapidot, T. Dependence of human stem cell engraftment and repopulation of NOD/SCID mice on CXCR4. *Science* **1999**, *283* (5403), 845−848.
(12) Zhu, B.; Xu, D.; Deng, X.; Chen, Q.; Huang, Y.; Peng, H.; Li, Y.; Jia, B.; Thoreson, W. B.; Ding, W.; Ding, J.; Zhao, L.; Wang, Y.; Wavrin, K. L.; Duan, S.; Zheng, J. CXCL12 enhances human neural progenitor cell survival through a CXCR7- and CXCR4-mediated endocytotic signaling pathway. *Stem Cells* **2012**, *30* (11), 2571−2583.
(13) Ceradini, D. J.; Kulkarni, A. R.; Callaghan, M. J.; Tepper, O. M.; Bastidas, N.; Kleinman, M. E.; Capla, J. M.; Galiano, R. D.; Levine, J. P.; Gurtner, G. C. Progenitor cell trafficking is regulated by hypoxic gradients through HIF-1 induction of SDF-1. *Nat. Med.* **2004**, *10* (8), 858−864.
(14) Tamamis, P.; Floudas, C. A. Molecular recognition of CXCR4 by a dual tropic HIV-1 gp120 V3 loop. *Biophys. J.* **2013**, *105* (6), 1502−1514.
(15) Oberlin, E.; Amara, A.; Bachelerie, F.; Bessia, C.; Virelizier, J. L.; Arenzana-Seisdedos, F.; Schwartz, O.; Heard, J. M.; Clark-Lewis, I.; Legler, D. F.; Loetscher, M.; Baggiolini, M.; Moser, B. The CXC chemokine SDF-1 is the ligand for LESTR/fusin and prevents infection by T-cell-line-adapted HIV-1. *Nature* **1996**, *382* (6594), 833−835.
(16) Balkwill, F. Cancer and the chemokine network. *Nat. Rev. Cancer* **2004**, *4* (7), 540−550.
(17) Richert, M. M.; Vaidya, K. S.; Mills, C. N.; Wong, D.; Korz, W.; Hurst, D. R.; Welch, D. R. Inhibition of CXCR4 by CTCE-9908 inhibits breast cancer metastasis to lung and bone. *Oncol. Rep.* **2009**, *21* (3), 761−767.
(18) Sun, Y. X.; Schneider, A.; Jung, Y.; Wang, J.; Dai, J.; Wang, J.; Cook, K.; Osman, N. I.; Koh-Paige, A. J.; Shim, H.; Pienta, K. J.; Keller, E. T.; McCauley, L. K.; Taichman, R. S. Skeletal localization and neutralization of the SDF-1(CXCL12)/CXCR4 axis blocks prostate cancer metastasis and growth in osseous sites in vivo. *J. Bone Miner. Res.* **2005**, *20* (2), 318−329.
(19) Huang, E. H.; Singh, B.; Cristofanilli, M.; Gelovani, J.; Wei, C.; Vincent, L.; Cook, K. R.; Lucci, A. A CXCR4 antagonist CTCE-9908 inhibits primary tumor growth and metastasis of breast cancer. *J. Surg. Res.* **2009**, *155* (2), 231−236.
(20) Matsusue, R.; Kubo, H.; Hisamori, S.; Okoshi, K.; Takagi, H.; Hida, K.; Nakano, K.; Itami, A.; Kawada, K.; Nagayama, S.; Sakai, Y. Hepatic stellate cells promote liver metastasis of colon cancer cells by the action of SDF-1/CXCR4 axis. *Ann. Surg. Oncol.* **2009**, *16* (9), 2645−2653.
(21) Ma, W. F.; Du, J.; Fu, L. P.; Fang, R.; Chen, H. Y.; Cai, S. H. Phenotypic knockout of CXCR4 by a novel recombinant protein TAT/54R/KDEL inhibits tumors metastasis. *Mol. Cancer Res.* **2009**, *7* (10), 1613−1621.
(22) Sciumè, G.; Santoni, A.; Bernardini, G. Chemokines and glioma: invasion and more. *J. Neuroimmunol.* **2010**, *224* (1−2), 8−12.
(23) Teicher, B. A.; Fricker, S. P. CXCL12 (SDF-1)/CXCR4 pathway in cancer. *Clin. Cancer Res.* **2010**, *16* (11), 2927−2931.
(24) Burger, J. A.; Tsukada, N.; Burger, M.; Zvaifler, N. J.; Dell'Aquila, M.; Kipps, T. J. Blood-derived nurse-like cells protect chronic lymphocytic leukemia B cells from spontaneous apoptosis through stromal cell-derived factor-1. *Blood* **2000**, *96* (8), 2655−2663.
(25) Burger, J. A.; Burger, M.; Kipps, T. J. Chronic lymphocytic leukemia B cells express functional CXCR4 chemokine receptors that mediate spontaneous migration beneath bone marrow stromal cells. *Blood* **1999**, *94* (11), 3658−3667.
(26) Burger, M.; Hartmann, T.; Krome, M.; Rawluk, J.; Tamamura, H.; Fujii, N.; Kipps, T. J.; Burger, J. A. Small peptide inhibitors of the CXCR4 chemokine receptor (CD184) antagonize the activation, migration, and antiapoptotic responses of CXCL12 in chronic lymphocytic leukemia B cells. *Blood* **2005**, *106* (5), 1824−1830.
(27) Nie, Y.; Waite, J.; Brewer, F.; Sunshine, M. J.; Littman, D. R.; Zou, Y. R. The role of CXCR4 in maintaining peripheral B cell compartments and humoral immunity. *J. Exp. Med.* **2004**, *200* (9), 1145−1156.
(28) Sanz-Rodríguez, F.; Hidalgo, A.; Teixidó, J. Chemokine stromal cell-derived factor-1alpha modulates VLA-4 integrin-mediated multiple myeloma cell adhesion to CS-1/fibronectin and VCAM-1. *Blood* **2001**, *97* (2), 346−351.
(29) Damiano, J. S.; Cress, A. E.; Hazlehurst, L. A.; Shtil, A. A.; Dalton, W. S. Cell adhesion mediated drug resistance (CAM-DR): role of integrins and resistance to apoptosis in human myeloma cell lines. *Blood* **1999**, *93* (5), 1658−1667.
(30) Zeelenberg, I. S.; Ruuls-Van Stalle, L.; Roos, E. Retention of CXCR4 in the endoplasmic reticulum blocks dissemination of a T cell hybridoma. *J. Clin. Invest.* **2001**, *108* (2), 269−277.
(31) Bertolini, F.; Dell'Agnola, C.; Mancuso, P.; Rabascio, C.; Burlini, A.; Monestiroli, S.; Gobbi, A.; Pruneri, G.; Martinelli, G. CXCR4 neutralization, a novel therapeutic approach for non-Hodgkin's lymphoma. *Cancer Res.* **2002**, *62* (11), 3106−3112.
(32) Bradstock, K. F.; Makrynikola, V.; Bianchi, A.; Shen, W.; Hewson, J.; Gottlieb, D. J. Effects of the chemokine stromal cell-derived factor-1 on the migration and localization of precursor-B acute lymphoblastic leukemia cells within bone marrow stromal layers. *Leukemia* **2000**, *14* (5), 882−888.
(33) Möhle, R.; Bautz, F.; Rafii, S.; Moore, M. A.; Brugger, W.; Kanz, L. The chemokine receptor CXCR-4 is expressed on CD34+ hematopoietic progenitors and leukemic cells and mediates transendothelial migration induced by stromal cell-derived factor-1. *Blood* **1998**, *91* (12), 4523−4530.
(34) Monaco, G.; Konopleva, M.; Munsell, M.; Leysath, C.; Wang, R. Y.; Jackson, C. E.; Korbling, M.; Estey, E.; Belmont, J.; Andreeff, M. Engraftment of acute myeloid leukemia in NOD/SCID mice is independent of CXCR4 and predicts poor patient survival. *Stem Cells* **2004**, *22* (2), 188−201.
(35) Tavor, S.; Petit, I.; Porozov, S.; Avigdor, A.; Dar, A.; Leider-Trejo, L.; Shemtov, N.; Deutsch, V.; Naparstek, E.; Nagler, A.; Lapidot, T. CXCR4 regulates migration and development of human acute myelogenous leukemia stem cells in transplanted NOD/SCID mice. *Cancer Res.* **2004**, *64* (8), 2817−2824.
(36) Allinen, M.; Beroukhim, R.; Cai, L.; Brennan, C.; Lahti-Domenici, J.; Huang, H.; Porter, D.; Hu, M.; Chin, L.; Richardson, A.; Schnitt, S.; Sellers, W. R.; Polyak, K. Molecular characterization of the tumor microenvironment in breast cancer. *Cancer Cell* **2004**, *6* (1), 17−32.
(37) Miao, Z.; Luker, K. E.; Summers, B. C.; Berahovich, R.; Bhojani, M. S.; Rehemtulla, A.; Kleer, C. G.; Essner, J. J.; Nasevicius, A.; Luker, G. D.; Howard, M. C.; Schall, T. J. CXCR7 (RDC1) promotes breast and lung tumor growth in vivo and is expressed on tumor-associated vasculature. *PNAS* **2007**, *104* (40), 15735−15740.
(38) Hernandez, L.; Magalhaes, M. A.; Coniglio, S. J.; Condeelis, J. S.; Segall, J. E. Opposing roles of CXCR4 and CXCR7 in breast cancer metastasis. *Breast Cancer Res.* **2011**, *13* (6), R128.
(39) Burger, M.; Glodek, A.; Hartmann, T.; Schmitt-Gräff, A.; Silberstein, L. E.; Fujii, N.; Kipps, T. J.; Burger, J. A. Functional expression of CXCR4 (CD184) on small-cell lung cancer cells mediates migration, integrin activation, and adhesion to stromal cells. *Oncogene* **2003**, *22* (50), 8093−8101.

(40) Kijima, T.; Maulik, G.; Ma, P. C.; Tibaldi, E. V.; Turner, R. E.; Rollins, B.; Sattler, M.; Johnson, B. E.; Salgia, R. Regulation of cellular proliferation, cytoskeletal function, and signal transduction through CXCR4 and c-Kit in small cell lung cancer cells. *Cancer Res.* **2002**, *62* (21), 6304−6311.
(41) Hartmann, T. N.; Burger, J. A.; Glodek, A.; Fujii, N.; Burger, M. CXCR4 chemokine receptor and integrin signaling co-operate in mediating adhesion and chemoresistance in small cell lung cancer (SCLC) cells. *Oncogene* **2005**, *24* (27), 4462−4471.
(42) Zhou, Y.; Larsen, P. H.; Hao, C.; Yong, V. W. CXCR4 is a major chemokine receptor on glioma cells and mediates their survival. *J. Biol. Chem.* **2002**, *277* (51), 49481−49487.
(43) Rubin, J. B.; Kung, A. L.; Klein, R. S.; Chan, J. A.; Sun, Y.; Schmidt, K.; Kieran, M. W.; Luster, A. D.; Segal, R. A. A small-molecule antagonist of CXCR4 inhibits intracranial growth of primary brain tumors. *PNAS* **2003**, *100* (23), 13513−13518.
(44) Geminder, H.; Sagi-Assif, O.; Goldberg, L.; Meshel, T.; Rechavi, G.; Witz, I. P.; Ben-Baruch, A. A possible role for CXCR4 and its ligand, the CXC chemokine stromal cell-derived factor-1, in the development of bone marrow metastases in neuroblastoma. *J. Immunol.* **2001**, *167* (8), 4747−4757.
(45) Zeelenberg, I. S.; Ruuls-Van Stalle, L.; Roos, E. The chemokine receptor CXCR4 is required for outgrowth of colon carcinoma micrometastases. *Cancer Res.* **2003**, *63* (13), 3833−3839.
(46) Wang, J.; Wang, J.; Sun, Y.; Song, W.; Nor, J. E.; Wang, C. Y.; Taichman, R. S. Diverse signaling pathways through the SDF-1/CXCR4 chemokine axis in prostate cancer cell lines leads to altered patterns of cytokine secretion and angiogenesis. *Cell Signal* **2005**, *17* (12), 1578−1592.
(47) Scala, S.; Ottaiano, A.; Ascierto, P. A.; Cavalli, M.; Simeone, E.; Giuliano, P.; Napolitano, M.; Franco, R.; Botti, G.; Castello, G. Expression of CXCR4 predicts poor prognosis in patients with malignant melanoma. *Clin. Cancer Res.* **2005**, *11* (5), 1835−1841.
(48) Scotton, C. J.; Wilson, J. L.; Scott, K.; Stamp, G.; Wilbanks, G. D.; Fricker, S.; Bridger, G.; Balkwill, F. R. Multiple actions of the chemokine CXCL12 on epithelial tumor cells in human ovarian cancer. *Cancer Res.* **2002**, *62* (20), 5930−5938.
(49) Kim, J.; Takeuchi, H.; Lam, S. T.; Turner, R. R.; Wang, H. J.; Kuo, C.; Foshag, L.; Bilchik, A. J.; Hoon, D. S. Chemokine receptor CXCR4 expression in colorectal cancer patients increases the risk for recurrence and for poor survival. *J. Clin. Oncol.* **2005**, *23* (12), 2744−2753.
(50) Veldkamp, C. T.; Seibert, C.; Peterson, F. C.; De la Cruz, N. B.; Haugner, J. C., 3rd; Basnet, H.; Sakmar, T. P.; Volkman, B. F. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci. Signal* **2008**, *1* (37), ra4.
(51) Wu, B.; Chien, E. Y.; Mol, C. D.; Fenalti, G.; Liu, W.; Katritch, V.; Abagyan, R.; Brooun, A.; Wells, P.; Bi, F. C.; Hamel, D. J.; Kuhn, P.; Handel, T. M.; Cherezov, V.; Stevens, R. C. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **2010**, *330* (6007), 1066−1071.
(52) Cormier, E. G.; Tran, D. N.; Yukhayeva, L.; Olson, W. C.; Dragic, T. Mapping the determinants of the CCR5 amino-terminal sulfopeptide interaction with soluble human immunodeficiency virus type 1 gp120-CD4 complexes. *J. Virol.* **2001**, *75* (12), 5541−5549.
(53) Zhou, N.; Luo, Z.; Luo, J.; Liu, D.; Hall, J. W.; Pomerantz, R. J.; Huang, Z. Structural and functional characterization of human CXCR4 as a chemokine receptor and HIV-1 co-receptor by mutagenesis and molecular modeling studies. *J. Biol. Chem.* **2001**, *276* (46), 42826−42833.
(54) Huang, X.; Shen, J.; Cui, M.; Shen, L.; Luo, X.; Ling, K.; Pei, G.; Jiang, H.; Chen, K. Molecular dynamics simulations on SDF-1alpha: binding with CXCR4 receptor. *Biophys. J.* **2003**, *84* (1), 171−184.
(55) Liu, D.; Madani, N.; Li, Y.; Cao, R.; Choi, W. T.; Kawatkar, S. P.; Lim, M. Y.; Kumar, S.; Dong, C. Z.; Wang, J.; Russell, J. D.; Lefebure, C. R.; An, J.; Wilson, S.; Gao, Y. G.; Pallansch, L. A.; Sodroski, J. G.; Huang, Z. Crystal structure and structural mechanism of a novel anti-human immunodeficiency virus and D-amino acid-containing chemokine. *J. Virol.* **2007**, *81* (20), 11489−11498.
(56) Xu, L.; Li, Y.; Sun, H.; Li, D.; Hou, T. Structural basis of the interactions between CXCR4 and CXCL12/SDF-1 revealed by theoretical approaches. *Mol. Biosyst.* **2013**, *9* (8), 2107−2117.
(57) Veldkamp, C. T.; Ziarek, J. J.; Su, J.; Basnet, H.; Lennertz, R.; Weiner, J. J.; Peterson, F. C.; Baker, J. E.; Volkman, B. F. Monomeric structure of the cardioprotective chemokine SDF-1/CXCL12. *Protein Sci.* **2009**, *18* (7), 1359−1369.
(58) Ryu, E. K.; Kim, T. G.; Kwon, T. H.; Jung, I. D.; Ryu, D.; Park, Y. M.; Kim, J.; Ahn, K. H.; Ban, C. Crystal structure of recombinant human stromal cell-derived factor-1alpha. *Proteins* **2007**, *67* (4), 1193−1197.
(59) 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.* **2003**, *85* (5), 2900−2918.
(60) Brooks, B. R.; Brooks, C. L., 3rd; Mackerell, A. D., Jr.; Nilsson, L.; Petrella, R. J.; Roux, B.; Won, Y.; Archontis, G.; Bartels, C.; Boresch, S.; Caflisch, A.; Caves, L.; Cui, Q.; Dinner, A. R.; Feig, M.; Fischer, S.; Gao, J.; Hodoscek, M.; Im, W.; Kuczera, K.; Lazaridis, T.; Ma, J.; Ovchinnikov, V.; Paci, E.; Pastor, R. W.; Post, C. B.; Pu, J. Z.; Schaefer, M.; Tidor, B.; Venable, R. M.; Woodcock, H. L.; Wu, X.; Yang, W.; York, D. M.; Karplus, M. CHARMM: the biomolecular simulation program. *J. Comput. Chem.* **2009**, *30* (10), 1545−1614.
(61) Im, W.; Lee, M. S.; Brooks, C. L., III Generalized Born model with a simple smoothing function. *J. Comput. Chem.* **2003**, *24*, 1691−1702.
(62) Pierce, B. G.; Hourai, Y.; Weng, Z. Accelerating protein docking in ZDOCK using an advanced 3D convolution library. *PLoS One* **2011**, *6* (9), e24657.
(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* **2010**, *78* (12), 2655−2667.
(64) Tamamis, P.; Pierou, P.; Mytidou, C.; Floudas, C. A.; Morikis, D.; Archontis, G. 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* **2011**, *79* (11), 3166−3179.
(65) Tamamis, P.; López de Victoria, A.; Gorham, R. D., Jr.; Bellows-Peterson, M. L.; Pierou, P.; Floudas, C. A.; Morikis, D.; Archontis, G. Molecular dynamics in drug design: new generations of compstatin analogs. *Chem. Biol. Drug. Des.* **2012**, *79* (5), 703−718.
(66) Singh, N.; Warshel, A. Absolute binding free energy calculations: on the accuracy of computational scoring of protein-ligand interactions. *Proteins* **2010**, *78*, 1705−1723.
(67) Crump, M. P.; Gong, J. H.; Loetscher, P.; Rajarathnam, K.; Amara, A.; Arenzana-Seisdedos, F.; Virelizier, J. L.; Baggiolini, M.; Sykes, B. D.; Clark-Lewis, I. Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J.* **1997**, *16* (23), 6996−7007.
(68) Kofuku, Y.; Yoshiura, C.; Ueda, T.; Terasawa, H.; Hirai, T.; Tominaga, S.; Hirose, M.; Maeda, Y.; Takahashi, H.; Terashima, Y.; Matsushima, K.; Shimada, I. Structural basis of the interaction between chemokine stromal cell-derived factor-1/CXCL12 and its G-protein-coupled receptor CXCR4. *J. Biol. Chem.* **2009**, *284* (50), 35240−35250.
(69) MacKerell, A. D., Jr.; Feig, M.; Brooks, C. L., III Extending the treatment of backbone energetics in protein force fields: limitations of gas-phase quantum mechanics in reproducing protein conformational distributions in molecular dynamics simulations. *J. Comput. Chem.* **2004**, *25*, 1400−1415.
(70) MacKerell, A. D., Jr.; Bashford, D.; Bellott, M.; Dunbrack, R. L., Jr.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E., III; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M. All-atom

empirical potential for molecular modeling and dynamics Studies of proteins. *J. Phys. Chem. B* **1998**, *102*, 3586−3616.
(71) Gorham, R. D., Jr.; Forest, D. L.; Tamamis, P.; López de Victoria, A.; Kraszni, M.; Kieslich, C. A.; Banna, C. D.; Bellows-Peterson, M. L.; Larive, C. K.; Floudas, C. A.; Archontis, G.; Johnson, L. V.; Morikis, D. Novel compstatin family peptides inhibit complement activation by drusen-like deposits in human retinal pigmented epithelial cell cultures. *Exp. Eye Res.* **2013**, *116*, 96−108.
(72) Kieslich, C. A.; Tamamis, P.; Gorham, R. D., Jr.; López de Victoria, A.; Sausman, N.; Archontis, G.; Morikis, D. Exploring protein-protein and protein-ligand interactions in the immune system using molecular dynamics and continuum electrostatics. *Curr. Phys. Chem.* **2012**, *2*, 324−343.
(73) Frishman, D.; Argos, P. Knowledge-based protein secondary structure assignment. *Proteins* **1995**, *23* (4), 566−579.
(74) Doranz, B. J.; Orsini, M. J.; Turner, J. D.; Hoffman, T. L.; Berson, J. F.; Hoxie, J. A.; Peiper, S. C.; Brass, L. F.; Doms, R. W. Identification of CXCR4 domains that support coreceptor and chemokine report funtions. *J. Virol.* **1999**, *73* (4), 2752−2761.
(75) Brelot, A.; Heveker, N.; Montes, M.; Alizon, M. Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. *J. Biol. Chem.* **2000**, *275* (31), 23736−23744.
(76) Choi, W. T.; Tian, S.; Dong, C. Z.; Kumar, S.; Liu, D.; Madani, N.; An, J.; Sodroski, J. G.; Huang, Z. Unique ligand binding sites on CXCR4 probed by a chemical biology approach: implications for the design ofselective human immunodeficiency virus type 1 inhibitors. *J. Virol.* **2005**, *79* (24), 15398−15404.
(77) Tian, S.; Choi, W. T.; Liu, D.; Pesavento, J.; Wang, Y.; An, J.; Sodroski, J. G.; Huang, Z. Distinct functional sites for human immunodeficiency virus type 1 and stromal cell-derived factor 1alpha on CXCR4 transmembrane helical domains. *J. Virol.* **2005**, *79* (20), 12667−12673.
(78) Murphy, J. W.; Cho, Y.; Sachpatzidis, A.; Fan, C.; Hodsdon, M. E.; Lolis, E. Structural and functional basis of CXCL12 (stromal cell-derived factor-1 alpha) binding to heparin. *J. Biol. Chem.* **2007**, *282* (13), 10018−10027.
(79) Ohnishi, Y.; Senda, T.; Nandhagopal, N.; Sugimoto, K.; Shioda, T.; Nagal, Y.; Mitsui, Y. Crystal structure of recombinant native SDF-1alpha with additional mutagenesis studies: an attempt at a more comprehensive interpretation of accumulated structure-activity relationship data. *J. Interferon Cytokine Res.* **2000**, *20* (8), 691−700.
(80) van Gunsteren, W. F.; Dolenc, J. Biomolecular simulation: historical picture and future perspectives. *Biochem. Soc. Trans.* **2008**, *36*, 11−15.
(81) Yuzlenko, O.; Lazaridis, T. Membrane protein native state discrimination by implicit membrane models. *J. Comput. Chem.* **2013**, *34*, 731−34738.
(82) Hansmann, U. H. E. Parallel tempering algorithm for conformational studies of biological molecules. *Chem. Phys. Lett.* **1997**, *281*, 140−150.
(83) Sugita, Y.; Okamoto, Y. Replica-exchange molecular dynamics method for protein folding. *Chem. Phys. Lett.* **1999**, *314*, 141−161.
(84) Pieridou, G.; Avgousti-Menelaou, C.; Tamamis, P.; Archontis, G.; Hayes, S. C. UV resonance Raman study of TTR(105−115) structural evolution as a function of temperature. *J. Phys. Chem. B* **2011**, *115* (14), 4088−4098.
(85) Haberthür, U.; Caflisch, A. FACTS: Fast analytical continuum treatment of solvation. *J. Comput. Chem.* **2008**, *29* (5), 701−715.
(86) 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* **2009**, *113* (47), 15639−15647.
(87) 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* **2011**, *357* (2), 717−722.
(88) Tamamis, P.; Terzaki, K.; Kassinopoulos, M.; Mastrogiannis, L.; Mossou, E.; Forsyth, T. V.; Mitchell, E. P.; Mitraki, A.; Archontis, G. Self-assembly of an aspartate-rich sequence from the adenovirus fibre shaft: insights from molecular dynamics simulations and experiments. *J. Phys. Chem. B* **2014**, *118* (7), 1765−1774.
(89) 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* **2012**, *7* (11), e49925.
(90) Huang, C. C.; Lam, S. N.; Acharya, P.; Tang, M.; Xiang, S. H.; Hussan, S. S.; Stanfield, R. L.; Robinson, J.; Sodroski, J.; Wilson, I. A.; Wyatt, R.; Bewley, C. A.; Kwong, P. D. Structures of the CCR5 N terminus and of a tyrosine-sulfated antibody with HIV-1 gp120 and CD4. *Science* **2007**, *317* (5846), 1930−1934.
(91) Tchernychev, B.; Ren, Y.; Sachdev, P.; Janz, J. M.; Haggis, L.; O'Shea, A.; McBride, E.; Looby, R.; Deng, Q.; McMurry, T.; Kazmi, M. A.; Sakmar, T. P.; Hunt, S., 3rd; Carlson, K. E. Discovery of a CXCR4 agonist pepducin that mobilizes bone marrow hematopoietic cells. *PNAS* **2010**, *107* (51), 22255−22259.
(92) Bellows-Peterson, M. L.; Fung, H. K.; Floudas, C. A.; Kieslich, C. A.; Zhang, L.; Morikis, D.; Wareham, K. J.; Monk, P. N.; Hawksworth, O. A.; Woodruff, T. M. De novo peptide design with C3a receptor agonist and antagonist activities: theoretical predictions and experimental validation. *J. Med. Chem.* **2012**, *55* (9), 4159−4168.
(93) Bellows, M. L.; Taylor, M. S.; Cole, P. A.; Shen, L.; Siliciano, R. F.; Fung, H. K.; Floudas, C. A. Discovery of entry inhibitors for HIV-1 via a new de novo protein design framework. *Biophys. J.* **2010**, *99* (10), 3445−3453.
(94) 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. Visualized Exp.* **2013**, *77*, e50476.
(95) Klepeis, J. L.; Floudas, C. A.; Morikis, D.; Tsokos, C. G.; Argyropoulos, E.; Spruce, L.; Lambris, J. D. Integrated computational and experimenal approach for lead optimization and design of compstatin variants with improved activity. *J. Am. Chem. Soc.* **2003**, *125* (28), 8422−8423.
(96) Morikis, D.; Soulika, A. M.; Mallik, B.; Klepeis, J. L.; Floudas, C. A.; Lambris, J. D. Improvement of the anti-C3 activity of complement using rational and combinatorial approaches. *Biochem. Soc. Trans.* **2004**, *32*, 28−32.
(97) Smadbeck, J.; Peterson, M. B.; Zee, B. M.; Garapaty, S.; Mago, A.; Lee, C.; Giannis, A.; Trojer, P.; Garcia, B. A.; Floudas, C. A. De novo peptide design and experimental validation of histone methyltransferase inhibitors. *PLoS One* **2014**, DOI: 10.1371/journal.-pone.0090095.