Published in final edited form as:

Sci Signal. ; 10(471): . doi:10.1126/scisignal.aah5756.

Structural basis for chemokine recognition by a G protein-coupled receptor and implications for receptor activation

Joshua J. Ziarek^{1,*}, Andrew B. Kleist¹, Nir London^{2,†}, Barak Raveh^{3,†}, Nicolas Montpas^{4,5}, Julien Bonnetterre^{4,5}, Geneviève St-Onge⁵, Crystal J. DiCosmo-Ponticello⁶, Chad A. Koplinski¹, Ishan Roy⁷, Bryan Stephens⁸, Sylvia Thelen⁹, Christopher T. Veldkamp¹⁰, Frederick D. Coffman¹¹, Marion C. Cohen⁶, Michael B. Dwinell⁷, Marcus Thelen⁹, Francis C. Peterson¹, Nikolaus Heveker^{4,5}, and Brian F. Volkman^{1,‡}

¹Department of Biochemistry, Medical College of Wisconsin, Milwaukee, WI 53226, USA

²Department of Organic Chemistry, Weizmann Institute of Science, Rehovot 76100, Israel

³Department of Bioengineering and Therapeutic Sciences, University of California, San Francisco, San Francisco, CA 94158, USA

⁴Centre de Recherche, Centre Hospitalier Universitaire Sainte-Justine, Montréal, Quebec H3T 1C5, Canada

⁵Department of Biochemistry and Molecular Medicine, Université de Montréal, Montréal, Quebec H3T 1J4, Canada

⁶Rutgers Graduate School of Biomedical Sciences, Newark, NJ 07101, USA

⁷Department of Microbiology and Molecular Genetics, Medical College of Wisconsin, Milwaukee, WI 53226, USA

⁸Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California, San Diego, La Jolla, CA 93093, USA

⁹Institute for Research in Biomedicine, Università della Svizzera Italiana, Via Vela 6, Bellinzona CH-6500, Switzerland

¹⁰Department of Chemistry, University of Wisconsin-Whitewater, Whitewater, WI 53190, USA

¹¹Department of Pathology and Laboratory Medicine and Center for Biophysical Pathology, Rutgers New Jersey Medical School, Newark, NJ 07103, USA

Abstract

Chemokines orchestrate cell migration for development, immune surveillance, and disease by binding to cell surface heterotrimeric guanine nucleotide-binding protein (G protein)-coupled receptors (GPCRs). The array of interactions between the nearly 50 chemokines and their 20

[‡]Corresponding author. bvolkman@mcw.edu.

^{*}Present address: Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115, USA.

[†]These authors contributed equally to this work.

SUPPLEMENTARY MATERIALS

www.sciencesignaling.org/cgi/content/full/10/471/eaah5756/DC1

Methods

References (68–74)

Author contributions: J.J.Z. and B.F.V. conceived the project. J.J.Z., N.M., J.B., G.S.-O., C.J.D.-P., C.A.K., I.R., B.S., S.T., F.D.C., M.C.C., M.T., and N.H. collected and analyzed functional data. C.T.V., and F.C.P. collected and analyzed NMR data. N.L. and B.R. designed the approach for and performed computational modeling of the full-length CXCL12: CXCR4 complex. J.J.Z., A.B.K., and B.F.V. analyzed the modeling results, wrote the manuscript, and generated the figures. B.F.V. supervised the research.

Competing interests: B.F.V., F.C.P., M.B.D., and C.A.K. each has a substantial financial interest in Protein Foundry, LLC (a manufacturer of recombinant proteins for research).

Data and materials availability: The solution NMR ensemble, and related data sets, of the LM form of CXCL12 in complex with CXCR4_{1–38} has been deposited in the Research Collaboratory for Structural Bioinformatics PDB and Biological Magnetic Resonance Bank (BMRB) database under the codes PDB 2N55 and BMRB 25694, respectively.

GPCR targets generates an extensive signaling network to which promiscuity and biased agonism add further complexity. The receptor CXCR4 recognizes both monomeric and dimeric forms of the chemokine CXCL12, which is a distinct example of ligand bias in the chemokine family. We demonstrated that a constitutively monomeric CXCL12 variant reproduced the G protein–dependent and β -arrestin–dependent responses that are associated with normal CXCR4 signaling and lead to cell migration. In addition, monomeric CXCL12 made specific contacts with CXCR4 that are not present in the structure of the receptor in complex with a dimeric form of CXCL12, a biased agonist that stimulates only G protein–dependent signaling. We produced an experimentally validated model of an agonist-bound chemokine receptor that merged a nuclear magnetic resonance–based structure of monomeric CXCL12 bound to the amino terminus of CXCR4 with a crystal structure of the transmembrane domains of CXCR4. The large CXCL12:CXCR4 protein–protein interface revealed by this structure identified previously uncharacterized functional interactions that fall outside of the classical “two-site model” for chemokine-receptor recognition. Our model suggests a mechanistic hypothesis for how interactions on the extracellular face of the receptor may stimulate the conformational changes required for chemokine receptor-mediated signal transduction.

INTRODUCTION

The past decade has witnessed major revisions to the classical model of G protein (heterotrimeric guanine nucleotide–binding protein)–coupled receptor (GPCR) signaling. Instead of there being a single type of agonist-driven intracellular response, it has been recognized that different ligands can stabilize distinct active states in a single receptor to shift the balance of functional outputs. The predominant modes of GPCR signaling originate with heterotrimeric G protein activation and β -arrestin recruitment. Agonists can selectively activate one (“biased agonists”) or both (“balanced agonists”) pathways (1). Among the nearly 50 ligands and 20 receptors that constitute the chemokine family, promiscuity is common and biased agonism signaling by GPCRs may provide the regulatory discrimination that orchestrates in vivo cell migration.

The chemokine CXCL12 (also known as stromal cell-derived factor-1) and its receptor CXCR4 have been the focus of intense study for more than two decades. CXCL12 and CXCR4 are essential in developmental and housekeeping roles, but they also participate in numerous pathologies, including HIV infection and more than 23 different types of cancer (2). Like most chemokines, CXCL12 forms dimers when present at increasing concentrations (3, 4), when crystallized (5–7), or when bound to glycosaminoglycans in the extracellular matrix (8, 9), but was nonetheless presumed to interact with CXCR4 exclusively as a monomer at chemotactic concentrations (~10 nM). We previously used a disulfide-locked, constitutively dimeric CXCL12 variant [locked dimer (LD); also known as CXCL12₂] to show that a dimeric ligand potentially activates a subset of wild-type (WT) CXCL12-induced intracellular signals, a distinct example of biased agonism arising from a change in the oligomeric state of a ligand (10, 11). We previously hypothesized that some of the contacts observed in the LD:CXCR4 N-terminal peptide (CXCR4_{1–38}) nuclear magnetic resonance (NMR) structure are absent in complexes formed with the monomeric chemokine. However, a soluble 1:1 complex representing the balanced agonist has been inaccessible to

structural analysis because interactions with the CXCR4N-terminal domain promote CXCL12 dimerization (12).

Here, we used a constitutively monomeric CXCL12 variant [locked monomer (LM); also known as CXCL12₁] (13) to explore the molecular basis of balanced CXCR4 signaling. We first confirmed that the strictly monomeric LM was a faithful analog of balanced signaling by WT CXCL12, activating both G protein–dependent and β -arrestin–dependent pathways. We then determined the structure of the LM form of CXCL12 bound to CXCR4_{1–38}. Apolar residues near the N terminus of CXCR4 docked into a cleft that is inaccessible in the dimeric chemokine, and this monomer-specific interaction was essential for full receptor activation. By merging our NMR structure of the CXCL12:CXCR4_{1–38} complex and the crystal structure of CXCR4 bound to an inhibitor, we modeled an intact 1:1 complex. Our hybrid model broadens the conceptually useful “two-site” representation (14) and suggests that receptor activation involves the formation of an extensive protein-protein interface encompassing nearly half of the surface of CXCL12. We postulate that multiple, spatially distinct chemokine:receptor contacts work in tandem to drive conformational changes in CXCR4 that initiate signal transduction.

RESULTS

The LM form of CXCL12 is a balanced CXCR4 agonist with enhanced G protein–dependent and β -arrestin–dependent signaling

At low concentrations, CXCL12 is hypothesized to interact with CXCR4 as a monomer to promote cellular migration. As local chemokine concentrations increase, the dimeric form predominates and stimulates nonmigratory signaling that we previously termed “cellular idling” (10). Similarly, a peptide corresponding to the CXCR4 N terminus can bind to CXCL12 in a 1:1 stoichiometry but also promotes the formation of a 2:2 complex at higher concentrations (10, 15). The binding of short CXCR4 peptides that have specific tyrosine sulfation patterns to chemokines can even allosterically modulate chemokine dimerization (13). To simplify this complex equilibrium, we engineered a disulfide-constrained CXCL12 variant (LM) that remains monomeric at millimolar concentrations (13).

To test whether the LM form of CXCL12 was functionally equivalent to the monomeric form of WT CXCL12, we compared their relative abilities to recognize and activate CXCR4. Radioligand displacement on CXCR4-expressing cell membrane preparations demonstrated similar affinities for WT and LM CXCL12, with dissociation constant (K_d) values of 1.4 ± 1.5 nM and 0.97 ± 1.5 nM, respectively (Fig. 1A). We next established the activity of LM as a CXCR4 agonist by measuring its ability to mobilize intracellular Ca^{2+} , a sensitive indicator of G protein activation. LM produced a robust Ca^{2+} flux response with a sigmoidal dose dependence and a potency indistinguishable from that of WT CXCL12 (Fig. 1B). Chemokines typically induce cellular migration over a narrow concentration range when measured with a Boyden chamber or a similar apparatus. Whereas LD CXCL12 does not promote chemotaxis at any concentration, a preferentially monomeric CXCL12 variant stimulates chemotaxis across a larger range of concentrations relative to that of the WT chemokine, producing a wider bell-shaped profile (12). We next tested the chemotaxis of NALM6 pre-B cells and MiaPaCa2 pancreatic cancer cells with Boyden and Transwell

migration chambers, respectively (Fig. 1, C and D). In both cases, the LM chemokine retained chemotactic activity at higher concentrations relative to the WT chemokine, but cell migration returned to baseline levels at the highest concentrations. To assess how CXCR4 activation was interpreted in the presence of other migratory signals, we incubated U-937 leukemia cells confined to an agarose droplet in serum-containing medium with increasing CXCL12 concentrations. Whereas the WT CXCL12 reduced the extent of U-937 cell migration at nearly all concentrations, with a maximal reduction of $40 \pm 18\%$, the LM CXCL12 substantially enhanced migration at all tested concentrations and generated a bell-shaped profile (Fig. 1E).

Chemotaxis is dependent on the recruitment of β -arrestin to CXCR4, which then stimulates lamellipodia formation and filamentous actin polymerization (10). We previously demonstrated that monomeric CXCL12 is primarily responsible for the recruitment of β -arrestin-2 to CXCR4 and the subsequent internalization of the receptor (10). Here, we performed dose-dependent bioluminescence resonance energy transfer (BRET) analysis to test whether the enhanced chemotactic profile of LM CXCL12 was reflected in increased β -arrestin signaling. At low concentrations, WT and LM CXCL12 both recruited β -arrestin-2 to CXCR4 with similar potencies and efficacies (Fig. 1F). However, at concentrations of 10 and 100 μ M, WT CXCL12-induced β -arrestin-2 recruitment was attenuated, reminiscent of the biphasic chemotaxis profile. In contrast, LM CXCL12 exhibited a sigmoidal dose-response profile, with a maximal BRET response maintained at the highest concentrations tested. However, the functional relevance of this divergent β -arrestin-2 response at concentrations of chemokine that are likely not physiologically relevant remains to be determined.

The N terminus of CXCR4 interacts differently with the LM and LD forms of CXCL12

We next used NMR spectroscopy to explore the structural mechanisms of receptor activation. Chemokine signaling is initiated by the formation of an extensive protein-protein interface, which is segregated into two distinct regions (14). First, the N terminus of the receptor engages the folded chemokine domain and contributes most of the binding energy (site 1). Subsequent docking of the flexible N terminus of the chemokine into a pocket within the transmembrane (TM) domain of the receptor activates receptor signaling (site 2). To assess the alteration in site 1 dynamics of CXCR4₁₋₃₈ upon chemokine binding, we measured $\{^1\text{H}\}$ - ^{15}N heteronuclear nuclear Overhauser effect (NOE) values, which reflect the backbone flexibility for each residue on picosecond to nanosecond time scales (Fig. 2A). In our previous study of the binding of CXCR4₁₋₃₈ to LD CXCL12 (10), we found that a substantial increase in $\{^1\text{H}\}$ - ^{15}N NOE values for residues 11 to 25 accompanied the transition from the highly flexible free peptide to a more ordered conformation observed for those residues in the complex with LD CXCL12, whereas negative or nearzero NOE values indicated that the first 10 amino acid residues of CXCR4₁₋₃₈ remain highly dynamic, consistent with an absence of stable intermolecular contacts. In contrast, we found that the binding of LM CXCL12 increased the NOE magnitude for residues 5 to 10 of CXCR4₁₋₃₈, in addition to that for residues 11 to 25 (Fig. 2A), suggesting that residues near the CXCR4 N terminus interacted to a greater extent with the LM form than with the LD form of CXCL12.

We previously showed that titrating [U-¹⁵N]-CXCR4₁₋₃₈ with WT CXCL12 caused a subset of heteronuclear single-quantum coherence (HSQC) resonances to shift in nonlinear trajectories. This suggests that complexes of different stoichiometries were present in a multistate equilibrium (10), which is consistent with our earlier observation that CXCR4₁₋₃₈ promotes CXCL12 dimerization (13, 15). Here, we used HSQC titrations with the LM and LD chemokines to probe the 1:1 and 2:2 interfaces in isolation. Titrating LM or LD CXCL12 into [U-¹⁵N]-CXCR4₁₋₃₈ produced chemical shift perturbations consistent with distinct chemical environments for CXCR4 residues 1 to 13 (Fig. 2B). In contrast to previously published spectra of the binding of [U-¹⁵N]-CXCR4₁₋₃₈ to WT CXCL12 (10, 15), all peaks in the receptor N terminus were visible throughout the titrations and traversed linear paths (table S1). Concatenation of the linear trajectories regenerated the complicated chemical shift perturbations caused by WT CXCL12, indicating that the LM and LD variants each formed nonexchanging complexes of defined stoichiometry that were equivalent to those formed by the WT protein (Fig. 2C). The LM variant also produced larger chemical shift perturbations in residues near the N terminus of the receptor, which together suggest that the LM CXCL12:receptor complex forms an interface that is somewhat distinct from that of the LD CXCL12:receptor complex.

The structure of CXCR4₁₋₃₈ bound to CXCL12 defines an extensive “site 1” interface

To understand how the N terminus of CXCR4 recognizes a CXCL12 monomer, we solved the structure of the LM CXCL12:CXCR4₁₋₃₈ complex by NMR (fig. S1 and tables S2 and S3). As expected, LM CXCL12 adopted the canonical chemokine fold, consisting of a flexible N terminus, followed by the N-loop, a three-stranded antiparallel β sheet, and a C-terminal helix (Fig. 3A and fig. S2). Determination of the CXCR4₁₋₃₈ contact surface required unambiguous identification of intermolecular NOEs, indicative of a stable interaction on the millisecond (or longer) time scale, which were observed from residues 4 to 27 along the peptide (fig. S3 and table S3). CXCR4₁₋₃₈ has a stable extended architecture, with a short β strand from Tyr⁷ to Ser⁹, which agrees well with previously published chemical shift perturbations (13).

The three CXCR4 tyrosine residues (at positions 21, 12, and 7) were previously identified as “hot spots” that make substantial contributions to the binding energy of site 1 (13). The hydrophobic contacts of Tyr²¹ appear to be primarily satisfied by Val⁴⁹ and the methylene of Glu¹⁵, whereas Arg⁴⁷ and, to a lesser extent, Asn⁴⁵ interact with the hydroxyl group of Tyr²¹ in CXCR4. Tyr¹² is buried into a deep cleft formed by Pro¹⁰, Lys²⁷, Leu²⁹, and Val³⁹ that has no obvious electrostatic or charge interactions for a sulfated tyrosine. An irregular turn involving Asp¹⁰ and Asn¹¹ of CXCR4₁₋₃₈ places Tyr⁷ in close proximity to Tyr¹². The hydroxyl group of Tyr⁷ is positioned toward His²⁵ and Lys²⁷, but no specific hydrogen bond interactions are evident. CXCR4₁₋₃₈ residues Tyr⁷ and Ser⁹ form backbone hydrogen bonds with Ile²⁸ and Asn³⁰ in LM CXCL12, adding a previously uncharacterized parallel strand to the three-stranded antiparallel β sheet of CXCL12 (Fig. 3B and fig. S4). The intermolecular β sheet is adjacent to a hydrophobic cleft between the β 1 strand and C-terminal helix of CXCL12. The side chains of Ile⁴ and Ile⁶ of CXCR4 occupy this cleft and interact with residues Leu²⁶, Trp⁵⁷, Tyr⁶¹, and Ala⁶⁵ of CXCL12 (Fig. 3C). Sulfation of Tyr⁷ reduces the affinity of the Ile⁴ to Asp¹⁰ heptapeptide of CXCR4 for LM CXCL12 sixfold compared to

that of the unsulfated heptapeptide (13). Rather than forming electrostatic contacts as a sulfotyrosine, Tyr⁷ may contribute to binding as part of a hydrophobic motif that includes Ile⁴ and Ile⁶. Exhaustive functional studies of mutants at the CXCL12: CXCR4 site 1 interface support the intermolecular contacts identified in our structure, including those involving Lys³⁷, Val³⁹, and Arg⁴⁷ of CXCL12 and Ile⁴, Ile⁶, Tyr⁷, Tyr¹², and Tyr²¹ of CXCR4 (table S4).

Comparison of the 1:1 and 2:2 stoichiometry NMR structures defines the basis for differential receptor recognition of monomeric and dimeric CXCL12

The interaction between hydrophobic CXCR4 residues and the LM helix (Fig. 3A) is consistent with previous NMR titration studies of CXCR4₁₋₃₈, as well as cross-saturation NMR experiments performed with full-length CXCR4 (10, 13, 15, 16). When monomeric CXCL12 binds to CXCR4, the newly formed intermolecular β strand (Tyr⁷ to Ser⁹; Fig. 3B) occupies space that would be filled by the β 1 strand of an opposing CXCL12 monomer. Because CXCL12 self-association buries the cleft formed by the β 1 strand and C-terminal α helix within the dimer interface, residues in the CXCR4 N terminus show the most pronounced differences between the 2:2 dimeric complex [Protein Data Bank (PDB) 2K04] and the 1:1 complex presented here (Fig. 3, D and E). In the 2:2 complex, CXCR4 is excluded from the same cleft and crosses the CXCL12 dimer interface at Tyr¹² to make electrostatic contacts with Lys²⁷ of one LD protomer and His²⁵ of the other (12). Tyr⁷ also engages the opposing LD subunit in a pocket formed by Val²³ and Arg²⁰ (12). These comparisons demonstrate that the LM CXCL12: CXCR4₁₋₃₈ complex contains a distinct interface for residues 1 to 12 that is incompatible with the dimerization of CXCL12 chemokines (Fig. 3, D and E).

N-terminal CXCR4 residues participate in chemokine recognition and receptor activation

To assess the functional contributions of Ile⁴ and Ile⁶ in CXCR4, we measured CXCL12 binding affinity and Ca²⁺ flux dose responses for a series of CXCR4 mutants. Ile⁴ and Ile⁶ were simultaneously mutated to either alanine or glutamic acid residues in FLAG-tagged CXCR4 (fig. S5A). Whereas substitution with alanines had no effect on binding, the affinity of WT and LM CXCL12 for the isoleucine-to-glutamate mutants was reduced 30- and 90-fold, respectively (Fig. 3F and fig. S5B). Similarly, receptor activation, as monitored by measurement of the dose-dependent Ca²⁺ response, was reduced upon mutagenesis to glutamic acid (fig. S5, C and D). Specifically, whereas the reduction in maximum Ca²⁺ flux likely reflects the decreased amounts of the isoleucine mutants, the increase in the EC₅₀ values of the glutamic acid mutants is likely due to the substitution itself. Together, these data suggest that specific apolar contacts at the CXCR4 N terminus observed in the NMR structure of LM: CXCR4₁₋₃₈ contribute to the binding of CXCL12 and receptor activation.

Structure-guided modeling of the full-length 1:1 receptor complex and mutagenesis studies support previously uncharacterized CXCL12: CXCR4 contacts

To date, no structure of an intact CXCL12: CXCR4 complex exists. By combining our NMR structure of LM CXCL12: CXCR4₁₋₃₈ with the crystal structure of CXCR4 bound to the small-molecule inhibitor IT1t (17), we assembled a 1:1 model for the complete CXCL12: CXCR4 complex at atomic resolution (see data file S1 for the model coordinates).

The docking of LM CXCL12 to CXCR4 proceeded in five steps, as elaborated in the Supplementary Materials. We first docked the LM Nterminal peptide (residues 1 to 8; KPVSLSYR) into the orthosteric site of CXCR4 (PDB 3ODU: chain A; residues 29 to 301) using the FlexPepDock ab initio protocol (18). In the second step, the N-terminal peptide model from the first step and CXCR4 Pro²⁷ [which was anchored by the disulfide between Cys28 and Cys274^{7,25}; superscripts refer to Ballesteros-Weinstein numbering (19)] were used to roughly guide the placement of the LM CXCL12: CXCR4₁₋₃₈ NMR structure. In the third and fourth steps, we optimized the model using the RosettaRelax protocol (20) and connected the structured domain of CXCR4 to the CXCR4₁₋₃₈ fragment using the Rosetta loop modeling protocol (21). Last, LM residues 1 to 8 were redocked to adjust to the relaxed complex using FlexPepDock ab initio.

Because the initial docking pose of the eight N-terminal residues of CXCL12 affected all subsequent modeling steps, we verified the stability of the N terminus in the final model by energy funnel analysis (fig. S6). To this end, we used FlexPepDock ab initio to redock residues 1 to 8 of CXCL12 to residues 25 to 298 of CXCR4 (as in step 1 of the Supplementary Materials), but this time, we initiated the docking runs using the coordinates of these residues from the final model. We generated 250,000 new models from this starting conformation, under the same binding site constraints used in step 5 of the Supplementary Materials, to force the known interactions with residues Asp⁹⁷ and Glu²⁸⁸ of CXCR4 (22, 23). We also generated 250,000 unconstrained models. The peptide pose was highly restricted regardless of the constraints to Asp⁹⁷ and Glu²⁸⁸. By plotting the sum of the Rosetta FlexPepDock reweighted score (18) and the binding site constraints (as described in step 5 of the Supplementary Materials), we observed that the N-terminal residues lie at the bottom of a deep energy funnel (fig. S6, A and B), which is supported by a large number of favorable contacts that are maintained in most low-energy models (fig. S6, C and D). We assume that this stability is due to the deep binding pocket in CXCR4, which restricts the conformational freedom of the peptide, as is characteristic of particularly stable peptide-protein complexes (24).

Similarly, to assess the stability of the chemokine position, we redocked the LM domain wrapped with CXCR4₄₋₂₆ to the rest of CXCR4, using as a starting pose the best model resulting from the RosettaRelax described in step 3 of the Supplementary Materials. No additional constraints were used in this docking process. Here, a pronounced energy funnel indicated that this docking pose is stable and is found at a local energy minimum (fig. S6E). As a control to determine the degree to which RosettaRelax influenced the TM architecture, 100 relaxation runs were performed with the IT1t-bound CXCR4 structure (PDB 3ODU:A). In comparison to the ensemble of IT1t control models, our model has distinct side-chain positions in the TM bundle of the receptor that likely result from the experimental origin of the site 1 interaction (fig. S7).

The resulting model defines a large, contiguous interface (~3300 Å²) that buries nearly 40% of the CXCL12 surface (Fig. 4A). The site 1 interface defined by the NMR structure of LM: CXCR4₁₋₃₈ and the site 2 contacts formed by insertion of the N terminus of CXCL12 into the orthosteric pocket of CXCR4 agree closely with previous NMR and cysteine cross-linking studies (16, 25). CXCL12 methyl signals reduced by 10% or more as a result of

transferred cross-saturation (TCS) from the detergent-solubilized receptor (16) trace the path of CXCR4 N-terminal residues as they wrap around the chemokine (Fig. 4B, green surface residues). Although Kofuku *et al.* (16) discounted the TCS effect observed for several methyl groups buried in the CXCL12 core, including Val18⁷¹, Leu26⁸², Ile51, and Ile58, our model shows that cross-saturation of Leu²⁶ and Ile⁵⁸ may be a result of their proximity to receptor residues Ile⁶ and Ile⁴, respectively.

Next, we inspected the model for known structure-activity relationships. The earliest CXCL12 structure-function analysis established a substantial role for Lys¹ and Pro² in receptor activation, with a complete loss of Ca²⁺ flux agonist activity upon deletion or substitution of either residue (14). Crump *et al.* (14) reported that these N-terminally modified CXCL12 proteins retained high affinity for the receptor and could function as potent CXCR4 antagonists, suggesting that site 2 contacts at the base of the orthosteric pocket participate in signal transduction but contribute little to the overall binding energy. In our model, Lys¹ interacts with CXCR4 Asp97^{2.63} and Glu288^{7.39} (Fig. 4C), which are essential for receptor activation (22, 23). Because these favorable electrostatic contacts appeared to be important for the binding of CXCL12, we measured the affinity of a CXCL12 variant lacking the first two residues (CXCL12_{3–68}). In contrast to Crump *et al.*, we found that CXCR4 binding to CXCL12_{3–68} was markedly reduced ($K_d = 464 \pm 2$ nM; fig. S8A) in comparison to its binding to WT CXCL12 ($K_d = 1.45 \pm 1.5$ nM; Fig. 1A). The nearly 300-fold reduction in affinity corresponds to a change of ~ 3.5 kcal/mol in the free energy of binding, consistent with the loss of key hydrogen bonds or salt-bridge interactions involving both amino groups of Lys¹. Another hypothesis is that removing the conformationally restrictive N-terminal Pro² increases the entropic penalty of stabilizing an unstructured peptide. Consistent with this loss of affinity, CXCL12_{3–68} functioned as a very weak antagonist [IC₅₀(median inhibitory concentration) = 4.5 ± 0.9 μ M] in measurements of CXCR4-mediated Ca²⁺ flux (Fig. 4C), and it did not stimulate Ca²⁺ flux at any concentration from 0.5 nM to 1 μ M (fig. S8B). Other previously described site 2 contacts include CXCR4 residues His281^{7.32}, which forms a polar contact with the carbonyl group of CXCL12 Pro², and Val196^{5.35}, which packs near to Val³ of CXCL12 (26).

We scanned the model for contacts that fell outside the definitions of either site 1 or site 2 and identified three CXCL12 residues (Arg⁸, Arg¹², and Asn³³; Fig. 4A, red) that contribute ~ 340 Å² to the interface with CXCR4. Arg⁸ and Arg¹², which flank the CXC motif and are homologous to a similar interface between vMIP-II and CXCR4 that was termed site 1.5 by Qin *et al.* (27), interact with CXCR4 residues Glu³² and Asp181^{ECL2}, respectively (Fig. 4D). Alanine substitution of either residue substantially reduced agonist potency in the Ca²⁺ flux assay (Fig. 4D), which was similar to previous studies that showed a loss of potency in CXCR4-mediated chemotaxis when either Arg⁸ or Arg¹² is mutated (5, 8). Asn³³, which is adjacent to Cys³⁴ in the CXCL12 β 1- β 2 loop, is positioned to form a hydrogen bond with CXCR4 Asn176^{ECL2}, and its mutation to glutamic acid reduced dose-dependent Ca²⁺ flux fourfold (Fig. 4E). Together, these results support the hypothesis that the previously uncharacterized contacts revealed by our model contribute to the functional interface between CXCL12 and CXCR4.

DISCUSSION

The oligomeric state responsible for directing migration has been debated since the discovery of chemokine oligomerization; however, the “active” structure of a chemokine is not generalizable because monomeric, dimeric, and even polymeric chemokines promote signaling (28). CXCL12 is the first example in an emerging group of CXC-type chemokines for which both the monomer and dimer potentially activate the same receptor to produce disparate cellular responses (12, 29). Our previous work established LD CXCL12 as a biased agonist that activates only the G protein–dependent subset of CXCR4-mediated signals (10–12). Here, we demonstrate that monomeric CXCL12 elicits the broadest range of WT CXCL12-stimulated signals (Fig. 1).

The functional equivalence of the LM form of CXCL12 to the WT form and the stoichiometric stability of the LM CXCL12: CXCR4 complex at concentrations required for NMR analysis enabled us to explore the structural mechanisms underlying the functional differences between the LM and LD forms of CXCL12. Our NMR structure attributes previously unclear chemical shift perturbations (10), heteronuclear NOE (10), and TCS (16) data collected on WT CXCL12 to contacts formed by the distal portion of the N terminus of CXCR4 with a hydrophobic cleft and β strand that are exposed on the surface of the CXCL12 monomer but are buried within the dimer interface. The LM-containing complex illustrates a clear interface between the globular domain of the chemokine and nearly 20 residues of CXCR4. The structures of three other monomeric chemokine: receptor complexes containing variable portions of site 1 contacts (27, 30, 31) have been solved to date (Fig. 5). A cysteine conserved in the N terminus of nearly all chemokine receptors (32) serves as a convenient landmark for comparison. The location of the cysteine, as well as the overall position of the receptor N terminus, varies widely among the four structures, which suggests that there is a distinct orientation for each chemokine globular domain on the full-length receptor (Fig. 5).

Comparison of the CXCL12: CXCR4 model with the vMIP-II: CXCR4 crystal structure

CXCR4 was previously crystallized in complex with vMIP-II (Fig. 6A), an antagonistic, broad-spectrum viral chemokine, by introducing a disulfide cross-link between extracellular loop 2 (ECL2) of the receptor and the N terminus of vMIP-II (27). The vMIP-II globular domain is positioned near TM helices TM1 and TM2 of the receptor and forms an intermolecular β sheet between the N-loop and CXCR4 Pro²⁷-Lys²⁸ (Fig. 6A). In contrast, the bent CXC motif, which is typical of CXC chemokines (27), prevents analogous contacts in CXCL12 (Fig. 6B). This shifts the globular domain toward TM4, TM5, and TM6 and modifies the specific site 2 interactions in our model (Fig. 6B). Despite the N terminus of CXCL12 having two residues less than that of vMIP-II, Lys¹ and the vMIP-II N terminus reach a similar depth in the CXCR4 orthosteric pocket, where they both make critical contacts with Asp97^{2,63} and Glu288^{7,39}. A CXCL12: CXCR4 model derived from the vMIP-II: CXCR4 crystal structure also has these electrostatic contacts with Lys¹, although the specific electrostatic interactions are reversed (27); there is no evidence to support one arrangement over the other. These differences in the site 2 interactions orient the plane of the

CXCL12 β sheet approximately parallel to the TM region with a $\sim 80^\circ$ rotation relative to the position of vMIP-II (Fig. 6C).

Because the electron density for CXCR4 residues 1 to 22 is absent from the crystal structure, presumably because of disorder, a detailed comparison of the vMIP-II: CXCR4 site 1 interaction with our CXCL12: CXCR4 model is not possible. Whereas the CXCL12: CXCR4 site 1 interaction contributes $\sim 66\%$ of the total binding energy (13, 15), the vMIP-II: CXCR4 complex is primarily driven by site 2 (27, 33). The less extensive vMIP-II site 1 interface and the comparatively small total contact surface ($\sim 1330 \text{ \AA}^2$ buried) are consistent with previous mutagenic studies and may reflect the capacity of vMIP-II to recognize seven different chemokine receptors (34–36).

Model-based insight into CXCR4 activation

Despite using an inactive CXCR4 conformation as a starting template, the site 1 contacts provided by our NMR structure appear to have stimulated the receptor to adopt a more active-like state during the model relaxation process (fig. S7). Although caution should be taken to not overanalyze structural models, we speculate that the hybrid nature of our model may provide uncharacteristic insight into chemokine receptor activation by comparison to the growing number of chemokine receptor structures now available (fig. S9). Chemokine recognition on the extracellular surface of our model induces conformational changes in the TM domain through three converging mechanisms: (i) ECL2 bringing TM2 and TM3 together, (ii) the inward deviation of the upper halves of TM6 and TM7, and (iii) the TM1 and TM7 movement toward the orthosteric pocket through the N-terminal loop of TM7 [also termed ECL4 (37)].

Structural studies of GPCRs have shown that agonist-induced changes at ECL2 can propagate conformational changes to the TM domain, thereby facilitating receptor activation (38, 39). With respect to CXCR4, our model suggests that Arg¹² brings Asp181^{ECL2} of the β hairpin toward the orthosteric pocket through an electrostatic interaction and that Asn³³ similarly attracts Asn176^{ECL2} at the base of the β 1 strand through a hydrogen bond (Fig. 7A). The “pulling” motion is transmitted to TM3 through the disulfide bond between Cys109^{2,25} and Cys186^{ECL2} and to TM2 through an electrostatic network consisting of Arg183^{ECL2}, Asp97^{2,63}, Glu288^{7,39}, and CXCL12 Lys¹. An analogous 30s loop–ECL2 interface occurs in the active-state crystal structure of the CX3CL1: US28 complex (31), but not in the inactive-state structure of the vMIP-II: CXCR4 complex (fig. S10A) (27). Collectively, the 30s loop–ECL2 interaction causes TM2 and TM3 to shift $\sim 1 \text{ \AA}$ toward each other relative to the three inactive CXCR4 crystal structures, consistent with previous predictions of chemokine receptor activation (40).

On the other side of the receptor, an extensive network of hydrogen bonds and salt bridges stabilizes the inward movement of the upper halves of TM6 and TM7 relative to CXCR4 crystal structures. The docking of CXCR4 Phe28^{N-term} into a pocket formed by CXCL12 Ser⁶, Phe¹³, and Phe¹⁴ brings TM7 inward through the C28^{N-term}–C274^{7,25} disulfide (Fig. 7B), which is a previously described microswitch for CXCR4 activation (41). The inwardly displaced helix is further anchored through a network of inter- and intramolecular hydrogen bonds including Ser6–Phe26^{N-term}, Ser4–His281^{7,32}, Ser4–Asp262^{6,58}, Pro2–His281^{7,32},

Val3-Asp262^{6,58}, and His281^{7,32}-Asp262^{6,58} (Fig. 7B). Only our model and the active-state CX3CL1:US28 complex share this inward position of TM7 relative to structures of chemokine receptors in their inactive state (Fig. 7B and fig. S10B). Allosteric communication between the N terminus and TM7 on one side and ECL2 on the other side could “pinch” the chemokine receptor together to “seesaw” TM6 inward at the extracellular surface and outward at the intracellular surface (42, 43). Consistent with this activation model, CXCL12 and CX3CL1 act as conduits by linking opposite sides of their receptors. Specifically, the chemokines simultaneously make contacts with TM3, TM6, and TM7 through extensive networks involving the two conserved receptor disulfides (TM3 to ECL2 and the N terminus to TM7) and the cysteine motif-30s loop disulfide that couples opposite sides of the chemokines (Fig. 7B and fig. S10A).

We hypothesize that these relatively subtle movements of TM2, TM3, TM6, and TM7 underlie substantial changes in a hydrophobic cluster of CXCR4, encompassing residues that belong to two conserved molecular switch motifs (Fig. 7B). Principally, Tyr116^{3,32} goes from being oriented toward the TM bundle in all three inactive-state CXCR4 structures to being oriented toward TM2 in our model, where it is positioned to form a hydrogen bond with Thr90^{2,56} and is aided by the displacement of TM3 in the same direction (Fig. 7, B to E, and fig. S11, A and B). In contrast, a hydrogen bond orients Tyr116^{3,32} toward the orthosteric pocket in two of the three antagonist-bound CXCR4 crystal structures, suggesting that CXCL12 destabilizes Tyr116^{3,32}, granting it greater conformational freedom to form new active-state contacts (17, 27). In this new position, Tyr116^{3,32} would clash with Trp94^{2,60}, perhaps explaining its ~120° upward rotation in our model relative to its position in all other CXCR4 structures (Fig. 7, C to E). Similar steric effects may cause the rearrangement of Phe292^{7,43} and Trp252^{6,48} such that all four residues adopt previously uncharacterized, favorable hydrophobic interactions in our model (Fig. 7, C to E). The downward movements of L120^{3,36} and E288^{7,39} may also stabilize new rotameric states of other core residues such as Trp252^{6,48}, which belongs to the conserved CWxP^{6,50} motif (44, 45), and Trp94^{2,60} of the less well-described TxP^{2,58}xW motif (46) that is enriched among chemokine receptors (fig. S11C). Furthermore, the conserved Trp94^{2,60} contacts the ligands in four of five antagonist-bound crystallized chemokine receptors (excluding CVX15 bound to CXCR4) (17, 27, 31, 47) and is implicated in small-molecule antagonist recognition by other chemokine receptors (48). Together, site 1 experimental restraints promote rearrangement in the hydrophobic core of our model, modifying the position of several residues that have been suggested to stabilize the inactive, antagonist-bound state.

Our model of CXCL12-mediated CXCR4 activation draws parallels to analogous models of β_2 -adrenergic receptor (β_2 AR) and μ -opioid receptor (μ OR) activation. In both instances, subtle TM movements at the orthosteric pocket trigger repacking of the hydrophobic “conserved core triad” (that is, Pro^{5,50}, Ile^{3,40}, and Phe^{6,44}), which facilitates a substantial (that is, 10 to 14 Å) outward deviation at the cytoplasmic domain of TM6, which is required for coupling to G proteins (49–51). We predict that the analogous helix movement in our model would likely be facilitated by concerted rearrangement of Phe292^{7,43} and Trp252^{6,48} (Fig. 7, C to E, and fig. S11, D and E). The conformational changes in the extracellular and core regions of our model are not completely propagated to the intracellular surface as evidenced by characteristic inactive-state conformations of the DRY motif, the NPxxY

motif, and the intracellular position of TM6 compared to those of the active-state structures of the β_2 AR, μ OR, and US28. This is unsurprising, because efforts to simulate (or in this case, model) receptor activation with inactive-state starting structures are met with considerable challenges (52). Although our model and the active-state CX3CL1:US28 complex show similarities at the extracellular surface and binding pocket, they become less similar in the TM region, which may reflect the orthosteric ligand pocket of constitutively active US28 being partially decoupled from intracellular G protein recognition and signaling (31). Nevertheless, we suggest that our model supports a previously uncharacterized mechanism by which chemokine receptor activation may proceed. Encouragingly, during the preparation of this manuscript, many of the residues that participate in our mechanistic explanation of CXCR4 activation (that is, W94^{2.60}, D97^{2.63}, Y116^{3.32}, D187^{ECL2}, W252^{6.48}, D262^{6.58}, H281^{7.32}, E288^{7.39}, and F292^{7.43}) were identified as being critical for CXCR4 activation by comprehensive mutagenesis of CXCR4 in Ca^{2+} flux studies (53).

Conclusion

Chemokines and their receptors assemble an active complex through a process that has been defined by the most easily investigated elements, the receptor N-terminal domain and the chemokine N terminus, which define chemokine recognition sites 1 and 2 in the canonical “two-step, two-site” model. Here, we put forward a model that combines details from the NMR and crystal structures of CXCL12 and CXCR4 complexes to provide an illustration of how these distinct structural elements are joined together in an extensive protein-protein interface. Exploiting this abundance of structural data, our model provides structural hypotheses for the effects that result from posttranslational modifications, such as sulfation of receptor tyrosine or chemokine citrullination (54, 55), and small-molecule or peptide-based inhibitors (fig. S12). Moreover, our model suggests a mechanism by which known and previously uncharacterized interactions between CXCL12 and CXCR4 may be translated into receptor activation. In light of emerging structural details, the two-site model, although useful, is likely an oversimplification of chemokine-receptor recognition. Knowledge of the entire chemokine:receptor interface could be exploited to develop inhibitors that disrupt contacts that are not apparent from either the site 1 NMR structures or site 2 crystal structures alone. Both the CXCR4 crystal structures and the LD CXCL12:CXCR4_{1–38} NMR structure have yielded higher hit rates than homology model drug discovery campaigns (56, 57), and as such, our hybrid model may serve as a useful structural template for homology-based drug discovery for the chemokine family.

MATERIALS AND METHODS

Protein expression and purification

The CXCR4_{1–38}, WT CXCL12, LD CXCL12, and LM CXCL12 proteins were produced as previously described (15, 58).

Radioligand binding competition assay

HEK 293E cells were seeded in Dulbecco’s modified Eagle’s medium supplemented with 10% fetal bovine serum (FBS) and penicillin/streptomycin (100 U/ml) (Life Technologies) in six-well plates and transiently transfected with polyethylenimine (PEI) (Polysciences Inc.)

with WT, Ile⁴Ala/Ile⁶Ala, or Ile⁴Glu/Ile⁶Glu FLAG-hCXCR4 complementary DNA vector (2 µg per well). Radioligand binding assays were performed 48 hours after transfection. Cells were washed twice in phosphate-buffered saline (PBS) and incubated for 5 min with 100 µM phenylarsine oxide (Sigma-Aldrich) in PBS at 37°C. Cells were washed twice and resuspended in binding buffer [50 mM Hepes (pH 7.4), 5 mM MgCl₂, 1 mM CaCl₂, and 0.2% (w/v) bovine serum albumin (BSA)], seeded in a 96-well flat-bottom plate at 20,000 cells per well, and incubated for 30 min at 37°C with 50 pM ¹²⁵I-CXCL12 (PerkinElmer) as a tracer and increasing concentrations of competing unlabeled chemokine. Bound radioactivity was separated from free ligands by filtration on borosilicate filter paper (Molecular Devices) treated with a 0.33% PEI solution. Receptor-bound radioactivity was quantified by gamma-radiation counting (PerkinElmer Life and Analytical Sciences). Binding experiments were carried out in duplicate.

THP-1 cell Ca²⁺ response

THP-1 cells (a human monocytic cell line) were washed twice and resuspended in 96-well dishes at 2×10^5 cells per well in an assay buffer [Hanks' balanced salt solution (HBSS), 20 mM Hepes (pH 7.4), 0.1% (w/v) BSA, and FLIPR Calcium 4 dye (Molecular Devices)] and then were incubated for 1 hour at 37°C and 5% CO₂. Fluorescence was measured at 37°C with a FlexStation3 Multimode Microplate Reader (Molecular Devices) with excitation and emission wavelengths of 485 and 515 nm, respectively. Chemokines were resuspended at the concentrations indicated in the figure legends and added to the cells after baseline fluorescence was measured for 20 s. The percentage Ca²⁺ flux was calculated from the maximum fluorescence minus the minimum fluorescence as a percent of baseline fluorescence. EC₅₀ values were determined by nonlinear fitting to a four-parameter logistic function. Experiments to determine the IC₅₀ values of truncated CXCL12 were performed as described earlier with the addition of 10 nM WT CXCL12.

Chemotaxis of NALM6 cells

NALM6 cells were cultured in RPMI 1640 medium supplemented with 10% FBS, penicillin-streptomycin (100 U/ml), 2 mM glutamine, 50 µM β-mercaptoethanol, nonessential amino acids, 1 mM sodium pyruvate, and 25 mM Hepes buffer (pH 7.3). Chemotaxis assays were performed in triplicate in 48-well Boyden chambers (Neuro Probe) with 5-µm pore-size polyvinylpyrrolidone-free polycarbonate membranes. Chemotaxis medium (RPMI 1640, 25 mM Hepes supplemented with 1% FBS) alone or chemotaxis medium containing increasing concentrations of CXCL12 variants was added to the lower wells. Cells (1×10^5 per well) resuspended in chemotaxis medium were added to the upper well and incubated for 90 min at 37°C in a 5% CO₂ atmosphere. Cells were removed from the upper part of the membrane with a rubber policeman. Cells attached to the lower side of the membrane were fixed and stained as described previously (59). Migrated cells were counted in five randomly selected fields of 1000-fold magnification.

Chemotaxis of MiaPaCa2 cells

Analysis of the chemotactic migration of pancreatic cancer MiaPaCa2 cells was performed as previously described (60) with Transwell plates coated with collagen (15 µg/ml). Briefly, MiaPaCa2 cells were serumstarved for 2 hours, removed from culture flasks with enzyme-

free dissociation buffer, washed, and then counted with a hemocytometer. Cells (100,000 cells in 10 μ l of serum-free medium) were plated into the top chamber of each Transwell. The bottom chamber of each Transwell contained each stimulant in 500 μ l of serum-free medium. MiaPaCa2 cells were allowed to migrate for 6 hours, after which the cells remaining on the top of the chamber were swabbed out. Plates were then fixed in 4% paraformaldehyde and stained with DAPI (4',6-diamidino-2-phenylindole). Migrated cells were visualized and counted by fluorescence microscopy, with five representative high-powered fields analyzed per well.

Analysis of β -arrestin-2 recruitment

β -Arrestin-2 recruitment was measured with an intermolecular BRET assay, which was performed as described previously (61, 62). Briefly, HEK 293E cells were cotransfected with 1 μ g of GFP₁₀- β -arrestin-2 construct and 0.05 μ g of CXCR4-RLuc3. All transfections were normalized to 2 μ g of DNA per well with empty vector. After overnight culture, the transiently transfected cells were seeded in poly-D-lysine-coated 96-well white clear-bottom microplates (ViewPlate, PerkinElmer Life and Analytical Sciences) and cultured for 24 hours. The medium of the cells was then exchanged for BRET buffer [PBS, 0.5 mM MgCl₂, and 0.1% (w/v) BSA]. β -Arrestin-2 recruitment was measured 15 min after ligand addition and 10 min after the addition of the RLuc3 substrate coelenterazine 400a (NanoLight Technology) at a final concentration of 5 mM. The values were corrected to BRETnet by subtracting the background BRET signal obtained from cells transfected with the luciferase construct alone.

Agarose microdroplet assay

The agarose microdroplet assay was performed to assess U-937 cellular migration, as previously described (63). U-937 target cells (0.5×10^6 cells/ml) were harvested and washed in HBSS. The cells were centrifuged at 800g and were transferred to a graduated 15-ml glass conical tube. The cell concentration was adjusted in agarose medium, prepared from 2% (w/v) low-melting temperature SeaPlaque agarose, and medium containing 15% FBS (1:4, v/v). A single agarose droplet (1 μ l, containing 1×10^5 target cells) was placed in the center of each well of a 96-well flat-bottom tissue culture plate, in triplicate, with a gastight 0.05-ml Hamilton syringe (Hamilton Company). Droplets were allowed to harden at 4°C for 20 min. Chilled test medium (200 μ l) was applied to each well. Test medium consisted of serum-free medium, 25% FBS, and the CXCL12 concentrations indicated in the figure legends. The plate was incubated for 18 to 24 hours at 37°C and 5% CO₂. After incubation, the radius of each droplet was determined, and target cell migration was measured at four directional points 90° from one another with an inverted light microscope equipped with a gridded eyepiece at $\times 40$ total magnification. The percentage inhibition of each sample was quantified. The plate was incubated for an additional 24 hours to measure recovery, and cell viability was determined by trypan blue exclusion.

NMR structure determination

All NMR spectra were acquired on a Bruker DRX 600-MHz spectrometer equipped with a ¹H, ¹⁵N, ¹³CTXI CryoProbe at 298 K. Experiments were performed in a solution containing 25 mM deuterated MES (pH 6.8), 10% (v/v) D₂O, and 0.02% (w/v) NaN₃. NOE

distance restraints were obtained from 3D ^{15}N -edited NOESY (NOE spectroscopy)-HSQC, aliphatic ^{13}C -edited NOESY-HSQC, and aromatic ^{13}C -edited NOESY-HSQC spectra ($\tau_{\text{mix}} = 80$ ms) collected on both $[\text{U-}^{15}\text{N}, ^{13}\text{C}]$ -LM CXCL12 saturated with CXCR4₁₋₃₈ and $[\text{U-}^{15}\text{N}, ^{13}\text{C}]$ -CXCR₁₋₃₈ saturated with LM CXCL12. Intermolecular NOEs were obtained from a3D F1- $^{13}\text{C}/^{15}\text{N}$ -filtered/F3- ^{13}C -edited NOESY-HSQC ($\tau_{\text{mix}} = 120$ ms) collected on both $[\text{U-}^{15}\text{N}, ^{13}\text{C}]$ -LM CXCL12 saturated with CXCR4₁₋₃₈ and $[\text{U-}^{15}\text{N}, ^{13}\text{C}]$ -CXCR₁₋₃₈ saturated with LM CXCL12. In addition to NOEs, backbone ϕ/ψ dihedral angle restraints were derived from $^1\text{H}^{\text{N}}$, $^1\text{H}^{\alpha}$, $^{13}\text{C}^{\alpha}$, $^{13}\text{C}^{\beta}$, $^{13}\text{C}'$, and ^{15}N chemical shift data with TALOS + (64). Both distance and dihedral restraints were used to generate initial NOE assignments and preliminary structures with the NOEASSIGN module of CYANA (65). Complete structure determination was undertaken as an iterative process of correcting and assigning NOEs and running structure calculations with CYANA (65). The 20 CYANA conformers with the lowest target function were further refined by a molecular dynamics protocol in explicit solvent (66) with XPLOR-NIH (67).

2D NMR characterization

All NMR spectra were acquired on a Bruker DRX 600-MHz spectrometer equipped with a ^1H , ^{15}N , $^{13}\text{CTXI}$ CryoProbe at 298 K. Experiments were performed in a solution containing 25 mM deuterated MES (pH 6.8), 10% (v/v) D_2O , and 0.02% (w/v) NaN_3 . Heteronuclear NOE experiments were collected on 250 μM $[\text{U-}^{15}\text{N}]$ -CXCR4₁₋₃₈ in the absence and presence of 500 μM LM CXCL12. ^{15}N -HSQC spectra were collected to monitor the interaction of 200 pM $[\text{U-}^{15}\text{N}]$ -CXCR4₁₋₃₈ titrated with 0, 50, 100, 150, 200, and 250 pM LD CXCL12. ^{15}N -HSQC spectra were collected to monitor the interaction of 750 pM $[\text{U-}^{15}\text{N}]$ -CXCR4₁₋₃₈ titrated with 0, 187.5, 375, 562.5, 750, and 843.75 pM LM CXCL12.

Ca^{2+} response of CXCR4 mutants

Cell culture, transfection, and Ca^{2+} flux assays of Chinese hamster ovary K1 cells were performed as previously described (13).

Statistical analysis

Data were analyzed by one-way analysis of variance (ANOVA) with GraphPad Prism 4.0 (GraphPad Software).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank T. Handel and I. Kufareva for insightful discussion and sharing the coordinates of their CXCL12:CXCR4 model.

Funding: This work was supported by a postdoctoral fellowship from the Medical College of Wisconsin Cancer Center (to J.J.Z.), NIH grants F30CA196040 (to A.B.K.) and R01 AI058072 (to B.F.V.), and Canadian Institutes of Health Research grant MOP-123421 (to N.H.). A.B.K. is a member of the NIH-supported (T32 GM080202) Medical Scientist Training Program at the Medical College of Wisconsin.

REFERENCES AND NOTES

1. Rajagopal S, Rajagopal K, Lefkowitz RJ. Teaching old receptors new tricks: Biasing seven-transmembrane receptors. *Nat Rev Drug Discov.* 2010; 9:373–386. [PubMed: 20431569]
2. Dell'Agnola C, Biragyn A. Clinical utilization of chemokines to combat cancer: The double-edged sword. *Expert Rev Vaccines.* 2007; 6:267–283. [PubMed: 17408375]
3. Veldkamp CT, Peterson FC, Pelzek AJ, Volkman BF. The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci.* 2005; 14:1071–1081. [PubMed: 15741341]
4. Baryshnikova OK, Sykes BD. Backbone dynamics of SDF-1 α determined by NMR: Interpretation in the presence of monomer-dimer equilibrium. *Protein Sci.* 2006; 15:2568–2578. [PubMed: 17075134]
5. Ohnishi Y, Senda T, Nandhagopal N, Sugimoto K, Shioda T, Nagai Y, Mitsui Y. Crystal structure of recombinant native SDF-1 α with additional mutagenesis studies: An attempt at a more comprehensive interpretation of accumulated structure-activity relationship data. *J Interferon Cytokine Res.* 2000; 20:691–700. [PubMed: 10954912]
6. Ryu EK, Kim TG, Kwon TH, Jung ID, Ryu D, Park YM, Kim J, Ahn KH, Ban C. Crystal structure of recombinant human stromal cell-derived factor-1 α . *Proteins.* 2007; 67:1193–1197. [PubMed: 17357154]
7. Murphy JW, Yuan H, Kong Y, Xiong Y, Lolis EJ. Heterologous quaternary structure of CXCL12 and its relationship to the CC chemokine family. *Proteins.* 2010; 78:1331–1337. [PubMed: 20077567]
8. Murphy JW, Cho Y, Sachpatzidis A, Fan C, Hodsdon ME, Lolis E. Structural and functional basis of CXCL12 (stromal cell-derived factor-1 α) binding to heparin. *J Biol Chem.* 2007; 282:10018–10027. [PubMed: 17264079]
9. Ziarek JJ, Veldkamp CT, Zhang F, Murray NJ, Kartz GA, Liang X, Su J, Baker JE, Linhardt RJ, Volkman BF. Heparin oligosaccharides inhibit chemokine (CXC motif) ligand 12 (CXCL12) cardioprotection by binding orthogonal to the dimerization interface, promoting oligomerization, and competing with the chemokine (CXC motif) receptor 4 (CXCR4) N terminus. *J Biol Chem.* 2013; 288:737–746. [PubMed: 23148226]
10. Drury LJ, Ziarek JJ, Gravel S, Veldkamp CT, Takekoshi T, Hwang ST, Heveker N, Volkman BF, Dwinell MB. Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways. *Proc Natl Acad Sci USA.* 2011; 108:17655–17660. [PubMed: 21990345]
11. Roy I, McAllister DM, Gorse E, Dixon K, Piper CT, Zimmerman NP, Getschman AE, Tsai S, Engle DD, Evans DB, Volkman BF, Kalyanaraman B, Dwinell MB. Pancreatic cancer cell migration and metastasis is regulated by chemokine-biased agonism and bioenergetic signaling. *Cancer Res.* 2015; 75:3529–3542. [PubMed: 26330165]
12. Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC III, Basnet H, Sakmar TP, Volkman BF. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci Signal.* 2008; 1:ra4. [PubMed: 18799424]
13. Ziarek JJ, Getschman AE, Butler SJ, Taleski D, Stephens B, Kufareva I, Handel TM, Payne RJ, Volkman BF. Sulfopeptide probes of the CXCR4/CXCL12 interface reveal oligomer-specific contacts and chemokine allostery. *ACS Chem Biol.* 2013; 8:1955–1963. [PubMed: 23802178]
14. Crump MP, Gong JH, Loetscher P, Rajarathnam K, Amara A, Arenzana-Seisdedos F, Virelizier JL, Baggiolini M, Sykes BD, 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; 16:6996–7007. [PubMed: 9384579]
15. Veldkamp CT, Seibert C, Peterson FC, Sakmar TP, Volkman BF. Recognition of a CXCR4 sulfotyrosine by the chemokine stromal cell-derived factor-1 α (SDF- α /CXCL12). *J Mol Biol.* 2006; 359:1400–1409. [PubMed: 16725153]
16. 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; 284:35240–35250. [PubMed: 19837984]

17. Wu B, Chien EY, Mol CD, Fenalti G, Liu W, Katritch V, Abagyan R, Brooun A, Wells P, Bi FC, Hamel DJ, Kuhn P, Handel TM, Cherezov V, Stevens RC. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. 2010; 330:1066–1071. [PubMed: 20929726]
18. Raveh B, London N, Zimmerman L, Schueler-Furman O. Rosetta FlexPepDock ab-initio: Simultaneous folding, docking and refinement of peptides onto their receptors. *PLOS ONE*. 2011; 6:e18934. [PubMed: 21572516]
19. Ballesteros JA, Weinstein H. Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. *Methods Neurosci*. 1995; 25:366–428.
20. Das R, Baker D. Macromolecular modeling with Rosetta. *Annu Rev Biochem*. 2008; 77:363–382. [PubMed: 18410248]
21. Stein A, Kortemme T. Improvements to robotics-inspired conformational sampling in Rosetta. *PLOS ONE*. 2013; 8:e63090. [PubMed: 23704889]
22. 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; 275:23736–23744. [PubMed: 10825158]
23. Choi WT, Tian S, Dong CZ, Kumar S, Liu D, Madani N, An J, Sodroski JG, Huang Z. Unique ligand binding sites on CXCR4 probed by a chemical biology approach: Implications for the design of selective human immunodeficiency virus type 1 inhibitors. *J Virol*. 2005; 79:15398–15404. [PubMed: 16306611]
24. London N, Movshovitz-Attias D, Schueler-Furman O. The structural basis of peptide-protein binding strategies. *Structure*. 2010; 18:188–199. [PubMed: 20159464]
25. Kufareva I, Stephens BS, Holden LG, Qin L, Zhao C, Kawamura T, Abagyan R, Handel TM. Stoichiometry and geometry of the CXC chemokine receptor 4 complex with CXC ligand 12: Molecular modeling and experimental validation. *Proc Natl Acad Sci USA*. 2014; 111:E5363–E5372. [PubMed: 25468967]
26. Wong RSY, Bodart V, Metz M, Labrecque J, Bridger G, Fricker SP. Comparison of the potential multiple binding modes of bicyclam, monocyclam, and noncyclam small-molecule CXC chemokine receptor 4 inhibitors. *Mol Pharmacol*. 2008; 74:1485–1495. [PubMed: 18768385]
27. Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, Fenalti G, Wu H, Han GW, Cherezov V, Abagyan R, Stevens RC, Handel TM. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science*. 2015; 347:1117–1122. [PubMed: 25612609]
28. Czaplewski LG, McKeating J, Craven CJ, Higgins LD, Appay V, Brown A, Dudgeon T, Howard LA, Meyers T, Owen J, Palan SR, Tan P, Wilson G, Woods NR, Heyworth CM, Lord BI, Brotherton D, Christison R, Craig S, Cribbes S, Edwards RM, Evans SJ, Gilbert R, Morgan P, Randle E, Schofield N, Varley PG, Fisher J, Waltho JP, Hunter MG. Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1 α , MIP-1 β , and RANTES. Characterization of active disaggregated chemokine variants. *J Biol Chem*. 1999; 274:16077–16084. [PubMed: 10347159]
29. Das ST, Rajagopalan L, Guerrero-Plata A, Sai J, Richmond A, Garofalo RP, Rajarathnam K. Monomeric and dimeric CXCL8 are both essential for in vivo neutrophil recruitment. *PLOS ONE*. 2010; 5:e11754. [PubMed: 20668677]
30. Millard CJ, Ludeman JP, Canals M, Bridgford JL, Hinds MG, Clayton DJ, Christopoulos A, Payne RJ, Stone MJ. Structural basis of receptor sulfotyrosine recognition by a CC chemokine: The N-terminal region of CCR3 bound to CCL11/eotaxin-1. *Structure*. 2014; 22:1571–1581. [PubMed: 25450766]
31. Burg JS, Ingram JR, Venkatakrishnan AJ, Jude KM, Dukkkipati A, Feinberg EN, Angelini A, Waghray D, Dror RO, Ploegh HL, Garcia KC. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science*. 2015; 347:1113–1117. [PubMed: 25745166]
32. Stone MJ, Chuang S, Hou X, Shoham M, Zhu JZ. Tyrosine sulfation: An increasingly recognised post-translational modification of secreted proteins. *N Biotechnol*. 2009; 25:299–317. [PubMed: 19658209]

33. Zhou N, Luo Z, Luo J, Hall JW, Huang Z. A novel peptide antagonist of CXCR4 derived from the N-terminus of viral chemokine vMIP-II. *Biochemistry*. 2000; 39:3782–3787. [PubMed: 10736178]
34. Kledal TN, Rosenkilde MM, Coulin F, Simmons G, Johnsen AH, Alouani S, Power CA, Luttichau HR, Gerstoft J, Clapham PR, Clark-Lewis I, Wells TNC, Schwartz TW. A broad-spectrum chemokine antagonist encoded by Kaposi's sarcoma-associated herpesvirus. *Science*. 1997; 277:1656–1659. [PubMed: 9287217]
35. Chen S, Bacon KB, Li L, Garcia GE, Xia Y, Lo D, Thompson DA, Siani MA, Yamamoto T, Harrison JK, Feng L. In vivo inhibition of CC and CX3C chemokine-induced leukocyte infiltration and attenuation of glomerulonephritis in Wistar-Kyoto (WKY) rats by vMIP-II. *J Exp Med*. 1998; 188:193–198. [PubMed: 9653095]
36. Luttichau HR, Lewis IC, Gerstoft J, Schwartz TW. The herpesvirus 8-encoded chemokine vMIP-II, but not the poxvirus-encoded chemokine MC148, inhibits the CCR10 receptor. *Eur J Immunol*. 2001; 31:1217–1220. [PubMed: 11298347]
37. Szpakowska M, Bercoff DPerez, Chevigne A. Closing the ring: A fourth extracellular loop in chemokine receptors. *Sci Signal*. 2014; 7:pe21. [PubMed: 25185155]
38. Ahuja S, Hornak V, Yan EC, Syrett N, Goncalves JA, Hirshfeld A, Ziliox M, Sakmar TP, Sheves M, Reeves PJ, Smith SO, Eilers M. Helix movement is coupled to displacement of the second extracellular loop in rhodopsin activation. *Nat Struct Mol Biol*. 2009; 16:168–175. [PubMed: 19182802]
39. Bokoch MP, Zou Y, Rasmussen SG, Liu CW, Nygaard R, Rosenbaum DM, Fung JJ, Choi HJ, Thian FS, Kobilka TS, Puglisi JD, Weis WI, Pardo L, Prosser RS, Mueller L, Kobilka BK. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. *Nature*. 2010; 463:108–112. [PubMed: 20054398]
40. Govaerts C, Bondue A, Springael JY, Olivella M, Deupi X, Le Poul E, Wodak SJ, Parmentier M, Pardo L, Blanpain C. Activation of CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3. *J Biol Chem*. 2003; 278:1892–1903. [PubMed: 12411445]
41. Rana S, Baranski TJ. Third extracellular loop (EC3)-N terminus interaction is important for seven-transmembrane domain receptor function: Implications for an activation microswitch region. *J Biol Chem*. 2010; 285:31472–31483. [PubMed: 20663868]
42. Rajagopalan L, Rajarathnam K. Structural basis of chemokine receptor function—A model for binding affinity and ligand selectivity. *Biosci Rep*. 2006; 26:325–339. [PubMed: 17024562]
43. Kleist AB, Getschman AE, Ziarek JJ, Nevins AM, Gauthier PA, Chevigne A, Szpakowska M, Volkman BF. New paradigms in chemokine receptor signal transduction: Moving beyond the two-site model. *Biochem Pharmacol*. 2016; 114:53–68. [PubMed: 27106080]
44. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, Pan AC, Liu CW, Fung JJ, Bokoch MP, Thian FS, Kobilka TS, Shaw DE, Mueller L, Prosser RS, Kobilka BK. The dynamic process of β_2 -adrenergic receptor activation. *Cell*. 2013; 152:532–542. [PubMed: 23374348]
45. Nomiyama H, Yoshie O. Functional roles of evolutionary conserved motifs and residues in vertebrate chemokine receptors. *J Leukoc Biol*. 2015; 97:39–47. [PubMed: 25416815]
46. Rosenkilde MM, Benned-Jensen T, Frimurer TM, Schwartz TW. The minor binding pocket: A major player in 7TM receptor activation. *Trends Pharmacol Sci*. 2010; 31:567–574. [PubMed: 20870300]
47. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, Li T, Ma L, Fenalti G, Zhang W, Xie X, Yang H, Jiang H, Cherezov V, Liu H, Stevens RC, Zhao Q, Wu B. Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science*. 2013; 341:1387–1390. [PubMed: 24030490]
48. Scholten DJ, Canals M, Maussang D, Roumen L, Smit MJ, Wijtmans M, de Graaf C, Vischer HF, Leurs R. Pharmacological modulation of chemokine receptor function. *Br J Pharmacol*. 2012; 165:1617–1643. [PubMed: 21699506]
49. Rasmussen SGF, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, Devree BT, Rosenbaum DM, Thian FS, Kobilka TS, Schnapp A, Konetzki I, Sunahara RK, Gellman SH, Pautsch A, Steyaert J, Weis WI, Kobilka BK. Structure of a nanobody-stabilized active state of the β_2 adrenoceptor. *Nature*. 2011; 469:175–180. [PubMed: 21228869]

50. Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, Thian FS, Chae PS, Pardon E, Calinski D, Mathiesen JM, Shah STA, Lyons JA, Caffrey M, Gellman SH, Steyaert J, Skiniotis G, Weis WI, Sunahara RK, Kobilka BK. Crystal structure of the β_2 adrenergic receptor–Gs protein complex. *Nature*. 2011; 477:549–555. [PubMed: 21772288]
51. Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, Kato HE, Livingston KE, Thorsen TS, Kling RC, Granier S, Gmeiner P, Husbands SM, Traynor JR, Weis WI, Steyaert J, Dror RO, Kobilka BK. Structural insights into μ -opioid receptor activation. *Nature*. 2015; 524:315–321. [PubMed: 26245379]
52. Dror RO, Arlow DH, Maragakis P, Mildorf TJ, Pan AC, Xu H, Borhani DW, Shaw DE. Activation mechanism of the β_2 -adrenergic receptor. *Proc Natl Acad Sci USA*. 2011; 108:18684–18689. [PubMed: 22031696]
53. Wescott MP, Kufareva I, Paes C, Goodman JR, Thaker Y, Puffer BA, Berdougou E, Rucker JB, Handel TM, Doranz BJ. Signal transmission through the CXCR4 chemokine receptor 4 (CXCR4) transmembrane helices. *Proc Natl Acad Sci USA*. 2016; 113:9928–9933. [PubMed: 27543332]
54. Farzan M, Babcock GJ, Vasilieva N, Wright PL, Kiprilov E, Mirzabekov T, Choe H. The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 α association and HIV-1 entry. *J Biol Chem*. 2002; 277:29484–29489. [PubMed: 12034737]
55. Struyf S, Noppen S, Loos T, Mortier A, Gouwy M, Verbeke H, Huskens D, Luangsay S, Parmentier M, Geboes K, Schols D, Van Damme J, Proost P. Citrullination of CXCL12 differentially reduces CXCR4 and CXCR7 binding with loss of inflammatory and anti-HIV-1 activity via CXCR4. *J Immunol*. 2009; 182:666–674. [PubMed: 19109200]
56. Veldkamp CT, Ziarek JJ, Peterson FC, Chen Y, Volkman BF. Targeting SDF-1/CXCL12 with a ligand that prevents activation of CXCR4 through structure-based drug design. *J Am Chem Soc*. 2010; 132:7242–7243. [PubMed: 20459090]
57. Mysinger MM, Weiss DR, Ziarek JJ, Gravel S, Doak AK, Karpik J, Heveker N, Shoichet BK, Volkman BF. Structure-based ligand discovery for the protein-protein interface of chemokine receptor CXCR4. *Proc Natl Acad Sci USA*. 2012; 109:5517–5522. [PubMed: 22431600]
58. Takekoshi T, Ziarek JJ, Volkman BF, Hwang ST. A locked, dimeric CXCL12 variant effectively inhibits pulmonary metastasis of CXCR4-expressing melanoma cells due to enhanced serum stability. *Mol Cancer Ther*. 2012; 11:2516–2525. [PubMed: 22869557]
59. Loetscher P, Seitz M, Clark-Lewis I, Baggiolini M, Moser B. Monocyte chemotactic proteins MCP-1, MCP-2, and MCP-3 are major attractants for human CD4⁺ and CD8⁺ T lymphocytes. *FASEB J*. 1994; 8:1055–1060. [PubMed: 7926371]
60. Roy I, Zimmerman NP, Mackinnon AC, Tsai S, Evans DB, Dwinell MB. CXCL12 chemokine expression suppresses human pancreatic cancer growth and metastasis. *PLOS ONE*. 2014; 9:e90400. [PubMed: 24594697]
61. Percherancier Y, Berchiche YA, Slight I, Volkmer-Engert R, Tamamura H, Fujii N, Bouvier M, Heveker N. Bioluminescence resonance energy transfer reveals ligand-induced conformational changes in CXCR4 homo- and heterodimers. *J Biol Chem*. 2005; 280:9895–9903. [PubMed: 15632118]
62. Gravel S, Malouf C, Boulais PE, Berchiche YA, Oishi S, Fujii N, Leduc R, Sinnett D, Heveker N. The peptidomimetic CXCR4 antagonist TC14012 recruits β -arrestin to CXCR7: Roles of receptor domains. *J Biol Chem*. 2010; 285:37939–37943. [PubMed: 20956518]
63. Adelman N, Hasson M, Masih N, Cohen MC. Correlation between agarose microdroplet and capillary tube procedures as assays for migration inhibition of target cells. *J Immunol Methods*. 1980; 34:235–242. [PubMed: 6995527]
64. Shen Y, Delaglio F, Cornilescu G, Bax A. TALOS+: A hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR*. 2009; 44:213–223. [PubMed: 19548092]
65. Güntert P. Automated NMR structure calculation with CYANA. *Methods Mol Biol*. 2004; 278:353–378. [PubMed: 15318003]
66. Linge JP, Williams MA, Spronk CA, Bonvin AMJJ, Nilges M. Refinement of protein structures in explicit solvent. *Proteins*. 2003; 50:496–506. [PubMed: 12557191]

67. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM. The Xplor-NIH NMR molecular structure determination package. *J Magn Reson.* 2003; 160:65–73. [PubMed: 12565051]
68. Rohl, CA., Strauss, CEM., Misura, KMS., Baker, D. *Methods in Enzymology.* Ludwig, B., Michael, LJ., editors. Vol. 383. Academic Press; 2004. p. 66-93.
69. Isberg V, Mordalski S, Munk C, Rataj K, Harpsøe K, Hauser AS, Vroiling B, Bojarski AJ, Vriend G, Gloriam DE. GPCRdb: An information system for G protein-coupled receptors. *Nucleic Acids Res.* 2016; 44:D356–D364. [PubMed: 26582914]
70. Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 Å crystal structure. *J Mol Biol.* 2004; 342:571–583. [PubMed: 15327956]
71. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SG, Thian FS, Kobilka TS, Choi HJ, Kuhn P, Weis WI, Kobilka BK, Stevens RC. High-resolution crystal structure of an engineered human β^2 -adrenergic G protein-coupled receptor. *Science.* 2007; 318:1258–1265. [PubMed: 17962520]
72. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krausz N, Choe HW, Hofmann KP, Ernst OP. Crystal structure of opsin in its G-protein-interacting conformation. *Nature.* 2008; 455:497–502. [PubMed: 18818650]
73. Manglik A, Kruse AC, Kobilka TS, Thian FS, Mathiesen JM, Sunahara RK, Pardo L, Weis WI, Kobilka BK, Granier S. Crystal structure of the μ -opioid receptor bound to a morphinan antagonist. *Nature.* 2012; 485:321–326. [PubMed: 22437502]
74. Zhou N, Luo Z, Luo J, Liu D, Hall JW, Pomerantz RJ, 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; 276:42826–42833. [PubMed: 11551942]

Fig. 1. LM enhances CXCR4-mediated Ca^{2+} flux, cell migration, and β -arrestin-2 recruitment (A) Binding of the indicated CXCL12 proteins was measured by radioligand displacement of ^{125}I -CXCL12 from CXCR4-containing membrane fragments made from human embryonic kidney (HEK) 293E cells. K_d values for CXCR4 binding of WT and LM CXCL12 were calculated as 1.45 ± 1.5 nM and 0.98 ± 1.5 nM (SD), respectively, from their corresponding log EC_{50} (median effective concentration) values of -8.84 ± 0.17 and -9.01 ± 0.17 (SD), respectively. Data are in duplicate from three experiments. cpm, counts per minute. (B) Dose-dependent treatment of THP-1 cells with either LM or WT CXCL12 induced CXCR4-dependent intracellular Ca^{2+} responses, with EC_{50} values of 7.1 ± 1.3 nM and 8.7 ± 1.7 nM (SD), respectively. Data are in triplicate from two experiments. (C) NALM6 cell migration in response to the indicated concentrations of WT or LM CXCL12 was quantified after 90 min of stimulation. Chemotaxis was determined by counting the number of migrated cells in five high-power magnification fields. Data are means $\pm$ SD of at least nine experiments per concentration. (D) Migration of MiaPaCa2 cells was monitored after 6 hours of stimulation with the indicated concentrations of WT and LM CXCL12 using Transwell migration chambers. Chemotaxis was determined by counting the number of

migrated cells in five high-power magnification fields. Data are means $\pm$ SD of five fields from four experiments. (E) U-937 cells were confined to 1- μ l agarose droplets, and migration was observed after 18 to 24 hours of incubation with test medium containing WT or LM CXCL12. The percentage of migration inhibition is presented as means $\pm$ SD; the chemokine-free control was normalized to zero. Data are means $\pm$ SD of four experiments. (F) HEK 293E cells transiently cotransfected with plasmids encoding GFP₁₀- β -arrestin-2 as a BRET donor and CXCR4-RLuc3 as a BRET acceptor were stimulated with increasing concentrations of WT and LM CXCL12, resulting in EC₅₀ values of 17.6 ± 1.1 nM for WT and 30.6 ± 1.1 nM (SD) for LM. Data are means $\pm$ SD of three experiments. * $P < 0.01$; ** $P < 0.001$.

Fig. 2. LM CXCL12 and LD CXCL12 have distinct interactions with the CXCR4 N terminus
(A) $\{^1\text{H}\}$ - ^{15}N heteronuclear NOE experiment of 250 mM $[\text{U}-^{15}\text{N}]$ -CXCR4₁₋₃₈ in the absence (green) and presence (blue) of 500 μM LM CXCL12. CXCR4₁₋₃₈ residues 4 to 7 exhibited a more stable interaction with LM CXCL12 than with LD CXCL12 (10). Data are from a single experiment. **(B)** Two-dimensional (2D) $^1\text{H}/^{15}\text{N}$ HSQC spectra of $[\text{U}-^{15}\text{N}]$ -CXCR4₁₋₃₈ titrated with increasing concentrations of LM CXCL12 (left) or LD CXCL12 (right). $[\text{U}-^{15}\text{N}]$ -CXCR4₁₋₃₈ (750 μM) was titrated with LM CXCL12 (0, 187.5, 375, 562.5, 750, and 843.75 μM). $[\text{U}-^{15}\text{N}]$ -CXCR4₁₋₃₈ (200 μM) was titrated with LD CXCL12 (0, 50, 100, 150, 200, and 250 μM). Data are from a single experiment. ppm, parts per million. **(C)** The disparate directions of chemical shift perturbations underscore that LM and LD form distinct interfaces with CXCR4₁₋₃₈. In some instances, as illustrated with Ser⁵, their trajectories can be visually concatenated to reproduce the progression of WT from a 1:1 to 2:2 complex.

Fig. 3. NMR structure of LM CXCL12 in complex with CXCR4₁₋₃₈

(A) Surface representation of LM (blue) in complex with CXCR4₁₋₃₈ (orange). To simplify the visualization, only CXCR4₁₋₃₈ residues 1 to 23 are visible, and tyrosine residues are shown in ball-and-stick representation. Previously published changes in LM ¹H/¹⁵N chemical shift upon CXCR4₁₋₃₈ addition are mapped onto the chemokine surface (yellow) (16). (B) CXCR4 residues 7 to 9 add an intermolecular strand parallel to β1 of the three-stranded antiparallel β sheet of the chemokine; the hydrogen bond network is represented by dashed lines. (C) CXCR4₁₋₃₈ residues Ile⁴ and Ile⁶ pack into a cleft between the β sheet and helix contacting LM residues Leu²⁶ and Tyr⁶¹. (D and E) Comparison of the NMR structures of the LM: CXCR4₁₋₃₈ (PDB 2N55) (D) and LD: CXCR4₁₋₃₈ (PDB 2K04) (E) complexes. (F) Binding of WT CXCL12 was measured by radioligand displacement of ¹²⁵I-CXCL12 from CXCR4_{WT} or CXCR4 (I4E/I6E) in membranes prepared from transiently transfected HEK 293E cells. *K_d* values for the binding of CXCR4_{WT} and CXCR4 (I4E/I6E) to CXCL12 were calculated as 1.45 ± 1.5 nM and 45.7 ± 1.9 nM (SD), respectively, from their corresponding log EC₅₀ values of -8.84 ± 0.17 and -7.34 ± 0.29 (SD), respectively. Data are means ± SD of three experiments.

Fig. 4. Hybrid model of the full-length CXCL12: CXCR4 complex and experimental validation (A) Combining the LM: CXCR4₁₋₃₈ NMR structure and the CXCR4 crystal structure enabled modeling of the intact 1:1 signaling complex. Model generation and coordinates are located in data file S1. CXCL12 is colored light blue, with site 1, 1.5, and 2 contacts shown in yellow, red, and dark blue, respectively. CXCR4 residues 4 to 28 are colored orange, and the TM region is shaded in gray. (B) CXCL12 methyl groups that exhibited NMR intensity reductions of at least 10% from CXCR4-mediated TCS (16) are highlighted in green. (C) The N-terminal residues of CXCL12 occupy the orthosteric pocket, where salt bridges from the Lys¹ α-amine and e-amino groups to Glu²⁸⁸ and Asp⁹⁷ of CXCR4 contribute substantially to the binding energy. N-terminal truncation of the first two residues abolishes the Ca²⁺ flux agonist activity of CXCL12 (fig. S8B), and the CXCL12₃₋₆₈ protein competes only weakly with 10 nM WT CXCL12 [$IC_{50} = 4.5 \pm 0.9 \mu M$ (SD)]. Data are means $\pm$ SD of four replicates from two experiments. (D) Arg⁸ and Arg¹² of CXCL12 form salt bridges with Glu³² and Asp¹⁸¹ of CXCR4, respectively. As predicted, mutagenesis reduced the Ca²⁺ flux response from 7.3 ± 2.2 nM for WT to 110 ± 11 nM for CXCL12(R8A) or 95 ± 10 nM for CXCL12(R12A). Data are means $\pm$ SD of four replicates from two experiments. (E) Our model suggests that CXCL12 Asn³³ contributes to binding and signaling but is not predicted to be a component of either site 1 or site 2. A fourfold change in the magnitude of Ca²⁺ flux

[$K_d = 21 \pm 5$ nM versus 5.2 ± 2 nM (SD)] confirms the contributions of Asn³³ to receptor activation. Data are means $\pm$ SD of four experiments.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Fig. 5. Comparison of site 1 structures

The CCL11:CCR3 NMR structure (PDB 2MPM; left), a portion of the vMIP-II:CXCR4 x-ray structure (PDB 4RWS; middle), and a portion of the CX3CL1 :US28 x-ray structure (PDB 4XT1 ; right) were aligned pairwise to the LM:CXCR4 NMR structure (PDB 2N55). The chemokines were aligned from the first cysteine to the last cysteine in each globular domain, yielding root mean square deviations (RMSDs) of 2.7 Å (CCL11), 2.2 Å (vMIP-II), and 1.9 Å (CX3CL1). For reference, the conserved cysteine in the receptor N terminus is at position 24 in CCR3, position 28 in CXCR4 (mutated to alanine in the CXCR4₁₋₃₈ peptide), and position 23 in US28.

Fig. 6. Comparison of the CXCL12:CXCR4 hybrid model to the vMIP-II:CXCR4 crystal structure

(A) vMIP-II:CXCR4 x-ray structure (PDB 4RWS). vMIP-II forms an intermolecular β strand between the CC-motif and CXCR4 residues Pro²⁷ and Cys²⁸, which positions the globular domain near TM1 and TM2. (B) CXCL12:CXCR4 hybrid model derived from docking the LM:CXCR4₁₋₃₈ NMR structure (PDB 2N55) to the CXCR4 x-ray structure (PDB 3ODU). The globular domain of the chemokine makes contacts with TM4, TM5, and TM6, making distinct site 1, 1.5, and 2 contacts relative to the vMIP-II:CXCR4 structure. (C) Pairwise alignment between CXCR4 residues 28 to 300 (1972 atoms) of the vMIP-II:CXCR4 structure and the CXCL12:CXCR4 model, yielding an RMSD of 2.2 Å.

Fig. 7. Receptor ECL2 and N terminus may translate CXCL12 binding into conformational changes of the TM bundle

(A) CXCL12: CXCR4 model derived from docking the LM CXCL12: CXCR4₁₋₃₈ NMR structure (PDB 2N55) to the CXCR4 x-ray structure (PDB 3ODU). Inset: Magnified view of ECL2, TM2, and TM3 with the interaction network labeled. Dashed yellow lines indicate likely hydrogen bond or electrostatic interactions. The structure of vMIP2: CXCR4 (PDB 4RWS) is shown in gray. (B) The CXCL12: CXCR4 model demonstrates extensive hydrogen bond and electrostatic interactions spanning sites 1, 1.5, and 2. Apolar and polar interactions at sites 1.5 and 2, formed between CXCR4 Cys28^{N-term}-CXCL12 Ser⁶ and CXCR4 Phe29^{N-term}-CXCL12 Arg¹², may work in tandem with an extensive site 1 network to pull the extracellular portion of TM6 and TM7 toward the bundle during receptor activation. (C and D) Comparison of key amino acid positions in the vMIP2: CXCR4 structure (PDB 4RWS) (C) and the CXCL12: CXCR4 model (D). (E) Overlay of the structures from (C) and (D). Green arrows indicate differences between the vMIP2: CXCR4 structure and the CXCL12: CXCR4 model.