



# Multiple mechanisms of self-association of chemokine receptors CXCR4 and CCR5 demonstrated by deep mutagenesis

Received for publication, March 25, 2023, and in revised form, August 24, 2023. Published, Papers in Press, September 9, 2023.

<https://doi.org/10.1016/j.jbc.2023.105229>

Kevin S. Gill<sup>1</sup> , Kritika Mehta<sup>1</sup> , Jeremiah D. Heredia<sup>1</sup>, Vishnu V. Krishnamurthy<sup>1</sup> , Kai Zhang<sup>1</sup>, and Erik Procko<sup>1,2,\*</sup> 

From the <sup>1</sup>Department of Biochemistry, University of Illinois, Urbana, Illinois, USA; <sup>2</sup>Cyrus Biotechnology, Seattle, Washington, USA

Reviewed by members of the JBC Editorial Board. Edited by Kirill Martemyanov

Chemokine receptors are members of the rhodopsin-like class A GPCRs whose signaling through G proteins drives the directional movement of cells in response to a chemokine gradient. Chemokine receptors CXCR4 and CCR5 have been extensively studied due to their roles in leukocyte development and inflammation and their status as coreceptors for HIV-1 infection, among other roles. Both receptors form dimers or oligomers of unclear function. While CXCR4 has been crystallized in a dimeric arrangement, available atomic resolution structures of CCR5 are monomeric. To investigate their dimerization interfaces, we used a bimolecular fluorescence complementation (BiFC)-based screen and deep mutational scanning to find mutations that change how the receptors self-associate, either *via* specific oligomer assembly or alternative mechanisms of clustering in close proximity. Many disruptive mutations promoted self-associations nonspecifically, suggesting they aggregated in the membrane. A mutationally intolerant region was found on CXCR4 that matched the crystallographic dimer interface, supporting this dimeric arrangement in living cells. A mutationally intolerant region was also observed on the surface of CCR5 by transmembrane helices 3 and 4. Mutations predicted from the scan to reduce BiFC were validated and were localized in the transmembrane domains as well as the C-terminal cytoplasmic tails where they reduced lipid microdomain localization. A mutation in the dimer interface of CXCR4 had increased binding to the ligand CXCL12 and yet diminished calcium signaling. There was no change in syncytia formation with cells expressing HIV-1 Env. The data highlight that multiple mechanisms are involved in self-association of chemokine receptor chains.

G protein-coupled receptors (GPCRs) are integral membrane proteins that activate heterotrimeric G protein signaling pathways, leading to diverse downstream signals. They are involved in sight, hormonal regulation, white blood cell trafficking, and a wide range of other physiological functions.

While it is accepted that many GPCRs dimerize, the functional consequences of dimerization on ligand interactions and downstream signaling remain incompletely understood, with structural experiments of class A rhodopsin-like GPCRs suggesting that they bind to ligand as monomers (1), and coimmunoprecipitation and FRET experiments suggesting that they are active as dimers (2–4). Ligand-receptor interactions may have additional complexities where the ligands may act as monomers or dimers with different effects (5) and there may be negative cooperativity between subunits in a GPCR dimer such that effectively only one of the receptors is ligand-bound (6). To determine whether dimerization is important for GPCR activity, researchers have tried different methods to identify dimerization interfaces for class A GPCRs with limited success. While class C GPCRs such as GABA<sub>B</sub> receptors, metabotropic glutamate receptors, and sweet and umami taste receptors are obligate dimers (7, 8), the dimerization of class A GPCRs is less understood, although it has been observed in multiple crystal structures including those of rhodopsin (9) and  $\mu$ -opioid receptor (10). Different transmembrane (TM) helices have been observed at dimer interfaces, including TM1 and TM4-6.

We have previously used conformation-dependent antibodies against the class A GPCRs CXCR4 and CCR5 to develop a method for deep mutational scanning of TM proteins expressed on human cells (11). CXCR4 and CCR5 are chemokine receptors, which in response to a chemokine gradient cause cell migration, often of white blood cells, but also of other cell types and some cancers (12, 13). These two chemokine receptors are also well studied for their role as coreceptors for HIV-1 infection (14–16). There is extensive evidence for CXCR4 dimerization and oligomerization (17–20), and the protein crystallizes as a dimer (21). Wu *et al.* showed a dimerization interface using TM3-5 and the extracellular tip of TM6 that was consistent across five different crystal forms. Structures of CCR5 have been determined with and without ligands as a monomer (22–24) but not yet as a homodimer. Models for CCR5 dimerization have been made through comparisons to other class A GPCR dimeric structures and molecular dynamics simulations (25, 26). These

\* For correspondence: Erik Procko, [procko@illinois.edu](mailto:procko@illinois.edu). Present address for Jeremiah D. Heredia: Codexis, Redwood City, California 94063.

## Self-association of chemokine receptors by deep mutagenesis

methods have suggested multiple different dimerization interfaces that collectively cover almost all sides of the protein. Given GPCRs are known to oligomerize and hetero-oligomerize with other GPCRs (27, 28), finding physiologically relevant dimerization interfaces has been challenging.

By examining the mutational landscapes of CXCR4 and CCR5 for binding to antibodies 12G5 and 2D7, respectively, we found conserved sites on the surface of the TM domains that, with respect to CXCR4, matched previous crystallographic dimers (21). To further elucidate how these chemokine receptors dimerize, we used bimolecular fluorescence complementation (BiFC) with deep mutational scanning to search for membrane-exposed surface patches that were intolerant of mutations for high-BiFC signal. BiFC is dependent upon proteins being in close proximity, whether through their physical binding or close association *via* other means, such as localized clustering. We found regions on the chemokine receptors that are hypothesized sites for dimerization, which on CXCR4 again matched closely with the crystallographic dimer conformation. We also found many structurally disruptive mutations that increased BiFC signal, including nonspecifically with an unrelated GPCR, suggesting over-expression of mutant chemokine receptors may lead to aggregation. Finally, mutations were found in the cytoplasmic tails that reduced BiFC between receptors and also reduced colocalization with a marker for lipid microdomains. We conclude that there are multiple mechanisms that mediate self-association of chemokine receptor polypeptides.

### Results

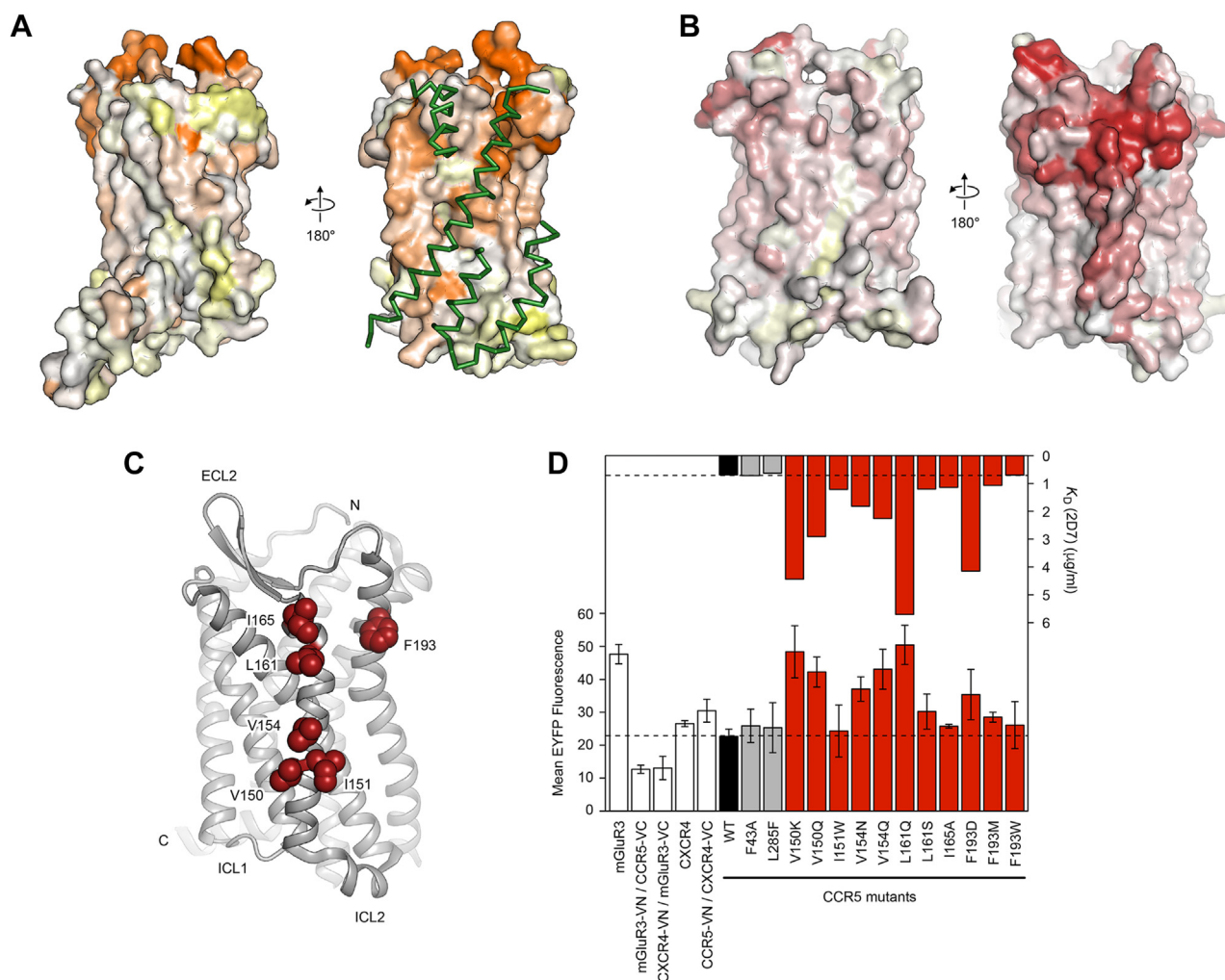
#### **Mutations to membrane-exposed surfaces of CXCR4 and CCR5 are deleterious for binding conformation-dependent antibodies**

We revisited published deep mutational scans of the chemokine receptors CXCR4 and CCR5 to search for evidence of receptor oligomers (11). These scans determined how mutations in the receptors impacted surface expression and recognition by conformation-dependent mAbs: 12G5 for CXCR4 and 2D7 for CCR5. These antibodies bind to extracellular loops of the receptors, and we hypothesized that mutations in TM regions that disrupt oligomeric organization may either alter surface expression or how the extracellular epitope is presented. Mapping conservation scores from the mutational scan of CXCR4 for binding to 12G5 onto the CXCR4 crystal structure (21) revealed that the dimer interface is indeed weakly but discernably more conserved than other membrane-exposed surfaces (Fig. 1A). We thus turned our attention to CCR5, for which a dimeric structure has not been determined at atomic resolution. Mapping conservation scores from the mutational scan of CCR5 for binding to 2D7 to the CCR5 monomer structure (22), we found that a surface primarily formed by TM4 was more conserved than other membrane-exposed surfaces and might therefore form the dimer interface (Fig. 1B). We thus introduced mutations into this region of CCR5 that were depleted in the scan for 2D7 binding (Fig. 1C). Two control mutations

located elsewhere on CCR5 and expected to have no effect were also evaluated. We found as expected that most of the mutations did indeed decrease 2D7 affinity (Fig. 1D). To understand whether the mutations decreased receptor associations as hypothesized, we coexpressed CCR5 mutants fused to N-terminal (VN) or C-terminal (VC) segments of split Venus (a variant of yellow fluorescent protein or YFP). If the receptors closely associate, the VN and VC polypeptides will fold together to produce fluorescent Venus. This method is known as BiFC (29). However, to our surprise, mutations that reduced CCR5 affinity for 2D7 caused increased BiFC signal (Fig. 1D). This suggested that the mutations to the hypothesized dimer interface of CCR5 were generally disruptive of structure, reducing 2D7 recognition and causing putative aggregation in the membrane, measured as increased BiFC. These initial observations inspired us to use BiFC as the basis for deep mutational scans of CXCR4 and CCR5 to see whether we could identify mutations that decreased receptor associations.

#### **Deep mutational scans of CXCR4 and CCR5 based on BiFC**

Deep mutational scanning couples an *in vitro* screen or selection of a diverse library of sequence variants with next generation sequencing. BiFC is well suited to illuminating how mutations in chemokine receptors influence receptor self-association, as BiFC facilitates the separation of cells using fluorescence-activated cell sorting (FACS). Using the split Venus BiFC method (29), single-site saturation mutagenesis libraries of CXCR4 and CCR5 (11) were modified by fusing the C-terminal half of split Venus to the C termini, creating CXCR4-VC and CCR5-VC libraries. These were expressed in an Expi293F cell line that was modified to (i) first remove endogenous CXCR4 using CRISPR-Cas9 (CXCR4-KO or X4-KO (11)) and (ii) second to stably integrate a gene encoding either CXCR4 or CCR5 fused at their C termini to the N-terminal half of split Venus, creating CXCR4-VN and CCR5-VN stable lines without endogenous CXCR4 expression. The VC- and VN-fused receptors had different N-terminal (and thus extracellular) epitope tags for their detection. The stable lines were transiently transfected with the matching VC-fused receptor libraries under conditions where the cells typically express no more than one gene per cell, providing a tight link from genotype to phenotype. The cell libraries were sorted by FACS in which collection gates were applied to cells that expressed CXCR4-VC or CCR5-VC mutants at the plasma membrane. In a second sorting experiment, cells were gated not only for surface CXCR4-VC or CCR5-VC expression but also for high levels of Venus fluorescence (Fig. 2), indicative of an association between a VC-fused mutant with a VN-fused WT receptor. RNA transcripts from the sorted libraries were analyzed by Illumina sequencing and compared to the naïve plasmid libraries to calculate an enrichment ratio for each mutation. The enrichment or depletion of single mutations in the CXCR4-VC and CCR5-VC genes are presented as heat maps that represent mutational landscapes for receptor associations (Figs. 3 and 4).



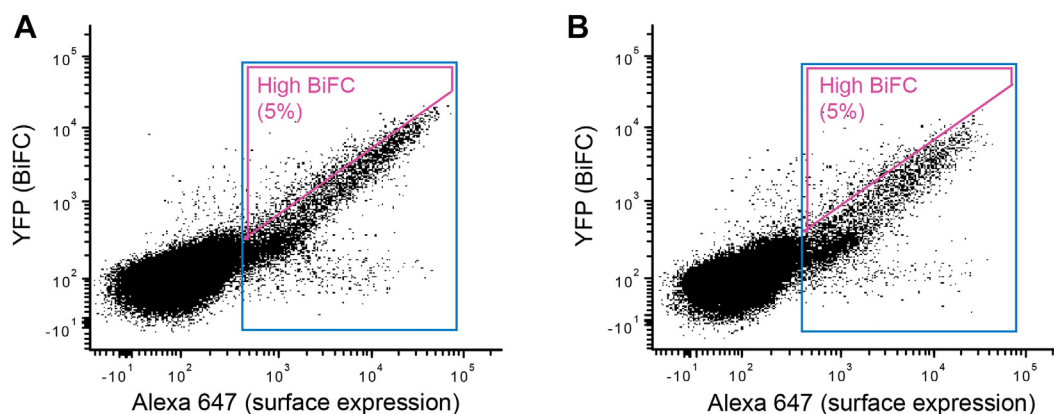
**Figure 1. Deep mutational scans of chemokine receptors based on recognition by conformation-dependent antibodies highlights conserved surfaces facing the membrane.** *A*, using the published deep mutational scan of CXCR4 for binding to monoclonal 12G5, experimental conservation scores are mapped to the CXCR4 surface (PDB 3ODU). Mutational tolerance of residues is shown from low in orange to high in white and yellow. At right, a more conserved surface exposed to the membrane corresponds to the interface for a second CXCR4 subunit in the crystal structure (regions in close contact are shown as green ribbons). *B*, conservation scores from the published deep mutational scan of CCR5 for binding to monoclonal 2D7 are mapped to the CCR5 crystal structure (PDB 4MBS), from conserved residues in red to mutationally tolerant residues in white and yellow. At right, one side of CCR5 exposed to the membrane is more conserved. *C*, cartoon of the CCR5 structure oriented with the conserved surface facing the reader. Residues targeted for mutagenesis are shown as red spheres. *D*, upper plot: CCR5 mutants were expressed on Expi293F cells and the apparent  $K_D$  of 2D7 was measured by flow cytometry. WT CCR5 and two control mutations not expected to impact 2D7 affinity are shown in black and gray, respectively. Mutations in the conserved membrane-exposed surface are shown in red. Lower plot: mutations were introduced into CCR5-VC and CCR5-VN constructs and BiFC measured in transfected Expi293F cells. Various controls are in white (see description in main text). Data are mean  $\pm$  SD,  $n = 3$  independent replicates. BiFC, bimolecular fluorescence complementation.

Sorting experiments were independently replicated. Enrichment ratios weakly agree between the replicate experiments and the agreement is higher when the libraries were sorted for expression only (Figs. 3A and 4A) versus sorting on expression and BiFC signal (Figs. 3B and 4B). Low-frequency mutations in the naïve library had weaker agreement between the replicate experiments, suggesting low-frequency mutations were not deeply sampled during sorting and thus yielded higher variation between replicates. Overall, we consider the “noise” in the data to be similar to earlier deep mutational scans of chemokine receptors (11) but considerably higher than in recent mutational scans of the sweet taste receptor, SARS-CoV-2 receptor ACE2, and serotonin transporter (30–32), which all use equivalent methods. The high noise in this data set may represent difficulties in achieving

consistent sort conditions on BiFC signal. Large variations in mutation frequencies in the naïve library is also expected to have contributed to uneven sampling of mutations during sorting of the transfected cell culture. The enrichment ratios for individual mutations should thus be considered as estimates or predictions until verified by targeted mutagenesis. The  $\log_2$  enrichment ratios for all substitutions at a given residue position may also be averaged to calculate a conservation score. Conservation scores show higher agreement between replicate sorting experiments and are considered a more reliable indicator of whether a particular residue is tolerant of mutations (Figs. 3, C and D and 4, C and D).

There are two key features in the mutational landscapes. First, for CCR5 but less prominently for CXCR4, mutations of TM residues to polar amino acids tend to reduce expression at

## Self-association of chemokine receptors by deep mutagenesis



**Figure 2. FACS collection gates for sorting human cells expressing SSM libraries of chemokine receptors CXCR4 and CCR5.** A and B, X4-KO Epi293F cells stably expressing (A) FLAG-CXCR4-VN or (B) FLAG-CCR5-VN were transfected with SSM libraries of (A) myc-CXCR4-VC or (B) myc-CCR5-VC. After gating the main population of viable cells based on forward-side scatter, cells were gated for surface expression of the VC-fused chemokine receptor mutants by detection of the myc tag (expression gate in blue). Within the expression gate, the top 5% of cells for Venus/YFP fluorescence were gated (high-BiFC gate in magenta). Cells in the expression gate or in the high-BiFC gate were collected in independent sorting experiments. BiFC, bimolecular fluorescence complementation; FACS, fluorescence-activated cell sorting; SSM, single-site saturation mutagenesis; YFP, yellow fluorescent protein.

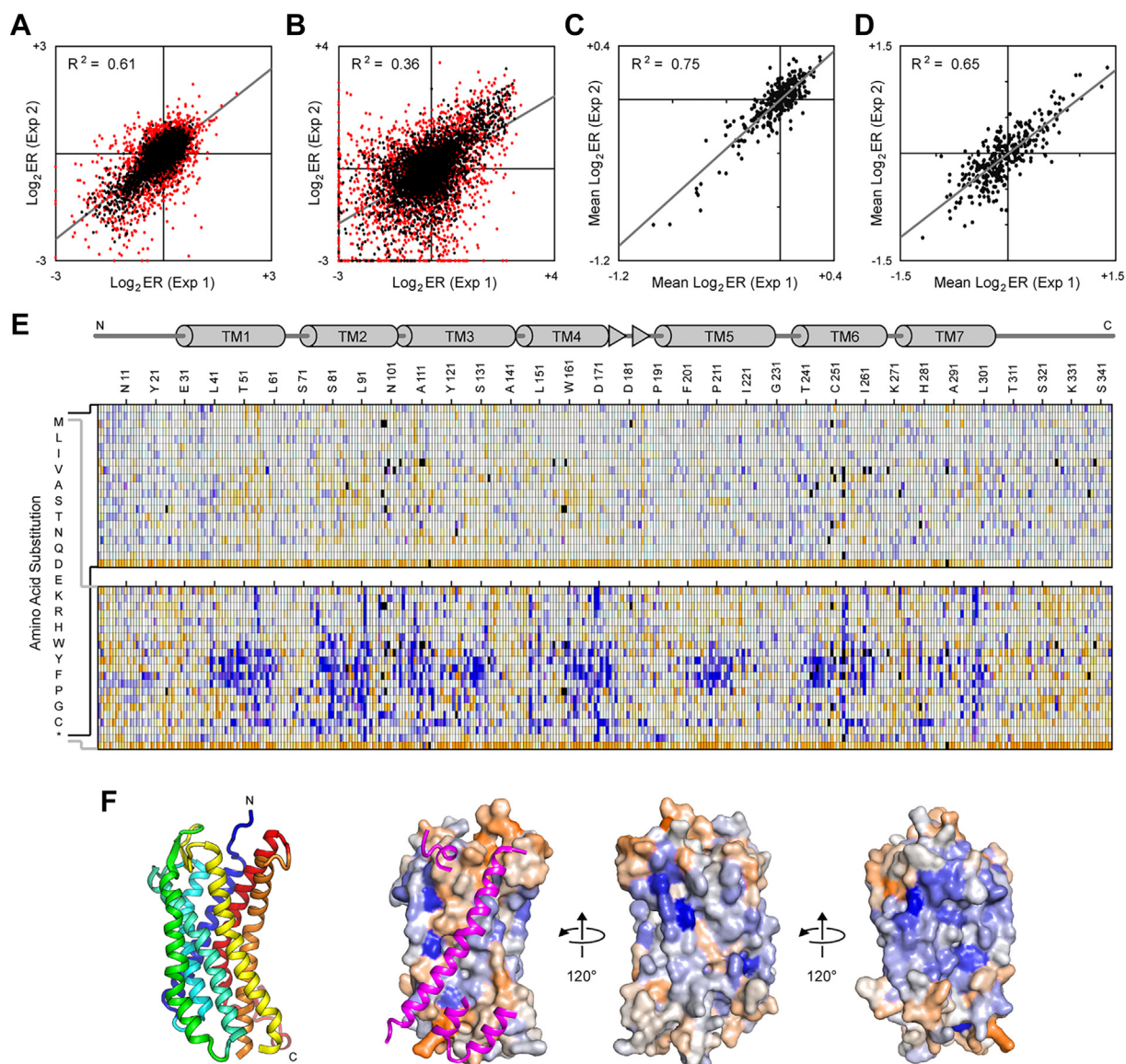
the cell surface (Figs. 3E and 4E). Such mutations are expected to adversely impact folding of the TM domain. This is consistent with previous deep mutational scans of both receptors (11). Second, there are a large number of mutations that are highly enriched for BiFC and these mutations are often found within the TM regions (Figs. 3E and 4E), where many are expected to be highly destabilizing for tertiary structure (for example, introduction of charged and polar residues in TM helices). It is likely that destabilizing mutations enhance BiFC through nonspecific aggregation of mutant proteins. This is consistent with our preliminary data showing that mutations in CCR5 that decreased binding to a conformation-dependent monoclonal antibody (and thus were damaging folded structure) were associated with elevated BiFC. These data indicate the challenges in using mutagenesis to identify dimerization sites in chemokine receptors; a mutation within a dimer interface may disrupt the native dimer but simultaneously increase aggregation and association of misfolded proteins.

To support these hypotheses, we chose 12 mutations for both CXCR4 and CCR5 that were enriched for high-BiFC signal. These mutations substitute amino acids for side chains with very different size and chemical properties. In some cases, they introduce prolines, which often disrupt secondary structure. We validated that the mutations within CXCR4-VC and CCR5-VC increased BiFC signal in the respective CXCR4-VN and CCR5-VN cell lines, above the BiFC signal of WT receptors (Fig. 5, A and C). As a control, the metabotropic glutamate receptor mGluR3, which has no known association with chemokine receptors, produced low-BiFC signals in these cell lines. Cells were gated based on surface levels of the different receptors to partially control for potential differences in expression, although we do not exclude the possibility that increased BiFC signal arises from elevated levels of aggregating protein trapped in intracellular compartments. Furthermore, we generated a stable mGluR3-VN expressing cell line that provides a high-BiFC signal when transfected with mGluR3-VC. mGluR3 is a class C GPCR that forms disulfide bonded dimers. The BiFC signal is low when

the mGluR3-VN line is transfected with WT CXCR4-VC or CCR5-VC. However, the mutant CXCR4 and CCR5 proteins had very high-BiFC signal with mGluR3 (Fig. 5, B and D), again suggesting they are nonspecifically associating with other proteins in the membrane due to destabilized structure.

### Mutations in CXCR4 and CCR5 that reduce receptor associations as measured by BiFC

To understand whether the data provides insights into specific dimer interactions surfaces, the conservation scores from the BiFC selection experiments were mapped to the crystal structures of CXCR4 and CCR5 (Figs. 3F and 4F). CXCR4 has been crystalized as a dimer, with the same dimeric arrangement being resolved in multiple crystal forms with different molecular packing, providing high confidence that the structure represents a relevant complex (21). Residues of CXCR4 known to be buried in the dimeric assembly were weakly more conserved in the BiFC-based selection, with lower mutational tolerance than the other exposed surfaces of the protein (Fig. 3F). These more conserved regions are principally on TM3, TM4, and TM5. We note that these regions approximately correspond to the TM surface that was more highly conserved in the published deep mutational scan of CXCR4 for binding to a conformation-dependent antibody (Fig. 1A). Overall, we conclude that the BiFC-based selection and deep mutational scan succeeded in identifying the known dimer interface of CXCR4. We thus turned our attention to CCR5, where the dimer interface has not been determined by structural methods at atomic resolution. Conservation scores from the BiFC-based deep mutational scan of CCR5 mapped to the protein's monomeric structure indicate that the TM surface formed by TM3-TM4 has reduced mutational tolerance compared to other surfaces (Fig. 4F). The TM3-TM4 surface of CCR5 was also more conserved in the published deep mutational scan of CCR5 for binding to a conformation-dependent antibody (Fig. 1). Accordingly, we hypothesize that this surface is buried in the CCR5 dimer.

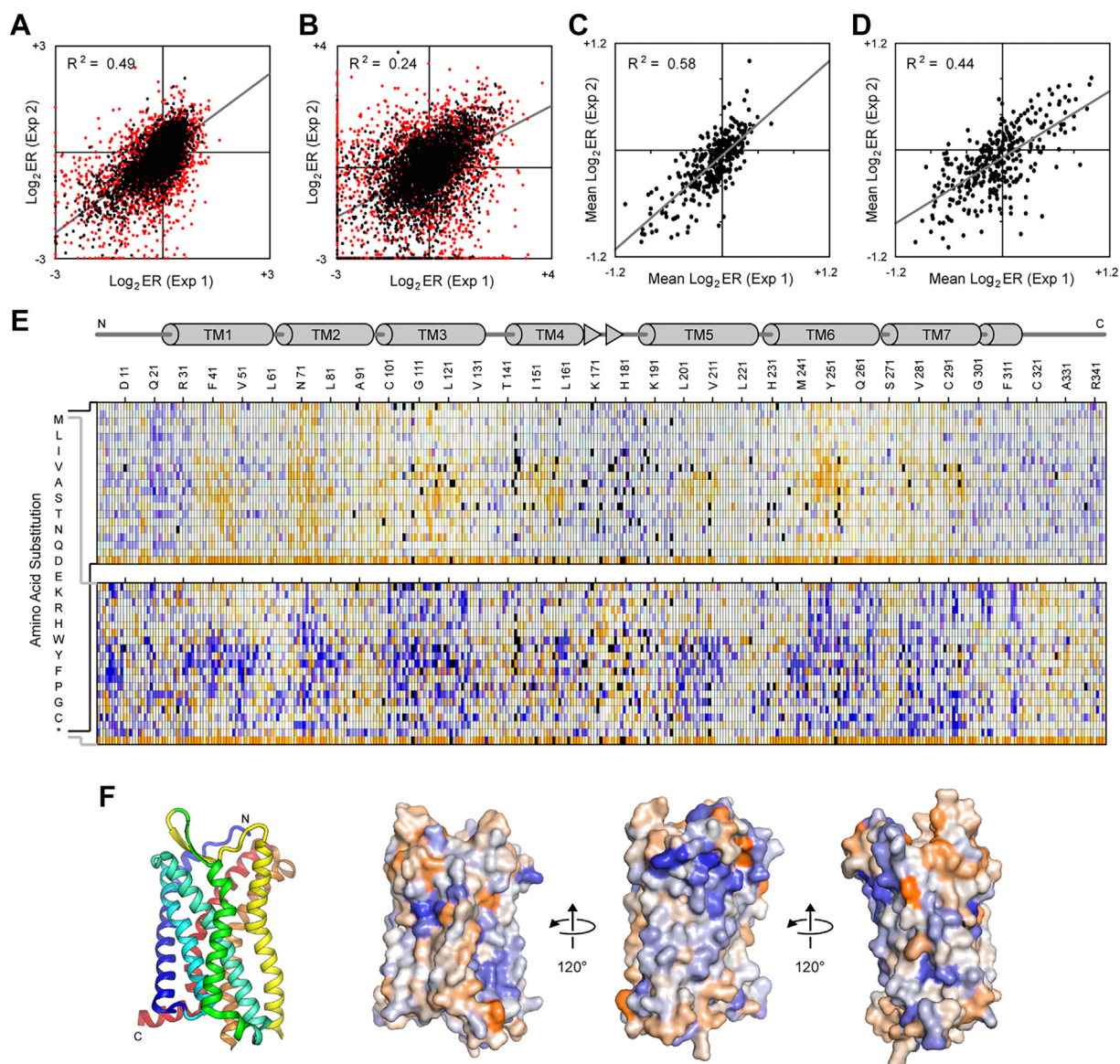


**Figure 3. Deep mutational scan and mutational landscape of CXCR4 based on receptor associations.** *A* and *B*, a single-site saturation mutagenesis (SSM) library of CXCR4-VC was expressed in CXCR4-VN Expi293F cells and sorted by FACS for CXCR4-VC surface expression and BiFC. Log<sub>2</sub> enrichment ratios for mutations were calculated following Illumina sequencing of the naïve and sorted libraries. Plots show agreement between mutation log<sub>2</sub> enrichment ratios after sorting for (*A*) expression or (*B*) expression and BiFC. R-squared values correspond to mutations with frequency  $\geq 5 \times 10^{-5}$  in the naïve library (shown in *black*). Rare mutations with frequency  $< 5 \times 10^{-5}$  in the naïve library are *red*. Mutations with frequency  $< 5 \times 10^{-6}$  were considered absent from the library. *C* and *D*, log<sub>2</sub> enrichment ratios for all amino acid substitutions at a given residue position were averaged to determine a mean conservation score. Plots show agreement of residue conservation scores between two independent sorting experiments for (*C*) expression or (*D*) expression and BiFC. *E*, mutational landscapes of CXCR4-VC sorted for surface expression (*top*) and BiFC (*bottom*). Log<sub>2</sub> enrichment ratios are plotted from depleted/deleterious ( $\leq -3$ , *orange*) to enriched ( $\geq +3$ , *dark blue*). Mutations missing in the naïve library are in *black*. The CXCR4 sequence is on the horizontal axis and amino acid substitutions are on the vertical axis. A schematic of CXCR4 secondary structure is shown at top, with cylinders representing  $\alpha$ -helices and *arrows* representing  $\beta$ -strands. *F*, ribbon structure of CXCR4 (PDB 3ODU) is shown at *left*, viewed from the plane of the membrane and colored *blue* to *red* from N terminus to C terminus. The helices forming the dimer interface are facing out. Adjacent on the *right*, the same orientation of CXCR4 is shown as a surface colored by conservation scores from the BiFC sorting experiment, with conserved residues in orange and residues that are hot spots for enriched mutations in *blue*. A second CXCR4 molecule is shown as *magenta ribbons*; for clarity, only regions of the second CXCR4 that are at the dimer interface are shown. To the *right*, the structure is rotated. BiFC, bimolecular fluorescence complementation; FACS, fluorescence-activated cell sorting.

We further inspected the mutational landscapes for mutations that were highly depleted in the BiFC-based selection and found a small number in the C-terminal cytoplasmic tails of the receptors. GPCRs are known to cluster in lipid microdomains (33, 34), and their clustering would be anticipated to increase BiFC signal. CXCR4 has a number of basic residues at its C terminus and CCR5 has palmitoylation sites (35), both of which are often associated with lipid microdomain localization.

We selected mutations of CXCR4 and CCR5 that were highly depleted, following the BiFC-based selection but were close to neutral in the selection for surface expression; low-BiFC signal was therefore not predicted to be due to decreased expression. We validated five to six mutations in CXCR4-VN and CCR5-VN as decreasing BiFC in the WT CXCR4-VC and CCR5-VC stable cell lines (Fig. 6, *B* and *E*). We considered that decreased BiFC between a mutant and WT receptor might

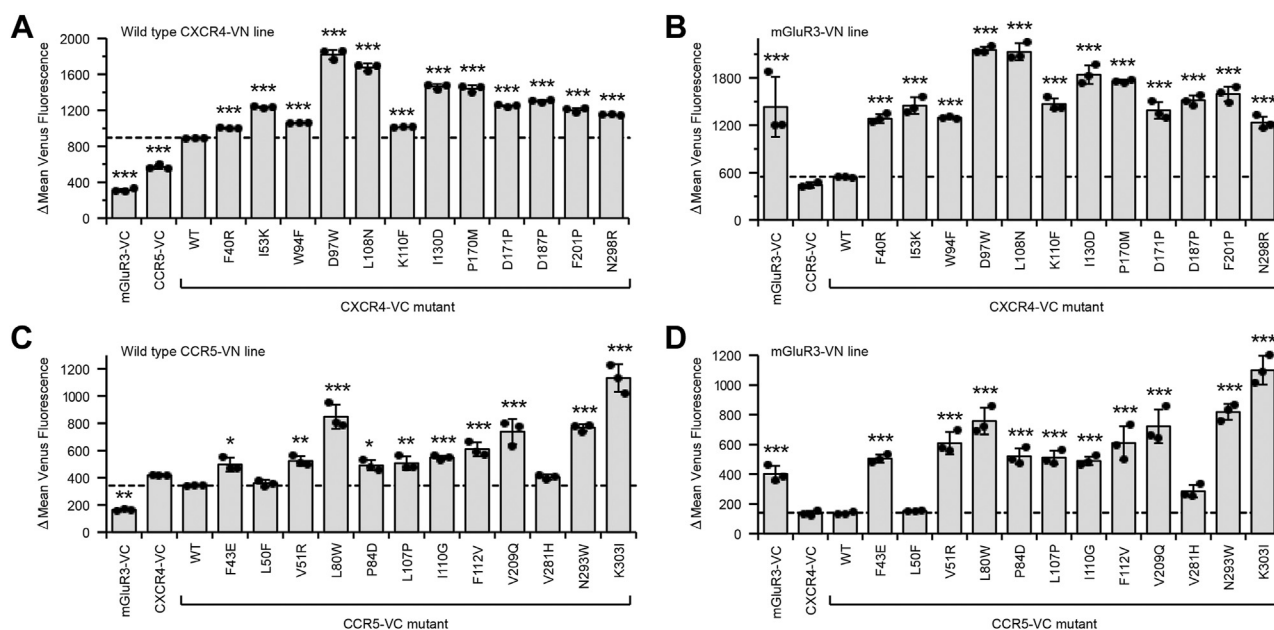
## Self-association of chemokine receptors by deep mutagenesis



**Figure 4. Deep mutational scan and mutational landscape of CCR5 based on receptor associations.** *A* and *B*, a SSM library of CCR5-VC was expressed in CCR5-VN Expi293F cells and sorted by FACS for CCR5-VC surface expression and BiFC. Agreement between  $\log_2$  enrichment ratios for all mutations is shown when the libraries were sorted for (*A*) expression or (*B*) expression and BiFC.  $R$ -squared values correspond to mutations with frequency  $\geq 5 \times 10^{-5}$  in the naïve library (*black*). Rare mutations with frequency  $< 5 \times 10^{-5}$  in the naïve library are *red*. Mutations with frequency  $< 5 \times 10^{-6}$  were considered absent from the library. *C* and *D*, agreement between residue conservation scores from two independent sorting experiments for (*C*) expression or (*D*) expression and BiFC. *E*, mutational landscapes of CCR5-VC sorted for surface expression (*top*) and BiFC (*bottom*).  $\log_2$  enrichment ratios are plotted from depleted/deleterious ( $\leq -3$ , *orange*) to enriched ( $\geq +3$ , *dark blue*). Missing mutations are in *black*. The CCR5 sequence is on the horizontal axis and amino acid substitutions are on the vertical axis. A schematic of CCR5 secondary structure is shown at *top*, with cylinders representing  $\alpha$ -helices and *arrows* representing  $\beta$ -strands. *F*, ribbon structure of CCR5 (PDB 4MBS) is shown at *left*, viewed from the plane of the membrane and colored *blue* to *red* from N terminus to C terminus. The proposed dimer interface of CCR5 is facing out. Adjacent on the *right*, the same orientation of CCR5 is shown as a surface colored by conservation scores from the BiFC sorting experiment, with conserved residues in *orange* and residues that are hot spots for enriched mutations in *blue*. To the *right*, the structure is rotated. BiFC, bimolecular fluorescence complementation; FACS, fluorescence-activated cell sorting; SSM, single-site saturation mutagenesis.

occur if the mutant or WT proteins preferentially self-associated, thus excluding VN- and VC-fused proteins from a dimer complex. We therefore introduced the mutations into VN and VC receptors, which were cotransfected (Fig. 6, C, F and G). In this arrangement, both receptor chains carry the same mutation. This led to the confirmation of four mutants in CXCR4 and two in CCR5 as decreasing self-association of receptor chains (Fig. 6J). In CXCR4, the mutations are P42W, which induces a kink in TM1 and likely has structural consequences;

I204W, which is located on TM5 in the crystallographic dimer interface and is thus predicted to disrupt CXCR4 dimers (Fig. 6A); and R322M and K327W, which change properties in the basic C-terminal tail of CXCR4 and are hypothesized to alter receptor clustering into lipid microdomains. In CCR5, the mutations are V157I, which is exposed to the membrane on TM4 (Fig. 6D) and is predicted to lie within a dimer interface; and C323I, which modifies a palmitoylation site in the C-terminal tail and is predicted to alter clustering into lipid microdomains.



**Figure 5. Validation of mutations in chemokine receptors that increase BiFC.** A and B, mutations of CXCR4 predicted from the deep mutational scan to increase BiFC with WT CXCR4 were validated. The mutations were introduced into CXCR4-VC and transfected into X4-KO Expi293F cell lines stably expressing either (A) CXCR4-VN or (B) mGluR3-VN. C and D, validation of CCR5 mutants enriched in the deep mutational scan for high BiFC signal. The mutations in the CCR5-VC construct were tested in cell lines stably expressing either (C) CCR5-VN or (D) mGluR3-VN. BiFC signal was measured by flow cytometry. Data are mean  $\pm$  SD,  $n = 3$  independent replicates. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; one-way ANOVA with Dunnett test. BiFC, bimolecular fluorescence complementation.

### Effects of CXCR4 and CCR5 mutations

To explore the mechanisms by which these mutations reduce receptor associations, we tested their colocalization with the matrix (MA) protein of HIV-1, which localizes to cholesterol-rich lipid microdomains (36). Expi293F cells that stably express MA fused at the C terminus to VN were transfected with VC-fused CXCR4 and CCR5 mutants (Fig. 6H). We found that mutations in the C-terminal tails of CXCR4 (R322M and K327W) and CCR5 (C323I) substantially reduced BiFC interactions with MA, consistent with these mutations reducing localization into lipid microdomains. By comparison, mutations at the putative dimer interfaces (I204W for CXCR4 and V157I for CCR5) had no impact on clustering with MA. Mutation P42W in CXCR4, which is anticipated to have substantial structural effects by targeting a kink-forming proline in TM1, also reduced association with the MA marker for lipid microdomains, indicating that the C-terminal tails are not solely responsible for membrane microdomain localization.

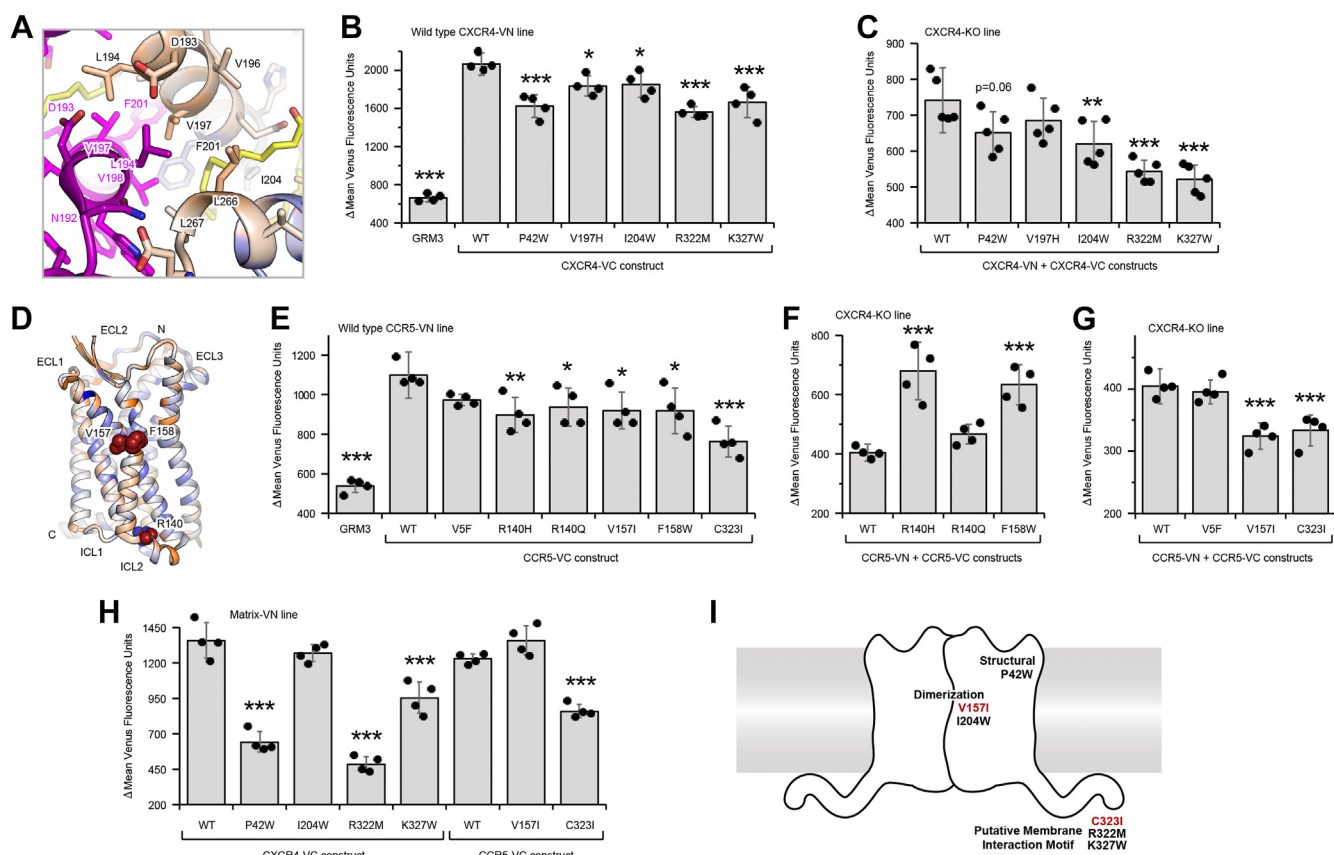
The dynamics of the CXCR4 mutants in the plasma membrane were analyzed by single-molecule total internal reflection (TIR) fluorescence microscopy. Halo-tagged CXCR4 proteins were labeled under conditions that permitted tracking of single-molecule fluorescence under TIR field. In addition to testing the aforementioned mutants of CXCR4, we included D97W as a control, a mutation located at the end of TM2 and oriented toward the ligand-binding cavity, which was shown to increase BiFC signal (Fig. 5A). With the exception of CXCR4-D97W, WT and mutant CXCR4 proteins were similar, with most molecules mobile within a confined space and with diffusion coefficients  $< 0.005 \mu\text{m}^2/\text{s}$ , although a small population of molecules were

more mobile (Fig. 7). Only CXCR4-K327W single molecules had substantially higher mobility based on increased diffusion coefficients. Molecules of the high-BiFC control, CXCR4-D97W, were nearly all immobile (approximately 80%) with low-diffusion coefficients, consistent with our speculation that mutations, which increase BiFC between receptors promote nonspecific aggregation in the membrane.

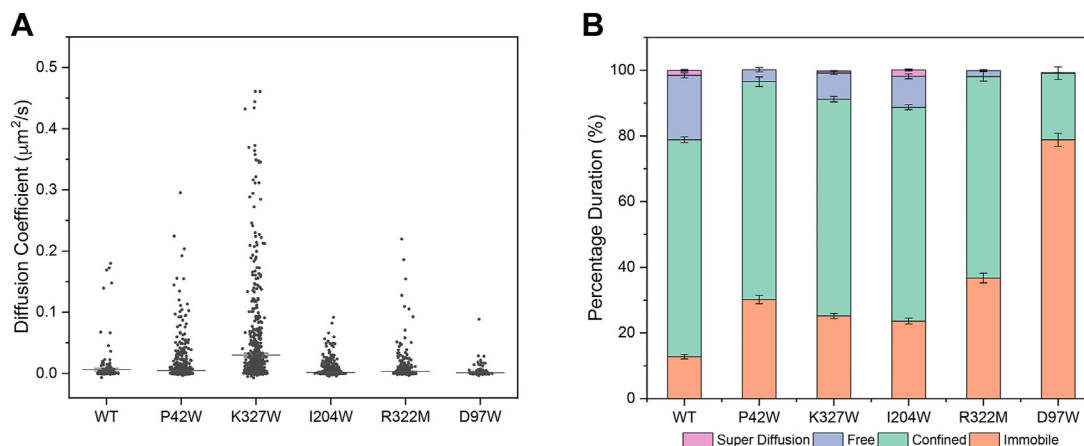
The effects of the mutations on ligand binding and calcium mobilization were measured. Using flow cytometry, binding of sfGFP-tagged CXCL12 significantly increased to the low-BiFC CXCR4 mutants P42W (2.5-fold) and I204W (5-fold) (Fig. 8A), despite no change or slightly reduced receptor expression at the plasma membrane (Fig. 8B). However, increased ligand binding was not associated with higher  $\text{Ca}^{2+}$  signaling but instead  $\text{Ca}^{2+}$  mobilization was decreased (Fig. 8C). The CXCR4-P42W mutant was inactive, consistent with P42 having an important structural role in inducing a kink in TM1, while CXCR4-I204W had reduced signaling. Hence while the mutations favor high ligand-binding states of CXCR4, increased activity for binding CXCL12 does not translate to increased  $\text{Ca}^{2+}$  signaling. The CCR5-V157I and C232I mutants were also trending toward decreased  $\text{Ca}^{2+}$  mobilization relative to WT CCR5 when stimulated with CCL5 (Fig. 8D; statistically significant decreases only at the 10 ng CCL5 concentration). The data suggest that receptor associations and/or localized clustering, either *via* oligomeric assemblies or through localization to lipid microdomains, facilitate  $\text{Ca}^{2+}$  signaling downstream of ligand binding.

CXCR4 and CCR5 are coreceptors with CD4 for HIV-1 entry into a host cell. The viral spike protein Env engages CD4 on the host cell, facilitating conformational changes that

## Self-association of chemokine receptors by deep mutagenesis



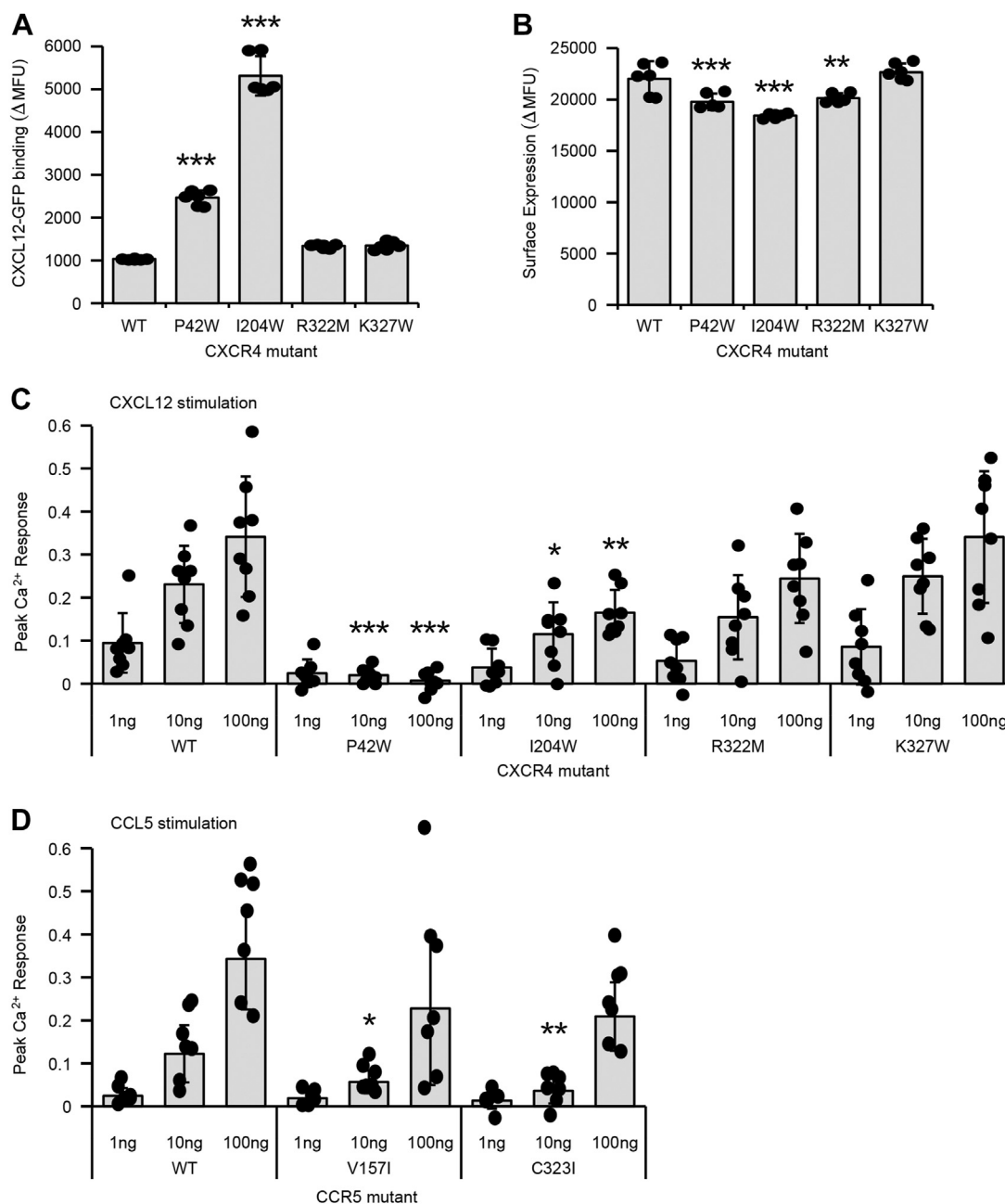
**Figure 6. Mutations in CXCR4 and CCR5 that decrease receptor associations based on BiFC signal.** *A*, dimerization interface of CXCR4 (PDB 30DU) showing one monomer in *magenta* and the other monomer colored by the averaged conservation scores from the BiFC-expression deep mutational scan (low-mutational tolerance in *orange* to high-mutational tolerance in *blue*). V197 and I204 are packed against a lipid (*yellow sticks*) near the interface. *B* and *C*, mutations in CXCR4-VC that were depleted in the BiFC-based deep mutational scan were validated in (*B*) a X4-KO Expi293F cell line stably expressing WT CXCR4-VN or (*C*) in X4-KO Expi293F cells in which mutant CXCR4-VC and CXCR4-VN constructs were transiently transfected. *D*, ribbon representation of CCR5 (PDB 4MB5) with conservation scores from the BiFC-expression mutational scan mapped to the structure. Residues of low-mutational tolerance are *orange*; residues of high-mutational tolerance are *blue*. A membrane-exposed surface patch that is less tolerant of mutations and is the proposed dimerization interface is oriented toward the reader, with mutations depleted in the mutational scan in *red spheres*. *E* and *F*, mutations in CCR5-VC that were depleted in the BiFC-based scan were transfected in (*E*) X4-KO Expi293F cells stably expressing WT CCR5-VN and BiFC measured by flow cytometry. The mutations were also tested in (*F* and *G*) X4-KO Expi293F cells in which both mutant CCR5-VC and CCR5-VN constructs were cotransfected. Mutations of CCR5 that increase BiFC when present in both receptor chains are in *panel F*, while mutations that decrease BiFC are in *panel G*. *H*, mutations in CXCR4-VC and CCR5-VC that reduced BiFC in *panels C* and *G* were transfected into a stable MA-VN Expi293F line as a marker for lipid microdomains and BiFC was measured. *I*, cartoon of a chemokine receptor dimer annotated with mutations that diminish BiFC between homodimers of CXCR4 (*black labels*) and CCR5 (*red labels*). In all data panels, plotted are means  $\pm$  SD,  $n = 4$  to 5 biological independent replicates. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; one-way ANOVA with Dunnett test. BiFC, bimolecular fluorescence complementation.



**Figure 7. Single-molecule imaging of CXCR4.** *A*, single-molecule imaging analysis of CXCR4 and its mutants. *A*, distribution of diffusion coefficients derived from DC-MSS analysis (54) for myc-Halo-CXCR4, comparing reduced self-association mutants (P42W, K327W, I204W, and R322M) to WT and one increased self-association mutant (D97W). *B*, DC-MSS analysis for the tracked trajectories of myc-Halo-CXCR4, dissecting transient motion in the tracks. DC-MSS, Divide and Conquer Moment Scaling spectrum transient diffusion analysis.



## Self-association of chemokine receptors by deep mutagenesis



**Figure 8. Ligand binding and signaling of chemokine receptor mutants with reduced self-association.** A, X4-KO Expi293F cells expressing CXCR4 mutants were incubated with 10  $\mu$ M CXCL12-sfGFP and bound ligand was detected by flow cytometry. B, surface expression of CXCR4 mutants measured by flow cytometry. In panels A and B, data are mean  $\pm$  SD, n = 6 biological independent replicates. \* $p$  < 0.05; \*\* $p$  < 0.01; \*\*\* $p$  < 0.001; one-way ANOVA with Dunnett test. C, calcium mobilization was recorded in Fluo-4-AM loaded X4-KO Expi293F cells transfected with the indicated CXCR4 mutants. Calcium mobilization at 1, 10, and 100 ng/ml CXCL12 concentrations are reported relative to maximum Ca<sup>2+</sup> response to 4  $\mu$ M ionomycin. Data are mean  $\pm$  SD, n = 8 biological independent replicates. D, as in panel C, with cells now expressing CCR5 proteins and stimulated with CCL5. Data are mean  $\pm$  SD, n = 6 to 7 biological independent replicates.

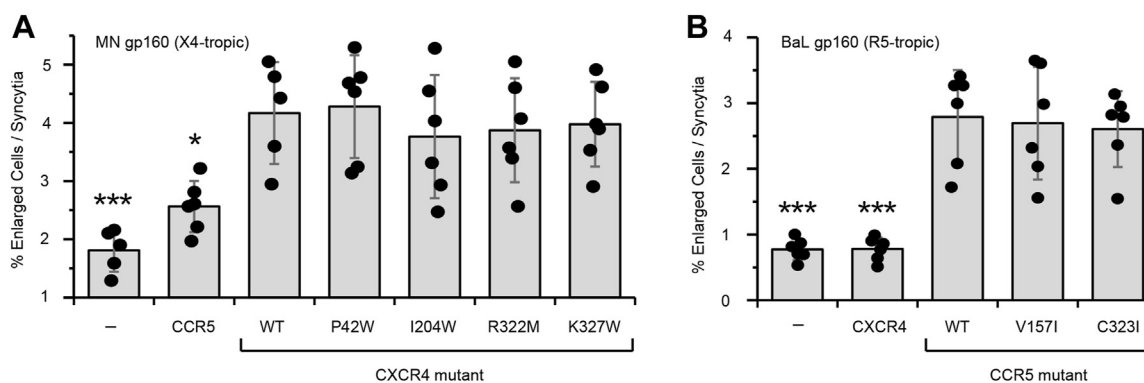
expose binding sites for CXCR4 (X4-tropic viruses) or CCR5 (R5-tropic viruses) (37). We tested whether the low-BiFC receptor mutations altered syncytia formation when Env-expressing cells were cultured with receptor-expressing cells. Cells expressing CD4 together with WT or mutant CXCR4 had no significant differences in syncytia formation when incubated with cells expressing Env from the X4-tropic MN strain (Fig. 9A), and likewise cells expressing CD4 together with WT or mutant CCR5 had no significant differences in

syncytia formation, when incubated with cells expressing Env from the R5-tropic BaL strain (16) (Fig. 9B).

### Discussion

Extensive research efforts have found that chemokine receptors CXCR4 and CCR5 homodimerize and heterodimerize (6, 20, 21, 27, 38, 39), a feature that is common to many other class A GPCRs (4, 9, 10). However, the organization of

## Self-association of chemokine receptors by deep mutagenesis



**Figure 9. HIV-1 Env-mediated syncytia formation is not affected by mutations in chemokine receptors that reduce self-associations.** Formation of syncytia was measured by flow cytometry after coculturing X4-KO Epi293F cells expressing either (A) MN gp160 or (B) BaL gp160 with cells expressing CD4 and the indicated CXCR4 or CCR5 mutants, respectively. Data are mean  $\pm$  SD,  $n = 6$  independent experiments. \* $p < 0.05$ ; \*\* $p < 0.01$ ; \*\*\* $p < 0.001$ ; one-way ANOVA with Dunnett test.

chemokine receptor dimers and their functional importance remain incompletely understood, in part because identifying mutations that alter dimer strength has been difficult. Here, we set out to define dimer interfaces and find mutations to modulate receptor associations through a BiFC-based deep mutational scan. The use of deep mutagenesis to identify dimerization interfaces was previously applied to the sweet taste receptor and informed the modeling of that receptor's dimeric architecture, in close agreement with a cryo-EM structure of a related class C GPCR (32). However, the mutational landscapes we report for CXCR4 and CCR5 unexpectedly revealed many hundreds of mutations that increased receptor associations based on increased BiFC signal. We show that these mutations not only increased BiFC between chemokine receptor chains but also with an unrelated class C GPCR, suggesting the mutations cause nonspecific associations and aggregation. This is supported by the observation that the mutations frequently altered chemical properties of amino acids within TM helices and may thus be deleterious for proper folding. Single-molecule imaging of CXCR4 with a high-BiFC mutation showed the mutant receptors were mostly immobile in the membrane.

The ease with which CCR5 and CXCR4 could be driven toward aggregation means overexpression studies of receptors must be approached with caution, lest nonspecific aggregation is falsely interpreted as biologically relevant dimer formation. Our data do not reveal the extent to which WT receptors might aggregate (although the  $\sim 10\%$  of WT CXCR4 observed to be immobile by single-molecule fluorescence microscopy may represent an aggregated population; see Fig. 7B), but with so many mutations increasing BiFC between chemokine receptors, we hypothesize that even overexpressed WT receptors will have an aggregated fraction or at a minimum are on the "precipice" of aggregating. GPCR oligomerization is generally demonstrated by overexpression of receptors that are fused with epitope tags, fluorescent proteins, or luciferase among other examples. Our data emphasizes the importance of controls to show that GPCR oligomers are specific and occur even at low expression levels.

Coimmunoprecipitation experiments, in which samples are rigorously washed prior to analysis, may be especially sensitive to false positives from aggregated receptor populations that remain tightly bound. By comparison, methods such as single-molecule fluorescence microscopy may be better suited for showing physiological associations at relevant receptor densities without being misled by aggregates; such studies have shown that representative GPCRs form transient dimers on the second timescale (40–42), although alternative methods have suggested GPCR dimers are more stable (43). Additional evidence based on functional properties, such as cointernalization (44) and cooperativity (45), would also provide strong evidence for meaningful receptor associations beyond nonspecific aggregation.

We compared mutations of human CXCR4 and CCR5 that are predicted to increase BiFC to genetic variants found in the gnomAD database (46). Six percent of missense mutations to CXCR4 are strongly predicted to increase aggregation based on enrichment in the BiFC-based selection (using as a criterion that the  $\log_2$  enrichment ratios  $>1$  in both independent replicates). However, only three of 129 missense variants observed in humans are strongly predicted to increase BiFC (CXCR4 mutants N56S, M205K, and P299L) and these variants are all exceedingly rare (single allele counts, allele frequencies  $\sim 4 \times 10^{-6}$ ). Likewise, 5% of missense mutations to CCR5 are strongly predicted to be aggregating based on enrichment in the BiFC-based selection, and yet only two of the predicted aggregation mutants are found among the 253 missense variants in the gnomAD database (CCR5 mutants A73D and L257Q). Again, the mutations are rare (allele counts of 2 and 1, allele frequencies  $\sim 8 \times 10^{-6}$  and  $\sim 4 \times 10^{-6}$ , respectively), despite CCR5 being a nonessential gene. We hypothesize that aggregating mutations are deleterious and thus rare in the human population. A detailed comparison of the deep mutational scanning results to human variant databases will require higher quality experimental data.

Despite the high prevalence of BiFC-enhancing mutations in the scans, there were surfaces on CXCR4 and CCR5 where mutations tended to be depleted following BiFC-based sorting.

We consider these relatively more conserved surfaces in the sorting experiments to be potential interfaces for dimer formation. CXCR4 has been crystallized as a symmetric dimer with intersubunit contacts formed primarily by TM3-5 (21). The TM3-5 surface exposed to the membrane is less tolerant of mutations in the BiFC-based deep mutational scan, and our data therefore support this dimeric arrangement of CXCR4 in living cells. Substitution I204W in TM5 caused a partial decrease in CXCR4 homodimerization based on BiFC. Cells expressing CXCR4-I204W had 5-fold enhanced binding to CXCL12, consistent with CXCL12/CXCR4 forming a 1:1 complex. However, Ca<sup>2+</sup> signaling by the mutant receptor was not increased at any of the CXCL12 concentrations tested, although we did not determine whether other signaling pathways or ultimately chemotaxis were affected. We also note that the data do not answer whether these changes in receptor properties were due to reduced dimerization or other structural perturbations caused by the mutation.

The structure of a CCR5 dimer has not been determined experimentally at atomic resolution, although chemical cross-linking has supported close contacts between TM5 helices across the interface (25). There is no surface on CCR5 that unequivocally stands out as the putative dimer interface in our BiFC-based deep mutational scan. Our data is most supportive of TM4 forming the CCR5 dimer interface, as the membrane-exposed surface of this helix was relatively more conserved than other TM regions, albeit weakly. It is difficult to compare our findings to the published cross-linking study of CCR5 that emphasized TM5, as cross-linking mutations were not explored in TM4. The substitution we confirmed as decreasing BiFC between CCR5 receptors, V157I, is located on TM4 and points outward toward the lipid phase of the membrane, appropriately situated to participate in a potential dimer interface.

We also found mutations in the cytoplasmic tails of CXCR4 and CCR5 that decreased BiFC. The cytoplasmic tails are known to contain features that promote lipid microdomain localization (35, 47–49) and decreased BiFC is readily explained by reduced clustering of receptors in membrane microdomains. In agreement, mutations in the C-terminal tails also reduced colocalization with MA, a marker for cholesterol-rich microdomains in cellular membranes.

Overall, deep mutational scanning of the chemokine receptors CXCR4 and CCR5 through the use of BiFC demonstrated that these receptors may associate through multiple mechanisms: (1) nonspecific associations or aggregation (2), dimerization, and (3) clustering into lipid microdomains. These findings have implications for studies into chemokine receptor associations, as results that are often interpreted as receptor dimerization or oligomerization may instead arise due to other mechanisms.

## Experimental procedures

### Cell culture

The X4-KO Expi293F and the MA-VN expressing X4-KO Expi293F cell lines are previously described (11, 50). Cells

(Thermo Fisher Scientific) were cultured in Expi293 Expression Media (Thermo Fisher Scientific) at 37 °C, 8% CO<sub>2</sub>, and 125 rpm. Cells were transfected at a density of 2 × 10<sup>6</sup> cells/ml with 500 ng plasmid DNA per ml of culture using ExpiFectamine (Thermo Fisher Scientific) according to the manufacturer's directions. To generate stable lines for BiFC, X4-KO Expi293F cells were transfected with linearized pCEP4-FLAG-CXCR4-VN or pCEP4-FLAG-CCR5-VN and selected with hygromycin B (200 µg/ml). The population of CXCR4-VN- or CCR5-VN-positive cells was enriched by FACS in which cells were washed in Dulbecco's PBS supplemented with 0.2% bovine serum albumin (BSA), stained with 1:200 chicken anti-FLAG-FITC (Immunology Consultants Laboratory) in PBS-BSA, washed, and resuspended for sorting on a BA FACS Aria II. The collected cells were expanded by culturing in Expi293 Expression Medium supplemented with penicillin-streptomycin.

### Plasmids

Plasmids for myc-tagged CCR5 and CXCR4, used in signaling and syncytia assays, are previously described (11) and available on Addgene (# 98948 and # 98946). Plasmids for myc-CXCR4-VC (Addgene # 98967), myc-CCR5-VC (# 98966), myc-GRM3-VC (# 98968), FLAG-CXCR4-VN (# 98964), FLAG-CCR5-VN (# 98963), and FLAG-GRM3-VN (# 98965) are previously described (11). Briefly, the N-terminal half (VN) of the yellow fluorescent protein Venus (a.a. 1–154; mutant I152L) or the C-terminal half (VC) of Venus (a.a. 155–238) were fused to the C-termini of tagged receptors and cloned into the NheI-XhoI sites of pCEP4 (Invitrogen). Plasmids for HIV-1 MN gp160 and BaL gp160 (# 100919) are previously described (11). pCMV3-CD4 was from Sino Biological (# HG10400-UT). For single-molecule imaging, a halo tag sequence was inserted into the myc-CXCR4 WT and low BiFC mutant sequences, in between the myc tag and the chemokine receptor sequence.

### Library generation

Single-site saturation mutagenesis libraries of human myc-tagged CXCR4 and CCR5 were generated previously (11). Here, the libraries were modified by using PCR-based fragment assembly to fuse the 3' ends of mutated chemokine receptor sequences to VC (a.a. 155–238 of Venus). The PCR products were digested and ligated into the NheI-XhoI sites of pCEP4 and electroporated into *Escherichia coli* DH5α cells. The number of transformants was >100× the theoretical sequence diversity. The myc-CXCR4-VC and myc-CCR5-VC plasmid libraries were transfected into stable FLAG-CXCR4-VN or FLAG-CCR5-VN Expi293F lines under conditions, where cells typically express no more than one mutant VC-fused chemokine receptor gene per cell: 1 ml culture at a density of 2 × 10<sup>6</sup> cells was transfected with 1 ng library DNA diluted with 1.5 µg pCEP4ΔCMV carrier DNA using ExpiFectamine (Thermo Fisher Scientific). The media was replaced 2 h posttransfection and cells were harvested for sorting 24 h posttransfection.

## Self-association of chemokine receptors by deep mutagenesis

### Sorting cells for surface expression and high-BiFC signal

Cells transfected with the libraries were washed with cold PBS-BSA and incubated with 1/300 anti-myc Alexa 647 (clone 9B11; Cell Signaling Technology) plus 1/300 anti-Flag Cy3 (clone M2; Sigma-Aldrich) in PBS-BSA for 30 min before being washed twice and resuspended in cold PBS-BSA supplemented with 1/100 fetal bovine serum. Cells were sorted on a BD FACS Aria II at the Roy J. Carver Biotechnology Center. Cells were first gated by forward-side scatter and side scatter for the main population of cells. For sorting based on surface expression, cells positive for the myc tag (Alexa 647 fluorescence) were collected. For sorting based on high-BiFC signal, within the expression gate, the top 5% of cells for YFP/Venus fluorescence relative to surface receptor expression were collected. Cells were sorted for no longer than 4 h to maintain high viability and were collected in tubes coated overnight in fetal bovine serum. Collected cells were centrifuged and the pellets stored at  $-80^{\circ}\text{C}$ .

### Illumina sequencing and analysis

RNA was extracted from frozen cells from FACS using the GeneJET RNA Purification Kit (Thermo Fisher Scientific) and reverse transcribed with Accuscript primed with the EBV reverse sequencing primer for first strand complementary DNA synthesis. The receptor complementary DNA was PCR amplified as three overlapping fragments to achieve full coverage of the gene. In a second round of PCR, experiment-specific barcodes and adaptors for annealing to the Illumina flow cell were added. Samples were sequenced ( $2 \times 250$  nt) using Illumina HiSeq 2500 at the UIUC Roy J. Carver Biotechnology Center. Data were analyzed using Enrich (<https://github.com/FowlerLab/Enrich2>) (51).  $\log_2$  enrichment ratios of mutants were adjusted by subtracting the enrichment of the WT sequence. Data and analysis commands are deposited with the National Center for Biotechnology Information's Gene Expression Omnibus under series accession number GSE125426.

### Validation of receptor mutants with changes in BiFC

pCEP4 plasmids containing WT VN- and VC-fused receptors were used as templates for targeted mutagenesis by overlap extension PCR. Mutated inserts in all plasmids were confirmed by Sanger sequencing. Plasmids were transiently transfected as described above into the relevant cell line and analyzed by flow cytometry for receptor expression and BiFC signal 24 h posttransfection. Cells were washed with PBS-BSA, incubated in 1:100 anti-myc Alexa 647 and 1:100 anti-FLAG Cy3 in PBS-BSA for 20 min, washed twice, and resuspended in PBS-BSA for analysis on a BD LSR II. Data were collected using instrument software and analyzed with FCS Express (*De Novo Software*). Cells were gated by forward-side scatter for the main population and then gated by Alexa 647/Cy3 fluorescence to control for a consistent level of surface expressed receptor across the samples. Mean YFP/Venus fluorescence was recorded.

### Colocalization with MA

Using a MA-VN cell line (50), cells were transiently transfected as described above with WT or mutant myc-tagged receptors fused to VC and cloned into pCEP4. Cells were processed 24 h posttransfection. Cells were washed PBS-BSA and incubated in a 1:200 dilution of anti-myc Alexa 647 (clone 9B11; Cell Signaling Technology). Cells were washed and analyzed on a BD LSR II flow cytometer.

### CXCL12 binding assay

Transfected X4-KO Expi293F cells were harvested 24 h posttransfection, washed with PBS-BSA, and incubated in 1:200 anti-myc-Alexa 647 (clone 9B11; Cell Signaling Technology) and 10  $\mu\text{M}$  CXCL12-sfGFP for 30 min. Cells were washed and analyzed on a BD LSR II. The preparation of CXCL12-sfGFP is described elsewhere (Heredia *et al.*, 2018).

### Calcium mobilization

X4-KO Expi293F cells were transiently transfected as described above with myc-tagged CXCR4 or CCR5 plasmids. Cells were prepared at room temperature 24 h posttransfection. Cells were washed with assay buffer (PBS containing 0.2% BSA and 1 mM  $\text{CaCl}_2$ ) and resuspended in assay buffer with 2  $\mu\text{M}$  Fluo-4-AM (Life Technologies). Fluo-4-AM was prepared as a 1 mM stock in dimethyl sulfoxide. Cells were incubated with 1:250 anti-myc-Alexa-647 (clone 9B11; Cell Signaling Technology) for 30 min with frequent mixing and washed and resuspended in assay buffer. Cells were analyzed on a BD Accuri C6 Cytometer. Cells were gated by forward-side scatter for the main population and the Fluo-4 fluorescence of the Alexa 647 positive population was monitored over time. To cells at baseline, CXCL12 (PeproTech) or CCL5 (PeproTech) were added at final concentrations of 1, 10, and 100 ng/ml. Fluorescence returned to baseline within 120 s, at which point ionomycin was added to a final concentration of 4  $\mu\text{M}$ . The spike in  $\text{Ca}^{2+}$ -dependent Fluo-4 fluorescence in response to the chemokine was normalized to the maximal response induced by ionomycin.

### Syncytia formation assay

X4-KO Expi293F cells ( $2 \times 10^6$  cells in 1 ml) were transfected with 50 ng pCMV3-CD4 and 450 ng pCEP4-myc-CXCR4 or pCEP4-myc-CCR5. A separate set of cells ( $2 \times 10^6$  cells in 1 ml) were transfected with 500 ng pCEP4-gp160 from the MN or BaL HIV-1 strains. Empty vector was used for control transfections. After 5 h incubation at  $37^{\circ}\text{C}$ , 8%  $\text{CO}_2$ , 125 rpm,  $0.2 \times 10^6$  receptor-expressing cells, and gp160-expressing cells were mixed and added to 0.6 ml Expi293 Expression Medium. The cells were added to wells of a 12-well tray that had been incubated for 30 min with 0.01% poly-L-lysine (Sigma) to coat the plastic surface. Cells were incubated without agitation for 20 h. Wells were washed with warm PBS and attached cells removed with 0.25% trypsin-2.21 mM EDTA for 15 min at  $37^{\circ}\text{C}$ . Detached cells were washed with cold PBS-BSA and analyzed on a BD LSR II flow cytometer to detect syncytia

based on high forward-side scattering. The positive gate was set at less than 1% of negative control cells.

### Single molecule imaging of CXCR4

Coverslips measuring 24 mm × 40 mm (VWR Cat. No. 48393230) were cleaned for single-molecule experiments as described previously (52). The washed coverslips were stored in molecular grade water, air-dried, and flamed for a few seconds. Coverslips were then immersed in a clean Petri dish with 50 µg/ml poly-L-lysine (Cat. No. P1274; Sigma-Aldrich) and incubated overnight. Coverslips were rinsed twice in molecular grade water and air dried. A large polydimethylsiloxane chamber was carefully assembled over a dried coverslip and placed in a 60-mm dish.

X4-KO Expi293F cells were transfected with wildtype or mutant myc-halo-CXCR4. Two hours posttransfection, cells were diluted 1:10 in Expi293 Expression Media and placed on the assembled polydimethylsiloxane coverslip chamber. Twenty-four hours posttransfection, cells were gently washed on the coverslip with PBS containing calcium and magnesium (Cat. No. 21-030CM; Corning) and incubated in 2 nM JF549-Halo-ligand (Cat. No. GA1110; Promega) for 15 min at 37 °C. Cells were washed three more times and incubated in PBS for another 15 min before four more washes and covered in PBS.

Single-molecule tracking was performed at room temperature in a custom-built TIR fluorescence microscope. A 100× oil immersion objective (100×, N.A. 1.49, oil immersion) was assembled on an inverted microscope. A 561 nm laser was used to excite the labeled molecules. Power of the laser was controlled using neutral density filters. Individual labeled molecules were seen as diffraction-limited spots on the cell surface in immediate contact with the coverslip. The fluorescence from the single spots was collected by the same objective, passing an emitter and captured by an Electron Multiplying Charge-Coupled Device camera. A total of 2400 frames/trajectory were acquired for each field-of-view with an integration time of 50 ms. The collected data was exported to FIJI (<https://imagej.net/software/fiji/>) and single spots were tracked using the plugin TrackMate (53). Individual spots were selected as 3×3 pixel bright features. Threshold was selected such that all the single spots were selected in a frame. For track generation, the linking distance was fixed to 2 pixels, merging was not allowed, gap closing distance was set as 2 pixels, and the Max frame gap was two frames. Data files were exported to MATLAB 2016b (MATLAB Release 2016b, The MathWorks, Inc). For Divide and Conquer Moment Scaling spectrum transient diffusion analysis, we used the MATLAB-based software (54). All the tracks were segregated into super diffusion, free, confined, and immobile motions using Divide and Conquer Moment Scaling spectrum transient diffusion analysis.

### Data availability statement

Illumina sequencing data are deposited with the National Center for Biotechnology Information's Gene Expression Omnibus under series accession number GSE125426. All

other data that support the findings are available from the corresponding author upon request.

**Acknowledgments**—The UIUC Roy J. Carver Biotechnology Center assisted with deep sequencing and cell sorting. The work described in this manuscript was conducted at the University of Illinois. Cyrus Biotechnology had no role in the project's design, execution, and analysis/interpretation. The views expressed in the manuscript are solely those of the authors, and they do not represent official views or opinions of the University of Illinois or Cyrus Biotechnology.

**Author contributions**—K. S. G. and E. P. conceptualization; K. S. G., K. M., V. V. K., K. Z., and E. P. methodology; K. S. G. and K. M. investigation; K. S. G. and K. M. writing-original draft; K. S. G., K. M., and E. P. formal analysis; J. D. H. resources; K. Z. and E. P. supervision; E. P. funding acquisition; E. P. writing-review and editing.

**Funding and additional information**—Research was supported by award number R01AI129719 from the National Institute of Allergy and Infectious Diseases of the National Institutes of Health to E. P. and by award number R01GM132438 from the National Institute of General Medical Sciences to K.Z. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

**Conflict of interest**—E. P. is an employee and shareholder of Cyrus Biotechnology, Inc. The other authors declare that they have no conflicts of interest with the contents of this article.

**Abbreviations**—The abbreviations used are: BiFC, bimolecular fluorescence complementation; BSA, bovine serum albumin; FACS, fluorescence-activated cell sorting; GPCR, G-protein-coupled receptor; MA, matrix; SSM, single-site saturation mutagenesis; TIR, total internal reflection; TM, transmembrane; YFP, yellow fluorescent protein.

### References

- Kufareva, I., Stephens, B. S., Holden, L. G., Qin, L., Zhao, C., Kawamura, T., *et al.* (2014) Stoichiometry and geometry of the CXC Chemokine receptor 4 complex with CXC ligand 12: molecular modeling and experimental validation. *Proc. Natl. Acad. Sci. U. S. A.* **111**, E5363–E5372
- Hebert, T. E., Moffett, S., Morello, J. P., Loisel, T. P., Bichet, D. G., Barret, C., *et al.* (1996) A peptide derived from a  $\beta$  2-adrenergic receptor transmembrane domain inhibits both receptor dimerization and activation. *J. Biol. Chem.* **271**, 16384–16392
- Albizu, L., Cottet, M., Kralikova, M., Stoev, S., Seyer, R., Brabet, I., *et al.* (2010) Time-resolved FRET between GPCR ligands reveals oligomers in native tissues. *Nat. Chem. Biol.* **6**, 587–594
- Kobayashi, D., Endo, M., Ochi, H., Hojo, H., Miyasaka, M., and Hayasaka, H. (2017) Regulation of CCR7-dependent cell migration through CCR7 homodimer formation. *Sci. Rep.* **7**, 1–14
- Drurya, L. J., Ziarek, J. J., Gravel, S., Veldkamp, C. T., Takekoshi, T., Hwang, S. T., *et al.* (2011) Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 17655–17660
- El-Asmar, L., Springael, J. Y., Ballet, S., Andrieu, E. U., Vassart, G., and Parmentier, M. (2005) Evidence for negative binding cooperativity within CCR5-CCR2b Heterodimers. *Mol. Pharmacol.* **67**, 460–469
- Jones, K. A., Borowsky, B., Tamm, J. A., Craig, D. A., Durkin, M. M., Dai, M., *et al.* (1998) GABAB receptors function as a heteromeric assembly of the subunits GABABR1 and GABABR2. *Nature* **396**, 674–679

## Self-association of chemokine receptors by deep mutagenesis

- El Moustaine, D., Granier, S., Doumazane, E., Scholler, P., Rahmeh, R., Bron, P., *et al.* (2012) Distinct roles of metabotropic glutamate receptor dimerization in agonist activation and G-protein coupling. *Proc. Natl. Acad. Sci. U. S. A.* **109**, 16342–16347
- Zhao, D. Y., Pöge, M., Morizumi, T., Gulati, S., Van Eps, N., Zhang, J., *et al.* (2019) Cryo-EM structure of the native rhodopsin dimer in nanodiscs. *J. Biol. Chem.* **294**, 14215–14230
- Manglik, A., Kruse, A. C., Kobilka, T. S., Thian, F. S., Mathiesen, J. M., Sunahara, R. K., *et al.* (2012) Crystal structure of the  $\mu$ -opioid receptor bound to a morphinan antagonist. *Nature* **485**, 321–326
- Heredia, J. D., Park, J., Brubaker, R. J., Szymanski, S. K., Gill, K. S., and Procko, E. (2018) Mapping interaction sites on human chemokine receptors by deep mutational scanning. *J. Immunol.* **200**, 3825–3839
- Gassenmaier, M., Chen, D., Buchner, A., Henkel, L., Schiemann, M., Mack, B., *et al.* (2013) CXCR4 chemokine receptor 4 is essential for maintenance of renal cell carcinoma-initiating cells and predicts metastasis. *Stem Cells* **31**, 1467–1476
- Müller, A., Homey, B., Soto, H., Ge, N., Catron, D., Buchanan, M. E., *et al.* (2001) Involvement of chemokine receptors in breast cancer metastasis. *Nature* **410**, 50–56
- Connor, R. I., Sheridan, K. E., Ceradini, D., Choe, S., and Landau, N. R. (1997) Change in coreceptor use correlates with disease progression in HIV-1–infected individuals. *J. Exp. Med.* **185**, 621–628
- Deng, H. K., Liu, R., Ellmeier, W., Choe, S., Unutmaz, D., Burkhart, M., *et al.* (1996) Identification of a major co-receptor for primary isolates of HIV-1. *Nature* **381**, 661–666
- Dragic, T., Litwin, V., Allaway, G. P., Martin, S. R., Huang, Y., Nagashima, K. A., *et al.* (1996) HIV-1 entry into CD4+ cells is mediated by the chemokine receptor CC-CKR-5. *Nature* **381**, 667–673
- Vila-Coro, A. J., Rodríguez-Frade, J. M., De Ana, A. M., Moreno-Ortiz, M. C., Martínez-A, C., and Mellado, M. (1999) The Chemokine SDF-1 $\alpha$  Triggers CXCR4 receptor dimerization and activates the JAK/STAT pathway. *FASEB J.* **13**, 1699–1710
- Lao, J., He, H., Wang, X., Wang, Z., Song, Y., Yang, B., *et al.* (2017) Single-molecule imaging Demonstrates ligand regulation of the oligomeric status of CXCR4 in living cells. *J. Phys. Chem. B* **121**, 1466–1474
- Martínez-Muñoz, L., Rodríguez-Frade, J. M., Barroso, R., Sorzano, C.Ó. S., Torreño-Pina, J. A., Santiago, C. A., *et al.* (2018) Separating Actin-dependent chemokine receptor Nanoclustering from dimerization indicates a role for clustering in CXCR4 signaling and function. *Mol. Cell* **70**, 106–119.e10
- Babcock, G. J., Farzan, M., and Sodroski, J. (2003) Ligand-independent dimerization of CXCR4, a principal HIV-1 coreceptor. *J. Biol. Chem.* **278**, 3378–3385
- Wu, B., Chien, E. Y. T., Mol, C. D., Fenalti, G., Liu, W., Katritch, V., *et al.* (2010) Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* **330**, 1066–1071
- Tan, Q., Zhu, Y., Li, J., Chen, Z., Han, G. W., Kufareva, I., *et al.* (2013) Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **341**, 1387–1390
- Peng, P., Chen, H., Zhu, Y., Wang, Z., Li, J., Luo, R. H., *et al.* (2018) Structure-based design of 1-Heteroaryl-1,3-propanediamine derivatives as a novel series of CC-chemokine receptor 5 antagonists. *J. Med. Chem.* **61**, 9621–9636
- Zhang, H., Chen, K., Tan, Q., Shao, Q., Han, S., Zhang, C., *et al.* (2021) Structural basis for chemokine recognition and receptor activation of chemokine receptor CCR5. *Nat. Commun.* **12**, 1–12
- Jin, J., Mombouisse, F., Boncompain, G., Koensgen, F., Zhou, Z., Cordeiro, N., *et al.* (2018) CCR5 adopts three homodimeric conformations that control cell surface delivery. *Sci. Signal.* **11**, 2869
- Zhang, F., Yuan, Y., Xiang, M., Guo, Y., Li, M., Liu, Y., *et al.* (2019) Molecular Mechanism Regarding Allosteric Modulation of Ligand Binding and the Impact of Mutations on Dimerization for CCR5 Homodimer. *J. Chem. Inf. Model.* **59**, 1965–1976
- Sohy, D., Yano, H., de Nadai, P., Urizar, E., Guillabert, A., Javitch, J. A., *et al.* (2009) Hetero-oligomerization of CCR2, CCR5, and CXCR4 and the protean effects of “selective” antagonists. *J. Biol. Chem.* **284**, 31270–31279
- Fonseca, J. M., and Lambert, N. A. (2009) Instability of a class A G protein-coupled receptor oligomer interface. *Mol. Pharmacol.* **75**, 1296–1299
- Ohashi, K., and Mizuno, K. (2014) A novel pair of split venus fragments to detect protein–protein interactions by in vitro and in vivo bimolecular fluorescence complementation assays. *Methods Mol. Biol.* **1174**, 247–262
- Chan, K. K., Dorosky, D., Sharma, P., Abbasi, S. A., Dye, J. M., Kranz, D. M., *et al.* (2020) Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. *Science* **369**, 1261–1265
- [preprint] Young, H. J., Chan, M., Selvam, B., Szymanski, S. K., Shukla, D., and Procko, E. (2021) Deep mutagenesis of a transporter for uptake of a non-native substrate identifies conformationally dynamic regions. *bioRxiv*. <https://doi.org/10.1101/2021.04.19.440442>
- Park, J., Selvam, B., Sanematsu, K., Shigemura, N., Shukla, D., and Procko, E. (2019) Structural architecture of a dimeric class C GPCR based on co-trafficking of sweet taste receptor subunits. *J. Biol. Chem.* **294**, 4759–4774
- Villar, V. A. M., Cuevas, S., Zheng, X., and Jose, P. A. (2016) Localization and signaling of GPCRs in lipid rafts. *Methods Cell Biol.* **132**, 3–23
- Goddard, A. D., and Watts, A. (2012) Regulation of G protein-coupled receptors by palmitoylation and cholesterol. *BMC Biol.* **10**, 1–3
- Blanpain, C., Wittamer, V., Vanderwinden, J. M., Boom, A., Renneboog, B., Lee, B., *et al.* (2001) Palmitoylation of CCR5 is Critical for receptor trafficking and efficient activation of intracellular signaling pathways. *J. Biol. Chem.* **276**, 23795–23804
- Holm, K., Weclawicz, K., Hewson, R., and Suomalainen, M. (2003) Human Immunodeficiency Virus type 1 assembly and lipid rafts: Pr55 gag associates with membrane domains that are largely Resistant to Brij98 but sensitive to Triton X-100. *J. Virol.* **77**, 4805–4817
- Chen, B. (2019) Molecular mechanism of HIV-1 entry. *Trends Microbiol.* **27**, 878–891
- Hernanz-Falcón, P., Rodríguez-Frade, J. M., Serrano, A., Juan, D., del Sol, A., Soriano, S. F., *et al.* (2004) Identification of amino acid residues crucial for chemokine receptor dimerization. *Nat. Immunol.* **5**, 216–223
- Issafras, H., Angers, S., Bulenger, S., Blanpain, C., Parmentier, M., Labbé-Jullié, C., *et al.* (2002) Constitutive agonist-independent CCR5 oligomerization and antibody-mediated clustering occurring at physiological levels of receptors. *J. Biol. Chem.* **277**, 34666–34673
- Kasai, R. S., Suzuki, K. G. N., Prossnitz, E. R., Koyama-Honda, I., Nakada, C., Fujiwara, T. K., *et al.* (2011) Full characterization of GPCR monomer-dimer dynamic equilibrium by single molecule imaging. *J. Cell Biol.* **192**, 463–480
- Hern, J. A., Baig, A. H., Mashanov, G. I., Birdsall, B., Corrie, J. E. T., Lazareno, S., *et al.* (2010) Formation and dissociation of M1 muscarinic receptor dimers seen by total internal reflection fluorescence imaging of single molecules. *Proc. Natl. Acad. Sci. U. S. A.* **107**, 2693–2698
- Calebiro, D., Rieken, F., Wagner, J., Sungkaworn, T., Zabel, U., Borzi, A., *et al.* (2013) Single-molecule analysis of fluorescently labeled G-protein-coupled receptors reveals complexes with distinct dynamics and organization. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 743–748
- Herrick-Davis, K., Grinde, E., Cowan, A., and Mazurkiewicz, J. E. (2013) Fluorescence correlation spectroscopy analysis of serotonin, adrenergic, muscarinic, and dopamine receptor dimerization: the oligomer number puzzle. *Mol. Pharmacol.* **84**, 630–642
- Tadagaki, K., Tudor, D., Gbahou, F., Tschische, P., Waldhoer, M., Bomsel, M., *et al.* (2012) Human cytomegalovirus-encoded UL33 and UL78 heteromerize with host CCR5 and CXCR4 impairing their HIV coreceptor activity. *Blood* **119**, 4908–4918
- Ferré, S., Casadó, V., Devi, L. A., Filizola, M., Jockers, R., Lohse, M. J., *et al.* (2014) G protein-coupled receptor oligomerization revisited: functional and pharmacological perspectives. *Pharmacol. Rev.* **66**, 413–434
- Karczewski, K. J., Francioli, L. C., Tiao, G., Cummings, B. B., Alföldi, J., Wang, Q., *et al.* (2020) The mutational constraint spectrum quantified from variation in 141,456 humans. *Nature* **581**, 434–443

## Self-association of chemokine receptors by deep mutagenesis

47. Hancock, J. F., Paterson, H., and Marshall, C. J. (1990) A polybasic domain or palmitoylation is required in addition to the CAAX motif to localize p21ras to the plasma membrane. *Cell* **63**, 133–139
48. Lindwasser, O. W., and Resh, M. D. (2001) Multimerization of human Immunodeficiency Virus type 1 gag promotes its localization to Barges, Raft-like membrane microdomains. *J. Virol.* **75**, 7913–7924
49. Nguyen, D. H., and Taub, D. (2002) CXCR4 function requires membrane cholesterol: implications for HIV infection. *J. Immunol.* **168**, 4121–4126
50. Heredia, J. D., Park, J., Choi, H., Gill, K. S., and Procko, E. (2019) Conformational engineering of HIV-1 Env based on mutational tolerance in the CD4 and PG16 bound states. *J. Virol.* **93**, e00219–e00319
51. Fowler, D. M., Araya, C. L., Gerard, W., and Fields, S. (2011) Enrich: software for analysis of protein function by enrichment and depletion of variants. *Bioinformatics* **27**, 3430–3431
52. Camp, T., Mehta, K., Sligar, S. G., and Zhang, K. (2020) Molecular orientation determination in nanodiscs at the single-molecule level. *Anal. Chem.* **92**, 2229–2236
53. Tinevez, J. Y., Perry, N., Schindelin, J., Hoopes, G. M., Reynolds, G. D., Laplantine, E., *et al.* (2017) TrackMate: an open and extensible platform for single-particle tracking. *Methods* **115**, 80–90
54. Vega, A. R., Freeman, S. A., Grinstein, S., and Jaqaman, K. (2018) Multistep track segmentation and motion classification for transient mobility analysis. *Biophys. J.* **114**, 1018