Noncanonical roles of chemokine regions in CCR9 activation revealed by structural modeling and mutational mapping

Ines De Magalhaes Pinheiro[#1], John R.D. Dawson[#2], Nicolas Calo[1,3], Marianne Paolini-Bertrand[1], Kalyana Bharati Akondi[#1], Gavin Tan[2], Tracy M. Handel[2], Irina Kufareva[2]*, Oliver Hartley[1,3]*

[1]Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland

[2]Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, La Jolla, CA, USA

[3]Orion Biotechnology, Campus Biotech Innovation Park, Geneva, Switzerland

*Corresponding authors:

Oliver Hartley, Ph.D., Associate Professor and Director, Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, 1 rue Michel Servet, 1206 Geneva, Switzerland, Phone: +41 22 379 54 75. Email: oliver.hartley@unige.ch

Irina Kufareva, Ph.D., Associate Adjunct Professor, Skaggs School of Pharmacy and Pharmaceutical Sciences, University of California San Diego, 9255 Pharmacy Lane, MC 0657, La Jolla, CA 92093. Phone: +1 858-822-4163. Email: ikufareva@ucsd.edu

[#]These authors contributed equally.

Abstract

The G protein-coupled CC chemokine receptor 9 (CCR9) plays a major role in inflammatory bowel disease and has been implicated in cancer. Despite its importance as a drug target, there is limited mechanistic understanding of how CCR9 engages and is activated by its endogenous chemokine agonist CCL25. Here, by combining structural modeling with multimodal pharmacological assessment of receptor mutants, we generated a functional map of the CCR9-CCL25 interaction interface and delineated key determinants of binding, agonism, constitutive activity, and G protein vs arrestin signaling. In contrast to all complexes studied to date, where chemokines drive receptor activation through their N-termini, we determined that CCL25 activates CCR9 via a distinct region, its 30s loop. In support of this non-canonical mechanism, CCR9 signaling is tolerant to alanine mutations in the N-terminus of CCL25 but strongly affected by modifications to the 30s loop. Through molecular evolution of the CCL25 N-terminus, we identified chemokine analogs with enhanced binding properties. However, in contrast to other receptor-chemokine systems, these analogs remained full agonists, consistent with the localization of CCL25 signaling determinants outside of the N-terminus. The non-canonical signature of CCR9 activation provides new insights to aid CCR9 drug discovery and may also inform structure-based design of drugs targeting other chemokine receptors.

Introduction

Chemokine receptors are members of the G protein-coupled (GPCR) chemokine receptor superfamily with a principal role in controlling the activation and trafficking of leukocytes [1]. They have been identified as key players in inflammation, infectious diseases and cancer [2], but developing effective medicines targeting these receptors has proven challenging, partly due to an incomplete molecular level understanding of how receptors are engaged and activated by chemokines, and what governs coupling to the principal intracellular effectors, G proteins and arrestins [3, 4].

Advances in Cryo-EM have led to a significant increase in the number of experimentally determined receptor-chemokine complexes (20 chemokine complex structures of 9 receptors in the PDB as of March 2024 [5], out of the total of 41 structures of 12 receptors) and has shed light on molecular activation mechanisms. A general two-site paradigm has been established in which the chemokine globular core provides binding affinity and specificity while its N-terminus drives activation [6-12]. However, alternative mechanisms have been suggested [13-18], and further study of receptor complexes will be necessary to refine the basis of both ligand engagement on the extracellular face and effector coupling on the intracellular face of chemokine receptors. Artificial-intelligence-powered computational modeling tools have the potential to complement experimental structure determination in this endeavor. For example, the breakthrough AlphaFold2 (AF2) technology [19, 20] not only delivers GPCR-peptide complex models with accuracy comparable to cryo-EM structures [21] but can also generate multi-state conformational ensembles [22-24].

C-C chemokine receptor 9 (CCR9) is a chemokine receptor that is expressed on subsets of developing thymocytes and intestinal lymphocytes [25]. Through interaction with its only chemokine ligand, CCL25, CCR9 promotes migration of these cells into their target organs (the thymus and small intestine, respectively) in the context of immune maintenance, surveillance and inflammation [26, 27]. The CCR9-CCL25 axis has attracted interest as a target for the treatment of inflammatory bowel disease [28], and has been studied in the context of tumor progression [29] and immuno-oncology [30]. However, no CCR9-targeting therapeutics have received regulatory approval. While an inactive structure of CCR9 in complex with the allosteric inhibitor vercirnon has been determined [31], the molecular mechanisms of CCR9 agonism remain unknown.

Here we map the determinants of CCR9 activation by CCL25 via AF2 structural modeling and pharmacological assessment of rationally selected CCR9 binding pocket mutants. We demonstrate that for CCR9, the main driver of activation is a structure located in the globular core of CCL25 – the 30s loop. Our data challenge the established two-site model for chemokine receptor engagement [6-12], highlight the diversity of molecular mechanisms by which chemokine receptor activation can occur, and inform rational structure-based targeting of chemokine receptors.

Materials and methods

Reagents

All commercially available reagents used in this study are listed in **Table S1**.

Chemokines

Chemokines were prepared by Fmoc solid phase peptide synthesis [32]. Following cleavage, peptides were ether-precipitated and folded using a glutathione redox buffer. Purity and integrity of products were routinely verified by analytical reversed-phase (RP-) HPLC and mass spectrometry.

The chemokines used in this study are based on a previously described C-terminally truncated version of CCL25 (CCL25(1-73) [33]). For preparations of CCL25 and [1P6]CCL25, Met^{64} was substituted for norleucine and an azidolysine residue was appended to the C-terminal end and chemokines were subjected to a column purification step (RP-HPLC) prior to and following the folding reaction. All other chemokine samples were produced using a previously described column-free method [32]. CCL25 samples prepared using the column purified and column-free methods were shown to have indistinguishable signaling activity from each other and from that of full-length (1-127) recombinant CCL25 (**Fig. S1**). Fluorescent versions of CCL25 and [1P6]CCL25 were generated by coupling an excess of TAMRA-PEG4-DBCO to the azido-lysine lateral chain. Excess dye was removed by 10 kDa size exclusion purification.

Plasmids

Previously generated plasmids used in this study are listed in **Table S2**. For the newly generated FUGW lentiviral vectors encoding CCR9 and CCR9 C-terminally fused to the *Renilla* luciferase variant RLuc8 (CCR9-RLuc8), and variants carrying single alanine mutations (K40A, R44A, Y126A, Y202A, S207A, T208A, K211A, Q267A, N271A, S292A, D296A and F299A), genes were synthesized and subcloned by GenScript using previously described template plasmids.

Cell culture

HEK293T parental cells and HEK293T cells expressing CCR9 (HEK-CCR9) and CCR9 Ala mutants were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 1% Penicillin-Streptomycin. MOLT-4 cells and CHO cells stably expressing CCR9 (CHO-CCR9) were cultured in RPMI-1640 supplemented with 10% FBS and 1% Penicillin-Streptomycin. Cells were grown in a humidified incubator at 37°C with 5% CO_2.

Chemokine phage display

Chemokine phage display was performed as described in [34]. Four phage libraries of CCL25 variants were generated, each incorporating full randomization of four residues in the N-terminus, with two libraries featuring a one-residue N-terminal extension. The four phage libraries were combined and subjected to selection on CHO-CCR9 cells. 48 enriched sequences identified after the third and fourth rounds of selection were chosen for further evaluation.

Generation of CCR9 expressing cell lines

HEK-CCR9 and HEK-CCR9-RLuc8 YFP-arrestin 3 (Arr3) cells, and respective alanine mutants, were obtained by lentiviral transduction as previously described [35], followed by selection of high-expressing populations via fluorescence-activated cell sorting (FACS) in a BD FACS Aria® Fusion flow cytometer using a fluorescent anti-CCR9 monoclonal antibody (mAb). For arrestin recruitment reporter cell lines, HEK293T parental cells were first lentivirally transduced with FUGW-YFP-Arr3, and a high-YFP-expressing population was selected by FACS. The selected FUGW-YFP-Arr3 cell population was then lentivirally transduced with appropriate FUGW-CCR9-RLuc8 vectors, followed by population selection by FACS with anti-CCR9 mAb.

Receptor surface level quantification via flow cytometry

Over the course of the work, CCR9 expression in HEK-CCR9 and HEK-CCR9-RLuc8 YFP-Arr3 cell lines was regularly measured using flow cytometry with a fluorescent anti-CCR9 mAb (**Fig. S2A-B**). For this, following detachment with 0.5 mM EDTA, 2 x 10^5 cells per sample were incubated with an anti-CCR9 mAb (1:100 dilution) in FACS buffer (1x PBS, 1 mM EDTA, 1% BSA) in 96-well plates. Following 1 h incubation, cells were washed once in FACS buffer and mAb binding was measured by flow cytometry on a Cytoflex instrument (Beckman Coulter). 10^4 events were collected per sample, in technical triplicates, and median fluorescence intensity (MFI) values of anti-CCR9 mAb were obtained using CytExpert (Beckman Coulter). Antibody binding signals were expressed as:

$$\frac{\log MFI_{mut} - \log MFI_{parental}}{\log MFI_{WT} - \log MFI_{parental}} \quad \text{(Eq.1)}$$

The antibody is directed against an unknown epitope on CCR9; to exclude the possibility of its binding being affected by individual Ala mutations, we assessed correlation of antibody binding between the untagged CCR9 mutants and their respective CCR9-RLuc8 counterparts. A mutation affecting antibody binding would be expected to alter surface detection levels for both receptor mutant variants; however, no correlation was found (**Fig. S2C**). Total expression in HEK-CCR9-RLuc8 cell lines was assessed by luminometry and compared with the surface expression in the same cell lines (**Fig. S2D**).

Flow cytometry-based chemokine binding assays

Following detachment with 0.5 mM EDTA, 2 x 10^5 cells per sample were incubated TAMRA-labelled chemokines diluted in FACS buffer (1x PBS, 1 mM EDTA, 1% BSA) in 96-well plates. Following 1 h incubation, cells were washed once in FACS buffer and ligand binding was measured by flow cytometry on a Cytoflex instrument (Beckman Coulter). 10^4 events were collected per sample, in technical triplicates, and median fluorescence intensity (MFI) values of CCL25-TAMRA and [1P6]CCL25-TAMRA were obtained using CytExpert (Beckman Coulter). Nonspecific binding was measured in parental HEK293T cells (**Fig. S3**). For binding evaluation of CCR9 mutants, we used the highest concentration (300 nM) of fluorescent chemokine and binding signals were expressed as:

$$\frac{\log MFI_{mut} - \log AF_{mut}}{\log MFI_{WT} - \log AF_{WT}} \quad \text{(Eq. 2)}$$

where AF (autofluorescence) is the MFI of the corresponding cell line in the same experiment in the absence of the fluorescent chemokine.

Calcium flux assays

MOLT-4, HEK-CCR9 and respective mutant cells were seeded at 3 x 10^4 cells/well in 384-well black clear flat bottom plates. 4 h later, cells were loaded with Fluo-8 calcium-sensitive fluorescent dye according to the manufacturer's instructions.

Fluorescence signals (excitation, 490 nm; emission, 525 nm) were recorded using an FDSS instrument (HAMAMATSU). In the agonist mode, signals were recorded before and after the addition of WT CCL25 or CCL25 analogs diluted to defined concentrations in assay buffer (1% BSA and 25 mM HEPES). For assessing the antagonist activity of phage display-derived CCL25 variants (or their capacity to desensitize the receptor), measurements were made following the addition of 100 nM CCL25 5 minutes later.

For each well, the recorded fluorescence signal was divided by the baseline signal of the same well acquired just before the corresponding treatment, and by the fluorescence values recorded for cells treated with vehicle only during the first (agonist mode) or both (antagonist mode) injections. Agonist responses were expressed as:

$$Ca^{2+}\ \text{signal} = \max\left(\left[\frac{Fluorescence\ (t)}{Mean\ Fluorescence_{\ t_{inje}}}\right]_{agonist} \Big/ \left[\frac{Fluorescence\ (t)}{Mean\ Fluorescence_{\ t_{inje}})}\right]_{Buffer}\right) \quad \text{(Eq. 3)}$$

where "t_inj" is the time of injection of agonist or buffer.

BRET assay for arrestin recruitment

HEK-CCR9-RLuc8 YFP-Arr3 and respective mutant cells were detached with 0.5 mM EDTA, seeded in 384-well black, clear flat-bottom plates at 2 x 10^4 cells per well in 30 µL per well of FluoroBrite™ DMEM, 10 % FBS, and incubated overnight at 37 °C, 5 % CO_2. Cells were then incubated for 10 minutes in BRET buffer (0.14 M NaCl, 6 mM KCl, 2 mM $MgSO_4$, 15 mM HEPES, 1 g/L glucose, 1 % BSA) containing 20 µM coelenterazine h and then stimulated with either chemokines at defined concentrations diluted in BRET buffer or BRET buffer alone. Luminescence was measured over 10 min using an FDSS instrument (HAMAMATSU) after agonist or buffer injection. Agonist responses were defined as area under the curve (AUC) of BRET signal:

$$BRET\ ratio(t) = \frac{em_{540}(t)}{em_{470}(t)}$$

$$\text{BRET signal} = \left(\left[\frac{BRET\ ratio\ (t)}{BRET\ ratio\ (t_{injection})}\right]_{agonist} \Big/ \left[\frac{BRET\ ratio\ (t)}{BRET\ ratio\ (t_{injection})}\right]_{Buffer}\right) \quad \text{(Eq. 4)}$$

Basal association (in the absence of agonist stimulation) was measured between the RLuc8-tagged receptor and YFP-tagged Arr3 by BRET (**Fig. S4A**). YFP-Arr3 acceptor expression levels

were assessed by flow cytometry and correlated with donor (WT or mutant CCR9-RLuc8) luminescence and basal BRET (**Fig. S4B**).

BRET Gαi/Gβγ dissociation assays

HEK-CCR9-WT, HEK-CCR9-T208A and HEK-CCR9-Q267A cells were co-transfected with Gαi(91)-Rluc2, mVenus-Gβ1 and untagged Gγ2 (in a 1:5:5 ratio) in a 6 well-plate with jetPRIME® reagent, according to the manufacturer's protocol. After 24 h, cells were detached with 0.5 mM EDTA and seeded at 2 x 10^4 cells/well in 384-well / flat bottom plates in 30 µL/well of FluoroBrite™ DMEM. Cells were then incubated for 10 minutes in 1X PBS containing 20 µM coelenterazine h followed by stimulation with either CCL25 at defined concentrations in BRET buffer or BRET buffer alone for 10 min. Signals were measured using an FDSS instrument (HAMAMATSU). Agonist responses were defined as the area over the curve (AOC) of BRET signal (eq. 4). Expression levels of mVenus-Gβ1 were quantified by flow cytometry.

Model building

Structural models of CCR9A complexes with WT CCL25 and [1P6]CCL25 were built by AlphaFold2 Multimer v2.3.2 [19, 20, 36] locally installed on the UCSD Triton Shared Computing Cluster (TSCC). For each complex, an ensemble of 25 models was built. Using ICM software version 3.9-3b [37] the WT CCL25 and [1P6]CCL25 models were modified to include a cyclized pGlu1 and a free positively charged N-terminus (NH_3^+), respectively. Chemokine molecules in all complexes were subjected to local gradient minimization with positional harmonic restraints on Cα atoms using ICM [37]. Complexes were then scored using the Radial and Topological Convolutional Neural Network (RTCNN), a deep learning-based scoring function implemented in ICM and trained to distinguish protein complexes with potent binders from similar decoy complexes [38-40]. High-scoring complexes were prioritized in the analysis.

For the best-scoring CCR9-[1P6]CCL25 model, additional refinement was performed in ICM in two stages: one employing 3D grid potentials and another full-atom representation of all components. During the first stage, the receptor binding pocket was represented with a set of grid interaction potentials, including those for van der Waals, electrostatic, hydrogen bonding and apolar surface interactions [41, 42]. The N-terminus and the 30s loop of [1P6]CCL25 (Tyr1-QASEDC-Cys8 and Ile28-QEVSGSCNL-Pro38 respectively) were built *ab initio*; an explicit disulfide bond was imposed between C7 and C35; and residues C8, I28, and P38 were tethered to the corresponding residues in the template. The conformational stack of the system including the N-terminus and 30s loop was initialized based on the AlphaFold2 model ensemble, and the system was then thoroughly sampled in the receptor potential grids, using biased probability Monte Carlo sampling in ICM, to optimize and expand on the conformational stack. For the second stage, the resulting conformational stack was merged with full atom models of the receptor, and at least 10^8 steps of Monte Carlo optimization were performed, allowing for the same level of flexibility in the chemokine fragments with added full flexibility of receptor binding pocket sidechains. For full-atom sampling, van der Waals, torsional, hydrogen bonding, electrostatic, and disulfide bond energy terms were used. The resulting conformations were

clustered, re-scored using RTCNN, inspected visually, and the binding geometry of the best scoring fragmented complex was transferred onto the full [1P6]CCL25 model for visualization.

Data analysis

Data normalization, collation, and statistical analyses were performed in GraphPad Prism 10.0. For chemokine binding to CCR9 mutants, binding ratios (eq.1) were log-transformed and evaluated using one-way ANOVA with post-hoc tests and Holm-Šídák's correction for multiple comparisons. For CCR9 mutant signaling assays, areas under concentration response curves (AUCRC) were calculated, divided by WT AUCRC from the same experiment, and similarly evaluated using one-way ANOVA with post-hoc tests and Holm-Šídák's correction for multiple comparisons. For calcium flux assays on CCL25 alanine mutants, concentration response curves (CRCs) were fitted with a four-parameter nonlinear regression and mutant models were compared one-by-one to the WT model using an F-test, followed by a Bonferroni correction for multiple testing [43]. **Tables S3**, **S4** and **S6** represent the p-values for all measured outputs. p-values of ≤0.05 were considered statistically significant. Radar plots were constructed using python and matplotlib [44].

Results

Structural features of the CCR9-CCL25 signaling complex

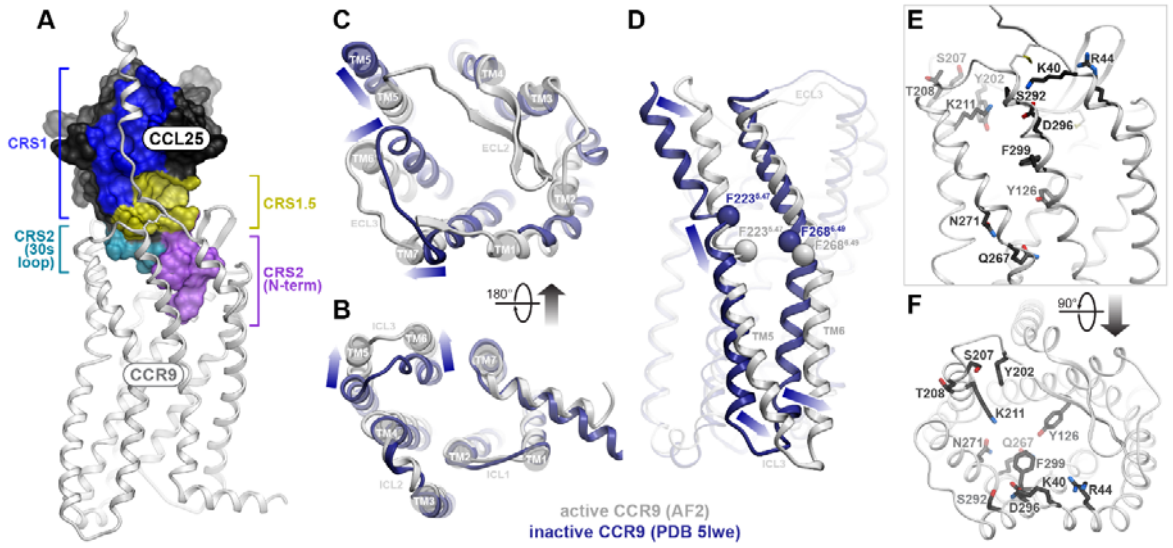

Figure 1. Structural model of CCR9 signaling complex with CCL25 and rationale for receptor mutagenesis. A. AlphaFold2 model of the CCR9-CCL25 signaling complex, viewed along the plane of the membrane. The receptor is shown as a white ribbon, the chemokine as a black mesh where surfaces interacting with the indicated receptor regions are colored. CRS: chemokine recognition site.

B-D. Structural superposition of CCR9 in its predicted CCL25-bound conformation (white) with the X-ray structure of CCR9 bound to the small molecule antagonist vercirnon (PDB entry 5LWE, navy) viewed perpendicular to the plane of the membrane from the intracellular or extracellular side (**B** and **C**, respectively), or parallel to the plane of the membrane in the TM5-to-TM2 direction (**D**). Arrows indicate the directions of the structural changes between the inactive and the predicted active states. In (**D**), the Cα atoms of reference residues close to the middle of TM5 and TM6 are shown in spheres.

E-F. CCR9 CRS2 residue positions selected for mutagenesis are shown as sticks and viewed parallel to the membrane (**E**) or perpendicular to the membrane from the extracellular side (**F**).

To gain insight into the interaction of CCR9 with CCL25, we constructed an AlphaFold2 model of the CCR9-CCL25 signaling complex (**Fig. 1A**). The complex features the overall architecture of all canonical receptor-chemokine complexes revealed by experimental structures determined to-date: the proximal N-terminus of the receptor (chemokine recognition site 1 or CRS1) binds in the surface groove between the N-loop and 40s loop of the chemokine core, the chemokine N-terminus is submerged in the orthosteric binding pocket (CRS2), and the areas surrounding the conserved disulfides of the chemokine and the receptor pack against each other and form the intermediate CRS1.5 (**Fig. 1A**).

Model superposition with the X-ray structure of inactive CCR9 bound to vercirnon [31] revealed that CCL25-bound CCR9 adopts a characteristic active conformation featuring a distinctive outward movement of the intracellular end of the transmembrane (TM) helices 5 and 6 (**Fig.**

1B). Prominent TM helix rearrangements are also apparent in the orthosteric binding pocket on the extracellular side of the receptor (**Fig. 1C**): compared to inactive CCR9, the CCL25-bound conformation features a large inward movement of the extracellular end of TM5 with a concurrent outward movement of extracellular loop (ECL) 3 and the adjoining ends of TM6 and TM7, providing an opening to accommodate the chemokine. These movements are consistent with activation-associated rearrangements on the extracellular side of other chemokine receptors [40]. However, unlike in these other receptors, the activation of CCR9 is also accompanied by a profound (~6 Å) downward (i.e in the intracellular direction) sliding-and-bending motion of TM5, and, to a lesser extent of TM6, relative to the rest of the TM bundle (**Fig. 1D**).

To better understand the shared and unique architectural features of the CCR9-CCL25 complex, we compared it to the available experimental structures of other active-state receptor-chemokine complexes [14, 17, 45-49]. CCR9 belongs to the same phylogenetic subfamily as CCR1, CCR2, CCR5 and CCR6, sharing 40-45% identity in the TM domains (CCR6 is the closest with 44.9% TM identity) and 30-35% identity in the TM regions and extracellular loops involved in chemokine binding (CCR5 is the closest with 34.7% identity). At the primary sequence level, the CCR6 ligand CCL20 is one of the most similar to CCL25, sharing the characteristics of a short N-terminal region (six residues in CCL25 versus five in CCL20; other chemokines have seven or more) and an unusually long 30s loop (11 residues in CCL25, 29-QEVSGSCNLPA-39 – versus 9 in CCL20; other chemokines have 7 or less, **Fig. S5A**). Despite these sequence similarities, the predicted binding mode of CCL25 to CCR9 is strikingly different from the experimentally determined CCR6-CCL20 complex [45] and more closely resembles the interactions of CCR1, CCR2 and CCR5 with their respective chemokines (**Fig. S5B-G**): whereas the 30s loop of CCL20 lies above the binding pocket of CCR6, the CCL25 30s loop enters the CCR9 binding pocket together with the N-terminal region and forms prominent interactions with CCR9 CRS2 (**Fig. 1A**).

To probe the functional significance of the observed CRS2 interactions between CCL25 and CCR9, we selected 10 residues in CCR9 CRS2 that make direct contact with the chemokine in the model (**Fig. 1E, F**) and generated alanine substitution mutants. We added two further alanine mutants at CCR9 Ballesteros-Weinstein (BW) [50] positions 6.48 and 6.51. Residues at these positions do not directly contact CCL25 in the model but are known to play critical roles in extracellular-to-intracellular signal transmission in other chemokine and non-chemokine GPCRs [51-54]. The 12 mutants were characterized (**Fig. S6**) with respect to their capacity to bind CCL25, to drive intracellular calcium (Ca^{2+}) mobilization (a G protein-mediated response), and to recruit arrestin 3 (Arr3).

Evaluation of the roles of CCR9 CRS2 residues in chemokine binding and signaling

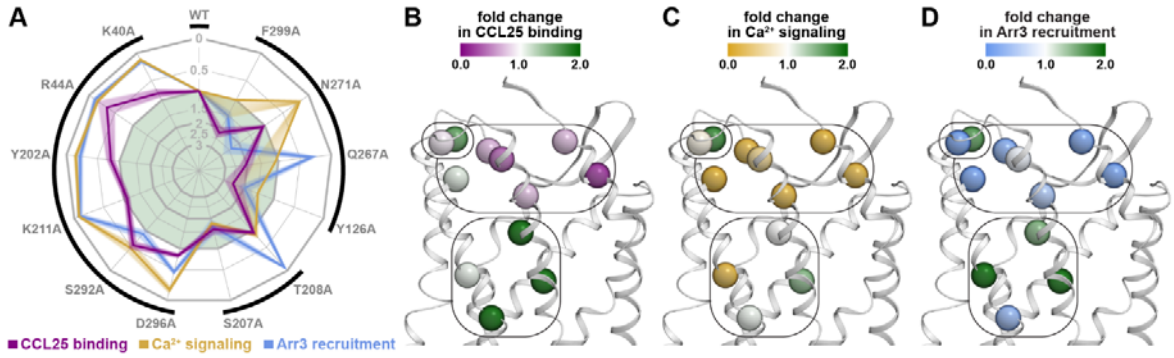

Figure 2. Overview of the impact of CCR9 mutations on CCL25 binding and CCL25-induced signaling.

A. Radar plot summarizing the impact of CCR9 mutations on CCL25 binding (purple), CCL25-induced intracellular Ca^{2+} mobilization (orange), and CCL25-induced Arr3 recruitment to the receptor (blue). Mutation impacts are expressed as ratios of mutant to WT responses in respective experiments; contours outside or inside the central light green area correspond to negative and positive impacts, respectively. Responses were determined as areas under the CCL25 concentration-response curves (AUCRCs) for Ca^{2+} mobilization and Arr3 recruitment, and as receptor-specific cell fluorescence increases in the presence of 300 nM of TAMRA-labeled CCL25 for binding. Data represent mean of 2-3 independent experiments; SEM values are represented by the contour width and transparency at the respective mutant axis. Black external brackets denote the groups of functionally and structurally related mutations as they are presented in this paper. Statistics are available in **Table S3.**

B-D. Detrimental or beneficial impact of mutations at selected residues is reflected in the color of their Cα atoms (spheres), in relation to CCL25 binding (**B**), CCL25-induced intracellular Ca^{2+} mobilization (**C**), and CCL25-induced Arr3 recruitment to CCR9 (**D**). The receptor is shown as a white ribbon and viewed parallel to the plane of the membrane. Rounded rectangles mark the groups of functionally and structurally related mutations, matching the outside brackets in (**A**).

The mutations had a broad range of effects, ranging from complete abrogation of binding or signaling to more than 2-fold enhancement compared to WT, as evidenced by the non-circular shapes of the assay contours in **Fig. 2A**. The binding and the two signaling readouts were not always affected in a consistent manner as reflected by divergent assay contour shapes (**Fig. 2A**): some mutants negatively impacted all three experimental readouts (e.g. R44A, **Fig. 2A**), others impacted signaling to a greater extent than binding (e.g. K211A, **Fig. 2A**), and some selectively abrogated one signaling response (Ca^{2+} mobilization or Arr3 recruitment) while making no difference in binding or the other signaling readout (e.g. T208A and N271A, **Fig. 2A**). When viewed on the receptor model, the mutations framing the opening of the pocket generally decreased all three readouts, whereas those located on the periphery or deeper in the binding pocket affected the readouts in divergent ways (**Fig. 2B-D**). Based on spatial location and impact, we tentatively grouped the mutations as shown in **Fig. 2B-D** and investigated the specific molecular interactions that are responsible for the observed alterations in chemokine binding and receptor signaling by mutations in each group.

The extracellular "rim" of CCR9 helical bundle is critical for binding CCL25 in a signaling-productive conformation

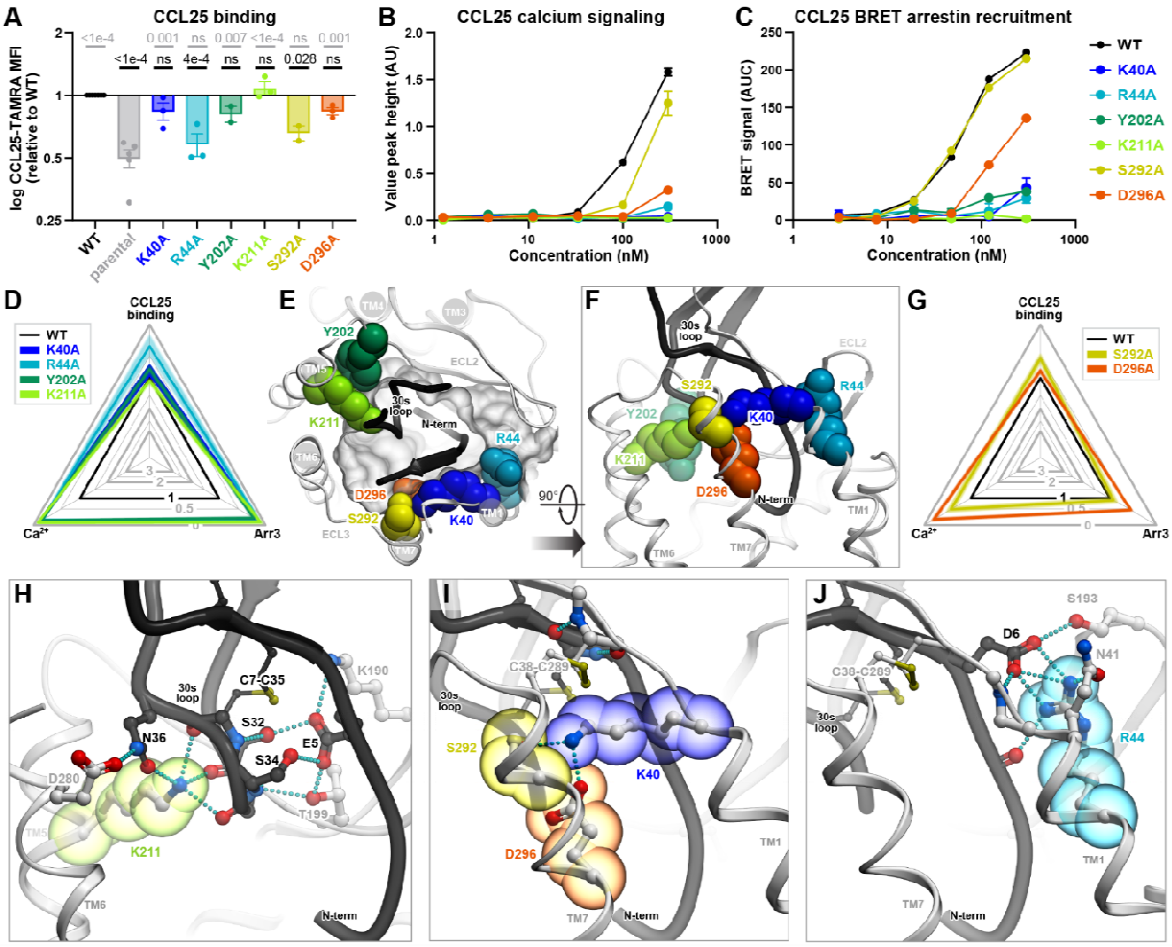

Figure 3. Pharmacological and structural evaluation of the mutations in the extracellular "rim" of the CCR9 helical bundle. Statistical analyses are shown in **Table S3.**

A. CCL25 binding to WT CCR9 and CCR9 mutants expressed in HEK293 cells. Cells were incubated at 4°C for 1h with 300 nM of TAMRA-labelled CCL25. Bars represent mean ± SEM of the ratio of specific binding signals (**eq. 2**) between the mutant and WT CCR9, measured in 2-3 independent experiments. P-values in comparison to HEK-CCR9 WT and HEK293 parental cells are shown for each mutant in black and grey, respectively (complete CCL25 binding CRCs are available in **Fig. S7**).

B. CCL25-induced Ca^{2+} signaling by WT CCR9 or CCR9 mutants expressed in HEK293 cells. Ca^{2+} signals in response to CCL25 at the indicated concentrations are shown as mean peak height (**eq. 3**) ± SEM from triplicate wells; the data shown are representative of 3 independent experiments (complete data set is available in **Fig. S8A-C**).

C. BRET assays for CCL25-induced Arr3 recruitment on WT CCR9 and CCR9 mutants. Data points represent mean ± SEM of BRET signal (**eq. 4**) obtained in triplicate wells; data shown is representative of 3 independent experiments (complete data set is available in **Fig. S8D-F**).

D, G. Radar plots summarizing the impact of the indicated mutants on CCL25 binding and signaling. Impact is quantified relative to WT CCR9 as in **Fig. 2A**. The black equilateral triangle denotes WT CCR9; mutant contour deviations from the equilateral shape indicate disproportionate impact in one or more assays compared to others.

E-F. The indicated residues are a part of a "rim" of the CCR9 helical bundle and interact with the proximal N-terminus and the 30s loop of the chemokine. The receptor and the chemokine are shown as a white and black ribbon, respectively; and the amino-acid residues of interest as colored spheres. In **E**, the complex is viewed across the plane of the membrane from the extracellular side, the binding pocket is shown as a transparent mesh, and most of the chemokine globular core is clipped for clarity. In **F**, the complex is viewed along the plane of the membrane.

H-J. Hydrogen-bonding networks of CCR9 K211$^{5.35}$ (**H**), K40$^{1.24}$ (**I**), and R44$^{1.28}$ (**J**) in the predicted CCR9-CCL25 complex. Hydrogen bonds are shown as cyan dotted lines. Receptor and chemokine ribbons, sticks, and spheres are colored as in **E-F**. The complex is viewed along the plane of the membrane as in (**F**).

Across the entire panel, four mutations stood out as being practically signaling-dead (less than 15% of WT signaling output): K40$^{1.24}$A, R44$^{1.28}$A, Y202^{ECL2}A, and K211$^{5.35}$A (**Fig. 3A-D** and **S8**). For R44$^{1.28}$A, the loss of signaling could be attributed to a complete loss of binding (**Fig. 3A**, the mutant is indistinguishable from the negative control, the parental cell line). However, for the remaining three mutants, the binding was fully or partially preserved (**Fig. 3A**).

In the CCR9-CCL25 model, K40$^{1.24}$, R44$^{1.28}$, Y202^{ECL2}, and K211$^{5.35}$ are part of the extracellular rim of the TM bundle that 'grips' the chemokine to hold it in the binding pocket (**Fig. 3E**). K40$^{1.24}$ and R44$^{1.28}$ primarily coordinate the proximal N-terminus of CCL25 while Y202^{ECL2} and K211$^{5.35}$ act on the 30s loop (**Fig. 3E**). Proximal to K40$^{1.24}$ are TM7 residues S292$^{7.28}$ and D296$^{7.32}$ that also contribute to the 'grip' (**Fig. 3E, F**). Functionally, the D296$^{7.32}$A mutation nearly eliminated the Ca^{2+} response and greatly reduced arrestin recruitment (only 15% and 46% of WT output retained, respectively), with no loss of binding, whereas S292$^{7.28}$A significantly weakened the binding and Ca^{2+} response, but retained arrestin recruitment (48% and 89%, respectively, **Fig. 3A-C, G, Fig. S7** and **S8**).

Closer examination of the predicted CCR9-CCL25 complex reveals that these residues are centers of three critical hydrogen bonding networks. First, K211$^{5.35}$ forms three hydrogen bonds with the backbones of the CCL25 30s loop residues S32, G33 and C35, and in addition bridges the chemokine residue N36 to D280$^{6.61}$, thus affixing the 30s loop to CCR9 TMs 5 and 6 (**Fig. 3H**). Second, K40$^{1.24}$ shapes the junction between CCR9 TMs 1 and 7 by hydrogen-bonding to both S292$^{7.28}$ and D296$^{7.32}$; this also improves packing of its side chain with the proximal N-terminus of CCL25 (**Fig. 3I**). Third, R44$^{1.28}$ is involved in an intricate network connecting the proximal N-terminus of the CCL25 (including residue D6), CCR9 TM1 (including the backbone of K40$^{1.24}$), and CCR9 ECL2 (S193^{ECL2}) (**Fig. 3J**). Completing the assembly, the proximal N-terminus of the chemokine and its 30s loop are connected not only through the C7-C35 disulfide bridge but also by hydrogen bonding between CCL25 S32, E5, and S34, with E5 also hydrogen-bonding to K190^{ECL2} and T199$^{45.51}$ in CCR9 ECL2 (**Fig. 3H**). Y202^{ECL2} provides steric interactions and structural support to the ECL2-30s loop-TM5 side of the assembly (**Fig. 3E**).

Altogether, these results delineate the role of the extracellular "rim" of the CCR9 orthosteric pocket in binding and positioning the proximal N-terminus and the 30s loop of the chemokine

in a signaling-productive conformation. Importantly, disrupting the hub of interactions involving K211$^{5.35}$ and Y202^{ECL2} had a disproportionate effect on signaling relative to binding (**Fig. 3A-D**); this suggests that the 30s loop of CCL25, including residue N36, drives CCR9 signaling through TM5 and TM6. The involvement of CCR9 TM5 and TM6 is consistent with the rearrangements of their extracellular ends observed in the active-inactive structure comparison (**Fig. 1C-D**). The central signaling role of the 30s loop is unusual in the chemokine receptor family [18] and may be a unique feature of the CCR9-CCL25 complex.

Top-of-TM5 mutations can enhance and bias CCR9 signaling in response to CCL25

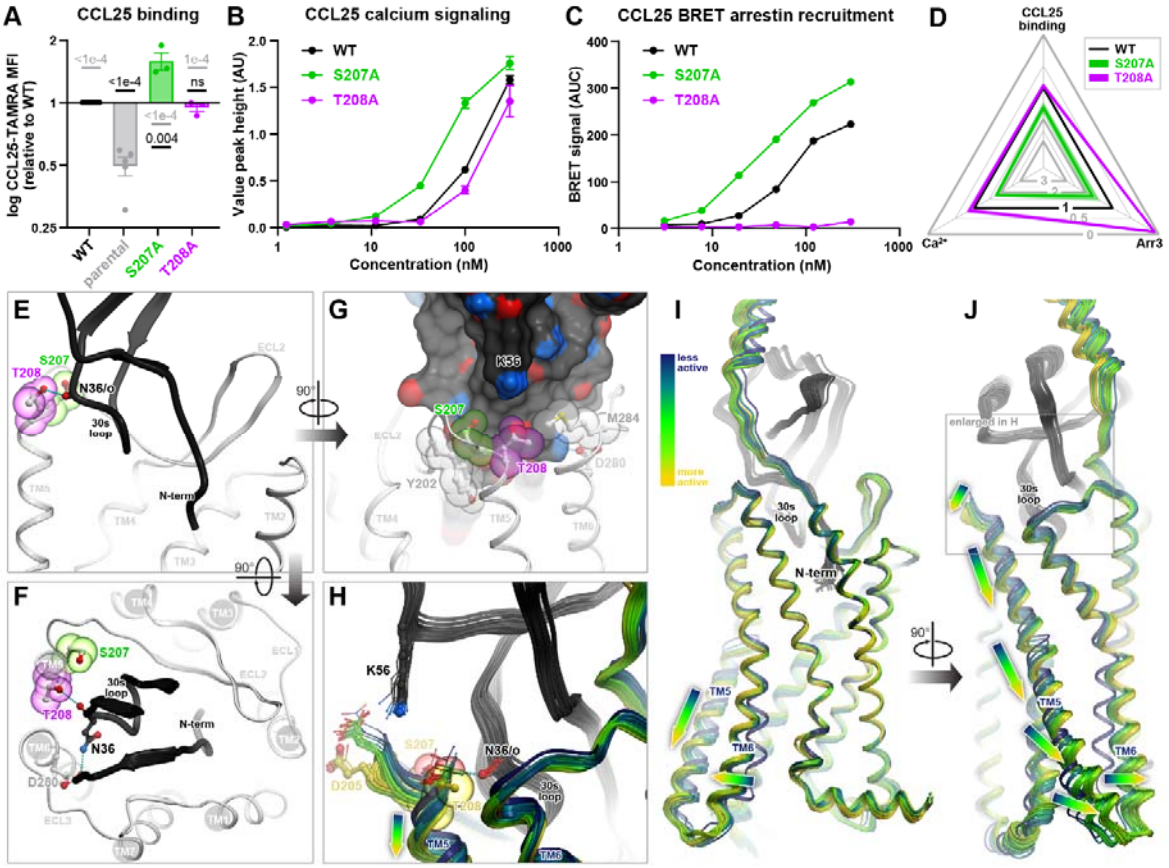

Figure 4. Pharmacological and structural evaluation of the peripheral top-of-TM5 CCR9 mutations. Statistical analyses are shown in **Table S3**.

A. CCL25 binding to WT CCR9 and CCR9 mutants expressed in HEK293 cells. Cells were incubated at 4°C for 1h with 300 nM of TAMRA-labelled CCL25. Bars represent mean ± SEM of the ratio of specific binding signals (**eq. 2**) between the mutant and WT CCR9, measured in 2-3 independent experiments. P-values in comparison to HEK-CCR9 WT and HEK293 parental cells are shown for each mutant in black and grey, respectively (complete CCL25 binding CRCs are available in **Fig. S9**).

B. CCL25-induced Ca^{2+} signaling on WT CCR9 or CCR9 mutants expressed in HEK293 cells. Ca^{2+} signals in response to CCL25 at the indicated concentrations are shown as mean peak height (**eq. 3**) ± SEM from triplicate wells; the data shown are representative of 3 independent experiments (complete data set is available in **Fig. S10A-C**).

C. BRET assays for CCL25-induced Arr3 recruitment on WT CCR9 and CCR9 mutants. Data points represent mean ± SEM of BRET signal (**eq. 4**) obtained in triplicate wells; data shown is representative of 3 independent experiments (complete data set is available in **Fig. S10D-F**).

D. Radar plot summarizing the relative impact of $S207^{5.31}A$ and $T208^{5.32}A$ on CCL25 binding and signaling. Impact is quantified relative to WT CCR9 as in **Fig. 2A**. Black equilateral triangle denotes WT CCR9; a nearly equilateral $S207^{5.31}A$ contour indicates a balanced enhancement in all responses whereas a scalene assay contour for $T208^{5.32}A$ emphasizes a disproportionate impact of this mutation on Arr3 recruitment compared to CCL25 binding and Ca^{2+} signaling.

E-F. Location of CCR9 $S207^{5.31}$ and $T208^{5.32}$ in relation to the CCL25 molecule in the complex. Receptor and chemokine are shown as white and black ribbons, respectively, residues of interest as sticks and

colored spheres, and hydrogen bonds as cyan dotted lines. In **E**, the complex is viewed along the plane of the membrane; receptor TM helices 1, 6, and 7 are removed for clarity. In **F**, the complex is viewed across the plane of the membrane from the extracellular side and most of the chemokine globular core is clipped for clarity.

G. Packing of CCR9 S207$^{5.31}$ and T208$^{5.32}$ against the predominantly non-polar surface of the chemokine 30s loop. The chemokine is shown as a black mesh with exposed hydrogen bond donors and acceptors colored in blue and red, respectively; all surface regions shown in black are nonpolar. Receptor residues of interest are shown in sticks and colored spheres. Complex is viewed along the plane of the membrane in the TM5-to-TM2 direction.

H-J. Analysis of the empirical "degree of receptor activation" in the ensemble of AF2 models of the CCR9-CCL25 complex. "Degree of activation" is quantified as the distance between the hydroxyl of CCR9 Y126$^{3.32}$ and the side chain amide of N271$^{6.52}$ (prioritized based on the comparison in **Fig. 1B-D**). The ensemble is shown in ribbons with the chemokine in black and the receptor colored by the "degree of activation". Arrows indicate the direction of conformational changes observed in the ensemble in concert with the Y126-N271 distance increase. In (**H**), CCR9 D205, S207$^{5.31}$, and T208$^{5.32}$ are shown as thin sticks; for the "most active" model (gold), S207$^{5.31}$ and T208$^{5.32}$ are shown as thicker sticks and spheres, and the hydrogen bond to CCL25 N36 as a cyan dotted line.

Two additional residues were mutated at the periphery of the CCR9 TM domain: S207$^{5.31}$ and T208$^{5.32}$. Despite being solvent-facing and adjacent to each other, these residues produced pronounced and strikingly distinct effects when mutated to alanine. S207$^{5.31}$A strongly enhanced CCL25 binding (**Fig. 4A** and **S9A-C**) together with CCL25-induced Ca^{2+} flux (**Fig. 4B** and **S10A-C**) and Arr3 recruitment (**Fig. 4C-D** and **S10D-F**). By contrast, T208$^{5.32}$A completely abrogated CCL25-mediated Arr3 association (**Fig. 4C-D** and **S10D-F**) while leaving CCL25 binding (**Fig. 4A** and **S9A-C**) and Ca^{2+} signaling (**Fig. 4B** and **S10A-C**) unchanged. The full G protein signaling capacity of CCR9 T208$^{5.32}$A was confirmed in a non-amplified G protein activation assay (Gαi/Gβγ dissociation BRET, **Fig. S11**).

In the complex model, S207$^{5.31}$ is positioned one helical turn above K211$^{5.35}$ and is proximal to an entirely hydrophobic surface on CCL25 30's loop (**Fig. 4E-G**). The elimination of S207$^{5.31}$ hydroxyl group via an alanine mutation would strengthen CCR9 hydrophobic packing against this surface, which explains the observed concerted increase in chemokine binding and agonism. T208$^{5.32}$ interacts with an adjacent part of the chemokine surface that features only a single polar atom, the backbone oxygen of residue N36. Interestingly, in the top-ranked (by pLDDT [19]) model of the complex, the rotamer state of T208$^{5.32}$ and its distance from N36 were not conducive to the formation of a hydrogen bond. However, a hydrogen-bond forming conformation was identified when examining the entire ensemble [22, 24] of AF2-generated CCR9-CCL25 models (**Fig. 4H**). Moreover, this conformation corresponds to what we interpret as the "most active" state of CCR9, based on three features emphasized by the active-inactive structure comparison in **Fig. 1B-D**: the largest outward movement of the intracellular end of TM6, the greatest intramolecular distance across the binding pocket in the TM3-to-TM6/7 direction, and the "deepest" position of TM5 relative to the rest of the TM helices (**Fig. 4H-J**). We hypothesize that Arr3 recruitment requires this "most active" state involving the T208$^{5.32}$-N36 hydrogen bond, whereas G protein association is permissive to a range of active-like

conformations of CCR9, as previously described for other GPCRs [55]. The loss of T208$^{5.32}$-N36 hydrogen bonding in the T208$^{5.32}$A mutant would therefore selectively abrogate Arr3 recruitment with minimal impact on G protein-mediated Ca^{2+} mobilization.

Altogether, these results establish distinct and nontrivial roles for two extracellularly facing, partially solvent-exposed residues S207$^{5.31}$ and T208$^{5.32}$ in controlling not only chemokine binding but also receptor signaling and bias. Worth noting is the proximity of these residues to K211$^{5.35}$ and the chemokine 30s loop, consistent with the role of TM5 as a driver of CCR9 activation and the 30s loop as a major signaling determinant in the CCR9-CCL25 complex.

Non-canonical roles of residues at and below the binding pocket floor in CCR9 activation

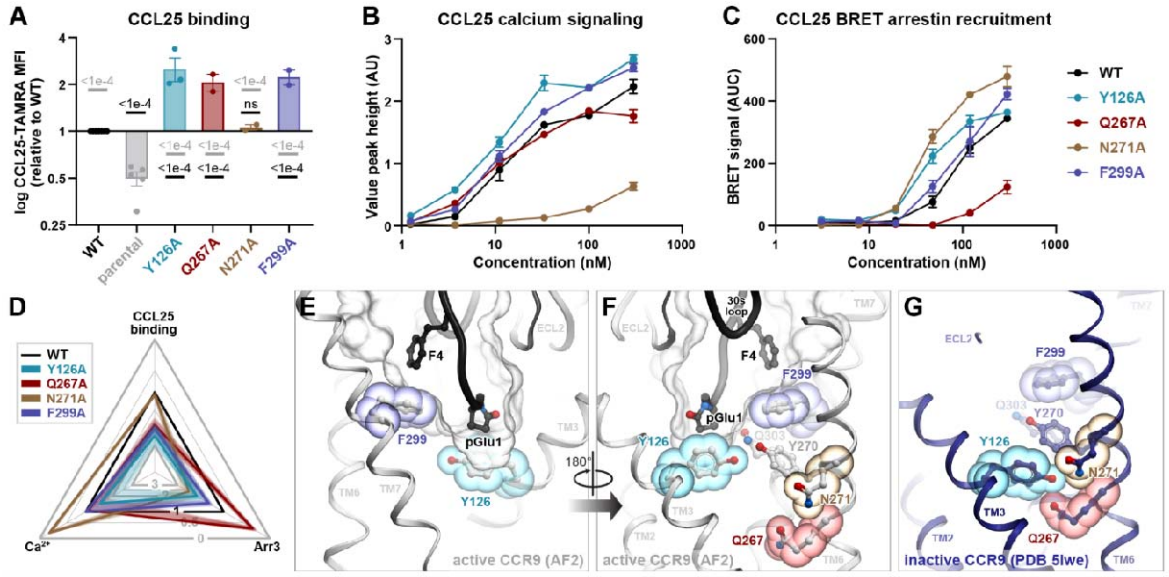

Figure 5. Pharmacological and structural evaluation of the CCR9 mutations at and below the binding pocket floor. Statistical analyses are shown in **Table S3**.

A. CCL25 binding to WT CCR9 and CCR9 mutants expressed in HEK293 cells. Cells were incubated at 4°C for 1h with 300 nM of TAMRA-labelled CCL25. Bars represent mean ± SEM of the ratio of specific binding signals (**eq. 2**) between the mutant and WT CCR9, measured in 2-3 independent experiments. P-values in comparison to HEK-CCR9 WT and HEK293 parental cells are shown for each mutant in black and grey, respectively (complete CCL25 binding CRCs are available in **Fig. S12**).

B. CCL25-induced Ca^{2+} signaling on WT CCR9 or CCR9 mutants expressed in HEK293 cells. Ca^{2+} signals in response to CCL25 at the indicated concentrations are shown as mean peak height (**eq. 3**) ± SEM from triplicate wells; the data shown are representative of 3 independent experiments (complete data set is available in **Fig. S13A-C**).

C. BRET assays for CCL25-induced Arr3 recruitment on WT CCR9 and CCR9 mutants. Data points represent mean ± SEM of BRET signal (**eq. 4**) obtained in triplicate wells; data shown is representative of 3 independent experiments (complete data set is available in **Fig. S13D-F**).

D. Radar plot summarizing the impact of Y126$^{3.32}$A, Q267$^{6.48}$A, N271$^{6.52}$A, and F299$^{7.35}$A on CCL25 binding and signaling. Impact is quantified relative to WT CCR9 as in **Fig. 2A**. Black equilateral triangle denotes WT CCR9; nearly equilateral (for Y126$^{3.32}$A and F299$^{7.35}$A) vs strongly asymmetric (for Q267$^{6.48}$A and N271$^{6.52}$A) mutant contours emphasize balanced vs biased impact of the corresponding mutations in the three assays.

E. Location of CCR9 Y126$^{3.32}$ and F299$^{7.35}$ (sticks and spheres colored in cyan and navy) in relation to the chemokine binding cavity (transparent mesh) and the N-terminus of the bound chemokine (black ribbon and sticks) in the predicted CCR9-CCL25 complex structure. The model is viewed along the plane of the membrane in the TM2-to-TM5 direction. TMs 1 and 2 are omitted for clarity.

F-G. Location of CCR9 Q267$^{6.48}$ and N271$^{6.52}$ (sticks and spheres colored in dark-red and brown) in relation to the chemokine binding cavity (transparent mesh, **F**), the 30s loop of the bound chemokine (black ribbon, **F**), and Y126$^{3.32}$/F299$^{7.35}$ in the predicted CCR9-CCL25 complex structure (**F**) or the X-ray structure of inactive vercirnon-bound CCR9 (PDB entry 5LWE [7], **G**). Both model and structure are viewed along the plane of the membrane in the TM5-to-TM2 direction. TMs 4 and 5 are omitted for clarity.

Next, we turned our attention to the residues deeper in the binding pocket - Y126$^{3.32}$ and F299$^{7.35}$, - and in the middle of the TM domain of CCR9 - Q267$^{6.48}$ and N271$^{6.52}$.

Y126$^{3.32}$ is located at the "floor" of the binding pocket, is highly conserved across the chemokine receptor family, and has been shown to be a critical signal initiation determinant as its mutations abrogate signaling in many receptors [54, 56-70]. Surprisingly, the CCR9 Y126$^{3.32}$A mutation did not have a negative impact; instead, it led to a striking increase (2.5-fold) in CCL25 binding (**Fig. 5A, D** and **Fig. S12**) accompanied by modest increases in Ca^{2+} signaling and arrestin recruitment (**Fig. 5B-D and Fig. S13**). A 2.2-fold increase in CCL25 binding was also observed for the F299$^{7.35}$A mutant (**Fig. 5A, D** and **Fig. S12**), with no effects on signaling (**Fig. 5B, D** and **Fig. S13**). In the model, Y126$^{3.32}$ is in direct contact with the N-terminal pyroglutamate (pGlu1) of CCL25 (expected to form from Q1 in the mature chemokine) while F299$^{7.35}$ T-stacks with CCL25 residue F4 (**Fig. 5E-F**). The observation that these two residues can be eliminated without negatively impacting CCL25 signaling suggests that, in contrast to other receptor-chemokine pairs [18], interactions between the distal N-terminus of CCL25 and the floor of CCR9 do not play a major role in either receptor binding or activation.

The remaining two residues, Q267$^{6.48}$ and N271$^{6.52}$, belong to TM6 and are positioned below the pocket "floor" (**Fig. 5F**). Q267$^{6.48}$ corresponds to the "toggle-switch" residue and is a tryptophan in most GPCRs. Its substitution by alanine substantially increased CCL25 binding and fully preserved Ca^{2+} signaling, while reducing Arr3 recruitment to 29% of the WT signal (**Fig. 5A-D, S11** and **S12**); the full G protein signaling capacity of this mutant was corroborated in a non-amplified G protein subunit dissociation assay (**Fig. S14**). N271$^{6.52}$A had no effect on CCL25 binding, but had an opposite effect compared to Q267$^{6.48}$, enhancing Arr3 recruitment and almost completely abrogating Ca^{2+} signaling (15% of WT response remaining, **Fig. 5A-D**, **Fig. S11** and **S13**).

The observed enhancement of chemokine binding and, in some mutants, signaling through one or both pathways could not be explained by improved intermolecular contacts: according to the structural model, the Y126$^{3.32}$A and F299$^{7.35}$A mutations eliminate contacts with the chemokine, and for residues Q267$^{6.48}$ and N271$^{6.52}$, direct chemokine contact is not possible at all. We therefore considered an alternative explanation in which the mutations promote chemokine binding and signaling by shifting the CCR9 conformational equilibrium towards the active, chemokine-compatible state. Indeed, in the inactive CCR9 structure [31], residues Y126$^{3.32}$, Q267$^{6.48}$, and N271$^{6.52}$ are proximal to each other and form numerous steric and polar contacts (**Fig. 5G**); their separation is exclusive to the active state of CCR9 (**Fig. 5F**), is concurrent with the outward movement of TM6 and TM7 **(Fig. 1C**), and thus can serve as a marker of CCR9 activation, as suggested by **Fig. 1B-D** and **Fig. 4**. By disrupting the Y126$^{3.32}$-Q267$^{6.48}$-N271$^{6.52}$ interaction cluster (**Fig. 5G**) and eliminating the inactive-state-specific cross-pocket coordination, alanine mutations of the participating residues are likely to destabilize the inactive state of CCR9 and make the active state more prevalent even in the absence of

chemokine, i.e. introduce receptor constitutive activity. Consistent with this, all four mutants showed significantly increased association with Arr3 in the absence of chemokine, as indicated by elevated basal BRET in the HEK-CCR9-RLuc8 cell lines (**Fig. S4**). The relative assessment of surface (by flow cytometry) vs total (by luminometry) receptor levels (**Fig. S2D**) suggested that Y126$^{3.32}$A and F299$^{7.35}$A, but not Q267$^{6.48}$A or N271$^{6.52}$A, are predominantly intracellular, indicative of constitutive internalization: a common feature of constitutively active, Arr3-associated receptors [71].

Beyond predicting this constitutive activity, the computational models were unable to explain the striking signaling bias of the Q276$^{6.48}$A and N271$^{6.52}$A mutants. However, the single helical turn that separates Q267$^{6.48}$ and N271$^{6.52}$ in the CCR9 structure harbors P269$^{6.50}$, the most conserved amino acid in TM6 of class A GPCRs and the core of the TM6 kink. Moreover, mutations of Q267$^{6.48}$ and N271$^{6.52}$ selectively eliminate TM6 contacts with TM7 and TM5, respectively (**Fig. S15**). Therefore, we hypothesize that the loss of these contacts alters the conformational coupling between the binding site and the intracellular effector interface in a manner that preferentially affects Arr3 or G protein.

Collectively, these results indicate that, in contrast to many other chemokine receptors, the residues at or below the floor of the CCR9 binding pocket are not the determinants of chemokine binding. However, they appear to subtly control the conformational preferences of the receptor, not only in the active-inactive spectrum but also in effector selectivity.

Partial alanine scanning of CCL25 confirms the key role of its 30s loop in CCR9 activation

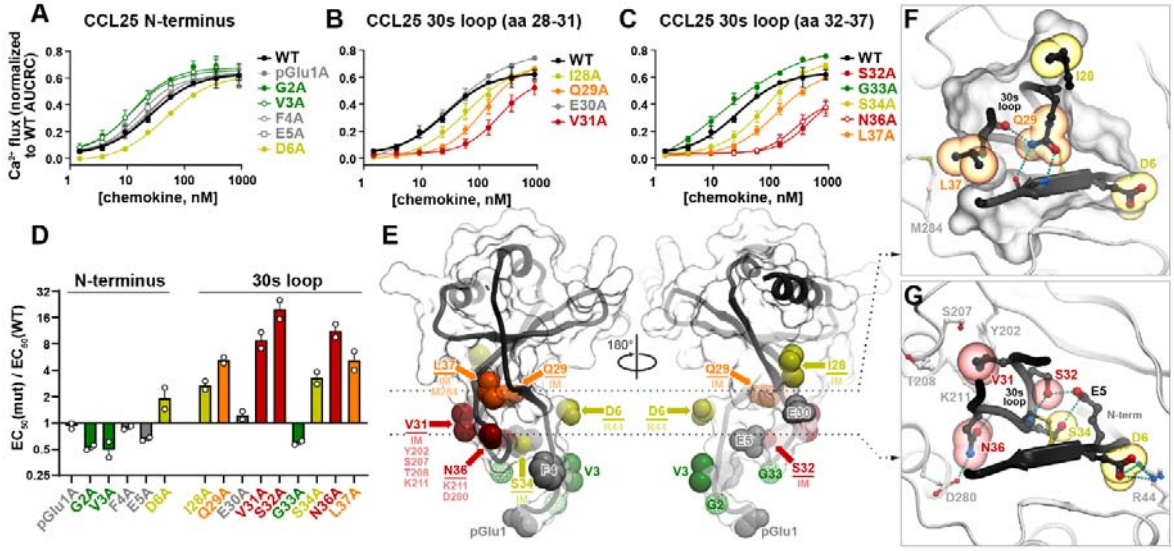

Figure 6. Pharmacological evaluation of CCL25 N-terminus and 30s loop mutants. Statistical analyses are shown in **Table S4**.

A-C. Concentration-dependent Ca^{2+} signaling responses of MOLT-4 cells to WT CCL25 and the indicated CCL25 mutants. Data points represent mean normalized peak height responses at the indicated ligand concentrations from 2 independent experiments (shown in **Fig. S16**).

D. Bar graph summarizing the data in **A-C**. Data points represent mean ratios (EC_{50} of CCL25 mutants / EC_{50} of WT CCL25) ± SEM from 2 independent experiments. Bars are colored red-to-green based on the degree of the mutation impact on chemokine signaling potency.

E. Mutagenesis results projected on the 3D model of CCL25 from the CCR9-CCL25 complex. The chemokine is shown as a black ribbon and a transparent surface mesh. For mutated residues, the atoms eliminated (or, for glycine residues, introduced) as a result of an alanine mutation are shown as spheres and colored in accordance with **A-D**. For each amino-acid residue with a measurable negative impact, its major intra- and intermolecular interactions are labeled. IM: intramolecular interaction.

F-G. Focused view of the intra- and intermolecular interactions involving the chemokine residues whose mutations have a negative impact on CCR9 signaling: D6, I28, Q29, and L37 (**F**) and V31, S32, S34, and N36 (**G**). The chemokine is shown in black ribbon and sticks (and a surface mesh in **F**), mutated residues in spheres colored as in **A-E**. The receptor is shown in white ribbon and sticks. The complex is viewed across the plane of the membrane from the extracellular side; **F** and **G** correspond to the cross-sectional planes indicated by dashed lines in **E**. Hydrogen bonds are shown as cyan dotted lines.

To further probe the key sites of interaction in the active CCR9-CCL25 complex, we generated a series of CCL25 alanine mutants targeting the N-terminal region (positions 1-6) and the 30s loop (positions 28-37). These mutants were assessed for their capacity to activate CCR9 using a Ca^{2+} flux assay on the MOLT-4 human T leukemia cell line, which endogenously expresses CCR9 [72] (**Fig. 6A-D, Fig. S16**).

Within the N-terminal region of CCL25, only mutation of D6 led to a decrease in signaling; G2A and V3A enhanced signaling and all other mutations had no impact (**Fig. 6A, D**). In contrast,

seven out of nine alanine mutations in the 30s loop reduced signaling, while only G33A enhanced signaling and E30A had no effect. The mutations only affected signaling potency (EC_{50}, **Fig. 6D**) but not efficacy (E_{max}, **Fig. 6A-C**), suggesting that CCR9 binding affinity was modulated but the ability to promote full agonism was preserved.

According to the structural model, CCL25 D6 and N36 are integral parts of the hydrogen bond networks surrounding CCR9 $R44^{1.28}$ (**Fig. 3J**) and $K211^{5.35}$ (**Fig. 3I**), respectively, while CCL25 V31 provide important steric packing interactions for CCR9 $Y202^{ECL2}$, $S207^{5.31}$, $T208^{5.32}$, and $K211^{5.35}$ (**Fig. 6E-G**). CCL25 S32 forms an intramolecular hydrogen bond network with CCL25 S34 and CCL25 E5 (**Fig. 6G**), likely contributing to the stabilization of CCL25 in a signaling-productive conformation. The aliphatic sidechain of L37 contributes to intramolecular packing and is also in contact with CCR9 residue $M284^{ECL3}$ (**Fig. 6F**), which is proximal to CCL25 N36 and CCR9 $T208^{5.32}$ and $D280^{6.61}$ (**Fig. 4G**). This suggests a role for L37 in stabilizing the TM5/6-30s loop interaction. The side chain of Q29 is buried in the chemokine core, forming two hydrogen bonds with the backbone of the proximal N-terminus C8 and one additional bond with the backbone of L37, thus stabilizing the 30s loop shape (**Fig. 6F**). The CCL25 30s loop residues with lower mutation impact also form prominent intramolecular interactions: the aliphatic side chain of I28 packs against the surrounding chemokine residues (**Fig. 6F**) and S34 hydrogen-bonds to CCL25 E5 (**Fig. 6G**), and together with S32 they likely stabilize CCL25 in a signaling-productive conformation.

These results provide further evidence that it is the 30s loop rather than the N-terminus of CCL25 that carries the critical determinants of CCR9 activation by forming key hydrogen bonding networks and hydrophobic interactions with residues in CCR9 TM5 and TM6.

Discovery of a 'super-binder' CCL25 analog [1P6]CCL25

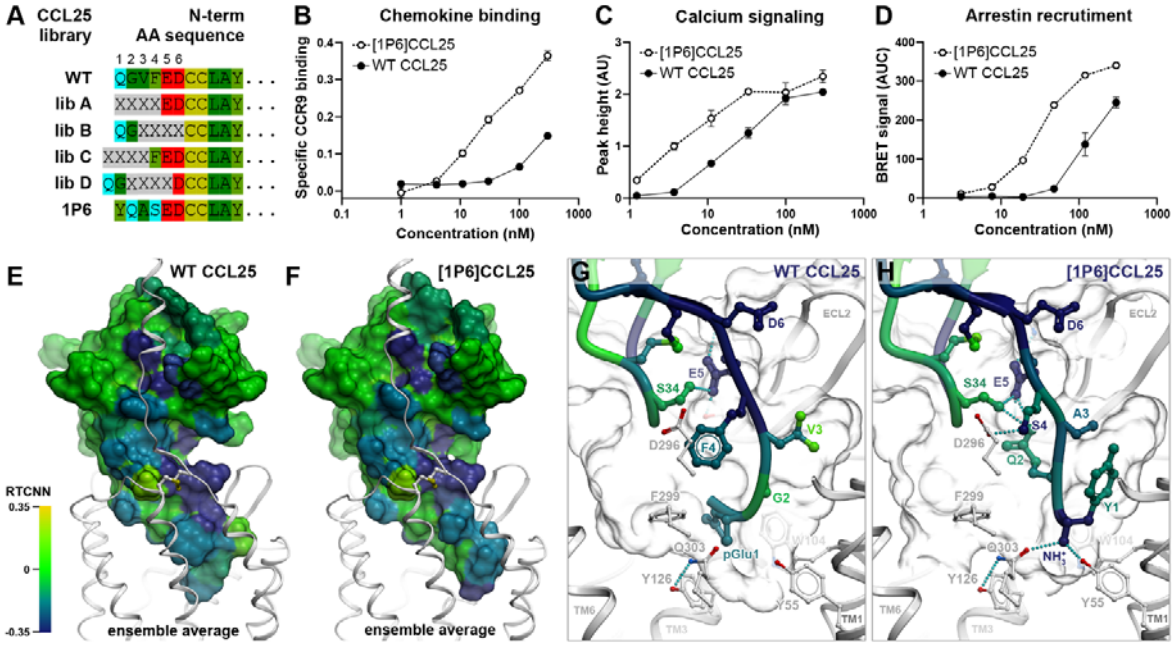

Figure 7. [1P6]CCL25 is an N-terminally engineered super-binder and super-agonist analog of CCL25.

A. CCL25 phage libraries used in this study, indicating N-terminal sequences of native CCL25 and the analog [1P6]CCL25. X = fully randomized residue; residues from Cys7 correspond to CCL25(7-73). N-terminal Q is expected to be cyclized to pGlu in the mature protein. Analogs sequences are available in **Table S5**.

B. CCL25 and [1P6]CCL25 binding to WT CCR9 stably expressed in HEK293 cells. HEK293 parental cell line and HEK293 cells stably expressing WT CCR9 were incubated at 4°C for 1h with the indicated concentrations of TAMRA-labelled CCL25 or [1P6]CCL25. Data represent mean ± SEM of specific binding signal from triplicate wells; the data shown are representative of 3 independent experiments (complete data set is available in **Fig. S17A-C**).

C. CCL25 and [1P6]-CCL25-induced Ca^{2+} signaling on WT CCR9 expressed in HEK293 cells. Ca^{2+} signals in response to CCL25 and [1P6]CCL25 at the indicated concentrations are shown as mean peak height (**eq. 3**) ± SEM from triplicate wells; the data shown are representative of 3 independent experiments (complete data set is available in **Fig. S17D-F**).

D. BRET assays for CCL25 and [1P6]-CCL25-induced Arr3 recruitment on WT CCR9. Data points represent mean ± SEM of BRET signal (**eq. 4**) obtained in triplicate wells; data shown is representative of 3 independent experiments (complete data set is available in **Fig. S17G-I**).

E-F. Overall view of top-scoring predicted conformations of CCL25 (**E**) and [1P6]CCL25 (**F**) (rainbow-colored surfaces) bound to CCR9 (white ribbons). Chemokines are colored by per-atom RTCNN scores averaged across the top 5 models from each model ensemble and aggregated to residue backbones and side chains. The complexes are viewed along the plane of the membrane.

G-H. Zoomed-in views of the chemokine N-termini in the top-scoring models of CCR9-CCL25 (**G**) and CCR9-[1P6]CCL25 (**H**) complexes. Chemokines are shown as ribbons, important residues discussed in the text as sticks. Color represents per-atom RTCNN scores aggregated to residue backbones (for ribbons and backbone sticks) and side chains (for side chain sticks). Ribbons for receptor TM helices 7 and (partially) 1 are omitted for clarity.

To further understand the structural regions that drive the capacity of CCL25 to bind and activate CCR9, we used a previously described phage display-based chemokine engineering strategy [34]. From libraries with diversity introduced into various N-terminal positions (**Table S5** and **Fig. S18**) we identified [1P6]CCL25, an analog with enhanced receptor engagement capacity, that contains N-terminal residues Y1-Q2-A3-S4 in place of pGlu1-G2-V3-F4. (**Fig. 7A**). Compared to WT CCL25, [1P6]CCL25 showed a substantial increase in CCR9 binding (**Fig. 7B** and **S17A-C**), similarly to previously reported phage-display-generated analogs of CCL5 [73-75] and CXCL12 [76]. [1P6]CCL25 also demonstrated enhanced CCR9 signaling activity, both in terms of Ca^{2+} mobilization and Arr3 recruitment (**Fig. 7C-D** and **S17D-I**).

The predicted binding mode of [1P6]CCL25 to CCR9 is similar to that of WT CCL25, with the residues shared between the two ligands making the same interactions with the receptor (**Fig. 7E-F**). However, notable differences are apparent in both the conformation of the distal N-terminus and the receptor interactions of the four substituted residues (**Fig. 7G-H**).

To provide a structural explanation for the enhanced binding and signaling activity of [1P6]CCL25, we used RTCNN [38-40], an AI-based scoring function for protein-ligand interactions. RTCNN scores of chemokine atoms were aggregated to residue backbones and side chains to estimate their respective contributions to the complex binding affinity. This suggested that the most favorable CCL25 N-terminus contributions are from residues E5 and D6 (the latter engaged with the CCR9 extracellular "rim" amino acid $R44^{1.28}$, **Fig. 3J**), which are shared by the two ligands (**Fig. 7G, H**). Per RTCNN, only two of the four N-terminal amino-acids, pGlu1 and F4, contribute favorably to binding of WT CCL25: pGlu1 stacks with $Y126^{3.32}$ and $W104^{2.60}$, whereas F4 T-stacks with $F299^{7.35}$. The roles of the remaining two WT CCL25 N-terminal residues, G2 and V3, are predicted to be neutral or minimally advantageous (**Fig. 7G**), in agreement with alanine mutagenesis results (**Fig. 6D**). In contrast, all four of the substituted N-terminal amino acids in [1P6]CCL25 contribute favorably to CCR9 binding. First, the positively charged N-terminal amine of Y1 (unavailable in WT CCL25 due to pGlu being cyclized) forms a cation-pi interaction with $W104^{2.60}$ and hydrogen-bonds to $Y55^{1.55}$ and $Q303^{7.39}$, while its side chain packs against TM1 ($A47^{1.31}$ and $L51^{1.35}$) and TM2 ($A107^{2.63}$ and $A108^{2.64}$) (**Fig. 7H**). Second, the side chain of Q2 is oriented towards the 30s loop and stabilizes it in a beneficial conformation through hydrogen bonding. Third, mutating V3 in WT CCL25 to A3 in [1P6]CCL25 reduces bulk to accommodate the Y1 sidechain. Finally, the side chain of [1P6]CCL25 S4 forms a hydrogen bond with $D296^{7.32}$.

These results demonstrate that engineering the N-terminal region of CCL25 can generate ligands with enhanced binding and signaling via the introduction of beneficial interactions with CRS2 of CCR9. However, in contrast with N-terminal molecular evolution studies of other chemokines [73, 76], our study did not yield receptor antagonists (**Table S6**, **Fig. S18**), suggesting that CCR9 activation determinants are located outside of the N-terminal domain of CCL25, likely in the 30s loop as suggested by the data in **Figs. 3-6**.

Super-binder [1P6]CCL25 is more tolerant to CCR9 mutations than WT CCL25

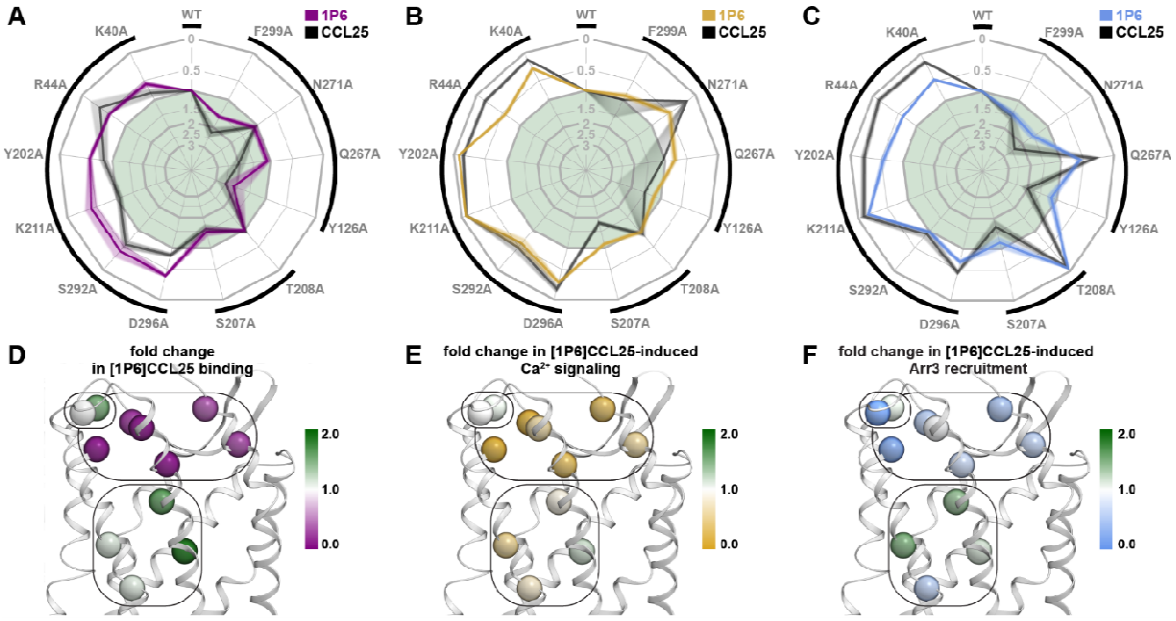

Figure 8. [1P6]CCL25 has increased tolerance to CRS2 mutations in CCR9.

A-C. Radar plots summarizing the impact of studied CCR9 mutations on [1P6]CCL25 binding (purple), [1P6]CCL25-induced intracellular Ca2+ mobilization (orange), and [1P6]CCL25-induced Arr3 recruitment to the receptor (blue). The impact of the same mutations on the binding and signaling of WT CCL25 (same as in **Fig. 2**) is shown in grayscale for reference. Mutation impacts are expressed as ratios of CCR9 mutant to WT responses in respective experiments; contours outside or inside the central light green area correspond to negative and positive impacts, respectively. Responses were determined as areas under the [1P6]CCL25 concentration-response curves (AUCRCs) for Ca^{2+} mobilization and Arr3 recruitment, and as receptor-specific cell fluorescence increases in the presence of 300 nM of TAMRA-labeled [1P6]CCL25 for binding. Data represent mean of 2-3 independent experiments; SEM values are represented by the contour width and transparency at the respective mutant axis. Black outside brackets denote the groups of functionally and structurally related mutations presented in **Figs. 3**, **4**, and **5**. Summary data and all individual replicates are shown from **Fig. S19** to **Fig. S25**. Statistics are available in **Table S6**.

D-F. Detrimental or beneficial impact of mutations at selected residues is reflected in the color of their Cα atoms (spheres), in relation to [1P6]CCL25 binding (**D**), [1P6]CCL25-induced intracellular Ca^{2+} mobilization (**E**), and [1P6]CCL25-inducedArr3 recruitment to CCR9 (**F**). The receptor is shown as white ribbons and viewed parallel to the plane of the membrane. Rounded rectangles mark the groups of functionally and structurally related mutations presented in this paper.

Next, we investigated the impact of the 12 CCR9 CRS2 point mutations (**Fig. 1E-F**) on the binding and signaling of the super-binder [1P6]CCL25 compared to WT CCL25. The mutations affected the two ligands in a broadly similar way, with the direction of impact (positive or negative) generally preserved for each mutation. However, ligand-specific differences in the magnitudes of the changes for certain mutations were observed (**Fig. 8**). In contrast to their varying negative effects on CCL25 binding (ranging from complete abrogation to no impact, **Fig. 3**), the six extracellular 'rim' mutations uniformly but only partially decreased [1P6]CCL25

binding (**Fig. 8A** and **Fig. S19**). For the remaining mutations, binding was either unaffected or enhanced for both ligands, with the relative enhancements less pronounced for [1P6]CCL25 (**Fig. 8A**). For Ca^{2+} mobilization and Arr3 recruitment, the positive and negative mutation impacts observed for WT CCL25 were either preserved or attenuated for [1P6]CCL25 (**Fig. 8B, C, E, F** and **Fig. S19**).

Collectively, these results suggest that the improved binding properties of [1P6]CCL25, mediated by the enhanced interactions of its distal N-terminus with CCR9 CRS2, not only make it a more potent agonist of the receptor but also more tolerant to CRS2 mutations.

Discussion

This study presents a comprehensive map of the interaction interface between CCR9 and its endogenous agonist CCL25, with delineation of determinants of binding, signaling, constitutive activity and bias. A key feature of our structural model is the depth to which the 30s loop of CCL25 accompanies the N-terminus into the TM domain of CCR9 (**Fig. 1A**). In this respect, the CCR9-CCL25 complex resembles the experimentally determined structures of CCR1, CCR2, and CCR5 with their respective chemokine agonists, but differs from those of CCR6, CXCR2, and CX3CR1, in which the 30s loops of the bound chemokines do not engage the TM domain, but instead interact with the extracellular loops [15, 77] (**Fig. S5E**). An extensive hydrogen bond network linking the 30s loop and the proximal N-terminus acts together with the first conserved disulfide bridge to fuse the two chemokine regions into a single structural unit.

Functional mapping of the CCR9-CCL25 interface revealed a noncanonical role for the CCL25 N-terminus. For many chemokines, the N-terminus serves as a 'message' for receptor activation [6-12]; however, in the case of CCL25 and CCR9, it appears to make a minimal contribution to either binding or signaling. Mutating receptor residues predicted to contact the N-terminus, or alanine substitution of CCL25 N-terminal residues, did not have any negative impact on chemokine binding or receptor activation (**Fig. 5-6**). However, increased binding could be achieved through molecular evolution of the CCL25 N-terminus, as illustrated by the potent chemokine analog [1P6]CCL25 (**Fig. 7**). Importantly, such molecular evolution did not produce changes in chemokine signaling unattributable to alterations in binding. This provides further evidence that the N-terminus of CCL25 is not the driver of signaling, and contrasts with findings for many other chemokines, where N-terminal modifications strongly affect signaling properties [8, 70, 73, 76, 78-80]. The only other known exception to the "N-terminal message" paradigm involves the atypical chemokine receptor ACKR3 which retains its functional response (Arr3 recruitment) to CXCL12 N-terminal mutants [76].

In support of a non-canonical signaling anatomy of the CCR9-CCL25 complex, we showed that the 30s loop of CCL25, rather than its N-terminus, is the principal determinant of receptor activation. The loop engages in extensive interactions with TM5, the domain that undergoes profound lateral and longitudinal motions upon receptor activation (**Fig. 1C-D**). Mutations of CCR9 TM5 and ECL2 residues that are in contact with the 30s loop abrogated signaling (K211$^{5.35}$ and Y202^{ECL2}) or made it strongly biased (T208$^{5.32}$), with no effect on chemokine binding. However, the 30s loop of CCL25 also plays a key role in receptor binding: the majority of 30s loop mutations led to significant decreases in Ca^{2+} flux potency but not efficacy, consistent with loss of receptor binding affinity.

Interpretation of GPCR structure-function studies is invariably complicated by the potential for individual receptor mutants to differentially impact aspects of receptor function such as constitutive activity, G protein and arrestin-coupling preferences, intracellular trafficking and signaling from subcellular compartments [73, 81-85]. In our study, we addressed this challenge

by assessing the impact of receptor and chemokine mutations on multiple aspects of receptor pharmacology. By systematically and quantitatively evaluating CCR9 mutation impacts on chemokine binding, calcium mobilization, and Arr3 recruitment we identified prominent examples of G protein-biased (T208$^{5.32}$A and Q267$^{6.48}$A) and Arr3-biased (N271$^{6.52}$A) mutants. Conscious of the potential amplification artefacts inherent to Ca^{2+} flux and other second messenger assays [70, 86-88] we confirmed full G protein competence of mutants (T208$^{5.32}$A and Q267$^{6.48}$A) in a non-amplified BRET-based Gαi/Gβγ dissociation assay. By measuring the basal arrestin association (**Fig. S4**), we identified mutants with high levels of constitutive activity (Y126$^{3.32}$A, F299$^{7.35}$A, Q267$^{6.48}$A and N271$^{6.52}$A), and by assessing surface and total receptor expression (**Fig. S2A-D**), inferred altered subcellular distribution for two of them (Y126$^{3.32}$A and F299$^{7.35}$A). Finally, all data was contextualized in an ensemble of structural models of the CCR9-CCL25 complex, allowing us to deconvolute molecular mechanisms underlying the altered pharmacology of the mutants.

The structural ensembles [22, 24] were made possible through the use of AF2, a technique known to generate near-experimental accuracy models for GPCR-peptide complexes [21]. Expanding on the previously solved X-ray structure of inactive, antagonist-bound CCR9 ([31], **Fig. 1B-D**), AF2 model ensemble provided a plausible basis for partial vs full activation of CCR9 (**Fig 4H-J**) and the requirements for Arr3 recruitment (**Fig. 4**). Nonetheless, certain conformational states of the complex remained inaccessible to AF2 modeling and so did the entropic component of binding. As a consequence, we were only able to provide tentative explanations for the biased activity of three receptor mutants (T208$^{5.32}$A, Q267$^{6.48}$A and N271$^{6.51}$A) and the affinity improvements of selected chemokine variants (G2A, V3A, and G33A). Future studies using molecular dynamics may provide more detailed answers to these questions.

In summary, this study reveals that CCR9-CCL25 is a receptor-chemokine pair with non-canonical mechanisms of engagement and signaling, adding diversity to the established two-site model of chemokine receptor activation. Our results suggest that engineering of the 30s loop of CCL25 can yield potent CCR9 modulators with tunable signaling activity and add to the growing portfolio of chemokine analogs suitable for clinical development [89]. They also provide insights for structure-based design of small molecule therapeutics for CCR9-related pathologies.

References

1. Griffith, J.W., C.L. Sokol, and A.D. Luster, *Chemokines and chemokine receptors: positioning cells for host defense and immunity.* Annual review of immunology, 2014. **32**: p. 659-702.
2. Proudfoot, A.E., *Chemokine receptors: multifaceted therapeutic targets.* Nat Rev Immunol, 2002. **2**(2): p. 106-15.
3. Weis, W.I. and B.K. Kobilka, *The molecular basis of G protein–coupled receptor activation.* Annual review of biochemistry, 2018. **87**: p. 897.
4. DeWire, S.M., et al., *β-Arrestins and Cell Signaling.* Annual Review of Physiology, 2007. **69**(Volume 69, 2007): p. 483-510.
5. Isberg, V., et al., *GPCRdb: an information system for G protein-coupled receptors.* Nucleic acids research, 2016. **44**(D1): p. D356-D364.
6. Monteclaro, F.S. and I.F. Charo, *The amino-terminal extracellular domain of the MCP-1 receptor, but not the RANTES/MIP-1alpha receptor, confers chemokine selectivity. Evidence for a two-step mechanism for MCP-1 receptor activation.* J Biol Chem, 1996. **271**(32): p. 19084-92.
7. Wells, T., et al., *The Molecular Basis of the Chemokine/Chemokine Receptor Interaction - Scope for Design of Chemokine Antagonists.* Methods, 1996. **10**(1): p. 126-34.
8. Crump, M.P., et al., *Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1.* The EMBO journal, 1997. **16**(23): p. 6996-7007.
9. Pease, J.E., et al., *The N-terminal extracellular segments of the chemokine receptors CCR1 and CCR3 are determinants for MIP-1alpha and eotaxin binding, respectively, but a second domain is essential for efficient receptor activation.* J Biol Chem, 1998. **273**(32): p. 19972-6.
10. Mayer, M.R. and M.J. Stone, *Identification of Receptor Binding and Activation Determinants in the N-terminal and N-loop Regions of the CC Chemokine Eotaxin*.* Journal of Biological Chemistry, 2001. **276**(17): p. 13911-13916.
11. Xanthou, G., T.J. Williams, and J.E. Pease, *Molecular characterization of the chemokine receptor CXCR3: evidence for the involvement of distinct extracellular domains in a multi-step model of ligand binding and receptor activation.* European Journal of Immunology, 2003. **33**(10): p. 2927-2936.
12. Kufareva, I., C.L. Salanga, and T.M. Handel, *Chemokine and chemokine receptor structure and interactions: implications for therapeutic strategies.* Immunology and cell biology, 2015. **93**(4): p. 372-383.
13. Sanchez, J., et al., *Evaluation and extension of the two-site, two-step model for binding and activation of the chemokine receptor CCR1.* Journal of Biological Chemistry, 2019. **294**(10): p. 3464-3475.
14. Shao, Z., et al., *Molecular insights into ligand recognition and activation of chemokine receptors CCR2 and CCR3.* Cell Discovery, 2022. **8**(1): p. 44.
15. Kleist, A.B., et al., *New paradigms in chemokine receptor signal transduction: Moving beyond the two-site model.* Biochem Pharmacol, 2016. **114**: p. 53-68.
16. Zheng, Y., et al., *Structure of CC Chemokine Receptor 5 with a Potent Chemokine Antagonist Reveals Mechanisms of Chemokine Recognition and Molecular Mimicry by HIV.* Immunity, 2017. **46**(6): p. 1005-1017.e5.
17. Isaikina, P., et al., *Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist.* Science advances, 2021. **7**(25): p. eabg8685.

18. Urvas, L. and E. Kellenberger, *Structural Insights into Molecular Recognition and Receptor Activation in Chemokine–Chemokine Receptor Complexes.* Journal of Medicinal Chemistry, 2023. **66**(11): p. 7070-7085.
19. Jumper, J., et al., *Highly accurate protein structure prediction with AlphaFold.* Nature, 2021. **596**(7873): p. 583-589.
20. Senior, A.W., et al., *Improved protein structure prediction using potentials from deep learning.* Nature, 2020. **577**(7792): p. 706-710.
21. Chitsazi, R., et al., *GPCR Dock 2021: A Blind Docking Competition in the AlphaFold2 Era.* in preparation.
22. Guo, H., A. Perminov, and S. Bekele, *AlphaFold2 models indicate that protein sequence determines both structure and dynamics. Sci Rep 12: 10696*. 2022.
23. Heo, L. and M. Feig, *Multi-state modeling of G-protein coupled receptors at experimental accuracy.* Proteins: Structure, Function, and Bioinformatics, 2022. **90**(11): p. 1873-1885.
24. Wayment-Steele, H.K., et al., *Predicting multiple conformations via sequence clustering and AlphaFold2.* Nature, 2023: p. 1-3.
25. Svensson, M. and W.W. Agace, *Role of CCL25/CCR9 in immune homeostasis and disease.* Expert review of clinical immunology, 2006. **2**(5): p. 759-773.
26. Wermers, J.D., et al., *The chemokine receptor CCR9 is required for the T-cell–mediated regulation of chronic ileitis in mice.* Gastroenterology, 2011. **140**(5): p. 1526-1535. e3.
27. Wu, W., et al., *Strong expression of chemokine receptor CCR9 in diffuse large B-cell lymphoma and follicular lymphoma strongly correlates with gastrointestinal involvement.* Human pathology, 2014. **45**(7): p. 1451-1458.
28. Igaki, K., et al., *MLN3126, an antagonist of the chemokine receptor CCR9, ameliorates inflammation in a T cell mediated mouse colitis model.* International immunopharmacology, 2018. **60**: p. 160-169.
29. Xu, B., et al., *CCR9 and CCL25: A review of their roles in tumor promotion.* J Cell Physiol, 2020. **235**(12): p. 9121-9132.
30. Chen, H., et al., *Intratumoral delivery of CCL25 enhances immunotherapy against triple-negative breast cancer by recruiting CCR9+ T cells.* Science advances, 2020. **6**(5): p. eaax4690.
31. Oswald, C., et al., *Intracellular allosteric antagonism of the CCR9 receptor.* Nature, 2016. **540**(7633): p. 462-465.
32. Paolini-Bertrand, M., et al., *Rapid and low-cost multiplex synthesis of chemokine analogs.* J Biol Chem, 2018. **293**(49): p. 19092-19100.
33. Eberhardson, M., et al., *Treatment of inflammatory bowel disease by chemokine receptor-targeted leukapheresis.* Clinical Immunology, 2013. **149**(1): p. 73-82.
34. Dorgham, K., et al., *Generating Chemokine Analogs with Enhanced Pharmacological Properties Using Phage Display.* Methods Enzymol, 2016. **570**: p. 47-72.
35. Salmon, P. and D. Trono, *Production and titration of lentiviral vectors.* Current protocols in human genetics, 2007. **54**(1): p. 12.10. 1-12.10. 24.
36. Evans, R., et al., *Protein complex prediction with AlphaFold-Multimer.* biorxiv, 2021: p. 2021.10. 04.463034.
37. Abagyan, R. and M. Totrov, *Biased probability Monte Carlo conformational searches and electrostatic calculations for peptides and proteins.* Journal of molecular biology, 1994. **235**(3): p. 983-1002.
38. Raush, E. and M. Totrov, *RTCNN Performance (CASF 2016 pose rank benchmark)*, in *Molsoft ICM User Group Meeting*. 2023: San Diego, CA.

39. Totrov, M., *New developments in ICM: neural networks and beyond*, in *Molsoft ICM User Group Meeting*. 2023: San Diego, CA, USA.
40. Dawson, J.R.D., et al., *Molecular determinants of antagonist interactions with chemokine receptors CCR2 and CCR5.* bioRxiv, 2024.
41. Abagyan, R. and M. Totrov, *High-throughput docking for lead generation.* Curr Opin Chem Biol, 2001. **5**(4): p. 375-82.
42. Neves, M.A., M. Totrov, and R. Abagyan, *Docking and scoring with ICM: the benchmarking results and strategies for improvement.* Journal of computer-aided molecular design, 2012. **26**: p. 675-686.
43. Abdi, H., *Bonferroni and Šidák corrections for multiple comparisons.* Encyclopedia of measurement and statistics, 2007. **3**(01): p. 2007.
44. Hunter, J.D., *Matplotlib: A 2D graphics environment.* Computing in science & engineering, 2007. **9**(03): p. 90-95.
45. Wasilko, D.J., et al., *Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20.* Nature Communications, 2020. **11**(1): p. 3031.
46. Liu, K., et al., *Structural basis of CXC chemokine receptor 2 activation and signalling.* Nature, 2020. **585**(7823): p. 135-140.
47. Zhang, H., et al., *Structural basis for chemokine recognition and receptor activation of chemokine receptor CCR5.* Nature communications, 2021. **12**(1): p. 4151.
48. Shao, Z., et al., *Identification and mechanism of G protein-biased ligands for chemokine receptor CCR1.* Nature chemical biology, 2022. **18**(3): p. 264-271.
49. Lu, M., et al., *Activation of the human chemokine receptor CX3CR1 regulated by cholesterol.* Science Advances, 2022. **8**(26): p. eabn8048.
50. Ballesteros, J.A. and H. Weinstein, *Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors*, in *Methods in neurosciences*. 1995, Elsevier. p. 366-428.
51. Filipek, S., *Molecular switches in GPCRs.* Current Opinion in Structural Biology, 2019. **55**: p. 114-120.
52. Holst, B., et al., *A Conserved Aromatic Lock for the Tryptophan Rotameric Switch in TM-VI of Seven-transmembrane Receptors 2.* Journal of Biological Chemistry, 2010. **285**(6): p. 3973-3985.
53. Schwartz, T.W., et al., *Molecular mechanism of 7TM receptor activation—a global toggle switch model.* Annu. Rev. Pharmacol. Toxicol., 2006. **46**: p. 481-519.
54. Wescott, M.P., et al., *Signal transmission through the CXC chemokine receptor 4 (CXCR4) transmembrane helices.* Proceedings of the National Academy of Sciences, 2016. **113**(35): p. 9928-9933.
55. Kato, H.E., et al., *Conformational transitions of a neurotensin receptor 1-Gi1 complex.* Nature, 2019. **572**(7767): p. 80-85.
56. Garcia-Perez, J., et al., *Allosteric model of maraviroc binding to CC chemokine receptor 5 (CCR5).* J Biol Chem, 2011. **286**(38): p. 33409-21.
57. Berkhout, T.A., et al., *CCR2: Characterization of the Antagonist Binding Site from a Combined Receptor Modeling/Mutagenesis Approach.* Journal of Medicinal Chemistry, 2003. **46**(19): p. 4070-4086.
58. Gavrilin, M.A., et al., *Site-directed mutagenesis of CCR2 identified amino acid residues in transmembrane helices 1, 2, and 7 important for MCP-1 binding and biological functions.* Biochemical and Biophysical Research Communications, 2005. **327**(2): p. 533-540.
59. Govaerts, C., et al., *Activation of CCR5 by Chemokines Involves an Aromatic Cluster between Transmembrane Helices 2 and 3 **.* Journal of Biological Chemistry, 2003.

278(3): p. 1892-1903.

60. Hall, S.E., et al., *Elucidation of Binding Sites of Dual Antagonists in the Human Chemokine Receptors CCR2 and CCR5.* Molecular Pharmacology, 2009. **75**(6): p. 1325.
61. Hébert, C.A., et al., *Partial functional mapping of the human interleukin-8 type A receptor. Identification of a major ligand binding domain.* Journal of Biological Chemistry, 1993. **268**(25): p. 18549-18553.
62. Jensen, P.C., et al., *Molecular Interaction of a Potent Nonpeptide Agonist with the Chemokine Receptor CCR8.* Molecular Pharmacology, 2007. **72**(2): p. 327.
63. Kondru, R., et al., *Molecular Interactions of CCR5 with Major Classes of Small-Molecule Anti-HIV CCR5 Antagonists.* Molecular Pharmacology, 2008. **73**(3): p. 789.
64. Rosenkilde, M.M., et al., *Activation of the CXCR3 Chemokine Receptor through Anchoring of a Small Molecule Chelator Ligand between TM-III, -IV, and -VI.* Molecular Pharmacology, 2007. **71**(3): p. 930.
65. Scholten, D.J., et al., *Identification of Overlapping but Differential Binding Sites for the High-Affinity CXCR3 Antagonists NBI-74330 and VUF11211.* Molecular Pharmacology, 2014. **85**(1): p. 116.
66. Thiele, S., et al., *Determination of the binding mode for the cyclopentapeptide CXCR4 antagonist FC131 using a dual approach of ligand modifications and receptor mutagenesis.* British Journal of Pharmacology, 2014. **171**(23): p. 5313-5329.
67. Wong, R.S.Y., et al., *Comparison of the Potential Multiple Binding Modes of Bicyclam, Monocylam, and Noncyclam Small-Molecule CXC Chemokine Receptor 4 Inhibitors.* Molecular Pharmacology, 2008. **74**(6): p. 1485.
68. Zachariassen, Z.G., et al., *Probing the Molecular Interactions between CXC Chemokine Receptor 4 (CXCR4) and an Arginine-Based Tripeptidomimetic Antagonist (KRH-1636).* Journal of Medicinal Chemistry, 2015. **58**(20): p. 8141-8153.
69. Nedjai, B., et al., *CXCR3 antagonist VUF10085 binds to an intrahelical site distinct from that of the broad spectrum antagonist TAK-779.* British Journal of Pharmacology, 2015. **172**(7): p. 1822-1833.
70. Stephens, B.S., et al., *Functional anatomy of the full-length CXCR4-CXCL12 complex systematically dissected by quantitative model-guided mutagenesis.* Science signaling, 2020. **13**(640): p. eaay5024.
71. Gilliland, C.T., et al., *The chemokine receptor CCR1 is constitutively active, which leads to G protein-independent, β-arrestin-mediated internalization.* J Biol Chem, 2013. **288**(45): p. 32194-32210.
72. Maciocia, P.M., et al., *Anti-CCR9 chimeric antigen receptor T cells for T-cell acute lymphoblastic leukemia.* Blood, The Journal of the American Society of Hematology, 2022. **140**(1): p. 25-37.
73. Gaertner, H., et al., *Highly potent, fully recombinant anti-HIV chemokines: reengineering a low-cost microbicide.* Proceedings of the National Academy of Sciences, 2008. **105**(46): p. 17706-17711.
74. Scurci, I., et al., *CCR5 tyrosine sulfation heterogeneity generates cell surface receptor subpopulations with different ligand binding properties.* Biochim Biophys Acta Gen Subj, 2021. **1865**(1): p. 129753.
75. Pinheiro, I., et al., *Arylsulfatases and neuraminidases modulate engagement of CCR5 by chemokines by removing key electrostatic interactions.* Scientific Reports, 2024. **14**(1): p. 292.
76. Hanes, M.S., et al., *Dual targeting of the chemokine receptors CXCR4 and ACKR3 with novel engineered chemokines.* Journal of Biological Chemistry, 2015. **290**(37): p. 22385-

22397.
77. Wedemeyer, M.J., et al., *The chemokine X-factor: Structure-function analysis of the CXC motif at CXCR4 and ACKR3.* J Biol Chem, 2020. **295**(40): p. 13927-13939.
78. Choi, W.-T., et al., *CCR5 mutations distinguish N-terminal modifications of RANTES (CCL5) with agonist versus antagonist activity.* Journal of Virology, 2012. **86**(18): p. 10218-10220.
79. Chevigné, A., et al., *Engineering and screening the N-terminus of chemokines for drug discovery.* Biochemical pharmacology, 2011. **82**(10): p. 1438-1456.
80. Paavola, C.D., et al., *Monomeric monocyte chemoattractant protein-1 (MCP-1) binds and activates the MCP-1 receptor CCR2B.* J Biol Chem, 1998. **273**(50): p. 33157-65.
81. Masuho, I., et al., *Distinct profiles of functional discrimination among G proteins determine the actions of G protein–coupled receptors.* Science signaling, 2015. **8**(405): p. ra123-ra123.
82. Irannejad, R., et al., *Conformational biosensors reveal GPCR signalling from endosomes.* Nature, 2013. **495**(7442): p. 534-8.
83. Gu, S., et al., *Ligand-Dependent Mechanisms of CC Chemokine Receptor 5 (CCR5) Trafficking Revealed by APEX2 Proximity Labeling Proteomics.* bioRxiv, 2023: p. 2023.11.01.565224.
84. Escola, J.-M., et al., *CC chemokine receptor 5 (CCR5) desensitization: cycling receptors accumulate in the trans-Golgi network.* The Journal of biological chemistry, 2010. **285**(53): p. 41772-41780.
85. Inoue, A., et al., *Illuminating G-protein-coupling selectivity of GPCRs.* Cell, 2019. **177**(7): p. 1933-1947. e25.
86. Ehlert, F.J., et al., *A Simple Method for Estimation of Agonist Activity at Receptor Subtypes: Comparison of Native and Cloned M₃ Muscarinic Receptors in Guinea Pig Ileum and Transfected Cells.* Journal of Pharmacology and Experimental Therapeutics, 1999. **289**(2): p. 981.
87. Sykes, D.A., M.R. Dowling, and S.J. Charlton, *Exploring the mechanism of agonist efficacy: a relationship between efficacy and agonist dissociation rate at the muscarinic M3 receptor.* Mol Pharmacol, 2009. **76**(3): p. 543-51.
88. Rajagopal, S., et al., *Quantifying ligand bias at seven-transmembrane receptors.* Molecular pharmacology, 2011. **80**(3): p. 367-377.
89. Julian, D.R., et al., *Chemokine-Based Therapeutics for the Treatment of Inflammatory and Fibrotic Convergent Pathways in COVID-19.* Curr Pathobiol Rep, 2021. **9**(4): p. 93-105.