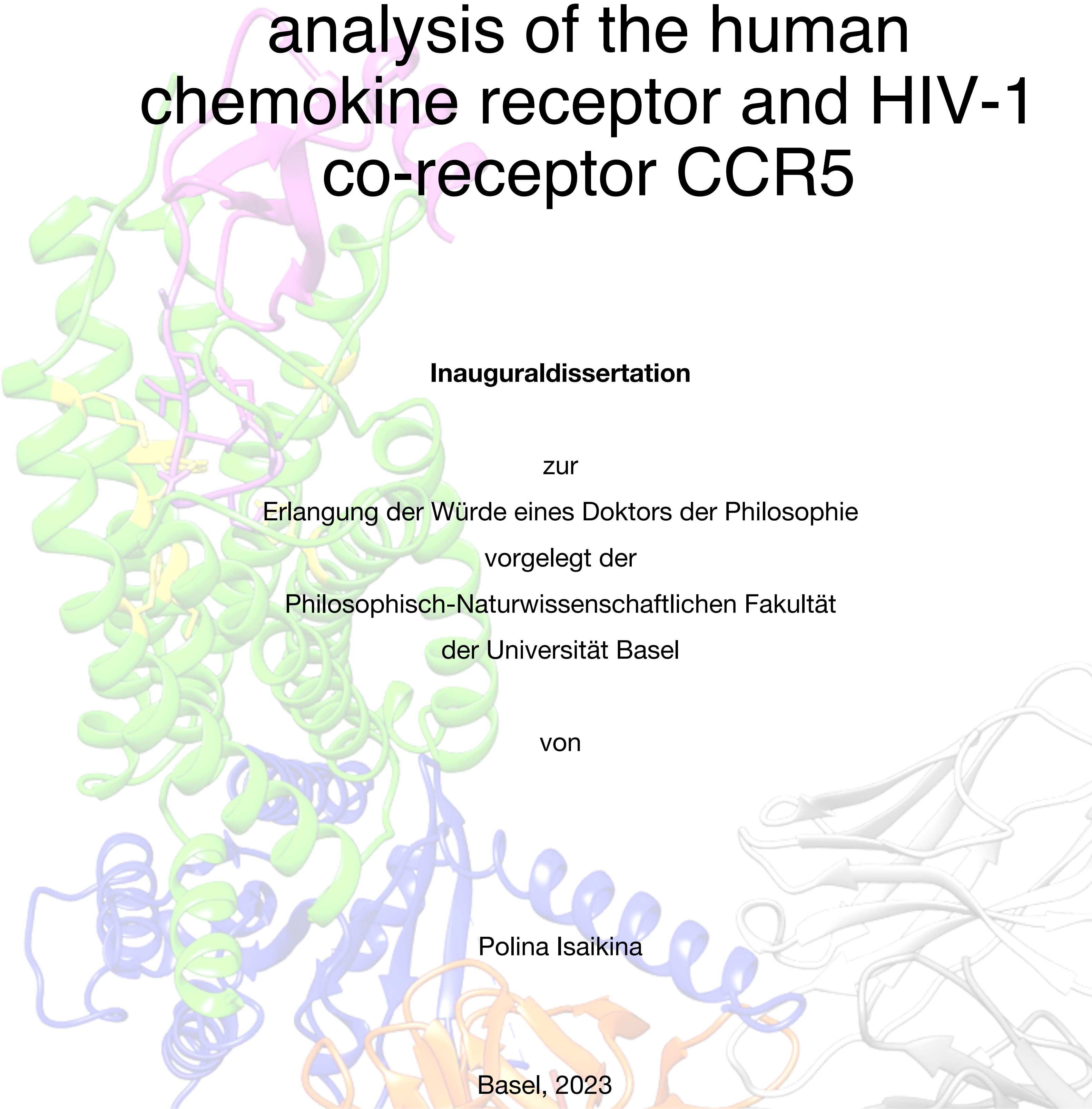
Structural and functional analysis of the human chemokine receptor and HIV-1 co-receptor CCR5



Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Polina Isaikina

Basel, 2023

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät auf Antrag von

Prof. Dr. Stephan Grzesiek
Prof. Dr. Henning Stahlberg
Prof. MD Brian K. Kobilka

Basel, den 21.02.23

Prof. Dr. Marcel Mayor
(Dekan)

Summary

This thesis describes the elucidation of the activation mechanism of the human CC chemokine receptor 5 (CCR5) by a chemokine analog, its signaling complexes with the heterotrimeric G_i protein, and subsequent interactions with arrestin2.

Chapter 1 presents a general introduction to G protein-coupled receptor (GPCR) biology and recent advances in the structure elucidation of GPCRs and their signaling complexes. Additionally, the chapter covers the biology of chemokine receptors, focusing on the role of CCR5 in the human immune system.

Chapter 2 describes the cryo-EM structure of wild-type human CCR5 in an active conformation bound to the super-agonist [6P4]CCL5 and the heterotrimeric G_i. The structure of the signaling complex allowed to describe the activation mechanism of CCR5 and to elucidate key elements of the variable pharmacology of CCL5 analogs. These results shed new light on the molecular pharmacology of chemokine receptors and show how a chemokine receptor can be activated by the 'deep' binding of the agonist N-terminus into the orthosteric receptor pocket.

Chapter 3 provides a detailed protocol for the biochemical preparation of the [6P4]CCL5•CCR5•G_i signaling complex and describes current advances in chemokine structure determination and associated challenges. An additional NMR characterization of the [5P14]CCL5 partial agonist and [5P12]CCL5 antagonist chemokines is given in the Appendix of this Chapter.

Chapter 4 focuses on the last step of the GPCR signaling cascade – the interaction of CCR5 with arrestin2. Two high-resolution X-ray structures of human arrestin2 in complex with two distinct CCR5 C-terminal phosphopeptides were solved. These structures, in combination with mass spectrometry, NMR, and biochemical and cellular assays, uncovered a key GPCR phosphomotif, which is recognized by arrestin2 and enables its tight association with a GPCR. A further analysis of available structural and functional data on GPCR•arrestin interactions suggests how a certain arrangement of phosphoresidues within the intracellular side of GPCRs define arrestin2 and arrestin3 isoforms specificities.

Chapter 5 describes the isolation of the full-length [6P4]CCL5•CCR5•arrestin2 complex, its biochemical and structural characterization. Furthermore, it discusses the challenges of the structural analysis of this complex.

The results from this thesis have been described in the following manuscripts and presented in the following talks at international conferences:

Published:

1. Isaikina P., Tsai C.-J., Dietz N., Pamula F., Grahl A., Goldie K.N., Guixà-González R., Branco C., Paolini-Bertrand M., Calo N., Cerini F., Schertler G.F.X., Hartley O.*, Stahlberg H., Maier T., Deupi X.*, Grzesiek S.* Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. *Science Advances*, **2021**; 7(25): eabg8685 (Chapter 2)
2. Isaikina P.*, Tsai C.-J., Petrovic I., Rogowski M., Meng Dürr A., Grzesiek S.* Preparation of a stable CCL5•CCR5•G_i signaling complex for Cryo-EM analysis. *Methods in Cell Biology, Elsevier*, **2022**; 169: 115–141 [*corresponding author], (Chapter 3)
3. Isaikina P.[#,*], Petrovic I.[#], Jacob R.P., Sarma P., Ranjan A., Baruah M., Panwalkar V., Maier T., Shukla A.K.*, Grzesiek S.* A key GPCR phosphorylation motif discovered in arrestin2•CCR5 phosphopeptide complexes. *bioRxiv,* **2022** [*corresponding author, [#] equal contribution], *(under review),* (Chapter 4)

In preparation:

1. Agustoni E., Mechaly, A., Isaikina P., Müntener T., Schirmer T., Buschiazzo A., Hiller, S., Structure and function of the Leptospira virulence switch protein LvrB.
2. Petrovic I., Isaikina P., Grzesiek S. Characterization of conformational selection of arrestin2 by NMR.

Invited talks:

1. Isaikina P.*, Tsai C.-J., Dietz N., Schertler G.F.X, Deupi X., Maier T., Grzesiek S. Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist at the Biozentrum and Pharmazentrum PhD Retreat, University of Basel, June 17-18, 2021, online
2. Isaikina P.*, Tsai C.-J., Branco C., Guixà-González R., Schertler G.F.X., Hartley I., Deupi X., Grzesiek S. Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist at the GPCR Keystone Symposium, April 06-09, 2022, Snowbird Resort, Utah, USA

3. Isaikina P.*, Grzesiek S. Structural and functional analysis of the CCR5 and its complexes at the Faculty of Medicine, University of Geneva, March 25, 2022, Geneva, Switzerland
4. Isaikina P.*, Petrovic I, Jacob R.P., Maier T., Grzesiek S. Structural and functional studies of CC chemokine receptor 5 phosphorylation and its interactions with arrestin2 at the 4GPCRnet International Symposium, September 26-29, 2022, Leipzig, Germany.

Acknowledgments

This thesis has been possible with many individuals’ continuous support and help. I would like to express gratitude to everyone who gave me encouragement and inspiration throughout the last 4 years. I was privileged to work with many very talented and skilled people and want to mention those of the highest impact specifically.

First and foremost, I thank my supervisor Prof. Dr. Stephan Grzesiek, for the opportunity to conduct my PhD thesis in his lab and for his trust in me. He enormously contributed not only to my scientific but also personal development. His broad fundamental knowledge, critical thinking, and genuine passion for science have enriched and inspired me. I have been fortunate to work on several exciting projects and learn crucial research aspects from him, including ethical principles such as transparency and precision while conducting and communicating science. I am incredibly grateful for the patience, freedom, and confidence Stephan has provided me. It made my PhD experience both fruitful and satisfying.

I am sincerely grateful to the thesis committee members and collaborators, Prof. Dr. Henning Stahlberg and Prof. Dr. Oliver Hartley, for their assistance, stimulating discussions, and the expertise they readily shared with me. I have highly benefited from their insightful comments on my work.

I would like to extend my profound gratitude to Prof. Dr. Brian K. Kobilka for kindly accepting to co-referee my thesis and for a fair evaluation of my work.

Part of this thesis results from a successful collaboration with Prof. Dr. Gebhard Schertler, Dr. Xavier Deupi, and Prof. Dr. Timm Maier. Gebhard Schertler and his group members, Dr. Ching-Ju Tsai and Dr. Fillip Pamula welcomed me at the PSI and generously shared their expertise on the GPCR•G protein complex preparations. Our late-night cryo-EM sample preparation of the CCR5•G_i complex yielded the beautiful structure.

Timm Maier provided tremendous support with cryo-EM data processing and atomic model building, patiently teaching me the fundamental steps of structural biology. I sincerely appreciate his availability for scientific discussion and advice. I am also grateful to Drs. Roman Peter Jakob, Niko Dietz, and other members of the Maier group. I am obliged to Roman Peter Jakob for sharing his practical knowledge about protein X-ray crystallography and for his outstanding contribution to the arrestin project.

I am thankful to Prof. Dr. Sebastian Hiller and Elia Agustoni for allowing me to be part of cryo-EM studies of histidine kinase LvrB and for many helpful discussions about science and career. I also want to thank other past and present members of the Hiller group, especially Dr. Stefan Bibow, for sharing his protocols and assisting with the first preparations of CCR5 in nanodiscs. I am obliged to Dr. Thomas Müntener for providing intelligent solutions to almost any technical problem.

I want to acknowledge the core facilities of the Biozentrum that make it a fantastic place to do science. In particular, I thank Dr. Kenneth Goldie for training me to operate microscopes and Drs. Lubomir Kovacik, Mohamed Chami, and David Kalbermatter for their assistance with cryo-EM data collections, which always went smoothly. I am also grateful to Dr. Alex Schmidt and Ulrike K. Lanner from the Proteomics Facility for contributing to the CCR5 phosphorylation measurements.

My sincerest gratitude goes to all past and present members of the Grzesiek group for all the great discussions, a supportive work environment, and a stable foundation for my work. Dr. Anne Grahl introduced me to many GPCR concepts, including handling the receptors in the lab. Moreover, her previous work provided an excellent basis for the CCR5 project. I want to thank Marco Rogowski for his assistance with some methods in the lab and his significant contribution regarding the expression and purification of several proteins essential for my projects, including various CCL5 analogs. Dr. Shin Isogai was willing to teach me about experiment planning and automation of purification procedures.

Ivana Petrovic largely contributed to the arrestin project with her extraordinary expertise. I have been lucky to share the project with her. Working with such a thoughtful and eager-to-get-things-done person has been a great pleasure. I want to thank Dr. Fengjie Wu and Alexandra Meng Dürr for proofreading parts of this thesis. Alexandra contributed to many other texts I have written with her originality and excellent language knowledge. I am also grateful to Fengjie for many fruitful discussions about GPCRs and for sharing his profound scientific knowledge. I learned a lot from Dr. Judith Habazettl about the practical aspects of protein NMR. She has always been eager to provide her competence in life and science. I would

like to thank Anna Golynski, a master student I had an opportunity to co-supervise. It was a bumpy ride with many technical challenges, but we both gained a lot of experience and finished the project.

Furthermore, I am deeply grateful for all the support of my friends and family. I thank my parents, Nina and Mikhail, and my brother Nikolai for their love, education, and unconditional endorsement. I also want to thank my friends from Russia with whom I stayed in contact for many years, and even >2000 km and other circumstances are not a problem for our friendship. My sincerest thanks go to Dr. Roman Solovov, Natasha Kalinchenkova, Margarita Sudakova, my cousin Varvara Parfilkina, and Anna Aidakova for all the stimulating discussions about science, politics, and other aspects of life.

I am particularly thankful to Dr. Karol Kaiser, Dr. Thomas Müntener, Raphael Bossart, Gabriel Oswald, Sophia Schmidt, Dr. Hugo Munoz-Hernandez, and Japhet Gnehm for their company, friendship, sense of humor, and the good times we have had together in Switzerland. I was blessed to meet you all and hope for many years to come. Special thanks go to Anita and Beat Leuch, who have always welcomed me at their farm and, with their empathy and care, supplied me with energy and hope.

Lastly, I am fortunate to share my life with my husband, Raphael Gnehm. His enormous positivity, love, and encouragement have made my life fulfilling. I have been warmly welcomed into his family, sharing many good memories of visiting the beautiful Thurgau area. Thank you all!

Abbreviations

Å	Angstrom (equal to 0.1 nanometer)
ACKR	Atypical chemokine receptor
AGC	Family of protein kinases A, G and C
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
AP	Adaptor protein
AF	AlphaFold
$A_{2A}R$	Adenosine 2A receptor
β1AR	Beta1-adrenergic receptor
cAMP	Cyclic adenosine monophosphate
CAAX	C-terminal tetrapeptide motif, where C cysteine, A aliphatic residue, X variable amino acid
CCD	Charge-coupled device
CCL	CC chemokine ligand
CCR	CC chemokine receptor
CHS	Cholesteryl hemisuccinate
CMOS	Complementary metal-oxide-semiconductor
COVID-19	Coronavirus disease of 2019
CV	Column volume
CRS	Chemokine recognition site
Cryo-EM	Cryogenic electron microscopy
CTF	Contrast transfer function
Da	Dalton, g/mol
DAG	Diacylglycerol
DEER	Doble electron-electron resonance
ECL	Extracellular loop
ERK	Extracellular signal-regulated kinase
EM	Electron microscopy

FRET	Fluorescence resonance energy transfer
GABA	Gamma-aminobutyric acid
GAG	Glycosaminoglycan
GDP	Guanosine diphosphate
GEF	Guanine nucleotide exchange factor
GLP	Glucagon-like peptide
Gp120	Glycoprotein 120
GPCR	G protein-coupled receptor
GRK	G protein-coupled receptor kinase
GTP	Guanosine triphosphate
h	hours
HIV	Human immunodeficiency virus
hpi	Hours post infection
HSQC	Heteronuclear single quantum coherence spectroscopy
ICL	Intracellular loop
IP_3	Inositol triphosphate
kV	kilovolt
LCP	Lipidic cubic phase
LMNG	Lauryl maltose neopentyl glycol
MD	Molecular dynamics
MIP-1α	Macrophage inflammatory protein-1α
mGlu	Metabotropic glutamate receptor
MS	Mass spectrometry
MIP-1α	Macrophage inflammatory protein-1α
MIP-1β	Macrophage inflammatory protein-1β
MWCO	Molecular weight cut-off
NAM	Negative allosteric modulator
NMR	Nuclear magnetic resonance
NOESY	Nuclear Overhauser enhancement spectroscopy

NTR	Neurotensin receptor
PAM	Positive allosteric modulator
PAGE	Polyacrylamide gel electrophoresis
PDB	RCSB Protein Data Bank
PIP2	Phosphatidylinositol 4,5-bisphosphate
PKA	Protein kinase A
PLC-β	Phospholipase Cβ
ppm	Parts per million
[PSC]	n-nonanoyl-thioprolyl-cyclohexylglycyl
PTM	Post-translation modification
RANTES	Regulated on activation, normal T cell expressed and secreted
RH	Regulator G protein signaling homology
SEC	Size-exclusion chromatography
SDS	Sodium dodecyl sulfate
Sf9	Spodoptera frugiperda
SFX	Serial femtosecond crystallography
SMO	Smoothened receptor
SSTR	Somatostatin receptor
Ste	Pheromone alpha factor receptor
T4L	T4 lysozyme
TEM	Transmission electron microscope
TM	Transmembrane
TMD	Transmembrane domain
TROSY	Transverse relaxation optimized spectroscopy
wt	Wild type
XFEL	X-ray free-electron laser
3DVA	Three-dimensional variability analysis

Table of contents

Introduction

1.1 G protein-coupled receptors

G protein-coupled receptors (GPCRs) represent one of the largest families of membrane receptors comprising more than 800 members in humans (1). GPCRs recognize diverse extracellular stimuli and transduce them into intracellular signals. As such they are involved in a wide array of physiological processes, which makes them major targets for drug development. Currently about one-third of drugs on the pharmaceutical market target a small fraction of all GPCRs (2,3).

Based on sequence homology and functional similarity, GPCRs are divided into the following classes: rhodopsin-like (class A), secretin (class B1) and adhesion (class B2), glutamate (class C), fungal mating pheromone (class D), cyclic AMP (class E) and frizzled and smoothened (class F) (4). Receptors from all classes are present in mammals with the exception of the classes D and E.

Although sequence homology is low between GPCRs from different classes, all of them share a universal seven-transmembrane (7TM) helical architecture consisting of the extracellular N-terminal sequence, seven helices connected by three intra- and three extracellular loops (ICLs and ECLs), as well as the intracellular C-terminal sequence (5,6) (Figure 1.1).

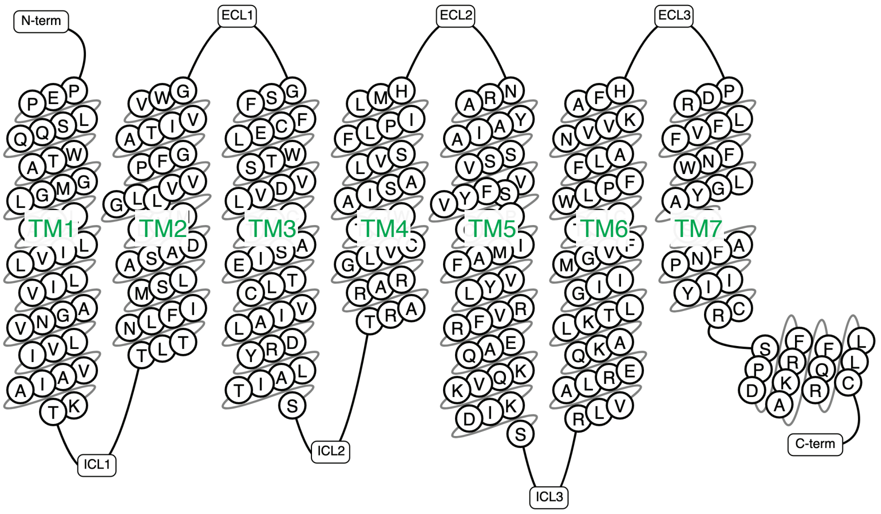


Figure 1.1. Schematic representation of GPCR topology (generated with GPCRdb, https://gpcrdb.org/)

Class A is the largest class comprising more than 80% of the GPCR genes, which includes adrenergic, dopamine, neurotensin, opioid, chemokine and many other receptor families. The receptors from this class share highly conserved motifs, which are proven to be functionally important (7). For convenience of receptor comparisons, the Ballesteros-Weinstein nomenclature was introduced for class A GPCRs (8). A particular residue is numbered within a helix relative to the most conserved residue of each helix, which is numbered 50. It is indicated by the superscript x.yy, where x is the TM helix number and yy is the residue number. This sequence annotation format has been adapted by the GPCRdb (9).

Although class A GPCRs are the best studied in terms of their structure and function, some of the receptor families from this class are far from being understood. One of the understudied families is the chemokine receptors. Their crucial regulation of the immune system and thus involvement in many diseases, e.g., HIV (10), has made them highly attractive targets for drug discovery. However, compared to other class A GPCRs the success has been modest and only three small-molecule drugs that target chemokine receptors have been clinically approved (11). The reason for this apparently comes from the complexity of the chemokine system, the promiscuity of the receptor-ligand interactions, and the fact that the endogenous ligands are full proteins (chemokines) and not small molecules. As such, the progress from other class A GPCR has not been easily transferable to the chemokine receptors.

1.2 GPCR signaling cascade

The canonical GPCR signaling paradigm includes interaction of an activated receptor with the three main intracellular partners: heterotrimeric guanine nucleotide-binding proteins (G proteins), G protein-coupled receptor kinases (GRKs), and arrestins (12).

Binding of an agonist to the extracellular pocket of a GPCR induces conformational changes, which propagate further to the intracellular part of the receptor. This allosteric conformational change of the receptor enables the coupling of the heterotrimeric G protein (13). In response, the G protein facilitates GDP/GTP exchange, promoting the dissociation of the G protein from the receptor and its subunits. To prevent overstimulation, the active receptor is phosphorylated by GRKs, enabling arrestin recruitment. The formation of the receptor•arrestin complex leads to receptor internalization, followed by lysosomal degradation of the receptor or its recycling back to the plasma membrane (14). The detailed molecular mechanisms of these complex and variable interactions involved in GPCR signaling remain to be elucidated (15). An overview of the signaling cascade is given in Figure 1.2.

There are some exceptions to this canonical GPCR regulation mechanism. For example, some GPCRs form so-called megaplexes, binding G proteins and arrestin simultaneously (15–17), while others lack functional G protein coupling, and still undergo agonist-induced arrestin recruitment (18).

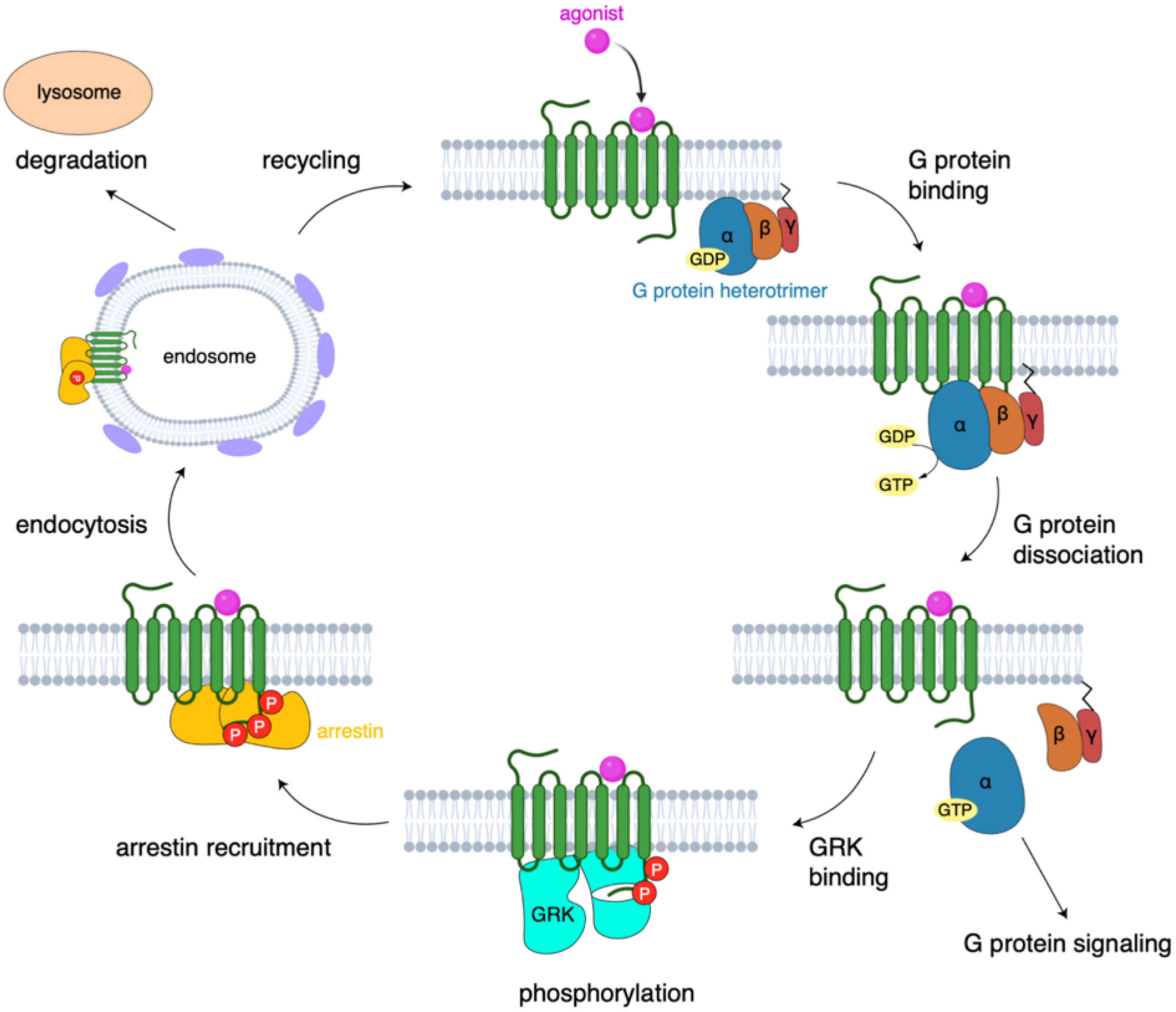


Figure 1.2. GPCR signaling cascade comprising the endogenous ligand, effector proteins and downstream signaling (see text).

1.3 G proteins

G proteins are heterotrimeric guanine nucleotide-binding proteins, which are located inside the cell and activated by ligand-stimulated GPCRs to induce further downstream signaling (19). They consist of α, β and γ subunits. The human genome encodes 33 G protein subunit genes, comprising 16 Gα, 5 Gβ, and 12 Gγ subunits. Based on the Gα isotype, G proteins are classified into four families: G_s, $G_{i/o}$, $G_{q/11}$ and $G_{12/13}$ (20). The $Gα_s$ family, where *s* means stimulation, has two members: $Gα_s$ and G_{olf} (*olf* – olfactory). The G_i (*i* stands for inhibition) family is the largest and includes the following members: $Gα_{i1}$, $Gα_{i2}$, $Gα_{i3}$, $Gα_o$, $Gα_{t1}$, $Gα_{t2}$, (*t* – transducin), $Gα_g$ (*g* – gustducin) and $Gα_z$. The G_q family consists of $Gα_q$, $Gα_{11}$, $Gα_{14}$ and $Gα_{15}$, and the $G_{12/13}$ family includes $Gα_{12}$ and $Gα_{13}$. GPCRs often couple to more than one G protein subtype (21). Recent advances in structure determination of GPCR•G protein complexes in combination with functional data have allowed to identify the main determinants of G protein selectivity (22).

The Gα subunit consists of two domains, a helical domain and a GTPase domain. The Gβ and Gγ subunits form a stable dimer and function as one unit. The Gβγ dimer is anchored to the inner part of the cell membrane via a prenylated C-terminal CAAX motif on the γ subunit. The heterotrimer assembles upon GDP binding to a cleft formed by the two domains of the Gα subunit (Figure 1.3A). Stimulation of a GPCR by an agonist promotes recruitment of a GDP-bound G protein heterotrimer to the receptor, which leads to the release of the GDP molecule (Figure 1.3B). Subsequently, GTP, which has an ~10 times higher cytosol concentration than GDP, binds to Gα (23). Once bound, GTP triggers the activation of the Gα subunit promoting its dissociation from both receptor and Gβγ. Active GTP-bound Gα and Gβγ continue downstream signaling via the engagement of various effector proteins. Depending on the G protein subtype, different effectors are recruited (Figure 1.3C) (24). For example, $Gα_s$ interacts with adenyl cyclase, which stimulates cAMP production. This further activates the protein kinase A (PKA) pathway. In contrast, GTP-bound $Gα_{i/o}$ reduces cAMP concentration by inhibiting adenyl cyclase. $Gα_{q/11}$ recruits phospholipase Cβ to hydrolyze phosphatidylinositol biphosphate (PIP2) to DAG and inositol triphosphate (IP_3), subsequently leading to Ca^{2+} release and protein kinase C activation. Both $Gα_{q/11}$ and $Gα_{12/13}$ are implicated in the activation of the Rho family of GTPases and Rho guanine nucleotide exchange factors (RhoGEFs) regulating actin dynamics. For all Gα subtypes, the signaling is terminated once GTP is hydrolyzed by the Gα GTPase activity. In contrast, after dissociation from the heterotrimer, Gβγ interacts with GRKs, promoting their association to the membrane, ion channels and PLC-β.

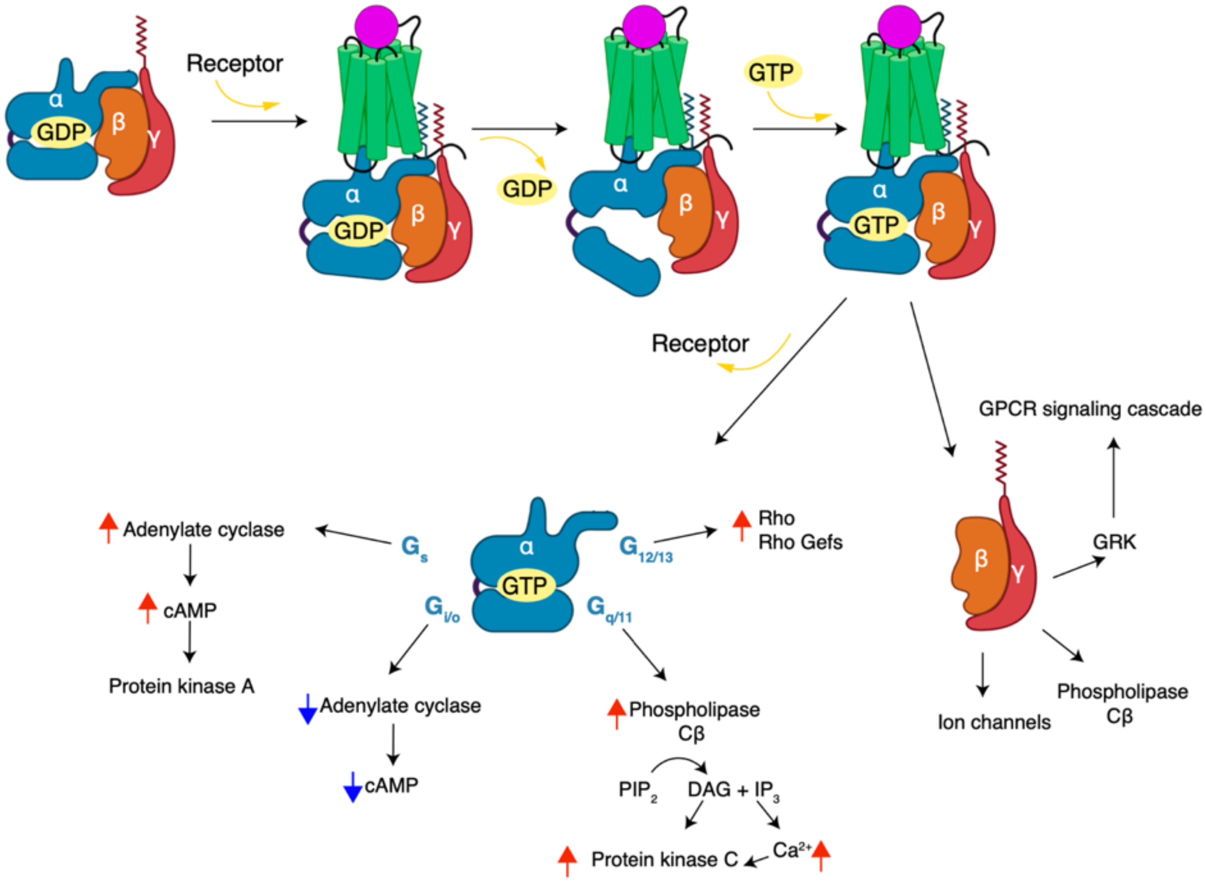


Figure 1.3. G protein signaling scheme. Left: inactive GDP-bound state of G protein heterotrimer. Middle: G protein heterotrimer coupling to GPCR and dissociation of GDP. Right: GTP binding to Gα, resulting in the dissociation of the G protein heterotrimer from the receptor and its subsequent splitting into Gα and Gβγ. Further downstream effector protein signaling pathways via both G protein subunits are shown at the bottom.

1.4 G protein-coupled receptor kinases

The initial step of agonist-bound GPCR desensitization is mediated by GRKs (Figure 1.2). GRKs are serine/threonine protein kinases that belong to the AGC family (protein kinases A, G and C). They recognize and phosphorylate multiple sites at intracellular parts of the GPCRs, i.e. mainly the C-terminal tail and intracellular loop 3 (ICL3) (25). The receptor phosphorylation promotes arrestin coupling and subsequent receptor internalization. There are seven GRKs which are subdivided into three groups based on their sequence homology: the visual GRK subfamily (GRK1 and

GRK7), the β-adrenergic receptor kinases subfamily (GRK2 and GRK3) and the GRK4 subfamily (GRK4, GRK5 and GRK6) (26).

All GRKs share well-conserved short N-terminal α-helical domains, regulator G protein signaling homology (RH) domains, catalytic domains as well as variable C-terminal regions (27). The C-terminal region of all GRKs is responsible for membrane localization of the kinases via different mechanisms. The visual GRKs are targeted to the membrane via prenylation sequences at their carboxy terminus (28). GRK2 and GRK3 have pleckstrin homology domains instead, which enables their binding to the dissociated and membrane-anchored Gβγ (29). Thereby GRK2 and GRK3 are directed to the membrane after receptor activation. Besides interacting with Gβγ, GRK2 binds $G\alpha_q$, promoting membrane association of the kinase (30–32). GRK5 contains a hydrophobic phospholipid-binding domain (33), whereas GRK4 and GRK6 harbor palmitoylation sites (34).

Several studies have provided insights into the sequence specificity of GRKs based on the ability of GRKs to phosphorylate synthetic peptides. For example, GRK1 and GRK2 prefer to phosphorylate serine/threonine residues localized near acidic residues, while GRK5 and GRK6 have a clear preference for serine/threonine preceded by basic residues (27). Despite these insights on the sequence preferences of GRKs, the molecular basis of GRK•GPCR specificity remains largely unknown. Nevertheless, it is clear that GRKs play an important role in GPCR bias (35), determining the diversity of the subsequent signaling pathways by differential receptor phosphorylation (36,37).

1.5 Arrestins

Following GPCR phosphorylation, cytoplasmic signaling proteins called arrestins are recruited to the receptor, promoting its desensitization (38). The arrestin family consists of four members in humans: visual arrestins (arrestin1 and arrestin4) and non-visual arrestins (arrestin2 and arrestin3, also known as β-arrestin1 and β-arrestin2, respectively) (39). The visual arrestins are expressed in the retina and regulate only photoreceptors. The two remaining non-visual arrestins are expressed in all other tissues and cell types and, thus, regulate hundreds of different GPCRs. These arrestins are genetically and structurally conserved and are composed of antiparallel β-sheet N and C domains.

While both non-visual isoforms mediate GPCR signaling, they often display significant functional divergence leading to different physiological outcomes (40). Arrestin-mediated signaling depends on phosphorylation levels and phosphorylation patterns, which are modulated by GRKs. Moreover, different phosphorylation patterns can induce distinct conformations of arrestin, thereby controlling the engagement of specific downstream signaling proteins (41–44). In terms of arrestin binding, GPCRs are categorized as “class A” when they bind arrestin in a transient manner and “class B” when they form stable complexes with both arrestin subtypes (45). Both of the receptor classes undergo arrestin-induced internalization forming early endosomes. However, the subsequent trafficking steps are distinct. Class A GPCRs are rapidly recycled back to the plasma membrane from early endosomes (46). In contrast, class B GPCRs traffic from early to late endosomes, resulting in lysosomal degradation (Figure 1.2) (45,47).

In general, arrestin recruitment to GPCR is a biphasic process. Initially, it requires arrestin activation by displacement of its autoinhibiting C-terminal tail with the phosphorylated receptor C-terminus (tail interaction). The tail displacement causes conformational changes in arrestin enabling its full engagement with the transmembrane helical bundle (core interaction) (48). Complexes of GPCRs, where arrestin is only bound to the phosphorylated receptor C-terminal tail, have been reported (49). While full arrestin engagement seems to be necessary for receptor desensitization (50), it is not required for its internalization (51). Recently, it has been reported that lipids, specifically phosphatidylinositol phosphates (PIPs), can play a crucial role in modulating the strength of GPCR-arrestin interactions and can by themselves catalytically activate arrestins (52).

Additionally, arrestins adopt multiple conformations to select different signaling partners. Receptor-bound arrestins involve downstream regulators such as clathrin, clathrin adaptor protein 2 (AP-2), and extracellular signal-regulated kinase (ERK). This promotes receptor transport via clathrin-coated pits to endosomes (14,53,54). Beyond this classic role, arrestins can interact with dozens of other signaling proteins, thereby activating G protein-independent signaling pathways, which play a role in transcription, cell proliferation, and apoptosis (55,56).

1.6 Ligands and biased signaling

GPCRs sense physically and chemically diverse ligands comprising photons, nucleotides, lipids, hormones, peptides, proteins, and others. Their properties can be described by two main parameters: affinity and efficacy, where efficacy is the capacity of the ligand to produce the receptor-mediated response at full receptor occupancy (57). Many GPCRs have basal activity even without any ligand stimulation. Based on their efficacy, ligands can be classified as full agonists, partial agonists, inverse agonists, and antagonists. An agonist activates a receptor above the basal level. A partial agonist induces weaker activity than the full agonist. In contrast, an inverse agonist lowers the activity below the basal state, leading to receptor inactivation, whereas an antagonist blocks the ligand-binding pocket with no impact on basal activity (Figure 1.4A).

The binding pocket which accommodates the endogenous GPCR ligands is termed orthosteric pocket. Additionally, GPCR can be regulated by molecules that bind at sites different from the orthosteric pocket. Such molecules are called allosteric modulators. They modulate the signaling of orthosteric ligands in a negative (negative allosteric modulator, NAM) or positive (PAM) way (58). Well-known endogenous allosteric modulators are sodium and cholesterol (59–61).

Some ligands of a given GPCR activate equally G protein and arrestin pathways, whereas others preferentially activate one pathway over the other (Figure 1.4B) (62). Selective receptor activation is referred to as biased agonism. Biased agonism provides an opportunity for developing drugs that are selective for desired physiological effects, but avoid undesired side effects (63).

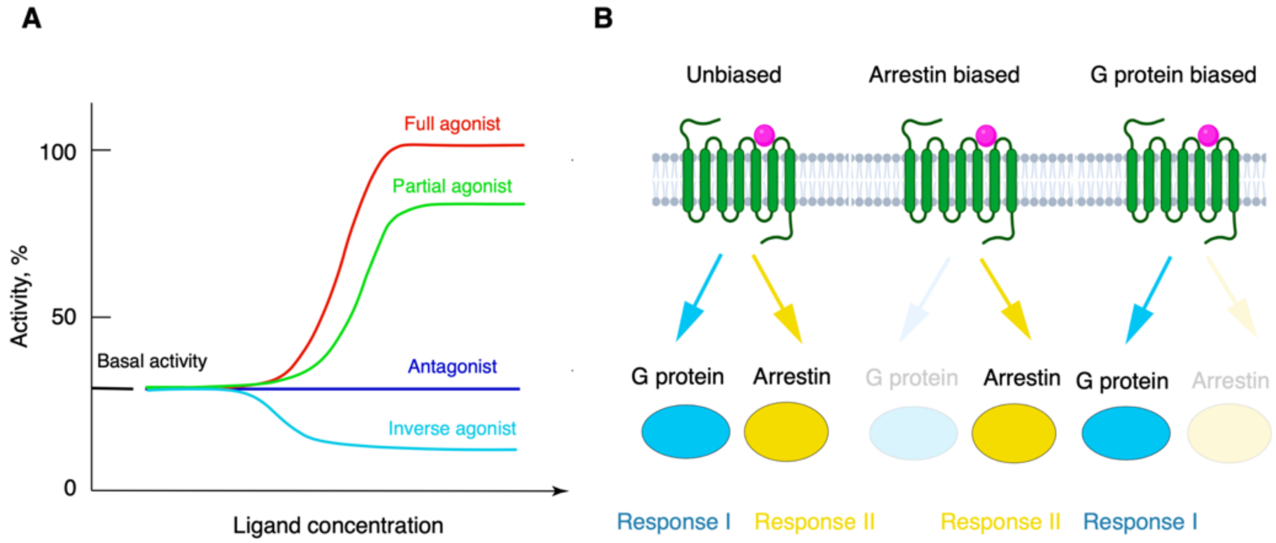


Figure 1.4. Characteristics of GPCR ligands. A. Efficacy-based ligand classification. B. Illustration of differential ligand-induced GPCR signaling and biased agonism.

Ligand bias is complex, and often experimental results are difficult to reproduce. Commonly, biased agonists are described using cellular signaling assays. However, the observed biased agonism depends on the cell type being used, thereby introducing system bias. The time points of the data collection are equally important since the ligand effect is time-dependent. Thus, distinct signaling responses of the ligands can potentially result from different binding kinetics (64). Therefore, precise and reproducible experiments are needed for ligand characterization. In addition, it seems essential for rational drug design to reveal the basics of biological processes underlying biased agonism by structure-function relationships.

1.7 Structure elucidation of GPCRs

Breakthroughs in protein engineering along with technological advancements in X-ray crystallography and recently in single-particle cryo-EM have resulted in an exponential increase of solved GPCR structures. At the time of writing (end of 2022), more than 140 unique GPCR structures have been reported, where 105 are from class A, 14 from secretin and six from adhesion families (class B), one from class D, and four from F GPCR families (GPCRdb, https://gpcrdb.org, Figure 1.5A). Although most GPCRs have been solved in their inactive conformation, for many of them, high-resolution structures of both states, active and inactive, are available (Figure 1.5B). The large majority of inactive GPCR structures have been solved by X-ray crystallography, while active GPCRs have mainly been solved in complexes with heterotrimeric G proteins by cryo-EM. Direct structural comparison of active and inactive states of the same receptor provides valuable insights into their activation mechanisms and has greatly advanced drug development (65).

For two GPCRs, namely, rhodopsin and NTR1, the structures in complex with their effector G proteins (66–69), arrestins (70–72) and GRKs (32,73) are available, covering the main steps of the GPCR signaling cascade. Despite the recent advances in receptor structure determination, structures of the majority of GPCRs remain unsolved. Especially challenging is capturing structures of the highly dynamic and often unstable GPCR•arrestin and GPCR•GRK complexes.

An increasingly powerful alternative for experimental structure determination is computational machine learning methods such as AlphaFold2 (AF2) (74) and RoseTTaFold (75), which may predict accurate 3D protein structures from sequences even without any homologs being available. These programs are based on neural network models trained on known experimental structures. While they provide outstanding results for globular proteins (76), and are extremely useful as starting coordinates for experiment-based model building and refinements, the accuracy of the predicted GPCR models has not been rigorously benchmarked, especially with respect to less rigid regions (77), GPCR conformational states (e.g., active/inactive) and effects of ligand-binding (78). Thus, further developments and comparative studies with experimentally determined structures are needed to increase the utility of these programs.

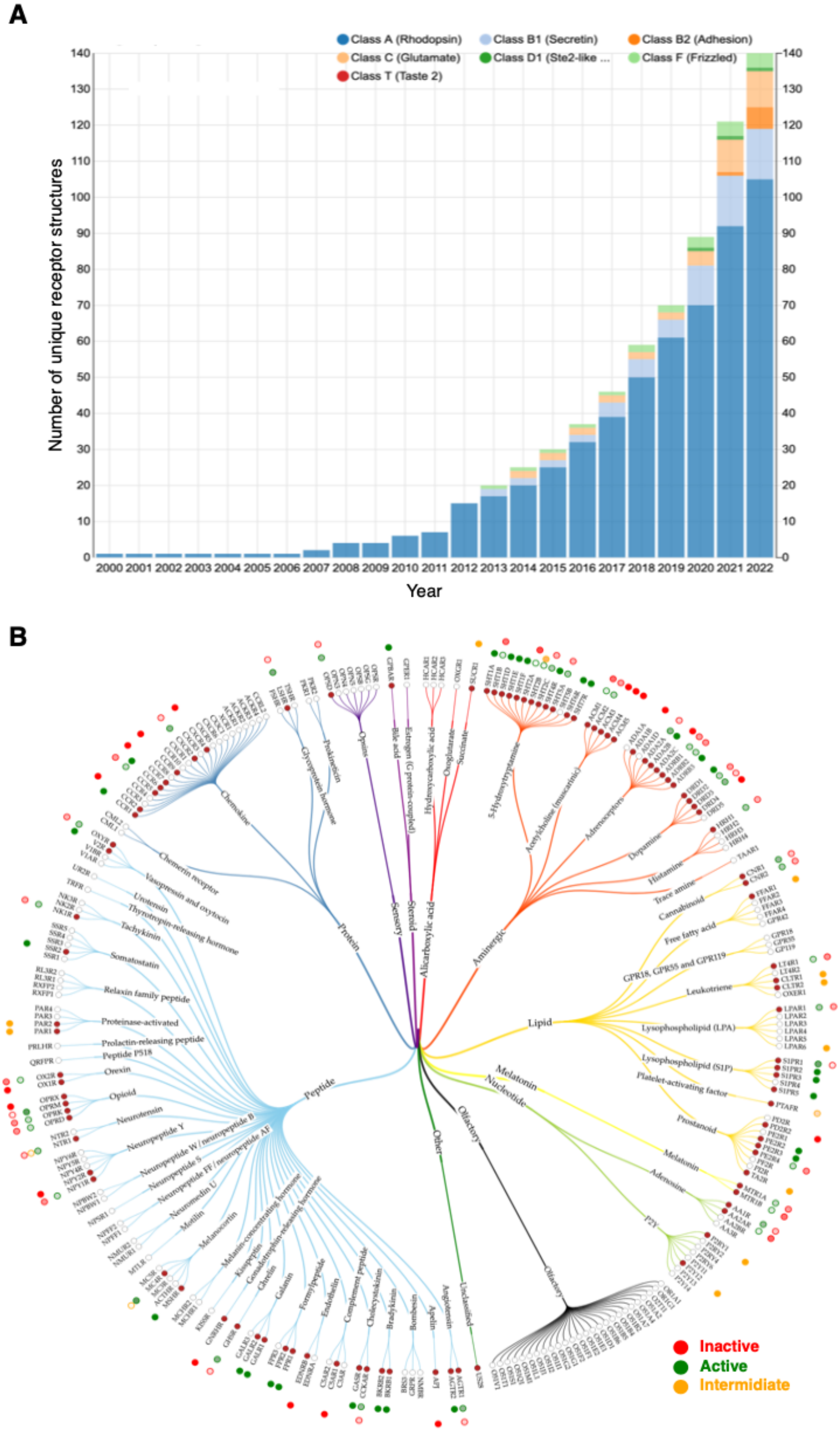


Figure 1.5. Structural coverage of GPCR classes from GPCRdb (https://gpcrdb.org/structure/statistics). A. Timeline of cumulative number of solved GPCR structures. B. Structural coverage of GPCRs solved in different conformational states (red: inactive; green: active; orange: intermediate).

1.7.1 Structural insights into GPCR architecture and activation by X-ray crystallography

The progress in the development of X-ray crystallographic techniques such as high-throughput screening for obtaining crystals, automated crystal handling, crystal soaking, developments of advanced X-ray sources and detectors have made crystallization and data collection highly automated and extremely efficient. However, X-ray structure determination still requires sufficiently large and well-diffracting crystals. The main obstacles to obtaining large and well-diffracting crystals of GPCRs have been difficulties in their recombinant expression, limited stability outside native membranes, and their inherently dynamic nature.

The first high-resolution structure of a GPCR was determined by X-ray crystallography in the year 2000 at a resolution of 2.8 Å (79). It was bovine rhodopsin isolated from its natural source, the bovine retina. The receptor was captured in its inactive state, confirming the previously predicted basic transmembrane architecture of seven α-helices followed by a short helix located at the cytoplasmic interface of the membrane. The extracellular side of the TM helices forms the orthosteric ligand-binding pocket. The ligand for rhodopsin is 11-*cis*-retinal, a chromophore molecule derived from vitamin A, which is covalently bound to the receptor. Upon light exposure, the 11-*cis*-retinal is converted to all-*trans*-retinal, causing a conformational change in the rhodopsin.

The first solved structure of a non-visual GPCR, determined seven years after the rhodopsin, was the recombinantly expressed β2 adrenergic receptor (β2AR) bound to high-affinity ligands (80–82). Protein engineering, namely the replacement of the long and flexible ICL3 with the better crystallizable T4 lysozyme, as well as breakthroughs in crystallization and X-ray data collection (83) significantly contributed to the β2AR structure determination. One important technological development has been the crystallization of membrane proteins in the lipidic cubic phase (LCP) (84). It provides a native-like environment and allows for hydrophobic contacts of membrane-embedded domains enhancing crystal lattice formation. About 90% of GPCRs have been crystallized using this method. The other 10% of GPCRs have been crystallized by vapor diffusion which is often associated with the use of harsher detergents (85).

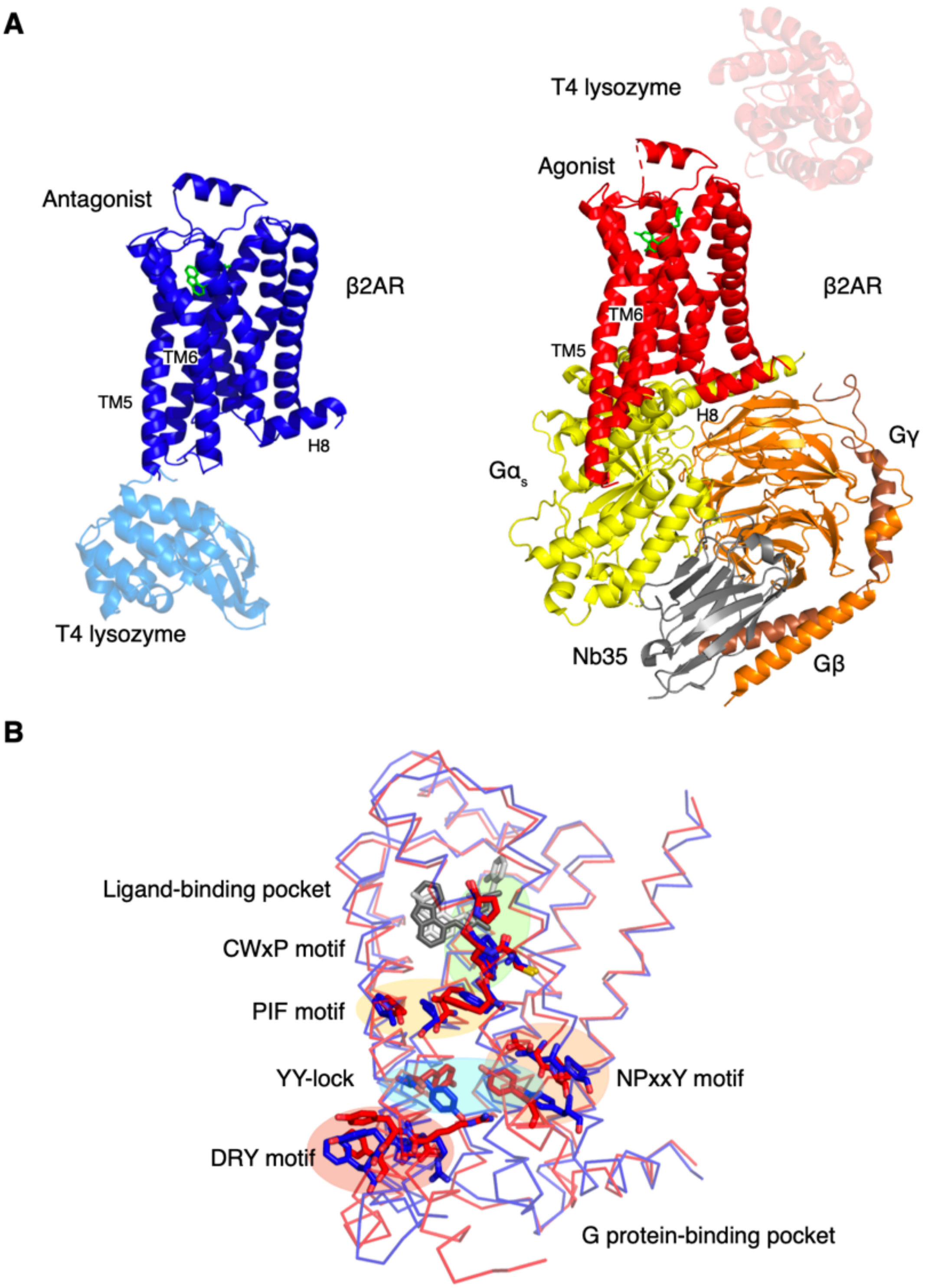


Figure 1.6. Representative structures of β2AR solved in inactive and active states. A. Inverse agonist-bound β2AR structure (inactive, PDB: 2RH1) and agonist-bound β2AR-G_s-Nb35 structure (active, PDB: 3SN6). B. Overlay of active (red) and inactive (blue) β2AR structures and highlighted conserved microswitch motifs.

A combination of the described methods has been applied to many other GPCRs, along with stabilizing point mutations, truncations of flexible regions, and binding of conformation-specific antibody fragments, which “lock” the receptor in a certain state. Additionally, the conformational heterogeneity of GPCRs can be reduced by binding of a high-affinity antagonist, which stabilizes the inactive receptor state. For this reason, most crystal structures of GPCRs have been solved in their inactive states. Interestingly, the agonist-bound β2AR structure is almost identical to the antagonist-bound one (86) (Figure 1.6A), and a shift towards a significantly different conformation was only observed with a bound effector protein (87–89), (Figure 1.6A). Certain GPCR conformations have been stabilized by other factors, e.g., point mutations (90), binding of specific lipids (91) or by applying pressure (92).

So far, only two GPCRs have been solved by X-ray crystallography in complex with a heterotrimeric G protein (89,93) (Figure 1.6A). The first was agonist-bound β2AR in complex with the heterotrimeric G$_s$ protein solved by the Kobilka lab (89). This structure allowed a comparison of the active and inactive states and has provided high-resolution insights into GPCR activation and signal transduction. For these breakthroughs in understanding GPCR structure and function, Robert J. Lefkowitz and Brian K. Kobilka were awarded the 2012 Nobel Prize in Chemistry (94,95).

Further structures have been solved with the active conformation stabilized by engineered G protein-mimicking nanobodies (96–98) and mini-G proteins (99,100). These studies have demonstrated a certain diversity of the activation mechanisms of GPCRs. Nevertheless, some of the features are conserved among class A GPCRs. In particular, the largest conformational change occurs in TM6, which undergoes an outward movement of several angstroms (14 Å in the case of β2AR) (Figure 1.6). Other TMs experience smaller conformational rearrangements. TM3 and TM5 experience rotation, while TM7 undergoes an inward displacement. All these rearrangements open up the intracellular part of the receptor in order to accommodate the α subunit of the G protein.

The receptor activation involves conformational rearrangements of four main conserved motifs among class A GPCR: C$^{6.47}$W$^{6.48}$xP$^{6.50}$, P$^{3.40}$I$^{5.50}$F$^{6.44}$, N$^{7.49}$P$^{7.50}$xxY$^{7.53}$, and D$^{3.49}$R$^{3.50}$Y$^{3.51}$ (Figure 1.6B). These ‘microswitch’ motifs are essential for the allosteric opening of the intracellular side of the receptor (101). The closest to the ligand-binding pocket is the CWxP motif, located on TM6. The W$^{6.48}$ (known as the “toggle switch”) of this motif connects the ligand binding site with the PIF motif, which is crucial for the outward movement of TM6. The DRY motif at the end of TM3 stabilizes the inactive state by forming a salt bridge with $E^{6.40}$. Characteristic water-mediated hydrogen bonds are formed between two highly conserved tyrosines $Y^{7.53}$ of NPxxY in TM7 and $Y^{5.58}$ in TM5 (102). This "YY-lock" is essential for stabilizing the active conformation of TM6 (103). In active structures, the NPxxY and DRY motifs are connected through the YY-lock and an interaction of $Y^{5.58}$ with $R^{3.50}$.

Structure determination of the GPCR•effector protein complexes remains extremely challenging by crystallography. For instance, the inherent instability and flexibility of the rhodopsin visual arrestin complex required extensive protein engineering, which led to the design of a fusion of these two proteins (71). A number of efforts have been undertaken to solve the structure of this rhodopsin-arrestin fusion complex. However, the crystals obtained through extensive optimization in various LCP conditions did not reach the necessary size (>20 um) and diffracted poorly (7–8 Å) (104). Application of serial femtosecond crystallography (SFX) using X-ray free-electron lasers (XFELs) and an LCP injection probe drastically improved the resolution of the same complex crystals to 3.3 Å (anisotropic), revealing for the first-time interactions of the fully engaged arrestin with the receptor core (72,104). In LCP-SFX the radiation damage is minimized since the data are collected from single exposures of many randomly oriented nanometer- to micrometer-size crystals, rather than collecting diffraction images from multiple exposures of a single crystal (105). Moreover, the data are collected at room temperature instead of the conventional freezing at 100 K, thereby capturing thermal motions close to native conditions. The method has been successfully applied to other challenging cases, for example, the first full-length structure of the class B glucagon receptor (106). Despite all these developments in X-ray crystallography, the progress in structure determination of full-length receptor signaling complexes has been modest.

1.7.2 Cryo-EM structures of diverse GPCR classes and complexes

Technical aspects

The situation has substantially changed with the so-called "resolution revolution" of single-particle cryo-EM, which has enabled the determination of hundreds of larger GPCR complex structures with limited stability (107–115). This progress was mostly achieved by the introduction of high-sensitivity direct electron detectors (116). Additional advances in the development of electron sources, energy filters (117,118), as well as data acquisition automation (119), and image processing software (120–123) have now significantly streamlined structure determination by cryo-EM.

Typically, a sample for cryo-EM is prepared from a 3–5 µl solution of 1–10 µg protein and applied to a metal grid covered with a thin layer of carbon or gold. The grid with the sample is rapidly plunged into liquid ethane (–182.8 °C) such that a thin layer (50–100 nm) of vitreous ice is created (124). Ideally, the protein is present in random orientations within the ice layer. However, some proteins tend to adopt preferred orientations or denature at the water-air interface. Although grid optimization can be challenging and time-consuming (125,126), the main advantage is that no crystallization is needed. Therefore, the receptor stability and minor flexibility are less of an issue for cryo-EM, since the grids with the protein can be prepared immediately after purification. This advantage, along with the low amount of sample needed, allows to avoid receptor engineering and sometimes to determine wild-type receptor structures. Moreover, a certain degree of heterogeneity is tolerable in cryo-EM, since the sample can be purified *in silico* during the iterative data processing.

A standard cryo-EM dataset consists of thousands of movies containing typically 40–60 individual images, so-called frames. The individual frames are subjected to motion correction and dose-weighted to remove the beam-induced motion and stage drift (127,128). Afterwards, the contrast transfer function (CTF) parameters are estimated to computationally correct each micrograph by the CTF and to evaluate the quality of the data. After these pre-processing steps, the protein particles are picked, typically identifying several millions of particles. The picked particles represent 2D projections of the protein and they are further classified based on their similarity by averaging them together to increase the signal-to-noise ratio. From the derived 2D classes, which should cover different orientations, the 3D map is generated based on the Fourier projection theorem (129). From the several millions of picked particles, usually only hundreds of thousands are of high-quality and are used to generate the final density map.

The resolution of cryo-EM structures of GPCR complexes, typically ranging from 2 to 4 Å, is comparable to X-ray structures (130). So far, the highest resolution of a GPCR structure achieved by cryo-EM is that of the $GABA_A$ receptor solved at 1.7 Å (118). It is important to mention that cryo-EM density maps often show significant resolution variations, reflecting the conformational heterogeneity of macromolecules, e.g. in loop regions. In contrast, the X-ray electron density maps often are more uniform in resolution due to the restricted flexibility of the molecules within the crystals (107).

Sometimes, distinct sub-conformations of macromolecules may be resolved from the subclasses of particles within a single dataset (131,132). To extract continuous motions of the macromolecules from the cryo-EM datasets, recently new computational approaches such as 3D variability analysis (3DVA) have been developed (133). The 3DVA is performed routinely on receptor complexes providing additional information on transient states and common motions present in GPCR•G protein complexes (134).

The fundamental limitation of cryo-EM structure determination is the particle size. For small molecules (less than 50–100 kDa), images have too low signal-to-noise ratios for particle picking. For membrane proteins with no distinct feature outside of the membrane, the particle alignment is the major obstacle. Therefore, the structure determination of inactive GPCR structures has been challenging for cryo-EM. The binding of antibody fragments, which interact with non-flexible parts of the receptor, increases the overall particle size and provides a source of alignment outside of the micelle. Recently, several studies have reported a universal antibody toolbox for structural studies of GPCRs (135,136)

A further limitation of cryo-EM is the high operational cost associated with high-end, high-voltage EMs. Most high-resolution structures have been obtained using microscopes operated at 300 kV. However, recent examples show that equivalent resolutions can be achieved at 200 kV, thereby significantly reducing the costs (115,137,138).

Overview of GPCR structures solved by cryo-EM

Single-particle cryo-EM has become the most common approach for the structure determination of GPCRs and enabled the solution of unique GPCRs from different subclasses (139). These receptor structures have been determined in different conformations in the presence of various ligands with distinct pharmacological properties. Figure 1.7 summarizes the diversity of GPCRs captured in these different states.

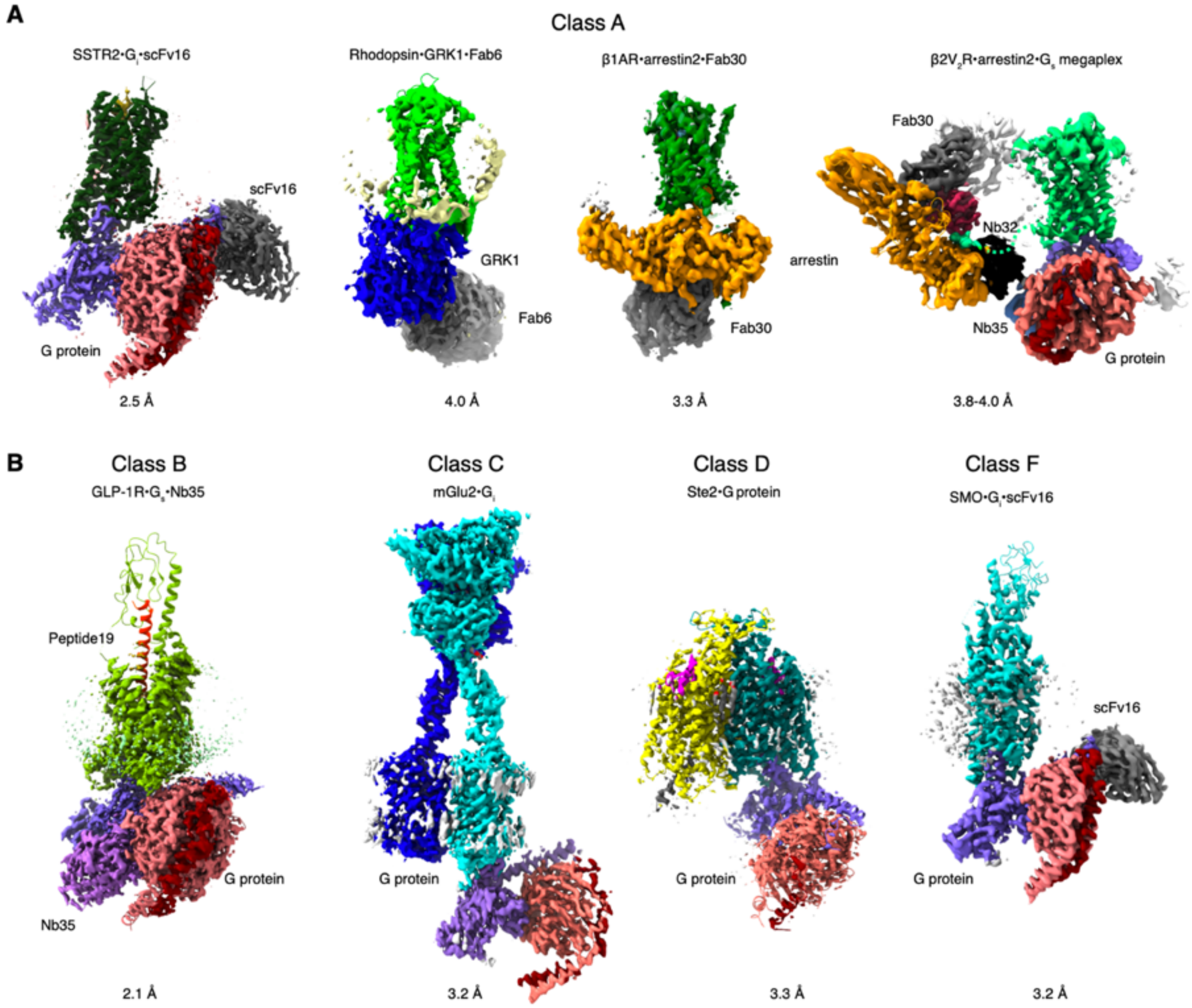


Figure 1.7. Examples of GPCR cryo-EM structures representing the major GPCR classes. A. Class A GPCR in complex with different transducer proteins: G_i heterotrimer (SSTR2•G_i•Fab16, EMDB: 25586), GRK (Rhodopsin•GRK1•Fab6, EMDB: 23980), arrestin (β1AR•arrestin2•Fab30, EMDB: 10515), G_s heterotrimer and arrestin megaplex (right, EMDB:9375 and EMDB: 9376). B. GPCR complex structures from classes B (GLP-1R•G_s, EMDB: 24680), C (mGlu2 in complex with G_i, EMDB: 23996), D (Ste2 dimer in complex with the G protein fragment and G protein heterotrimer, EMDB: 11720) and F (SMO•G_i, EMDB: 22118).

The class A GPCR constitutes the largest and most diverse subfamily from which many receptor structures have been reported, showing novel and diverse activation mechanisms across this class. Especially insightful have been the cryo-EM structures of agonist-bound GPCRs in the active state coupled to different G proteins. Most of the class A GPCRs were solved in complexes with $G_{i/o}$ (120 PDB entries) or G_s (62 PDB entries), 16 new structures with G_q, and only one structure of sphingosine phosphate receptor in complex with $G_{12/13}$ (140). The overall architecture of the GPCR•G protein complexes is similar. The α-helix H5 (αH5) of Gα establishes extensive interactions with receptors and has been reported to be the primary determinant of the G protein-coupling selectivity (141). However, there is apparent variability in the extent of the transmembrane domain opening, the insertion depth of G protein αH5 into the 7TM core, and other receptor-specific interactions within the cytoplasmic region, e.g., ICL3 (142–147). Although these interactions vary for different receptors, generally, the G_s and G_q coupling is mediated mainly by ionic interactions, whereas G_i and G_{12} rely more on hydrophobic interactions.

Solely for class A GPCR, the structures of other signaling partners, such as GRKs (73,148) and arrestins (16,149–153) have been reported (Figure 1.7A). These structures have revealed the major steps of the GPCR signaling and have largely advanced our knowledge of GPCR function and biased agonism. Moreover, the structure of the megacomplex, where heterotrimeric G_s protein and arrestin together bind to the receptor, has introduced a new paradigm into the classical view of the GPCR signaling cascade (16,154). Within the class A GPCRs, also several structures of 'orphan' receptors with unknown endogenous ligands have been solved, including the itch GPCRs (155–157). These structures have revealed unique aspects of the activation of these orphan receptors and have provided valuable information for potential drug design. Despite the tremendous progress in structure determination of the class A GPCRs, so far, no structure of an olfactory receptor has been reported.

The first GPCR•G protein complex structures solved by cryo-EM were complexes of class B1 receptors, namely of the calcitonin (158) and GLP-1 receptors (159). These have provided a model for structural studies of other GPCR complexes. Until now, structures of almost all class B1 receptors have been solved (Figure 1.7B), either in the presence of peptide hormones and G proteins (160–162) or in the apo state (160).

Recently, significant progress has also been achieved in the structural elucidation of class B2 GPCRs, also known as adhesion GPCRs (aGPCRs) (147,163–166). The aGPCRs have distinct molecular features and are activated by unique mechanisms.

These receptors possess large extracellular N-terminal domains containing the GAIN domain, which undergoes autoproteolysis producing N- and C-terminal fragments (165). Subsequently, a part of the N-terminal fragment binds into the GPCR transmembrane domain, thereby activating the receptor. The cryo-EM structures of adhesion receptors bound to their N-terminal fragments (164,167,168) or steroid-like molecules (166) have revealed details of their activation mechanisms. Also structures of the remaining receptor classes, such as C (169–171), D (172) and F (173,174), have been recently solved in their native forms, highlighting the unique architectures across different GPCR classes.

1.7.3 Addressing GPCR dynamics

Cryo-EM and X-ray crystallography have provided fundamental knowledge on GPCR architecture, details of ligand binding, and their activation mechanisms. However, the solved structures represent only snapshots of single conformations. Other methods such as nuclear magnetic resonance (NMR) (90,175,176), double electron-electron resonance (DEER) (177), fluorescence resonance energy transfer (FRET) (178), and molecular dynamic (MD) simulations (179) have revealed that GPCRs are highly dynamic proteins which sample multiple conformations in any ligand-bound state. GPCR function follows from this interchange, which occurs on timescales ranging from picoseconds to milliseconds and even seconds (180). The comprehensive understanding of GPCR function, including allosteric modulation and biased agonism, requires high-resolution data on these dynamic equilibria.

Early DEER and single-molecule FRET experiments have revealed various conformational transitions induced by ligands, in particular the outward movement of the intracellular part of the TM6 upon GPCR activation (181–185). However, these methods, which detect single distances between introduced labels, are limited in their resolution to several angstroms. Furthermore, the attachment of the labels may potentially interfere with the native GPCR dynamics.

Solution-state NMR spectroscopy has the potential to obtain the highest resolution for the description of GPCR dynamics, since in principle dynamic information can be obtained for any atom with a magnetically active nucleus. This dynamic information can be obtained on several timescales comprising nanoseconds, the micro- to millisecond range, and the larger than seconds range for real-time observation. The observations can capture local and global conformational changes at atomic resolution at close to physiological conditions. Although GPCRs are challenging targets for solution-state NMR due to the large size of the receptor-detergent micelle (>100 kDa), recent methodological advances, including new isotopic labeling strategies (^{2}H, ^{13}C, ^{15}N, and ^{19}F) (186,187), high-field NMR instruments, cryogenic probes, transverse relaxation-optimized spectroscopy (TROSY) (188), and new residue-assignment strategies (189–191) have enabled a precise description of various aspects of receptor function.

The majority of NMR studies have been carried out on stabilized β-adrenergic and adenosine receptors using side chain labeling, e.g., $^{13}CH_3$-methione (192,193) as well as ^{13}C-methyl-modified lysines (194) and ^{19}F-tagged cysteines (195–199). These side chain NMR studies have provided valuable insights into the basal activity of GPCRs, effects of ions and lipids, exchange kinetics and other aspects of receptor activation dynamics. However, the information is limited to local conformational changes, and the spectra are often not well resolved. In contrast, the detection of amide ^{1}H-^{15}N backbone resonances offers unique information on backbone structure, H-bonding and long-range motions, and the spectra are much better resolved. However, fewer studies have utilized backbone NMR since the detection of ^{1}H-^{15}N backbone resonances is less sensitive than that of $^{13}CH_3$-methyl groups or ^{19}F-labeled side chains.

Nevertheless, detailed studies of ^{1}H-^{15}N backbone resonances resolving 37 residues have been carried out on ^{15}N-valine-labeled or ^{15}N-tyrosine-labeled turkey β1AR, reporting on the structure and dynamics of the apo receptor, in complex with six different orthosteric ligands as well as with the G protein mimicking nanobody Nb80 (90,92,103,191). This allowed a precise quantification of the dynamic equilibria, thereby identifying crucial details of allosteric signaling (176). Recently, NMR experiments on this receptor revealed that high pressure shifts the conformational equilibria towards the active conformation, providing evidence of a volume reduction upon activation (92). This must be due to the compression of empty cavities. Some of the cavities could be localized by anomalous X-ray scattering of xenon-derivatized receptor crystals (91). One of the cavities colocalizes with the cholesterol-binding site, a known negative allosteric modulator of the adrenergic receptor, and indeed the

NMR experiments showed that the cholesterol derivative CHS abolishes activation. Well-resolved ^1H-^{15}N spectra have also been obtained on the uniformly $^2H/^{15}N$-labeled $A_{2A}R$, receptor, where ^1H-^{15}N resonances of tryptophane side chains and the glycine backbone were assigned (200), and on leucine ^1H-^{15}N backbone resonances of selectively labeled β2AR (201).

Other methods, such as hydrogen-deuterium exchange or native mass spectrometry (MS) are also used to describe the structural dynamics of GPCR signaling in the solution state (202–205). Lately, native MS combined with MD simulations has captured the allosteric modulation of the turkey β1AR induced by 14 different ligands and mini-G proteins (205). These findings explain the mechanism of biased agonism and the role of endogenous zinc ions in stabilizing the receptor•G protein complex.

Noticeable progress has also been made in developing time-resolved cryo-EM based on microfluidics combined with rapid freezing. However, the accessible time scales are in the range of milliseconds to minutes (206). This only allows the detection of large amplitude motions of protein domains. A further time-resolved method, which covers exceptionally broad ranges of protein dynamics at femtosecond resolution, is the time-resolved SFX (tr-SFX). It requires protein crystals, advanced instrumentation, in particular XFEL (207), and the synchronized detection of induced conformational changes. Therefore, it is almost exclusively used for studying light-activated proteins such as rhodopsin, where conformational changes have been followed from several femtoseconds to milliseconds (208). However, such fast dynamics cannot be studied for normal GPCRs, which are activated by the binding of diffusing ligands.

1.8 Chemokines and their receptors

Chemokines and their cognate receptors are widely expressed in the immune system and can also be found in non-immune cells, e.g., the central nervous system. They regulate a wide range of physiological processes, such as the development and homeostasis of the immune system, as well as inflammation and infection (209). The chemokine system is implicated in various pathologies, including cancer, rheumatoid arthritis, and HIV infection (210). Thus, chemokine receptors constitute highly relevant therapeutic targets. However, the success rate of drug discovery has been modest, showing that more efforts to understand their complexity are required to target this system effectively.

The chemokine system comprises more than 50 chemokines and 24 receptors in humans. While 20 chemokines typically signal via $G_{i/o}$ protein and subsequently arrestin pathways (211), the remaining four receptors (ACKR1–ACKR4) signal exclusively via the arrestin pathway and are therefore named atypical decoy receptors. The chemokines and their cognate receptors are divided into 4 classes according to the organization of the first two cysteine residues in the N-terminal region of the chemokine, resulting in CC, CXC, CX3C and XC groups. Some chemokine receptors are able to bind multiple chemokine ligands from the same group and vice versa. Additionally, chemokine receptors can exhibit signaling bias induced by specific chemokines, leading to unique functions of each ligand•receptor pair (212).

Chemokines are soluble proteins of 8–10 kDa molecular weight, which are secreted in response to inflammation. They share very similar three-dimensional structures, although their sequence homology is relatively low. Chemokines can have several functional forms: monomeric, dimeric and oligomeric. They bind their cognate receptors mainly as a monomer. However, when the dimer interface does not compete with binding to the receptor, they can also bind as a dimer (213,214). In addition, further intermolecular interactions may occur leading to higher oligomerization. There are several types of chemokine dimer formation (215). One occurs in CC chemokines via intermolecular β-sheet formation of the flexible N-termini, which prevents its binding to the receptor (216). A further mechanism found in CXC chemokines involves an intermolecular extension of the anti-parallel β-sheet, which is compatible with receptor binding (213,214). In addition, higher-order oligomerization may occur, which seems mainly due to electrostatic interactions. The propensity for oligomer formation varies among chemokines. In their oligomeric forms, chemokines interact with glycosaminoglycans (GAGs), thus mediating leucocyte recruitment (217).

Interestingly, some of the chemokine receptors are prone to form functional homo- and heterodimers and even oligomers (218–221). An additional level of complexity arises from post-translational modifications (PTMs) of the receptor N-termini, namely, tyrosine sulfation, glycosylation, and sialylation of the glycans (222). These PTMs contribute to the chemokine•receptor affinity via electrostatic interactions with the chemokine core (216,223–225).

The complexity of the chemokine signaling system and the high dynamics of the chemokine•receptor complexes make structural studies challenging. Therefore, most chemokine receptor structures have been solved in their inactive state in complex with small molecule ligands. In recent years, with cryo-EM development, considerable progress has been made in the structural elucidation of active chemokine receptors bound to agonist chemokines and G proteins. The solved structures include CCR1 (226), CCR2 (227), CCR5 (138,228), CCR6 (229) and CXCR2 (214). Despite this success, only three receptors (CCR2, CCR5 and CXCR2) have been captured in both inactive and active states, which revealed a diversity of activation mechanisms. It is worth mentioning that many of the receptor complexes were forced by a fusion of the chemokine with the receptor N-terminus along with other significant modifications (226–228), thereby potentially altering their native interactions. Moreover, the densities of the N-terminal portion of the chemokine are poorly defined in many cases, which impedes confident model building (Figure 1.8). This poor density apparently indicates a certain flexibility of the chemokine N-termini, adopting multiple conformations within the TM bundle of the receptor. Additional efforts and complementary approaches are required to obtain a full understanding of these intricate chemokine•receptor interactions.

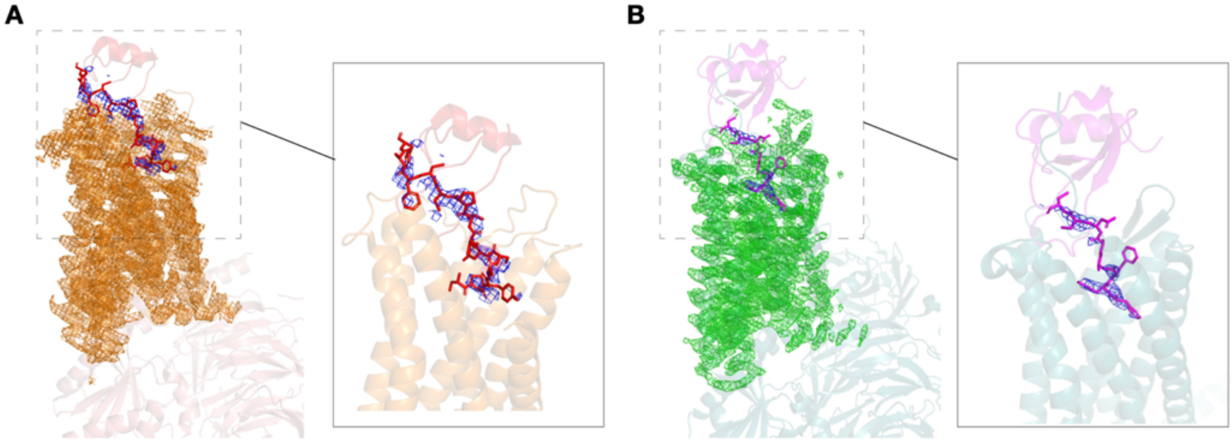


Figure 1.8. Examples of the modeled N-termini of the chemokines with the density being displayed. A. CCL5•CCR5 structure (PDB: 7F1R) with cryo-EM density being displayed at a 6 σ cut-off within 2 Å of the CCL5 N-terminus and the CCR5. B. CCL15 (26-92)•CCR1 (PDB: 7VL9) with cryo-EM density being displayed at a 6σ cut-off within 2 Å of the CCL15 N-terminus and the CCR1.

The human CC chemokine receptor 5 and its chemokine ligands

CCR5 is a medically highly relevant chemokine receptor, expressed in various immune cell types and other cell types, e.g., neurons. It is the major HIV-1 co-receptor. Since its discovery in 1996 (230–232), CCR5 has been under extensive investigation not only due to its major role in HIV infection (233) but also because it is involved in inflammation (234), the pathology of cancer (235,236), and COVID-19 (237). Several endogenous chemokine ligands bind CCR5, namely CCL3 (also known as MIP-1α), CCL4 (MIP-1β) and CCL5 (RANTES), causing its activation and subsequent immune responses.

Even before CCR5 had been identified, its chemokine ligands were found to inhibit HIV entry, with CCL5 being the most efficient in blocking the binding site for the viral glycoprotein gp120 and in promoting CCR5 endocytosis (238). However, the use of wild-type CCL5 as an anti-HIV agent is problematic, since CCL5 induces CCR5 signaling and activates T cells with the increased risk of severe side effects. Moreover, its anti-HIV potency is in the micromolar range (239), which is relatively low for therapeutic purposes. Therefore, several CCL5 analogs have been developed using phage display screening of mutations covering its N-terminal 9 amino acids (240) or chemical modifications of this N-terminal part (241). These modifications have yielded a set of CCL5 variants with enhanced anti-HIV potency and changed G protein and arrestin signaling properties (Table 1.1).

Table 1.1. List of the most relevant CCL5 analogs with diverse pharmacological properties (239–241,243,245).

Name	**N-terminal sequence**	**Anti-HIV potency** [pM]	**Pharmacological property**
CCL5	`.SPYSSDTTPCC-`	>10’000	agonist
[PSC]CCL5	`***SSDTTPCC-`	25	super-agonist
[6P4]CCL5	`QGPPGDIVLACC-`	21	super-agonist
[5P14]CCL5	`QGPPLMSLQVCC-`	26	partial agonist
[5P12]CCL5	`QGPPLMATQSCC-`	28	antagonist
[5P7]CCL5	`QGPPLMALQSCC-`	17	antagonist

One of the most promising CCL5 analogs as anti-HIV drug is [5P12]CCL5 (242). It inhibits HIV at low picomolar concentrations and exhibits antagonistic properties, reducing the risk of side effects, e.g., inflammatory responses (243). Recently [5P12]CCL5 has passed Phase 1 clinical trials (244). Further interesting CCL5 variants with high anti-HIV activity are listed in Table 1.1. [6P4]CCL5 and [PSC]CCL5 are strong agonists and capable of internalizing the receptor with longer retention times than wild-type CCL5. [5P14]CCL5 has been characterized as an arrestin-biased agonist (243). However, the observed effects seem to be cell-line dependent, and subsequent studies showed that it is instead a partial unbiased agonist (245). Finally, similar to [5P12]CCL5, [5P7]CCL5 also is a potent antagonist and also has picomolar anti-HIV potency. The inactive, chemokine-bound structure of CCR5 has been solved in complex with this ligand (246).

So far, maraviroc is the only CCR5-targeting drug that has been approved as an anti-HIV-1 agent. However, CCR5 is also critically involved in cancer and autoimmune disorders. Yet no other CCR5 drugs against these diseases have been approved. An advanced structural and functional understanding of chemokine receptor biology is expected to assist future drug discovery endeavors.

1.9 Aims of the thesis

The goals of this thesis were to investigate the activation mechanism of CCR5 by its cognate chemokines, to provide a structural explanation for the variable pharmacology of CCL5 N-terminal variants, and to investigate the structural basis of the functional interactions of CCR5 with effector proteins. By a combination of various structural and biophysical methods the following goals could be reached:

- Development of conditions for the assembly of a stable complex between wild-type CCR5, the super-agonist chemokine [6P4]CCL5, and the G_i heterotrimer
- Structure elucidation of the active wild-type CCR5 in complex with [6P4]CCL5 and the G_i heterotrimer by cryo-EM, which revealed the activation mechanism of CCR5
- Characterization of the mechanism of arrestin2 recruitment using CCR5 and its C-terminal phosphopeptides as a model system
- Development of conditions for stable complex formation of wild-type CCR5 with [6P4]CCL5 and arrestin2 and respective sample preparation for cryo-EM analysis

2 Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist

2.1 Original Manuscript

Reference: Isaikina P., Tsai C.-J., Dietz N., Pamula F., Grahl A., Goldie K.N., Guixà-González R., Branco C., Paolini-Bertrand M., Calo N., Cerini F., Schertler G.F.X., Hartley O.*, Stahlberg H., Maier T., Deupi X.*, Grzesiek S.* Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. *Science Advances*, **2021**
https://www.science.org/doi/10.1126/sciadv.abg8685
*Corresponding author

Short description: This publication describes the structure elucidation of CCR5 in complex with the chemokine super-agonist [6P4]CCL5 and the G_i protein, which revealed the CCR5 activation mechanism.

STRUCTURAL BIOLOGY

Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist

Polina Isaikina[1], Ching-Ju Tsai[2], Nikolaus Dietz[1], Filip Pamula[2,3], Anne Grahl[1], Kenneth N. Goldie[4], Ramon Guixà-González[2], Camila Branco[5], Marianne Paolini-Bertrand[5], Nicolas Calo[5], Fabrice Cerini[5], Gebhard F. X. Schertler[2,3]*, Oliver Hartley[5,6]*, Henning Stahlberg[4†], Timm Maier[1], Xavier Deupi[2]*, Stephan Grzesiek[1]*

The human CC chemokine receptor 5 (CCR5) is a G protein–coupled receptor (GPCR) that plays a major role in inflammation and is involved in cancer, HIV, and COVID-19. Despite its importance as a drug target, the molecular activation mechanism of CCR5, i.e., how chemokine agonists transduce the activation signal through the receptor, is yet unknown. Here, we report the cryo-EM structure of wild-type CCR5 in an active conformation bound to the chemokine super-agonist [6P4]CCL5 and the heterotrimeric G_i protein. The structure provides the rationale for the sequence-activity relation of agonist and antagonist chemokines. The N terminus of agonist chemokines pushes onto specific structural motifs at the bottom of the orthosteric pocket that activate the canonical GPCR microswitch network. This activation mechanism differs substantially from other CC chemokine receptors that bind chemokines with shorter N termini in a shallow binding mode involving unique sequence signatures and a specialized activation mechanism.

INTRODUCTION

The human CC chemokine receptor 5 (CCR5) is a G protein–coupled receptor (GPCR) that plays a major role in inflammation by recruiting and activating leukocytes (*1*). CCR5 is also the principal HIV coreceptor (*2*), is involved in the pathology of both cancer (*3*) and neuroinflammation (*4*), and has been implicated in the inflammatory complications of coronavirus disease 2019 (COVID-19) (*5*, *6*). Soon after the discovery of CCR5, it became evident that its natural chemokine ligands inhibit HIV entry (*7*), with CCL5 (RANTES) being most efficient, acting both by blocking the binding site for the viral glycoprotein gp120 and by promoting CCR5 endocytosis (*8*). Modifications of the N-terminal region of CCL5 preceding residue C10 yielded HIV entry inhibitors with significantly higher potency (*9*–*11*). These analogs belong to a group of over 100 engineered CCL5 N-terminal variants that show notable differences in their anti-HIV, endocytotic, affinity, and signaling properties ranging, e.g., from super-agonist to strong antagonist behavior (*10*, *11*). The molecular basis of these N-terminal structure–related activity differences is currently unclear.

Whereas a good structural understanding has been reached of the activation mechanisms of class A GPCRs by small-molecule ligands (*12*), the activation mechanism of the chemokine receptor subclass is not yet well understood. Inactive structures of a number of chemokine receptors have been solved, including complexes of CCR5 with the engineered chemokine antagonist [5P7]CCL5 (*13*), the viral gp120•human CD4 complex (*14*), the HIV inhibitor maraviroc (*15*), and other small-molecule antagonists (*16*). In contrast, only two active-state human chemokine receptor complex structures are currently available: CCL20•CCR6•G_o (*17*) and CXCL8•CXCR2•G_i (*18*). In these structures, CCL20 and CXCL8 adopt a shallow binding mode in which the chemokine N terminus is not deeply inserted into the orthosteric pocket and activation apparently involves transmission of forces directly from the extracellular domain of the receptor. There are also two inverse-agonist–bound (*19*), one agonist-bound, and one apo (*20*) structures available of the viral chemokine receptor US28. All these structures are in active conformation. However, US28 is constitutively active and can engage thousands of distinct chemokine sequences, some of them leading to a moderate increase in activity (*20*). Thus, chemokine-induced activation of US28 is thought to result from a rather sequence-insensitive mechanism, in which the steric bulk of the ligand is more important than specific interactions between the chemokine and the receptor (*20*).

In contrast to CCR6 and CXCR2, many native human chemokine receptors such as CCR5 have chemokine ligands with longer N termini, which likely insert more deeply into the orthosteric pocket of the receptor. As many agonist and antagonist CCL5 ligand variants have been identified for CCR5 that differ only in the composition of the first ~10 residues but not in their length (*11*). The decisive contacts for CCR5 signaling interactions must be located at the bottom of the orthosteric pocket rather than at the extracellular surface of the receptor.

With the aim of elucidating the apparently different activation mechanisms of CC chemokine receptors and to provide a general structural explanation for the variable pharmacology of CCL5 N-terminal variants, we solved the structure of wild-type human CCR5 in complex with the super-agonist [6P4]CCL5 and the G_i heterotrimer.

RESULTS

Overall structure of the [6P4]CCL5•CCR5•G_i complex

A stable [6P4]CCL5•CCR5•G_i complex was obtained by incubating detergent-solubilized human wild-type full-length CCR5 with the

---

[1]Focal Area Structural Biology and Biophysics, Biozentrum, University of Basel, CH-4056 Basel, Switzerland. [2]Paul Scherrer Institute, CH-5232 Villigen PSI, Switzerland. [3]Department of Biology, ETH Zurich, CH-8093 Zurich, Switzerland. [4]Center for Cellular Imaging and NanoAnalytics, Biozentrum, University of Basel, CH-4058 Basel, Switzerland. [5]Department of Pathology and Immunology, Faculty of Medicine, University of Geneva, Geneva, Switzerland. [6]Orion Biotechnology, Ottawa, Canada.
*Corresponding author. Email: stephan.grzesiek@unibas.ch (S.G.); xavier.deupi@psi.ch (X.D.); oliver.hartley@unige.ch (O.H.); gebhard.schertler@psi.ch (G.F.X.S.)
†Present address: Faculty of Biology and Medicine, UNIL, and Laboratory of Biological Electron Microscopy (LBEM), IPHYS, SB, EPFL, Route de la Sorge, CH-1015 Lausanne, Switzerland.

G_i heterotrimer and [6P4]CCL5 (fig. S1). The complex was treated with apyrase to hydrolyze guanosine diphosphate and was further stabilized by addition of the Fab fragment Fab16 (*21*, *22*), which recognizes an interface between the Gα and Gβγ subunits of the G_i heterotrimer (fig. S2). Single-particle cryo–electron microscopy (cryo-EM) analysis with extensive particle classification yielded a three-dimensional (3D) density map with a nominal global resolution of 3.15 Å (Fig. 1A, fig. S3, and table S1). The map is well resolved for most of parts of CCR5, the [6P4]CCL5 N terminus, the G_i heterotrimer (fig. S4), and Fab16. The density of the globular core of [6P4]CCL5 and the adjacent CCR5 N terminus and extracellular parts of the receptor have less defined density, indicating relative flexibility in these parts of the structure. A 3D variability analysis of the cryo-EM data (movie S1) and molecular dynamics (MD) simulations of the atomic model (fig. S5A) reveal a certain degree of mobility of the [6P4]CCL5 core, the receptor N terminus, the extra- and intracellular loops, and transmembrane (TM) helices 5, 6, and 7. Still, the MD simulations indicate persistent interactions mediated by the N-terminal residues 0 to 8, the β1/β3-strands, and the 30s loop of the chemokine (fig. S5B). Apart from a small ~5° difference in the orientation, the position of the [6P4]CCL5 core is very similar to that of [5P7]CCL5 in the inactive [5P7]CCL5•CCR5 complex (Figs. 1C and 2A). Nevertheless, this minor change in orientation also leads to small (1 to 2 Å) but noticeable movements at the extracellular ends of TM1 and TM7 (Fig. 2B).

The open conformation of the intracellular part of the active CCR5 differs from all inactive CCR5 structures, thereby enabling binding of the G protein (Fig. 1C): TM6 is moved outward from the heptahelical bundle accompanied by further rearrangements of TM5, TM7, and intracellular loop 4 (ICL4). The moderate outward movement of TM6 and the arrangement of G_i relative to CCR5 (Fig. 1, A to C) agree with previous GPCR•G_i complexes (*23*).

CRS1 interactions

The interactions between [6P4]CCL5 and CCR5 can be separated into the three canonical chemokine recognition sites (CRS): CRS 1, 1.5, and 2 (Fig. 2C) (*13*, *24*). CRS1 consists of the contacts of the chemokine core with the extracellular side of the receptor and is dominated by electrostatic interactions. The core of [6P4]CCL5 sits on top of a wide opening in the extracellular part of the CCR5 TM bundle,

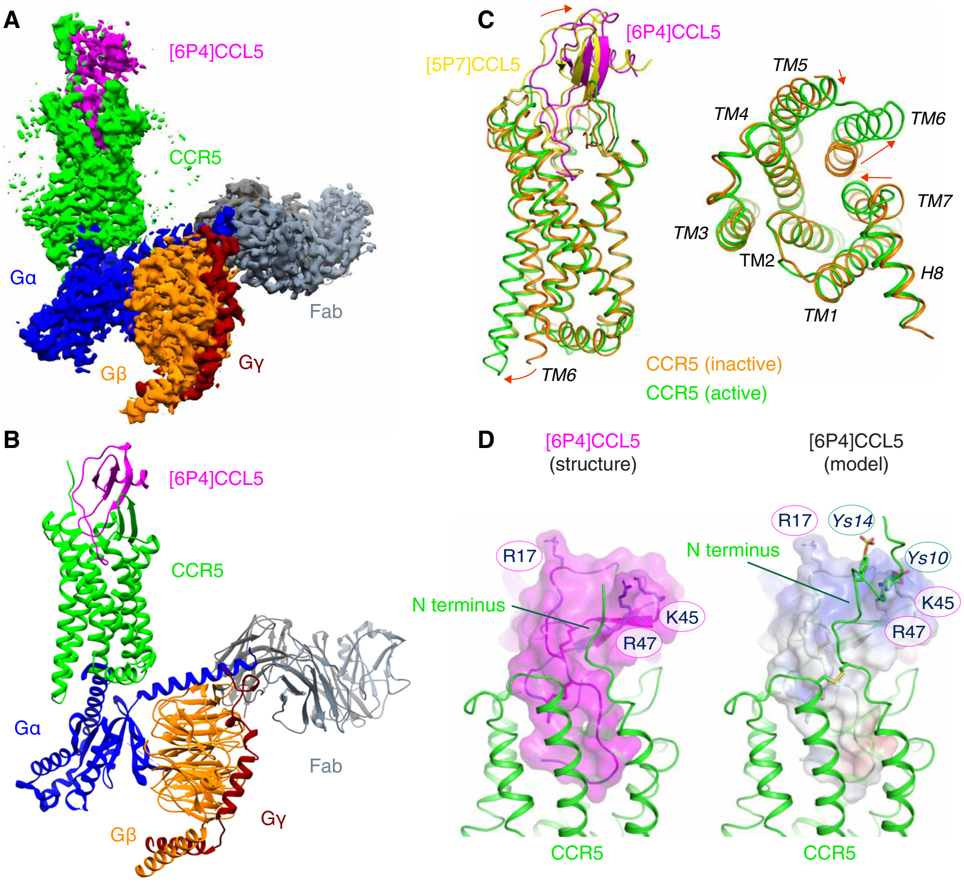


Fig. 1. Cryo-EM structure of the [6P4]CCL5•CCR5•Fab16 complex. (**A**) Cryo-EM map of the [6P4]CCL5•CCR5•G_i•Fab16 complex colored by subunits ([6P4]CCL5, magenta; CCR5, green; $Gα_i$, blue; Gβ, orange; Gγ, maroon; and Fab16, gray). (**B**) Atomic model of the [6P4]CCL5•CCR5•G_i•Fab16 complex in the same view and color scheme as shown in (A). (**C**) Side and cytoplasmic views of the structural overlay of active CCR5 (green) in complex [6P4]CCL5 (magenta) and inactive CCR5 [orange; Protein Data Bank (PDB) ID: 5UIW] in complex with [5P7]CCL5 (yellow). Substantial structural changes between two conformations are indicated by red arrows. The C101$^{3.25}$-C178^{ECL2}, C20$^{N\text{-}term}$-C269$^{7.25}$ disulfide bridges conserved in chemokine receptors are shown in dark yellow. (**D**) Interactions between the [6P4]CCL5 core and the CCR5 N terminus at the CRS1 site (left, cryo-EM structure; right, cryo-EM/NMR–based model). In the model, sulfo-tyrosines ^{S}Y10 and ^{S}Y14 are depicted as sticks, and the [6P4]CCL5 surface is colored according to its electrostatic potential (negative, red; positive, blue).

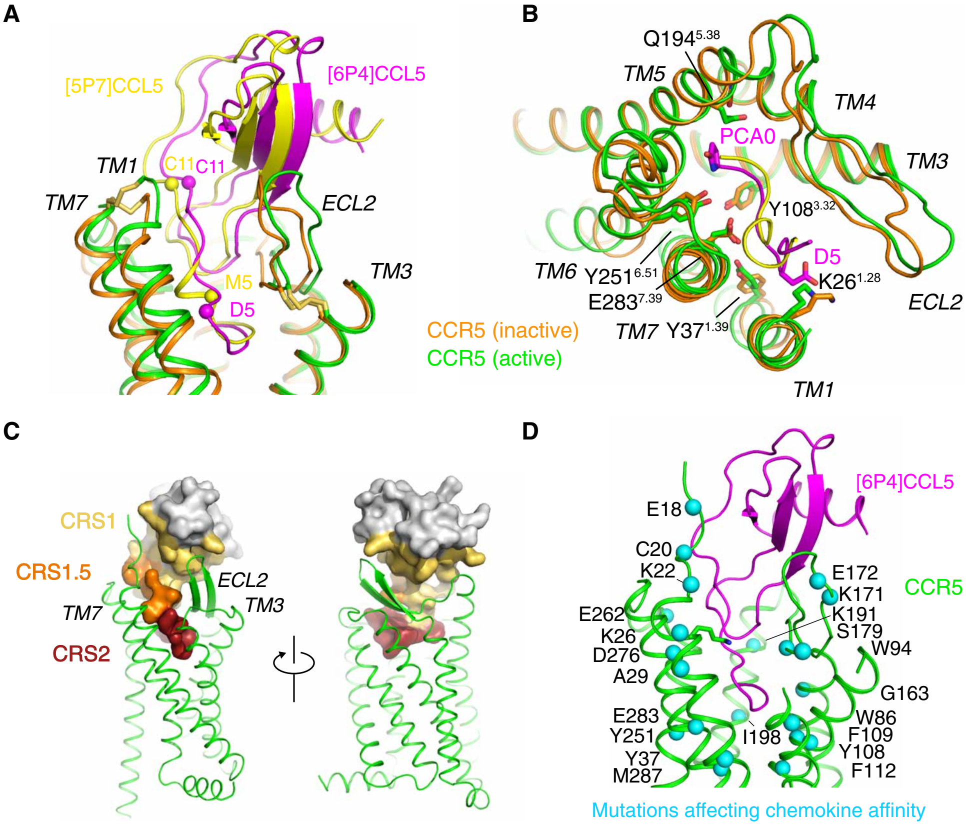


Fig. 2. Deep insertion of CCL5 chemokines into the orthosteric CCR5 ligand pocket. (**A** and **B**) Comparison between the insertion of the agonist [6P4]CCL5 (magenta) into active CCR5 (green) and the antagonist [5P7]CCL5 (yellow) into inactive CCR5 (orange; PDB ID: 5UIW) (A, side view; B, top view). Only the CCL5 N-terminal residues are shown in (B). Important residues participating in the CCL5-CCR5 interaction are marked. (**C**) Location of CRS1, CRS1.5, and CRS2 chemokine recognition sites in the [6P4]CCL5•CCR5•G_i complex. (**D**) Locations of previously identified CCR5 point mutants [as reviewed by (*13*)] that affect chemokine affinity within CCR5. Respective residues are shown in the [6P4]CCL5•CCR5•G_i complex as blue spheres. [6P4]CCL5 (magenta) and CCR5 (green) are shown in cartoon.

which is shaped by two disulfide bridges [C101$^{3.25}$-C178^{ECL2}, conserved in class A GPCRs, and C20$^{N\text{-term}}$-C269$^{7.25}$, specific to chemokine receptors; superscripts indicate the GPCRdb numbering scheme (*25*)] (Figs. 1C and 2A). The [6P4]CCL5 strand β_1 makes extensive contacts with polar residues in extracellular loop (ECL) 2, while the CCR5 N terminus directs toward a shallow groove between the chemokine N-loop and 40s loop forming further extensive ionic and polar interactions. Interactions between CCR5 residues S17 and E18 and the chemokine residues R47 and Q48 are visible in the density. However, poorly defined density prevented building a model of the CCR5 N-terminal residues 1 to 15 with confidence. To gain insights into this region, these CCR5 residues were modeled (Fig. 1D) on the basis of the nuclear magnetic resonance (NMR) structure of CCL5 in complex with an N-terminal fragment (residues 1 to 27) of CCR5 sulfated at residues Y10 and Y14 (*26*). Sulfation at Y10 and Y14 of CCR5 is important for chemokine affinity (*26*–*29*) and is expected to be present also in the insect cell–expressed CCR5 used in the current study (*30*). The stability of the modeled interactions between the sulfated CCR5 N terminus and [6P4]CCL5 was assessed by MD (movie S2). The simulations reveal persistent interactions between ^{s}Y10 and ^{s}Y14 of CCR5 and residues in the N-loop, 40s loop, and β3-strand (fig. S6), including K45, R47, and R17 of [6P4]CCL5, in complete agreement with the NMR-observed nuclear Overhauser effect (NOE) contacts (*29*). In addition, comparison of MD trajectories between sulfated and nonsulfated CCR5 indicates that sulfation induces a higher number of contacts between the chemokine and the receptor N terminus (fig. S6), consistent with the higher affinity of the sulfated form (*28*).

CRS2 interactions and activation

The N terminus of [6P4]CCL5 reaches deep into the orthosteric pocket (CRS2) between the CCR5 7TM bundle (Fig. 2). In complete agreement with this deep binding mode, point mutations of many CCR5 residues lining the CRS2 site, the CRS1.5 site at the rim of the orthosteric pocket, and the extracellular N-terminal CRS1 site have been shown previously to affect the affinity of CCR5 for chemokines (Fig. 2D). This deep insertion contrasts with the shallow binding modes observed for the chemokines in the CCL20•CCR6•G_o (*17*) and CXCL8•CXCR2•G_i (*18*) complexes (see below). Of note, the N-terminal residues preceding C10 and C11 of monomeric CCL5 in solution undergo large amplitude motions on the nanosecond time scale, as revealed by ^{15}N relaxation data (*31*). However, they adopt a fixed conformation in the CCR5 complex.

The [6P4]CCL5 residues 0 to 3 form the distal N terminus, which is located at the bottom of CRS2 (Figs. 2, A and C, and 3, A and B). As compared to [5P7]CCL5, the deeper binding pose of [6P4]CCL5 slightly relocates the N-terminal pyroglutamate (PCA) group (Fig. 3A). The packing of CCR5 against the PCA group is not very

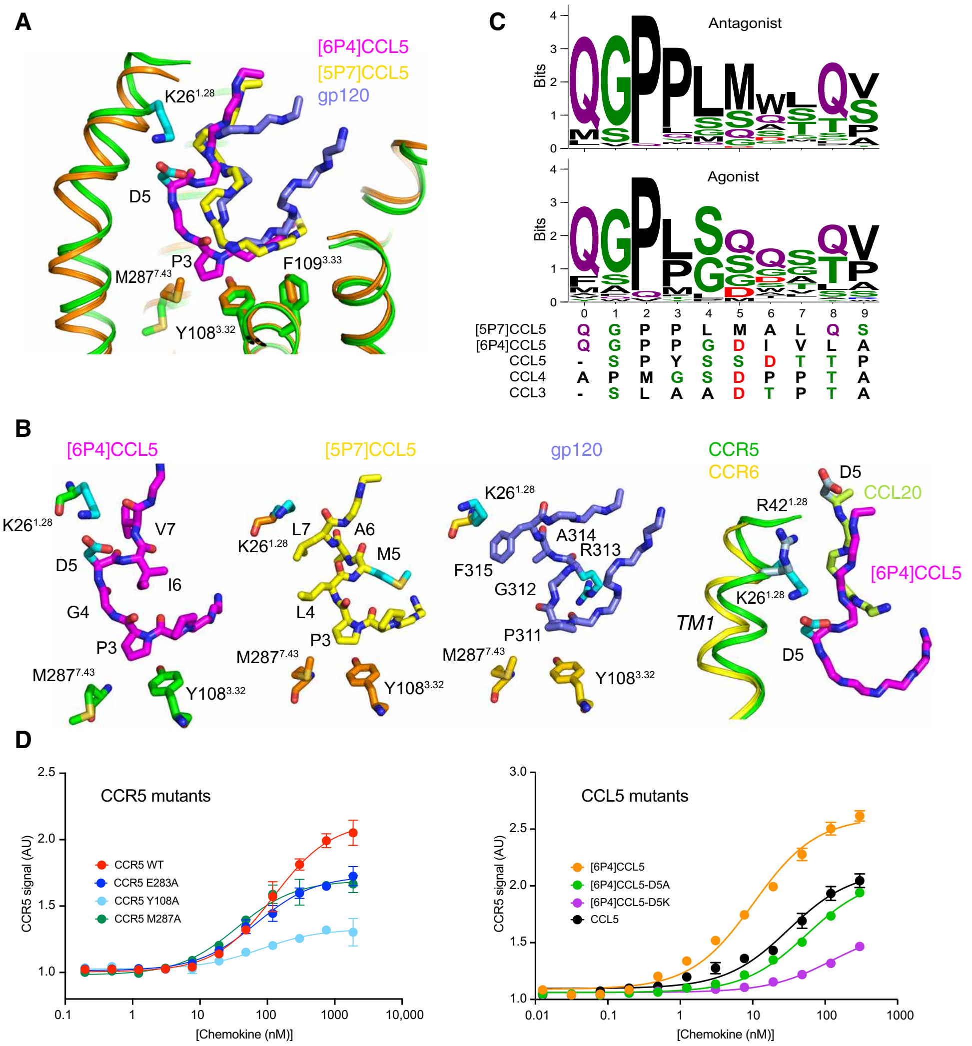


Fig. 3. Activation mechanism of CCR5 by [6P4]CCL5 at CRS2. (**A**) Comparison of insertion depths of agonist [6P4]CCL5 (magenta), antagonist [5P7]CCL5 (yellow), and the antagonist V3 loop of gp120 (slate; PDB ID: 6MEO) into active CCR5 (green) and inactive CCR5 (orange). Important interacting residues are shown as sticks. (**B**) Detailed view of [6P4]CCL5 N terminus inserted into active CCR5, [5P7]CCL5 N terminus, and gp120 V3 loop inserted into inactive CCR5; and comparison of the insertion of [6P4]CCL5 N terminus into active CCR5 and of CCL20 N terminus into active CCR6 (PDB ID: 6WWZ). (**C**) Sequence composition of the N termini of CCL5 natural amino acid variants (table S2) with low (≤20%, $N = 83$, top) and high (≥50%, $N = 34$, bottom) calcium signaling. The N termini of [5P7]CCL5, [6P4]CCL5, and wild-type CCL3 to CCL5 (all agonists) are shown below. (**D**) Effect of CCR5 (left) and CCL5 (right) mutations on CCR5 signaling activity as monitored by Ca^{2+} flux measurements in human embryonic kidney (HEK) cells. Data points represent mean peak height ± SD ($n = 3$). Data shown are representative of two (left) or three (right) independent experiments. Fits to data points by a three-parameter response model are shown as solid lines. Respective fit values for E_{max} and EC_{50} are listed in table S3. AU, arbitrary units.

tight, and our MD simulations of the [6P4]CCL5•CCR5 complex show that the PCA group interacts through dynamic water-mediated hydrogen bonds with nearby residues Q194$^{5.38}$ and Y251$^{6.51}$ (fig. S7A). An inspection of the electron density of the [5P7]CCL5•CCR5 crystal structure (*13*) allowed us to also model additional water molecules in the vicinity of the [5P7]CCL5-PCA, which similarly connect to surrounding residues K191$^{5.35}$, Q194$^{5.38}$, Y251$^{6.51}$, N258$^{6.58}$, and T259$^{6.59}$ (fig. S7B). Thus, the PCA group of the chemokine does not appear to have well-defined contacts to CCR5. Rather, the surrounding CCR5 cavity may accommodate even larger moieties such as the alkyl chains of AOP- (*9*) or PSC-CCL5 (*32*), thereby increasing the potency of these ligands.

[6P4]CCL5 residues 4 to 9 form the proximal N terminus, which acts as a hinge between the chemokine core and the distal N terminus (Figs. 2, A and C, and 3, A and B). Conspicuously, residues P3 to D5, which constitute the turn between the proximal and the distal [6P4]CCL5 N terminus, insert several angstroms deeper into the CCR5 orthosteric pocket than the corresponding residues of [5P7]CCL5 or the V3 loop of gp120 in the respective inactive complexes with CCR5 (Fig. 3A). Because of this deeper insertion, P3 of [6P4]CCL5 can displace CCR5 M287$^{7.43}$ and Y108$^{3.32}$ (Fig. 3, A and B), thereby apparently activating the canonical GPCR microswitch network (see below), which remains in the inactive conformation in the [5P7]CCL5 or gp120 complexes. A similar CCR5 activation mechanism involving M287$^{7.43}$ has been suggested on the basis of a computational model of the CCR5/CCL5 complex (*13*).

The deeper pose of the [6P4]CCL5 N terminus partially overlaps with that of the antagonist maraviroc (fig. S8) (*15*). However, maraviroc inserts its phenyl ring between Y108$^{3.32}$ and F109$^{3.33}$ of the "aromatic connector" (see below), thereby apparently blocking the conformational rearrangement necessary for activation. A comparison of the chemokine N termini in the [6P4]CCL5•CCR5 versus the CCL20•CCR6 complexes (Fig. 3B) also shows the much reduced insertion depth of the latter, which prevents it from reaching sites corresponding to the activation switches identified at the bottom of the orthosteric pocket of CCR5.

The different insertion depth of the [6P4]CCL5 and [5P7]CCL5 N termini into CRS2 is caused by a markedly different structure of their proximal N-terminal residues 5 to 8, a short helix in [5P7]CCL5 and an extended coil in [6P4]CCL5 (Fig. 2, A and B, and 3, A and B). The hinge function of this structure is presumably key to the control of receptor activation. In [6P4]CCL5, D5 of the extended hinge forms an ionic interaction with K26$^{1.28}$ in TM1, whereas the side chain of the equivalent M5 of [5P7]CCL5 points in the opposite direction forming a helical turn (Fig. 3B). Apparently, this helical turn is also pushed sideways by unfavorable interactions between [5P7]CCL5 L7 and K26$^{1.28}$. Very similar interactions and conformations are present in the inactive gp120•CCR5 complex, with F315, R313, and P311 taking the roles of [5P7]CCL5 L7, M5, and P3, respectively. Besides the ionic D5-K26$^{1.28}$ interactions, the extended backbone at residues 2, 4, and 5 of the active [6P4]CCL5 is further stabilized by contacts to E283$^{7.39}$.

The structural finding that the CCL5 the N-terminal hinge conformation controls the insertion depth of its residues 3 to 5 and thereby the activation state of CCR5 is corroborated by a statistical analysis of the pharmacological properties of CCL5 N-terminal amino acid variants. Currently, ~140 of these have been characterized for G protein signaling (table S2), CCR5 internalization, and anti-HIV activity (*11*). Sequence analysis shows that residues 0 to 3 (highest abundance: QGPL, distal N terminus) and 8 and 9 (highest abundance: QV, proximal N terminus) are highly similar between N-terminal variants with low ($N = 83$) and high ($N = 34$) signaling activity (Fig. 3C). The latter is expected since the panel of tested variants was to some extent biased toward these residues (*11*). In contrast, strong differences are observed for residues 4 to 7 in the proximal N terminus: In agonist variants, the small, hydrophilic, or negatively charged amino acids S, Q, G, and D dominate, whereas antagonist variants contain mostly the large hydrophobic amino acids L, M, and W. Apparently, the small hydrophilic residues direct the hinge toward K26$^{1.28}$ in TM1, whereas the large hydrophobic residues make the hinge collapse to a helical turn. In agreement with their agonist pharmacology, both [6P4]CCL5 and wild-type CCL5 as well as the other major CCR5 agonist chemokines CCL3 (MIP1α) and CCL4 (MIP1β) contain an aspartic acid residue at positions 5 or 6 (Fig. 3C), which presumably stabilizes the extended hinge by forming a salt bridge to K26$^{1.28}$.

Essential parts of the proposed CCR5 activation mechanism were tested by CCR5 and [6P4]CCL5 point mutants using cellular Ca^{2+} flux activation assays in human embryonic kidney (HEK) cells (Fig. 3D and table S3). Consistent with our model, the CCR5 M287$^{7.43}$A, Y108$^{3.32}$A, and E283$^{7.39}$A mutations all reduced E_{max} for [6P4]CCL5 activation by ~40 to 70% without affecting the half-maximal effective concentration (EC_{50}). This indicates that these mutations reduce signaling without modifying chemokine affinity. In contrast, the [6P4]CCL5 D5A and D5K mutations decreased E_{max} by 30 and 60%, respectively and increased EC_{50} ~5- to 10-fold. Similar observations were made in Chinese hamster ovary (CHO) cells expressing CCR5 (table S3). The fact that these [6P4]CCL5 mutations reduce both signaling and binding affinity agrees with our proposed mechanism, since the unfavorable interactions of the CCR5 residue K26$^{1.28}$ with an alanine or lysine at position 5 in [6P4]CCL5 are expected to alter the hinge structure and reduce the binding enthalpy.

Signal transmission from CRS2 to the microswitch network

As described above, the straight conformation of the [6P4]CCL5 proximal N terminus pushes the region around residue P3 toward the bottom of CRS2, with the backbone of residues 2, 4, and 5 of [6P4]CCL5 interacting with E283$^{7.39}$ (Fig. 4A). The straight hinge appears to be further stabilized by contacts of [6P4]CCL5 residues P3 to D5 to a cluster of the hydrophobic CCR5 residues F85$^{2.59}$, W86$^{2.60}$, Y89$^{2.63}$, and L104$^{3.28}$. These contacts may act as "counter bearing" to promote the force of the [6P4]CCL5 N terminus toward the bottom of CRS2.

The deeper placement of residue P3 of [6P4]CCL5 as compared to [5P7]CCL5 forces a relocation of M287$^{7.43}$ in the receptor (Figs. 3A and 4A), which is accompanied by noticeable local changes in the backbone of TM7 (fig. S9) that bring the intracellular half of this helix toward the receptor core. This movement allows H289$^{7.45}$ to push onto W248$^{6.48}$, possibly assisting the relocation of TM6 (Fig. 4A). P3 also lies on top of an aromatic connector formed by CCR5 residues Y108$^{3.32}$, F109$^{3.33}$, and F112$^{3.36}$ (Fig. 4B), forcing the movement of Y108$^{3.32}$ and resulting in a cascade of aromatic side chain relocations that transmit the activation signal to the receptor core. This apparently switches the PIF motif (P206$^{5.50}$, I116$^{3.40}$, and Y244$^{6.44}$) to an active conformation (Fig. 4D) and induces the large-scale movement of TM6. The relocation of TM6 and TM7 coincides with local structural changes in the NPxxY motif (Fig. 4C), leading to the formation of the conserved water-mediated interaction between Y297$^{7.53}$ and Y214$^{5.58}$ (*33*) and the opening of the binding pocket for H5 of G_i, which includes R126$^{3.50}$ in the open conformation of the intrahelical ionic lock of the DRY motif (Fig. 4E). The MD simulations show that all these mentioned residues maintain stable contacts corresponding to the active conformation of the receptor (fig. S10).

An overview of all CCR5 point mutations (either previously described or generated in this study) that hinder signaling but do not reduce chemokine binding (Fig. 4F) confirms the essential aspects of the global signal transmission from the chemokine binding site to the canonical GPCR microswitch network. Such mutations comprise (i) F85$^{2.59}$, Y89$^{2.63}$, and L104$^{3.28}$ (*34*, *35*), suggesting that shaping and clasping of the hinge by the TM2/TM3 counter bearing is necessary for agonist efficacy; (ii) Y108$^{3.32}$, F109$^{3.33}$, and F112$^{3.36}$ [(*35*) and this study], proving the importance of the aromatic connector; and (iii) E283$^{7.39}$ and M287$^{7.43}$ (this study), showing the involvement of residues in TM7 in the shaping of the agonist conformation of the chemokine and as a possible route to the rearrangement of W248$^{6.48}$.

G_i interactions

The binding interface of G_i to CCR5 is mediated exclusively by the $G\alpha$ subunit (fig. S11) and can be divided into two main regions: the rim and the core (Fig. 5A and fig. S12). The rim contains two

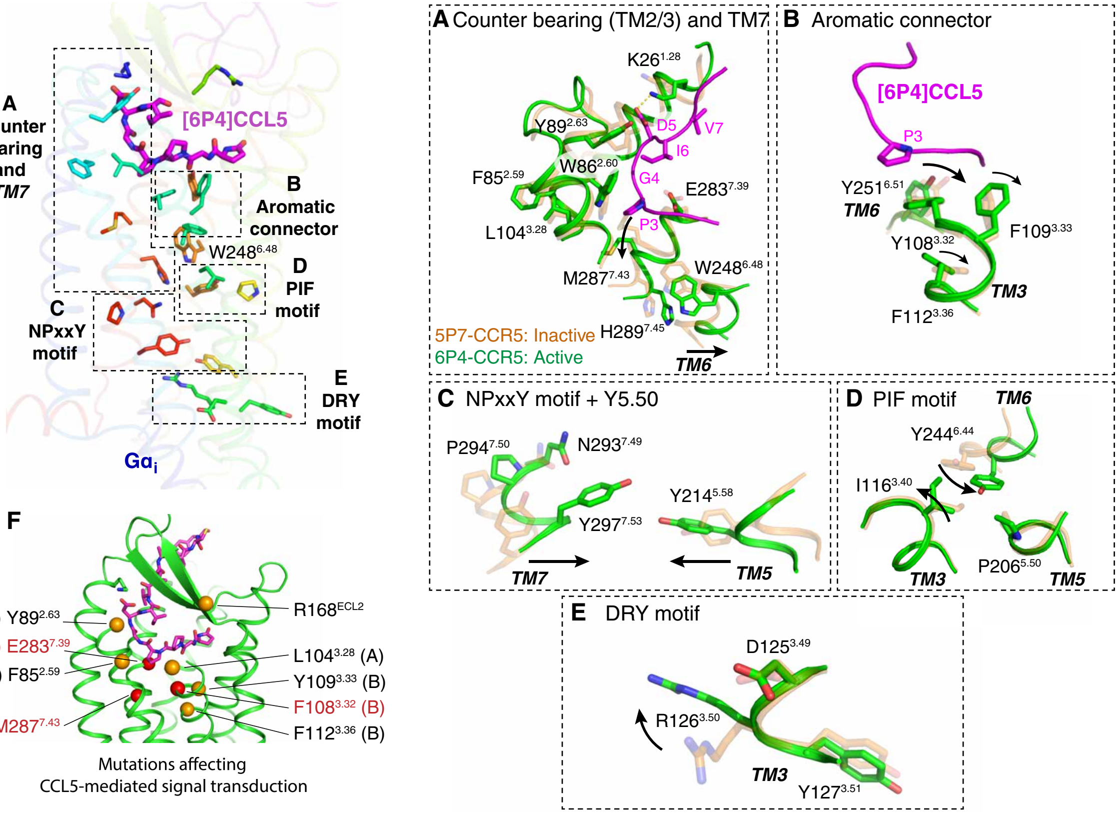


Fig. 4. Transmission of the chemokine agonist signal to the receptor activation switches. Top left: Residue groups connecting the N-terminal region of [6P4]CCL5 (magenta) to key CCR5 activation switches. Relevant residues are shown as sticks. For clarity, only part of the receptor structure is shown. (**A**) Counter bearing in TM2/3 and transmission through TM7. (**B**) Transmission through the aromatic connector. (**C**) Activation of the NPxxY motif and Y214$^{5.58}$ in TM5. (**D**) Activation of the PIF motif. (**E**) Activation of the DRY motif. (**F**) Point mutations shown to affect CCR5 activation but not chemokine affinity [orange, mutations from literature (see main text); red, mutations performed in this work]. The letters next to the residue numbers direct to the panel where these residues are shown in more detail.

clearly separated parts (Fig. 5B): a proximal side formed by the end of the αN helix (αNβ1) and nearby β strands (β2β3) in Gα and ICL2 in the receptor and a distal side formed by β strands (α4β6) in Gα and ICL3 in the receptor. In addition, the rim also includes interactions between α5 in G$_i$ and ICL2/3 of the receptor. The core of the CCR5•G$_i$ complex interface is formed exclusively by interactions of α5 in G$_i$ (Fig. 5A). Here, the G$_i$ α5 helix interacts with the cytoplasmic sides of TM2, TM3, and TM5 in one side of the core binding pocket, while the C-terminal hook of α5 (residues 352 to 354) leans toward TM6 and ICL4.

These interfaces are common to all GPCR/G protein complexes, as they arise from the common overall relative orientation of the bound components. However, analysis of the currently available complexes reveals that the precise location and nature of the individual interface contacts vary to a certain degree (fig. S13). At the proximal rim of the interface, contacts are mostly hydrophobic and consistent with other G$_i$ complexes. At the distal rim, we observe several ionic interactions absent in other structures. However, the most noticeable differences lie in the core region of the binding interface, where we observe different contacts between the hook of α5 (the last three C-terminal residues of G$_i$) and ICL4 of CCR5. This is due to a distinct conformation of ICL4 of CCR5 in which G301$^{8.47}$ and E302$^{8.48}$ slightly relocate compared to, e.g., the G$_i$ complexes of the neurotensin type 1 (NT1R) or μ-opioid (μOPR) receptors, resulting in a different set of interactions between E302$^{8.48}$ and the hook of α5 (Fig. 5C). A 3D variability analysis of the cryo-EM density reveals structural heterogeneity around ICL4 that allows to model a main conformation as shown in Fig. 5C and a second minor conformation that is similar to the NTR1 and μOPR complexes (fig. S14, A to C). The MD simulations indicate that ICL4 reverts to a preferred conformation in the absence of Gα$_i$ (fig. S14D). An analysis of further solved GPCR•G$_i$ structures also highlights the structural plasticity of ICL4 (fig. S15).

Structure-activity relationship of CCR5 chemokine ligands

The comparison between our structure and the inactive [5P7] CCL5•CCR5 complex (*13*) allows us to precisely pinpoint the activation mechanism of CCR5 by a chemokine agonist (Fig. 6A). The overall binding poses of the [5P7]CCL5 antagonist (*13*) and the [6P4]CCL5 agonist are similar, with the globular core of the chemokine held by the receptor N terminus and ECL2 and the chemokine N terminus reaching deep into the receptor TM bundle. However, despite having the same 10-residue length, the N termini of the two CCL5 derivatives differ in their amino acid sequences. This results in different chemokine/receptor interactions in this region: small, hydrophilic, or negatively charged residues in sequence positions 4 and 5 of [6P4]CCL5 lead to a straight conformation of the proximal N terminus that pushes residue P3 against the bottom of CRS2. Thus, P3 exerts a force that is bolstered on the counter bearing hydrophobic residues W86$^{2.60}$ and Y89$^{2.63}$ onto the aromatic connector

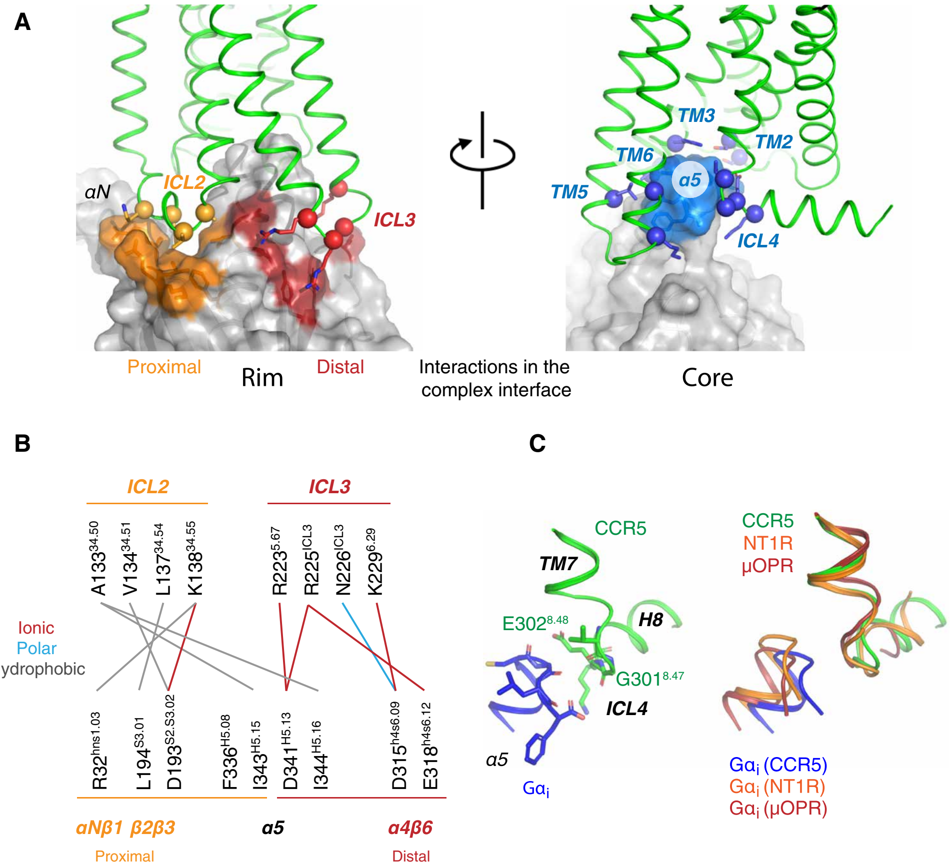


Fig. 5. Binding interfaces between CCR5 and $G\alpha_i$. (**A**) Binding interfaces at the rim (left) and core (right) of the complex. Each interface is colored on the surface of the G protein and interacting residues in the receptor are shown as spheres (Cα) and sticks (side chains). (**B**) Residue-residue interactions at the rim of the binding interface. (**C**) Left: Structure of the α5 hook and ICL4 in the [6P4]CCL5•CCR5 complex. Key residues are shown as sticks. Right: Comparison between the α5 hook and ICL4 in the G_i-bound receptors CCR5, neurotensin type 1 (NT1R), and μ-opioid (μOPR).

and residue M287$^{7.43}$. This triggers the canonical GPCR activation switches resulting in the relocation of TM5/6/7 and the stabilization of the receptor active conformation. In contrast, the large hydrophobic residues at positions 4 and 5 in [5P7]CCL5 force the proximal N-terminal hinge into a turn structure making P3 recede (Figs. 6A and 3B) and thereby leaving the receptor in the inactive state. The highly conserved (~70% in nonolfactory human class A GPCRs) residue W248$^{6.48}$ lies at the center of these activating conformational changes, connecting the rearrangements at H289$^{7.45}$ and Y244$^{6.44}$ and, thus, the large-scale relocation of TM7 and TM6.

On the basis of their N-terminal sequence, we expect that other identified CCR5 agonist or antagonist chemokines feature respective similar deeper (6P4[CCL5]-active-like) or less deep (5P7[CCL5]-inactive-like) positions of their N-terminal turns within CRS2. Using our structure as a template, we modeled the wild-type agonist CCL5 bound to CCR5 (fig. S16 and movie S3). CCL5, as the [6P4] CCL5 agonist, features an aspartate in its N terminus (D6) able to interact with K26$^{1.28}$. The MD simulations reveal a similar deep binding pose of the CCL5 N terminus with a straight-hinge conformation of residues 5 to 8 where Y3 could be playing the role of P3 in [6P4]CCL5 to engage the aromatic connector and M287$^{7.43}$ (Fig. 6A). As expected, a previous model of the active CCL5•CCR5 complex based on the inactive [5P7]CCL5•CCR5 structure (*13*) does not show this straight-hinge conformation but rather the inactive helical turn. The CCR5 chemokine agonist ligands CCL3 and CCL4 are closely related to CCL5 having similar N-terminal sequence lengths and compositions (Fig. 3C). We therefore expect that these chemokines also adopt the straight-hinge conformation and use the same activation mechanism as [6P4]CCL5 or CCL5 with the aspartates at position 5 and the bulky residues at position 2 carrying out analogous functions.

DISCUSSION

The activation mechanism in CCL5/CCR5, in which the N terminus of the chemokine reaches deep into the TM bundle, differs substantially from that of CCL20/CCR6, where a much shorter CCL20 adopts a shallower binding pose and engages a noncanonical activation mechanism (Figs. 3B and 6B) (*17*). Thus, CC chemokine receptors can apparently be activated through two very different mechanisms by "long" and "short" chemokines. But what are the molecular features in the receptor that determine the type of activation? A phylogenetic analysis of CC chemokine receptors (fig. S17) puts CCR5 and CCR6 into distinct subgroups. A more detailed sequence comparison of key residues in the activation mechanism shows that CC chemokine receptors can be divided into two main groups according to the nature of the residue at position 6.48 (W versus Q) and, to some extent, of the aromatic connector (Fig. 6C). CC chemokine

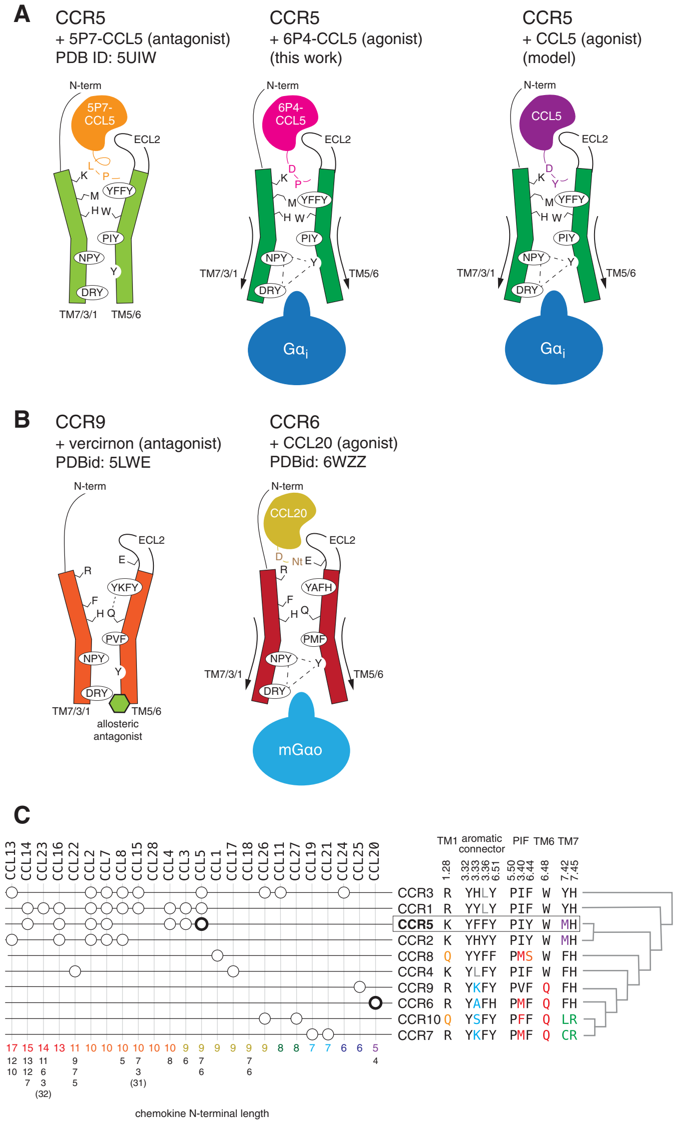


Fig. 6. Activation mechanism of CC chemokine receptors. (**A**) CCR5 bound to the antagonist [5P7]CCL5 (left, orange), the super-agonist [6P4]CCL5 (center, magenta), and the natural agonist CCL5 (right, purple). Key residues in the activation mechanism of CCR5 are shown. (**B**) Proposed activation mechanism for CCR6 by CCL20 (yellow, right) by comparing with the structure of CCR9 (left). (**C**) Pairing between CC chemokine receptors and CCL chemokines (*59*). At the right, the sequence composition of key positions is shown, together with the phylogenetic relationship between the receptors. The lengths of the CCL chemokine N termini according to UniProt (*60*) are shown at the bottom. The available active CCR/CCL complex structures are shown in bold.

receptors featuring the conserved W$^{6.48}$ (CCR1, CCR2, CCR3, CCR4, CCR5, and CCR8) tend to be more promiscuous and preferentially recognize chemokines with longer N termini (9 to 14 residues) (Fig. 6C). On the other hand, CC chemokine receptors featuring Q$^{6.48}$ (CCR6, CCR9, CCR7, and CCR10) bind to only a few (1 to 2) chemokines with short N termini (4 to 9 residues). Although position 6.48 allows for a certain degree of variability in human class A GPCRs (70% W, 15% F, 5% Y, and 10% other), a Q at this position is exclusive of this subgroup of chemokine human receptors, supporting the uniqueness of this "shallow" activation mechanism.

It is interesting to observe that many chemokines undergo post-translational proteolytic processing leading to different N-terminal lengths, which may constitute a layer of regulation (*36*, *37*). Thus, a CCL5 variant lacking the first two N-terminal residues (CCL5$^{3-68}$) behaves as a natural chemotaxis inhibitor, and a 10-fold higher concentration compared to wild-type CCL5 is required to induce a significant calcium response (*37*). Similarly, CCL5$^{4-68}$ has an about 10-fold lower affinity for CCR5 compared to CCL5$^{3-68}$ or wild-type CCL5 and is less potent in stimulating lymphocyte chemotaxis or inhibiting HIV infection (*36*). These findings are in complete agreement with the lack of contacts at the bottom of the CCR5 CRS2 region expected for such CCL5 truncations.

The structure of CCR5 in an active conformation allows us to elucidate a novel activation pathway of CC chemokine receptors by a chemokine agonist. In CCR5 and related receptors (CCR1, CCR2, CCR3, and CCR4), the respective cognate chemokines have long N termini and bind deep into the orthosteric pocket (CRS2), thereby triggering the rearrangement of an aromatic connector in TM3 and TM6 and of the TM7 backbone. The activating force exerted by the deep binding [6P4]CCL5 N terminus appears to be stabilized by a cluster of hydrophobic CCR5 residues in TM2 and TM3 that line the extended N-terminal hinge of this agonist chemokine. Understanding this force balance may help in the design of small-molecule agonists, which could activate the connector region at the bottom of CRS2 by pushing against this counter bearing.

W$^{6.48}$ lies at the center of these conformational changes connecting the receptor activation pathways through TM7 and TM6. In contrast, a subgroup of CC chemokine receptors (CCR6, CCR7, CCR9, and CCR10) harbors a Q residue at this position, a unique feature in human class A GPCRs. The cognate chemokines of these receptors have shorter N termini featuring a shallow binding mode and a specialized mode of activation. We expect that our findings will help to rationalize the relationship between sequence, structure, and activity of chemokines and their receptors and aid drug discovery.

METHODS

Protein expression and purification

The wild-type human CCR5 gene containing a C-terminal 3C cleavage site followed by a FLAG-tag was cloned into the pFastBac1 vector and expressed in *Spodoptera frugiperda Sf9* insect cells using the baculoviral infection system. CCR5 expression and membrane preparation were performed as described (*15*). Membranes from a 1-liter culture of Sf9 cells were resuspended in 10 ml of lysis buffer containing iodoacetamide (2 mg/ml), and EDTA-free complete protease inhibitor cocktail tablets, and incubated at 4°C for 1 hour. Then, membranes were solubilized by supplementing 0.5% lauryl maltose neopentyl glycol (LMNG) at 4°C for 3 hours. The soluble fraction was isolated by centrifugation at 140,000*g* and incubated with 1 ml of M2 anti-FLAG affinity resin overnight at 4°C. The latter column was washed with 10 column volumes (CV) of washing buffer 1 [25 mM Hepes, 400 mM NaCl, 10% glycerol, and 0.1% LMNG (w/v), pH 7.5], followed by 10 CV of washing buffer 2 (25 mM Hepes, 400 mM NaCl, 2 mM adenosine 5′-triphosphate, 5 mM $MgCl_2$, 10% glycerol, and 0.1% LMNG, pH 7.5) and subsequently washed with another 6 CV of washing buffer 1. The receptor was eluted with 3 CV of elution buffer consisting of 25 mM Hepes, 400 mM NaCl, 0.01% LMNG, and FLAG peptide (200 μg/ml; DYKDDDDK; pH 7.5).

A DNA construct of [5P14]CCL5 cloned into a pET32a vector was a gift of P. LiWang. The DNA sequence of [6P4]CCL5 was obtained by mutating this [5P14]CCL5 construct using standard QuickChange polymerase chain reaction. [6P4]CCL5 with enterokinase-cleavable N-terminal thioredoxin fusion and hexa-histidine tags was expressed in the *Escherichia coli* BL21 (DE3) strain cultured in Lysogeny broth media. Protein production was induced with 1 mM isopropyl β-D-thiogalactopyranoside when the optical density at 600 nm reached 0.7 to 0.8. After induction, cells were grown for 20 hours at 22°C and then harvested by centrifugation. Ten grams of the cell pellet was resuspended in 50 ml of resuspension buffer (50 mM tris, 6 M guanidinium HCl, and 200 mM NaCl, pH 8.0) and lysed using a French press. The supernatant was isolated by centrifugation at 27,000*g* for 1 hour and applied to a 5-ml HisTrap column. The column was washed with 10 CV of resuspension buffer and eluted with 3 CV of 60 mM NaOAc, 200 mM NaCl, and 6 M guanidinium HCl. β-Mercaptoethanol (20 mM) was added to the elution fraction and incubated for 1 hour. The denatured protein was added dropwise into 250 ml of folding buffer (550 mM L-arginine hydrochloride, 20 mM tris, 200 mM NaCl, 1 mM EDTA, 1 mM reduced glutathione, and 0.1 oxidized glutathione, pH 8.0) and incubated overnight at 4°C. The solution was concentrated [molecular weight cutoff (MWCO), 10 kDa] and dialyzed in 20 mM tris, 200 mM NaCl, and 2 mM $CaCl_2$ (pH 8.0). To cleave the fusion tags, enterokinase (New England Biolabs) was added, and the solution was incubated for 24 hours at room temperature. The protein was separated from the fusion tag using an acetonitrile gradient on a C4 reversed-phase chromatography column (Vydac, Hesperia, CA) and then lyophilized. The lyophilizate was resuspended in 25 mM phosphate buffer (pH 4). The N-terminal amino acid of [6P4]CCL5 glutamine (Q0) was cyclized at 37°C for 48 hours.

The human $G\alpha_i$ subunit ($G\alpha_{i1}$) with an N-terminal tobacco etch virus (TEV) protease–cleavable deca-histidine tag was expressed in the *E. coli* BL21 (DE3) strain and purified as described (*22*).

The transducin heterotrimer was isolated from the rod outer segment of bovine retina (W L Lawson Company) and $G\beta_1\gamma_1$ was separated from $G\alpha_t$ with Blue Sepharose 6 Fast Flow (GE Healthcare) as described (*22*). The $G\alpha_{i1}\beta_1\gamma_1$ heterotrimer (G_i) was prepared by mixing equimolar amounts of $G\alpha_{i1}$ and $G\beta_1\gamma_1$ and incubated at 4°C for 1 hour shortly before use for CCR5-G_i complex formation. Fab16 was produced by papain digestion of immunoglobulin G16 as described (*22*).

Formation of the [6P4]CCL5•CCR5•G_i•Fab16 complex

Pooled fractions of CCR5 eluted from the anti-FLAG resin and a molar excess of G_i heterotrimer were mixed together and incubated for 30 min. Then, an equimolar amount of [6P4]CCL5, together with apyrase (25 mU/ml), was added and incubated for another 2 hours. The complex was mixed with molar excess (1:1.4) of Fab16 and further incubated for at least 1 hour. The mixture of [6P4]CCL5•CCR5•G_i and Fab16 was concentrated using an Amicon Ultra concentrator (MWCO, 100 kDa) and loaded onto a Superdex 200 Increase 10/300 GL column for size exclusion chromatography (SEC) with buffer containing 25 mM Hepes, 150 mM NaCl, and 0.01% LMNG (pH 7.5). The protein quality of each fraction was evaluated by SDS–polyacrylamide gel electrophoresis (fig. S1, A and B). Fractions showing good purity and complex integrity were pooled together and concentrated for EM grid preparation.

Cryo-EM sample preparation and image acquisition

For cryo-EM, 3.5 μl sample (2.5 mg/ml) was directly applied to glow-discharged 200-mesh carbon grids (Quantifoil Cu R1.2/1.3, 200 mesh). Grids were immediately plunge-frozen in liquid ethane using an FEI Vitrobot Mark IV (Thermo Fisher Scientific) with a blotting time of 3 s. The grids were screened for ice thickness and particle distribution using a Glacios Cryo-TEM operated at 200 kV. Images were acquired from the selected grid using a Glacios Cryo-TEM (Thermo Fisher Scientific) operated at 200 kV equipped with a Gatan K3 Summit direct electron detector (Gatan Inc.). Automated data collection was carried out using SerialEM with a set of customized scripts enabling automated low-dose image acquisition (*38*, *39*) and online prescreened during data collection using FOCUS (*40*). Movie stacks of 40 frames were obtained with a defocus range of −1.0 to −2.0 μm at a magnification of ×45,000 (nominally ×36,000) and the K3 detector operated in super-resolution mode (super-resolution pixel size, 0.556 Å). Each movie had a total accumulated dose exposure of ~49 e/Å^2. A total of 2586 image stacks were collected for the [6P4]CCL5•CCR5•G_i•Fab16 complex.

Cryo-EM data processing

Contaminated micrographs were removed manually. Patch motion correction and patch contrast transfer function (CTF) parameter estimation were performed using algorithms implemented in cryoSPARC v2.15.0 (*41*). After sorting, micrographs with estimated resolution worse than 6.0 Å were discarded. The remaining motion-corrected images summed with dose weighting were used for all further image processing in cryoSPARC. Approximately 2.6 million particles were auto-picked and subjected to several rounds of reference-free 2D classification to remove false-positive particles. A total of 345,458 particles from 3D classes that demonstrated clear structural features were combined and subjected to 3D refinement, which led to a reconstruction at 3.6-Å resolution. Nonuniform refinement (*42*) with subsequent local refinements was performed in cryoSPARC v3.1.0 and improved the overall resolution to 3.15 Å [Fourier shell correlation (FSC) = 0.143].

The final set of homogeneous [6P4]CCL5•CCR5•G_i•Fab16 complex particles was subjected to 3D variability analysis implemented in CryoSPARC (*43*) using three variability components and a low-pass filter resolution of 4 Å after applying a soft mask to exclude solvent and micelle.

Reported resolutions calculated with a soft shape mask are based on the gold-standard FSC using the 0.143 criterion. The local resolution was determined using ResMap (*44*).

Model building and refinement

The crystal structures of the G_i heterotrimer [Protein Data Bank (PDB) ID: 5KDO], Fab16 (PDB ID: 6QNK), and the [5P7]CCL5•CCR5 complex (PDB ID: 5UIW) were used as initial templates for model building. The models were docked into the 3D map as rigid bodies in Chimera (*45*). The [6P4]CCL5 N terminus (up to the residue 9) was built ab initio. As compared to residues 1 to 9, a lower definition of the density was observed in the region of the N-terminal pyroglutamate (PCA0). The remaining part of [6P4]CCL5 was taken from the 5UIW structure. Several rounds of manual building were performed in Coot (*46*). The model was finalized by refinement in Phenix 1.18.2. (*47*) against the 3.15-Å cryo-EM map. Structural figures were prepared in Chimera and PyMOL (https://pymol.org/2/). The refinement statistics are summarized in table S1.

Amino acid sequence analysis

The analysis of N-terminal sequence similarity of the natural amino acid CCL5 variants (table S2) was carried using WebLogo (*48*).

Characterization of resistance to GTPγS

To assess the stability of the purified [6P4]CCL5•CCR5•G_i complexes with or without Fab16, they were incubated with 100 μM guanosine 5′-*O*-(3′-thiotriphosphate) (GTPγS) in 25 mM Hepes, 150 mM NaCl, and 0.01% LMNG (pH 7.5) for 1 hour at 4°C and analyzed by SEC with a Superdex 200 Increase 10/300 GL column monitoring the protein intrinsic tryptophan fluorescence (λ_{ex} = 280 nm; λ_{em} = 350 nm). As standards, SEC analyses were also carried out on purified Fab16, Gβγ, $G\alpha_i$, and mixtures thereof using identical buffer conditions. The respective chromatograms are shown in fig. S1C.

Cellular Ca^{2+} flux assays for receptor activation

Human CCL5 and reference standard [6P4]CCL5 were prepared by chemical synthesis as previously described (*10*, *11*). [6P4]CCL5 variants (D5A and D5K), as well as a sample of unmodified [6P4] CCL5, were prepared using a previously described multiplex chemical synthesis approach (*49*).

For experiments involving CCR5 mutants, HEK cells were transiently transfected with expression vectors obtained by site-directed mutagenesis (Q5 Site-Directed Mutagenesis kit, New England Biolabs) of the parent FUGW-CCR5 vector, which was generated by Gibson Assembly (New England Biolabs) as previously described (*50*). HEK cells (1.25×10^6) were seeded overnight in 10-cm dishes and transfected with CCR5 expression vectors (jetPRIME, Polyplus Transfection) according to the manufacturer's instructions. Cells were used in Ca^{2+} flux experiments 24 hours later. For experiments involving [6P4]CCL5 variants, HEK (*10*) and CHO (*49*) cell clones stably expressing CCR5 (HEK-CCR5 and CHO-CCR5, respectively) were used.

Ca^{2+} flux measurements were performed using a Functional Drug Screening System (FDSS) microcell device (HAMAMATSU). On the day of the experiment, cells were detached in phosphate-buffered saline (PBS) supplemented with 0.48 mM EDTA and added (20,000 cells per well) to wells of black-walled clear-bottom 384-well plates. Cells were then loaded with a calcium-sensitive fluorescent dye (Screen Quest Fluo-8 No Wash Calcium Assay Kit, AAT Bioquest) according to the manufacturer's instructions. Fluorescence signals (excitation, 490 nm; emission, 525 nm) were recorded before and after addition of agonist (dissolved in PBS supplemented with 1% bovine serum albumin and 25 mM Hepes) at defined concentrations. Agonist responses were defined as the maximum Ca^{2+} flux fluorescence signal divided by that of a control well with cells treated with buffer only. Dose-response curves were fitted (GraphPad Prism) to the agonist responses *R* at each concentration using a three-parameter agonist versus response model, $R = R_0 + \frac{[\text{agonist}] \times (R_{\max} - R_0)}{EC_{50} + [\text{agonist}]}$, where R_0 presents the baseline and $R_{\max} - R_0 = E_{\max}$.

Modeling and MD simulations

CCR5 N-terminal residues 1 to 19 were built using as template residues 1 to 14 of the NMR solution structure of a doubly sulfated (at Y10 and Y14) N-terminal segment of CCR5 bound to CCL5 (PDB ID: 6FGP). The chemokine in the latter structure was then used as a guide for the structural alignment to our cryo-EM structure. Remaining residues 15 to 19 of CCR5 were then connected to the rest of the cryo-EM model using Modeller v9.16 (*51*). All models derived from Modeller were then subjected to 300 iterations of variable target function method optimization and MD and simulated annealing optimization (within Modeller), scored using the discrete optimized protein energy potential, and the best-scoring model was selected (Fig. 1D, right).

This model of CCR5 (residues 1 to 320) bound to [6P4]CCL5 was used for MD simulations of the nonsulfated and sulfated (Y10 and Y14) forms. Coordinates were first preprocessed using VMD1.9.3 (*52*). The receptor-ligand complex (i.e., CCR5-[6P4]CCL5 or CCR5-CCL5) was then embedded into a 90 Å × 90 Å lipid bilayer composed of 80% 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine and 20% cholesterol. The system was solvated with explicit water molecules, neutralized, and its ionic strength was adjusted using the CHARMM-GUI builder (*53*). Disulfide bridges were explicitly defined between C50-C11 and C34-C10 in CCL5 or [6P4]CCR5 and $C101^{3.25}$-C178 and $C20\text{-}C269^{7.25}$ in CCR5. Except for CCR5 residues $D76^{2.50}$, $E283^{7.39}$, and $E302^{8.48}$, which were protonated, all titratable residues of CCR5 and CCL5 were left in their dominant protonation state at pH 7.0. Before production runs, the geometry of the system was optimized by energy minimization and further relaxed by a sequence of equilibration steps where harmonic positional restraints were applied to all Cα atoms of the protein and gradually released throughout the equilibration. In the last equilibration step (i.e., before completely releasing all protein restraints), water, ion, and lipids were allowed to diffuse without restraints during 50 ns to allow for adequate equilibration of the lipid mixture. After equilibration was completed, five independent trajectories of each system were spawned from the last snapshot of the equilibrated trajectory using a random seed. Production simulations for each replica were run in the isothermal-isobaric (NPT) ensemble at 1013 bar and 310 K for 500 ns each. All simulations were run using Gromacs v2020 (*54*) with the CHARMM36m force field (*55*). Gromacs v2020 and VMD1.9.31 were used to postprocess and analyze all trajectories. MD simulation figures were rendered using VMD1.9.3 and the R ggplot2 library (*56*). Circular plots of residue contacts were generated using the mdciao library (*57*).

The equilibrated model [6P4]CCL5 bound to CCR5 was used to model the binding pose of the wild-type CCL5. The sequence of CCL5 was threaded on [6P4]CCL5 (6P4: QGPPGDIVLACC/CCL5: SPYSSDTTP-CC) and steric clashes were relieved using the molecular graphics software PyMOL. Using this structure as a template, residues 1 to 9 of CCL5 and all residues within 8 Å around Y3 of CCL5 were remodeled with Modeller v9.16 using the protocol described above. The stability of the resulting binding pose was assessed by MD simulations using the protocol described above.

A list of simulations performed in this work is given in table S4. MD simulations were performed at the Paul Scherrer Institute computing cluster and at the Swiss National Supercomputing Centre (CSCS).

Electrostatic potentials were calculated using the Adaptive Poisson-Boltzmann Solver (APBS) method (*58*) as implemented in PyMOL using a concentration of 0.150 M for the +1 and 1 ion species. The biomolecular surface is colored from red (5 kT/e) to blue (+5 kT/e) according to the potential on the soluble accessible surface.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/7/25/eabg8685/DC1

View/request a protocol for this paper from *Bio-protocol*.

REFERENCES AND NOTES

1. I. Scurci, E. Martins, O. Hartley, CCR5: Established paradigms and new frontiers for a ‘celebrity’ chemokine receptor. *Cytokine* **109**, 81–93 (2018).
2. G. Alkhatib, The biology of CCR5 and CXCR4. *Curr. Opin. HIV AIDS* **4**, 96–103 (2009).
3. D. Aldinucci, N. Casagrande, Inhibition of the CCL5/CCR5 axis against the progression of gastric cancer. *Int. J. Mol. Sci.* **19**, 1477 (2018).
4. G. Martin-Blondel, D. Brassat, J. Bauer, H. Lassmann, R. S. Liblau, CCR5 blockade for neuroinflammatory diseases — Beyond control of HIV. *Nat. Rev. Neurol.* **12**, 95–105 (2016).
5. R. L. Chua, S. Lukassen, S. Trump, B. P. Hennig, D. Wendisch, F. Pott, O. Debnath, L. Thürmann, F. Kurth, M. T. Völker, J. Kazmierski, B. Timmermann, S. Twardziok, S. Schneider, F. Machleidt, H. Müller-Redetzky, M. Maier, A. Krannich, S. Schmidt, F. Balzer, J. Liebig, J. Loske, N. Suttorp, J. Eils, N. Ishaque, U. G. Liebert, C. von Kalle, A. Hocke, M. Witzenrath, C. Goffinet, C. Drosten, S. Laudi, I. Lehmann, C. Conrad, L.-E. Sander, R. Eils, COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. *Nat. Biotechnol.* **38**, 970–979 (2020).
6. B. K. Patterson, H. Seethamraju, K. Dhody, M. J. Corley, K. Kazempour, J. P. Lalezari, A. P. Pang, C. Sugai, E. B. Francisco, A. Pise, H. Rodrigues, M. Ryou, H. L. Wu, G. M. Webb, B. S. Park, S. Kelly, N. Pourhassan, A. Lelic, L. Kdouh, M. Herrera, E. Hall, E. Aklin, L. Ndhlovu, J. B. Sacha, Disruption of the CCL5/RANTES-CCR5 pathway restores immune homeostasis and reduces plasma viral load in critical COVID-19. *medRxiv* 2020.05.02.20084673 (2020).
7. F. Cocchi, A. DeVico, A. Garzinodemo, S. Arya, R. Gallo, P. Lusso, Identification of RANTES, MIP-1α, and MIP-1β as the major HIV-suppressive factors produced by $Cd8^+$ T-cells. *Science* **270**, 1811–1815 (1995).
8. A. Trkola, W. A. Paxton, S. P. Monard, J. A. Hoxie, M. A. Siani, D. A. Thompson, L. Wu, C. R. Mackay, R. Horuk, J. P. Moore, Genetic subtype-independent inhibition of human immunodeficiency virus type 1 replication by CC and CXC chemokines. *J. Virol.* **72**, 396–404 (1998).
9. G. Simmons, P. R. Clapham, L. Picard, R. E. Offord, M. M. Rosenkilde, T. W. Schwartz, R. Buser, T. N. C. Wells, A. E. I. Proudfoot, Potent Inhibition of HIV-1 infectivity in macrophages and lymphocytes by a novel CCR5 antagonist. *Science* **276**, 276–279 (1997).
10. O. Hartley, H. Gaertner, J. Wilken, D. Thompson, R. Fish, A. Ramos, C. Pastore, B. Dufour, F. Cerini, A. Melotti, N. Heveker, L. Picard, M. Alizon, D. Mosier, S. Kent, R. Offord, Medicinal chemistry applied to a synthetic protein: Development of highly potent HIV entry inhibitors. *Proc. Natl. Acad. Sci. U.S.A.* **101**, 16460–16465 (2004).
11. H. Gaertner, F. Cerini, J. M. Escola, G. Kuenzi, A. Melotti, R. Offord, I. Rossitto-Borlat, R. Nedellec, J. Salkowitz, G. Gorochov, D. Mosier, O. Hartley, Highly potent, fully recombinant anti-HIV chemokines: Reengineering a low-cost microbicide. *Proc. Natl. Acad. Sci. U.S.A.* **105**, 17706–17711 (2008).
12. W. I. Weis, B. K. Kobilka, The molecular basis of G protein-coupled receptor activation. *Annu. Rev. Biochem.* **87**, 897–919 (2018).
13. Y. Zheng, G. W. Han, R. Abagyan, B. Wu, R. C. Stevens, V. Cherezov, I. Kufareva, T. M. Handel, Structure of CC chemokine receptor 5 with a potent chemokine antagonist reveals mechanisms of chemokine recognition and molecular mimicry by HIV. *Immunity* **46**, 1005–1017.e5 (2017).
14. M. M. Shaik, H. Peng, J. Lu, S. Rits-Volloch, C. Xu, M. Liao, B. Chen, Structural basis of coreceptor recognition by HIV-1 envelope spike. *Nature* **565**, 318–323 (2019).
15. Q. Tan, Y. Zhu, J. Li, Z. Chen, G. W. Han, I. Kufareva, T. Li, L. Ma, G. Fenalti, J. Li, W. Zhang, X. Xie, H. Yang, H. Jiang, V. Cherezov, H. Liu, R. C. Stevens, Q. Zhao, B. Wu, Structure of the CCR5 chemokine receptor-HIV entry inhibitor maraviroc complex. *Science* **341**, 1387–1390 (2013).
16. P. Peng, H. Chen, Y. Zhu, Z. Wang, J. Li, R.-H. Luo, J. Wang, L. Chen, L.-M. Yang, H. Jiang, X. Xie, B. Wu, Y.-T. Zheng, H. Liu, 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 (2018).
17. D. J. Wasilko, Z. L. Johnson, M. Ammirati, Y. Che, M. C. Griffor, S. Han, H. Wu, Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20. *Nat. Commun.* **11**, 3031 (2020).
18. K. Liu, L. Wu, S. Yuan, M. Wu, Y. Xu, Q. Sun, S. Li, S. Zhao, T. Hua, Z.-J. Liu, Structural basis of CXC chemokine receptor 2 activation and signalling. *Nature* **585**, 135–140 (2020).
19. J. S. Burg, J. R. Ingram, A. J. Venkatakrishnan, K. M. Jude, A. Dukkipati, E. N. Feinberg, A. Angelini, D. Waghray, R. O. Dror, H. L. Ploegh, K. C. Garcia, Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science* **347**, 1113–1117 (2015).
20. T. F. Miles, K. Spiess, K. M. Jude, N. Tsutsumi, J. S. Burg, J. R. Ingram, D. Waghray, G. M. Hjorto, O. Larsen, H. L. Ploegh, M. M. Rosenkilde, K. C. Garcia, Viral GPCR US28 can signal in response to chemokine agonists of nearly unlimited structural degeneracy. *eLife* **7**, e35850 (2018).
21. S. Maeda, A. Koehl, H. Matile, H. Hu, D. Hilger, G. F. X. Schertler, A. Manglik, G. Skiniotis, R. J. P. Dawson, B. K. Kobilka, Development of an antibody fragment that stabilizes GPCR/G-protein complexes. *Nat. Commun.* **9**, 3712 (2018).
22. C.-J. Tsai, J. Marino, R. Adaixo, F. Pamula, J. Muehle, S. Maeda, T. Flock, N. M. Taylor, I. Mohammed, H. Matile, R. J. Dawson, X. Deupi, H. Stahlberg, G. Schertler, Cryo-EM structure of the rhodopsin-Gαi-$\beta\gamma$ complex reveals binding of the rhodopsin C-terminal tail to the gβ subunit. *eLife* **8**, e46041 (2019).
23. A. Glukhova, C. J. Draper-Joyce, R. K. Sunahara, A. Christopoulos, D. Wootten, P. M. Sexton, Rules of engagement: GPCRs and G proteins. *ACS Pharmacol. Transl. Sci.* **1**, 73–83 (2018).
24. D. J. Scholten, M. Canals, D. Maussang, L. Roumen, M. J. Smit, M. Wijtmans, C. de Graaf, H. F. Vischer, R. Leurs, Pharmacological modulation of chemokine receptor function. *Br. J. Pharmacol.* **165**, 1617–1643 (2012).
25. V. Isberg, C. de Graaf, A. Bortolato, V. Cherezov, V. Katritch, F. H. Marshall, S. Mordalski, J.-P. Pin, R. C. Stevens, G. Vriend, D. E. Gloriam, Generic GPCR residue numbers – Aligning topology maps while minding the gaps. *Trends Pharmacol. Sci.* **36**, 22–31 (2015).
26. M. Abayev, J. P. G. L. M. Rodrigues, G. Srivastava, B. Arshava, Ł. Jaremko, M. Jaremko, F. Naider, M. Levitt, J. Anglister, The solution structure of monomeric CCL5 in complex with a doubly sulfated N-terminal segment of CCR5. *FEBS J.* **285**, 1988–2003 (2018).
27. M. Farzan, T. Mirzabekov, P. Kolchinsky, R. Wyatt, M. Cayabyab, N. P. Gerard, C. Gerard, J. Sodroski, H. Choe, Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell* **96**, 667–676 (1999).
28. L. Duma, D. Häussinger, M. Rogowski, P. Lusso, S. Grzesiek, Recognition of RANTES by extracellular parts of the CCR5 receptor. *J. Mol. Biol.* **365**, 1063–1075 (2007).
29. N. Kessler, S. R. Akabayov, A. Moseri, L. S. Cohen, D. Sakhapov, D. Bolton, B. Fridman, L. E. Kay, F. Naider, J. Anglister, Allovalency observed by transferred NOE: Interactions of sulfated tyrosine residues in the N-terminal segment of CCR5 with the CCL5 chemokine. *FEBS J.* **288**, 1648–1663 (2021).
30. L. Nisius, M. Rogowski, L. Vangelista, S. Grzesiek, Large-scale expression and purification of the major HIV-1 coreceptor CCR5 and characterization of its interaction with RANTES. *Protein Expr. Purif.* **61**, 155–162 (2008).
31. M. Wiktor, O. Hartley, S. Grzesiek, Characterization of structure, dynamics, and detergent interactions of the anti-hiv chemokine variant 5P12-RANTES. *Biophys. J.* **105**, 2586–2597 (2013).
32. M. Lederman, R. Veazey, R. Offord, D. Mosier, J. Dufour, M. Mefford, M. Piatak, J. Lifson, J. Salkowitz, B. Rodriguez, A. Blauvelt, O. Hartley, Prevention of vaginal SHIV transmission in rhesus macaques through inhibition of CCR5. *Science* **306**, 485–487 (2004).
33. A. Grahl, L. A. Abiko, S. Isogai, T. Sharpe, S. Grzesiek, A high-resolution description of β1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR. *Nat. Commun.* **11**, 2216 (2020).
34. C. Blanpain, B. Doranz, A. Bondue, C. Govaerts, A. De Leener, G. Vassart, R. Doms, A. Proudfoot, M. Parmentier, The core domain of chemokines binds CCR5 extracellular domains while their amino terminus interacts with the transmembrane helix bundle. *J. Biol. Chem.* **278**, 5179–5187 (2003).
35. C. Govaerts, A. Bondue, J.-Y. Springael, M. Olivella, X. Deupi, E. L. Poul, S. J. Wodak, M. Parmentier, L. Pardo, C. Blanpain, Activation of CCR5 by chemokines involves an aromatic cluster between transmembrane helices 2 and 3. *J. Biol. Chem.* **278**, 1892–1903 (2003).
36. J. K. Lim, J. M. Burns, W. Lu, A. L. DeVico, Multiple pathways of amino terminal processing produce two truncated variants of RANTES/CCL5. *J. Leukoc. Biol.* **78**, 442–452 (2005).
37. P. Proost, I. D. Meester, D. Schols, S. Struyf, A.-M. Lambeir, A. Wuyts, G. Opdenakker, E. D. Clercq, S. Scharpé, J. V. Damme, Amino-terminal truncation of chemokines by CD26/dipeptidyl-peptidase IV. Conversion of RANTES into a potent inhibitor of monocyte chemotaxis and HIV-1-infection. *J. Biol. Chem.* **273**, 7222–7227 (1998).
38. D. N. Mastronarde, Automated electron microscope tomography using robust prediction of specimen movements. *J. Struct. Biol.* **152**, 36–51 (2005).
39. M. Schorb, I. Haberbosch, W. J. H. Hagen, Y. Schwab, D. N. Mastronarde, Software tools for automated transmission electron microscopy. *Nat. Methods* **16**, 471–477 (2019).
40. N. Biyani, R. D. Righetto, R. McLeod, D. Caujolle-Bert, D. Castano-Diez, K. N. Goldie, H. Stahlberg, Focus: The interface between data collection and data processing in cryo-EM. *J. Struct. Biol.* **198**, 124–133 (2017).
41. A. Punjani, J. L. Rubinstein, D. J. Fleet, M. A. Brubaker, cryoSPARC: Algorithms for rapid unsupervised cryo-EM structure determination. *Nat. Methods* **14**, 290–296 (2017).
42. A. Punjani, H. Zhang, D. J. Fleet, Non-uniform refinement: Adaptive regularization improves single-particle cryo-EM reconstruction. *Nat. Methods* **17**, 1214–1221 (2020).
43. A. Punjani, D. J. Fleet, 3D variability analysis: Resolving continuous flexibility and discrete heterogeneity from single particle Cryo-EM. *J. Struct. Biol.* **213**, 107702 (2021).
44. A. Kucukelbir, F. J. Sigworth, H. D. Tagare, Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
45. E. F. Pettersen, T. D. Goddard, C. C. Huang, G. S. Couch, D. M. Greenblatt, E. C. Meng, T. E. Ferrin, UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
46. P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, Features and development of *Coot*. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 486–501 (2010).

47. D. Liebschner, P. V. Afonine, M. L. Baker, G. Bunkóczi, V. B. Chen, T. I. Croll, B. Hintze, L.-W. Hung, S. Jain, A. J. McCoy, N. W. Moriarty, R. D. Oeffner, B. K. Poon, M. G. Prisant, R. J. Read, J. S. Richardson, D. C. Richardson, M. D. Sammito, O. V. Sobolev, D. H. Stockwell, T. C. Terwilliger, A. G. Urzhumtsev, L. L. Videau, C. J. Williams, P. D. Adams, Macromolecular structure determination using x-rays, neutrons and electrons: Recent developments in Phenix. *Acta Crystallogr. D Biol. Crystallogr.* **75**, 861–877 (2019).
48. G. E. Crooks, G. Hon, J.-M. Chandonia, S. E. Brenner, WebLogo: A sequence logo generator. *Genome Res.* **14**, 1188–1190 (2004).
49. M. Paolini-Bertrand, F. Cerini, E. Martins, I. Scurci, O. Hartley, Rapid and low-cost multiplex synthesis of chemokine analogs. *J. Biol. Chem.* **293**, 19092–19100 (2018).
50. I. Scurci, K. B. Akondi, I. Pinheiro, M. Paolini-Bertrand, A. Borgeat, F. Cerini, O. Hartley, CCR5 tyrosine sulfation heterogeneity generates cell surface receptor subpopulations with different ligand binding properties. *Biochim. Biophys. Acta Gen. Subj.* **1865**, 129753 (2021).
51. B. Webb, A. Sali, Comparative protein structure modeling using MODELLER. *Curr. Protoc. Bioinformatics* **54**, 5.6.1–5.6.37 (2016).
52. W. Humphrey, A. Dalke, K. Schulten, VMD: Visual molecular dynamics. *J. Mol. Graph.* **14**, 33–38 (1996).
53. S. Jo, T. Kim, W. Im, Automated builder and database of protein/membrane complexes for molecular dynamics simulations. *PLOS ONE* **2**, e880 (2007).
54. M. J. Abraham, T. Murtola, R. Schulz, S. Páll, J. C. Smith, B. Hess, E. Lindahl, GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* **1–2**, 19–25 (2015).
55. J. Huang, S. Rauscher, G. Nawrocki, T. Ran, M. Feig, B. L. de Groot, H. Grubmüller, A. D. MacKerell Jr., CHARMM36m: An improved force field for folded and intrinsically disordered proteins. *Nat. Methods* **14**, 71–73 (2017).
56. H. Wickham, *Ggplot2: Elegant Graphics for Data Analysis* (Springer, New York, 2009).
57. G. Pérez-Hernández, *gph82/mdciao* (2021); https://github.com/gph82/mdciao.
58. N. A. Baker, D. Sept, S. Joseph, M. J. Holst, J. A. McCammon, Electrostatics of nanosystems: Application to microtubules and the ribosome. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 10037–10041 (2001).
59. F. Bachelerie, A. Ben-Baruch, A. M. Burkhardt, C. Combadiere, J. M. Farber, G. J. Graham, R. Horuk, A. H. Sparre-Ulrich, M. Locati, A. D. Luster, A. Mantovani, K. Matsushima, P. M. Murphy, R. Nibbs, H. Nomiyama, C. A. Power, A. E. I. Proudfoot, M. M. Rosenkilde, A. Rot, S. Sozzani, M. Thelen, O. Yoshie, A. Zlotnik, International union of basic and clinical pharmacology. LXXXIX. Update on the extended family of chemokine receptors and introducing a new nomenclature for atypical chemokine receptors. *Pharmacol. Rev.* **66**, 1–79 (2014).
60. UniProt Consortium, UniProt: A worldwide hub of protein knowledge. *Nucleic Acids Res.* **47**, D506–D515 (2019).

Acknowledgments: This work was supported by the Swiss National Science Foundation (grants 149927 and 173089 to S.G.; grant 192780 to X.D.; grant 184828 to O.H.; SNF R'EQUIP 177084 to T.M.; NCCR TransCure to H.S.; and SNF Sinergia 183563 to G.F.X.S.), the Synapsis Foundation (grant 2018-P104 to G.F.X.S.), and the European Union (grants FP6-EMPRO and FP7-CHAARM to S.G.). We acknowledge M. Chami and L. Kovacik (Biozentrum BioEM Lab) for help with cryo-EM data collection, P. LiWang (University of California) for a gift of the [5P14] CCL5 DNA construct, M. Rogowski (Biozentrum) for preparation of [6P4]CCL5, J. Muehle (Paul Scherrer Institute) for preparation of Fab16, M. Caubet (High Performance Computing and Emerging Technologies Group, Paul Scherrer Institute) for technical support with MD simulations, S. Isogai for helpful discussion, and the sciCORE facility of the University of Basel for the computer infrastructure. **Author contributions:** P.I., A.G., and S.G. conceived the study. P.I., C.-J.T., and F.P. expressed and purified proteins and developed the protocol for forming the [6P4]CCL5•CCR5•G_i•Fab16 complex. A.G. assisted with the expression and purification of CCR5. P.I. and C.-J.T. prepared cryo-EM grids. K.N.G. and P.I. collected the cryo-EM data. P.I. and N.D. processed the cryo-EM data and built and refined the model. X.D. and R.G.-G. designed and performed MD simulations. H.S. and T.M. provided guidance on EM sample preparation, data collection, and model refinement. C.B. generated the CCR5 mutants used in cellular Ca^{2+} flux assays. M.P.-B. synthesized chemokine analogs for use in the signaling assays, which were performed and analyzed by N.C. and F.C. P.I., X.D., O.H., and S.G. analyzed the structure and wrote the manuscript, with input from C.-J.T., A.G., and G.F.X.S. **Competing interests:** G.F.X.S. is a cofounder and scientific advisor of the companies leadXpro AG and InterAx Biotech AG. O.H. is the inventor of 6P4[CCL5] and is both a shareholder and an affiliate of Orion Biotechnology Switzerland, to whom the rights to the invention have been assigned. The other authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The cryo-EM map of [6P4]CCL5•CCR5•Gi•Fab16 has been deposited in the Electron Microscopy Data Bank as entry EMD-12746 and the corresponding model in the Protein Data Bank as entry 7O7F.

Submitted 1 February 2021
Accepted 29 April 2021
Published 16 June 2021
10.1126/sciadv.abg8685

Citation: P. Isaikina, C.-J. Tsai, N. Dietz, F. Pamula, A. Grahl, K. N. Goldie, R. Guixà-González, C. Branco, M. Paolini-Bertrand, N. Calo, F. Cerini, G. F. X. Schertler, O. Hartley, H. Stahlberg, T. Maier, X. Deupi, S. Grzesiek, Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. *Sci. Adv.* **7**, eabg8685 (2021).

Science Advances

Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist

Polina Isaikina, Ching-Ju Tsai, Nikolaus Dietz, Filip Pamula, Anne Grahl, Kenneth N. Goldie, Ramon Guix-Gonzlez, Camila Branco, Marianne Paolini-Bertrand, Nicolas Calo, Fabrice Cerini, Gebhard F. X. Schertler, Oliver Hartley, Henning Stahlberg, Timm Maier, Xavier Deupi, and Stephan Grzesiek

Sci. Adv., **7** (25), eabg8685.
DOI: 10.1126/sciadv.abg8685

Science Advances (ISSN 2375-2548) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science Advances* is a registered trademark of AAAS.

2.2 Supplementary Material

Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/7/25/eabg8685/DC1

Supplementary Materials for

Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist

Polina Isaikina, Ching-Ju Tsai, Nikolaus Dietz, Filip Pamula, Anne Grahl, Kenneth N. Goldie, Ramon Guixà-González, Camila Branco, Marianne Paolini-Bertrand, Nicolas Calo, Fabrice Cerini, Gebhard F. X. Schertler*, Oliver Hartley*, Henning Stahlberg, Timm Maier, Xavier Deupi*, Stephan Grzesiek*

*Corresponding author. Email: stephan.grzesiek@unibas.ch (S.G.); xavier.deupi@psi.ch (X.D.); oliver.hartley@unige.ch (O.H.); gebhard.schertler@psi.ch (G.F.X.S.)

Published 16 June 2021, *Sci. Adv.* **7**, eabg8685 (2021)
DOI: 10.1126/sciadv.abg8685

The PDF file includes:

Figs. S1 to S17
Legends for movies S1 to S3

Other Supplementary Material for this manuscript includes the following:

(available at advances.sciencemag.org/cgi/content/full/7/25/eabg8685/DC1)

Movies S1 to S3
Tables S1 to S4

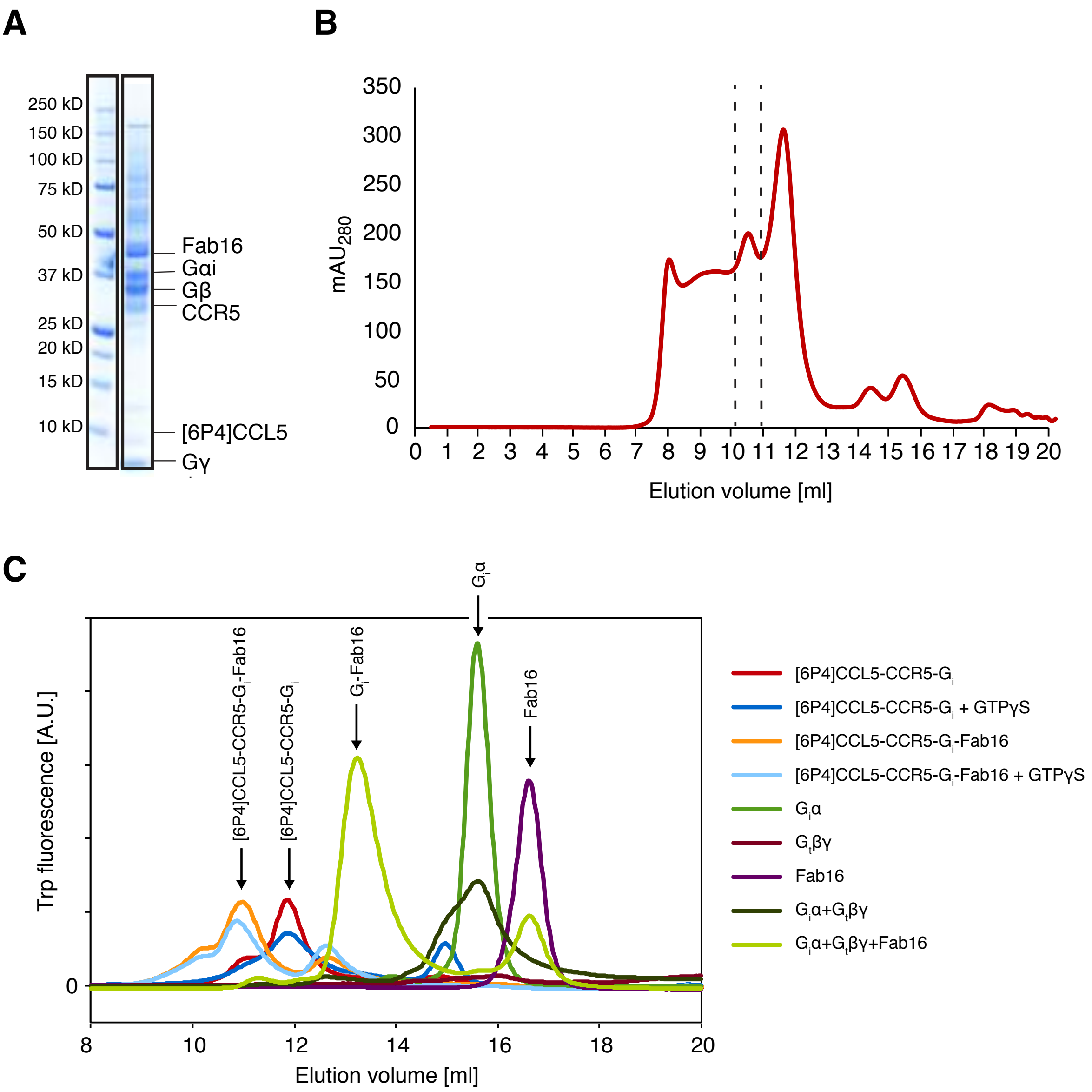
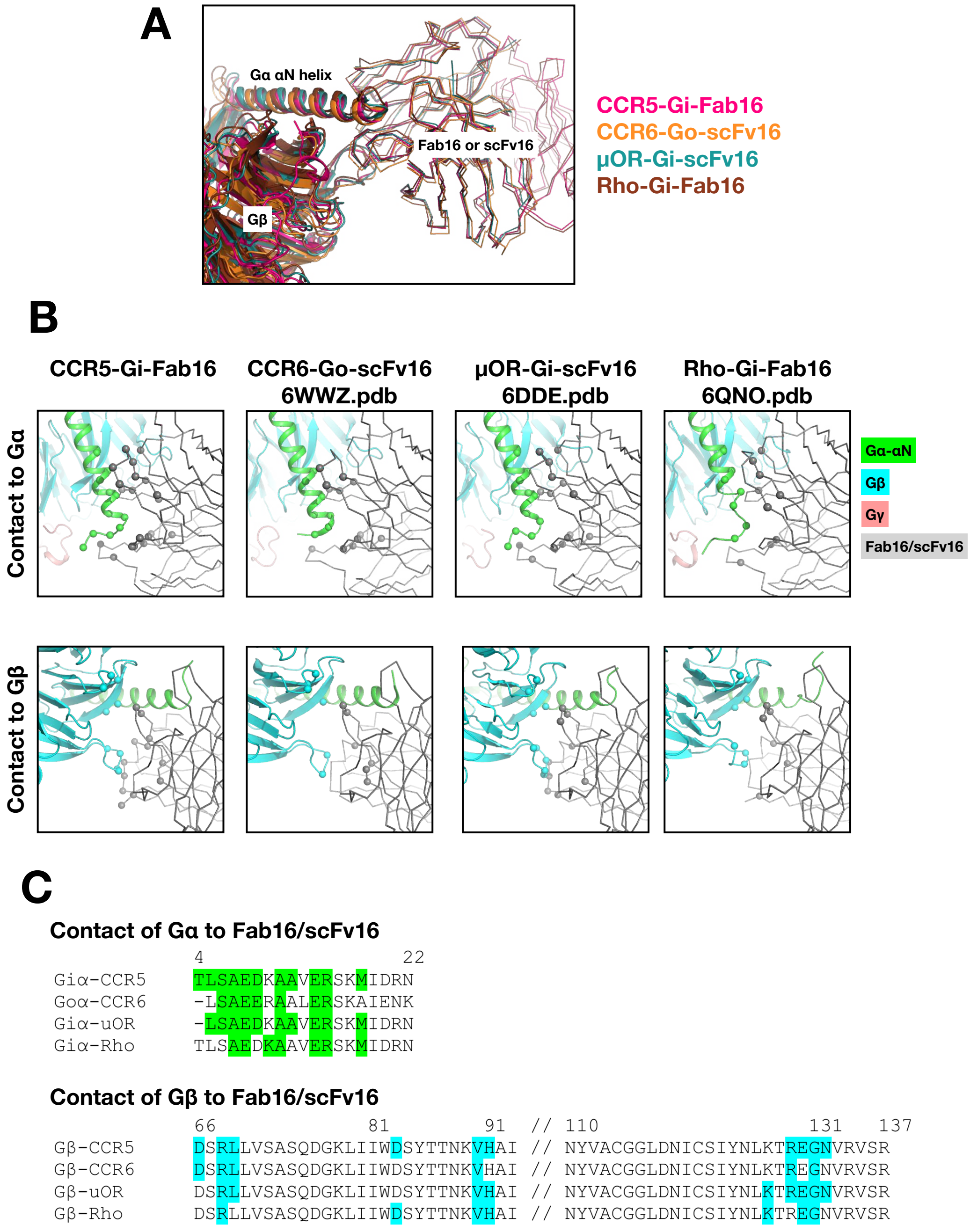


Figure S1. Purification of the [6P4]CCL5•CCR5•G$_i$•Fab16 complex (A) SDS-PAGE analysis and (B) Superdex200 size-exclusion chromatography elution profile of the [6P4]CCL5•CCR5•G$_i$•Fab16 complex. The dashed lines in B indicate the part of the elution volume used for cryo-EM sample preparation. (C) Analytical size-exclusion chromatography was used follow complex formation at different stages as well as the stability of the complexes in the presence and absence of Fab16 and GTPγS. The integrity of the [6P4]CCL5•CCR5-G$_i$ complex (red) is disturbed in the presence of GTPγS (dark blue). Addition of Fab16 (orange) stabilizes the complex in the presence of GTPγS (light blue), as the antibody fragment constrains the conformational flexibility of the G$_i$ heterotrimer.



Contact of Gα to Fab16/scFv16

```
            4                 22
Giα-CCR5    TLSAEDKAAVERSKMIDRN
Goα-CCR6    -LSAEERAALERSKAIENK
Giα-uOR     -LSAEDKAAVERSKMIDRN
Giα-Rho     TLSAEDKAAVERSKMIDRN
```

Contact of Gβ to Fab16/scFv16

```
            66             81        91  // 110                    131   137
Gβ-CCR5     DSRLLVSASQDGKLIIWDSYTTNKVHAI // NYVACGGLDNICSIYNLKTREGNVRVSR
Gβ-CCR6     DSRLLVSASQDGKLIIWDSYTTNKVHAI // NYVACGGLDNICSIYNLKTREGNVRVSR
Gβ-uOR      DSRLLVSASQDGKLIIWDSYTTNKVHAI // NYVACGGLDNICSIYNLKTREGNVRVSR
Gβ-Rho      DSRLLVSASQDGKLIIWDSYTTNKVHAI // NYVACGGLDNICSIYNLKTREGNVRVSR
```

Figure S2. Contacts between Fab16 or scFv16 and Gα/β in existing GPCR/G$_{i/o}$ protein complexes. (A) Structural superposition of GPCR-G$_{i/o}$ complex structures at the interface between G protein α/β subunits and Fab16 or scFv16. Structures of CCR5-G$_i$ (this structure), CCR6-G$_o$ (PDB ID: 6WWZ), μ-opioid receptor-G$_i$ (PDB ID: 6DDE), and rhodopsin-G$_i$ (PDB ID: 6QNO) are aligned to the scFv16 chain. (B) Detail of the contact interfaces between Fab16 or scFv16 and the αN helix of Gα (top) and Gβ (bottom). Contact residues were selected within 4 Å of the neighboring molecule and are shown as spheres (only Cα). (C) Sequence alignment of the Gα and Gβ residues forming contacts with the antibody fragments. Residues of Gα (top panel) and Gβ (bottom panel) within 4 Å of Fab16 or scFv16 are highlighted in green and cyan, respectively. The comparison shows that the binding interfaces are very similar in all complexes.

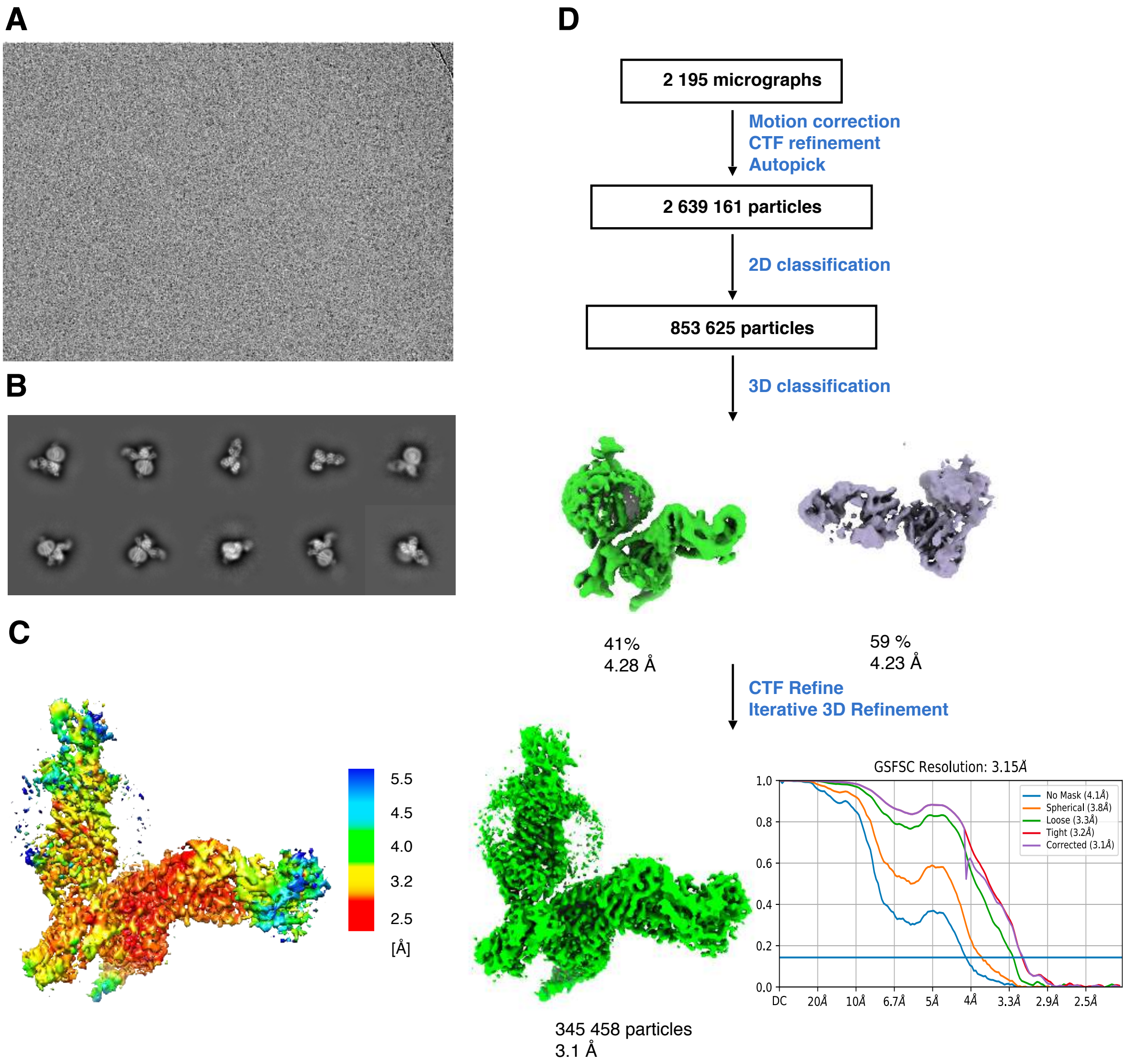


Figure S3. Cryo-EM data processing. (A) Representative cryo-EM micrograph of the [6P4]CCL5•CCR5•G$_i$•Fab16 complex sample and (B) representative 2D class averages showing distinct views and structural features. (C) Density map colored by local resolution. (D) Workflow of cryo-EM data processing. The Fourier shell correlation (FSC) curve indicates an overall nominal resolution of 3.1 Å using the Gold-standard FSC=0.143 criterion.

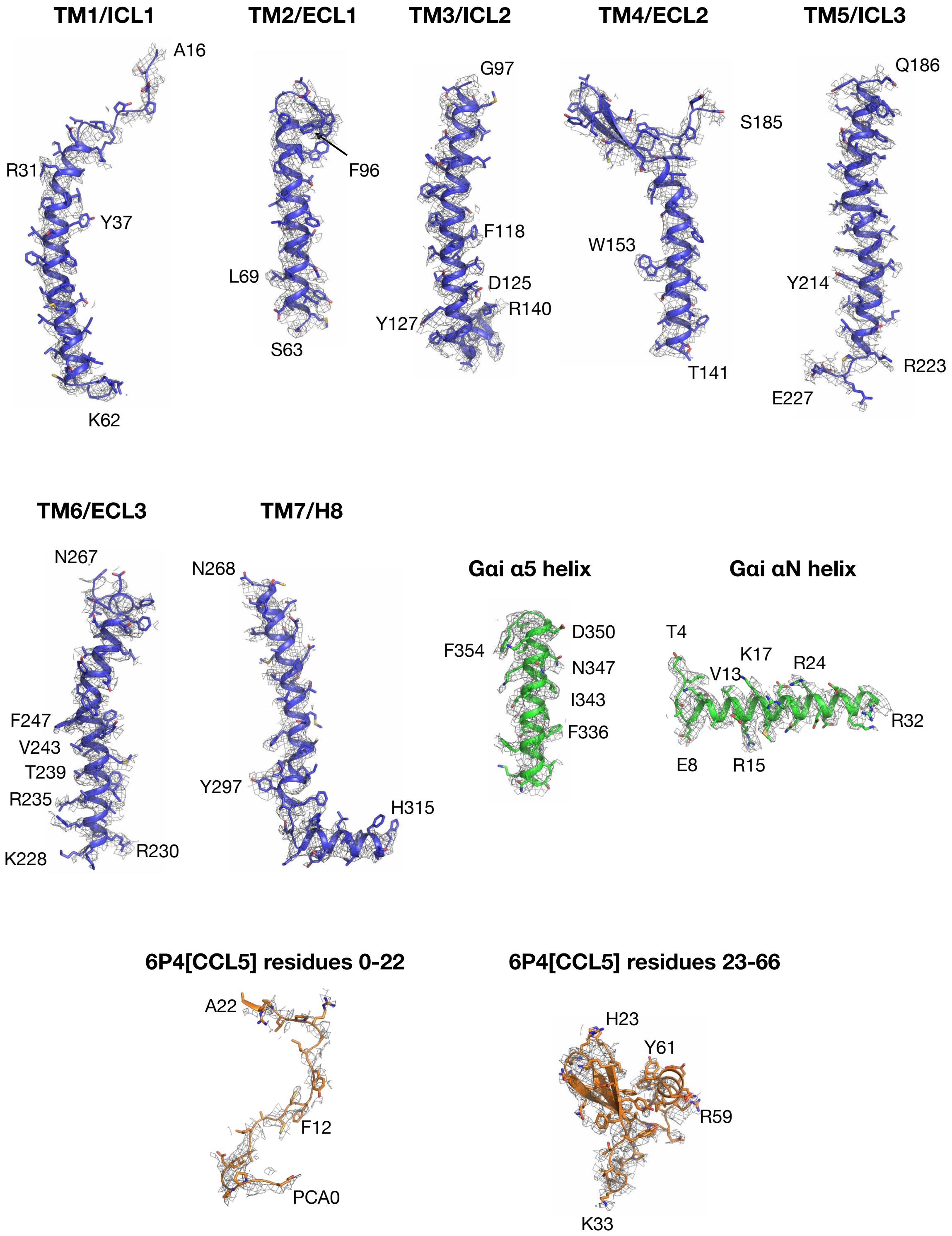
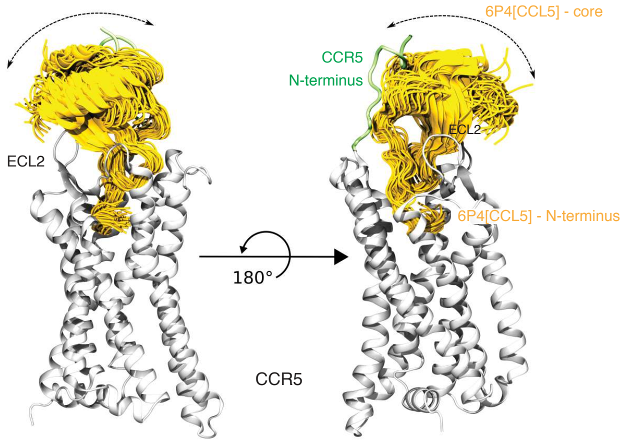


Figure S4. Atomic model of [6P4]CCL5•CCR5•Gi-Fab16 in the cryo-EM density map. Transmembrane helices TM1-TM7 and H8 of CCR5 (blue cartoons) are shown together with the cryo-EM density map displayed at a 7.5 σ cut-off within 2 Å of the model. The αN and α5 helices of $G\alpha_i$ (green) and [6P4]CCL5 residues 0-22 and 23-66 (orange) are shown in the same representation.

A Conformational dynamics of 6P4[CCL5] in MD simulations



B Residue-residue contacts between CCR5 (sulfated at Y10 and Y14) and [6P4]CCL5 in the cryo-EM structure and during MD simulations

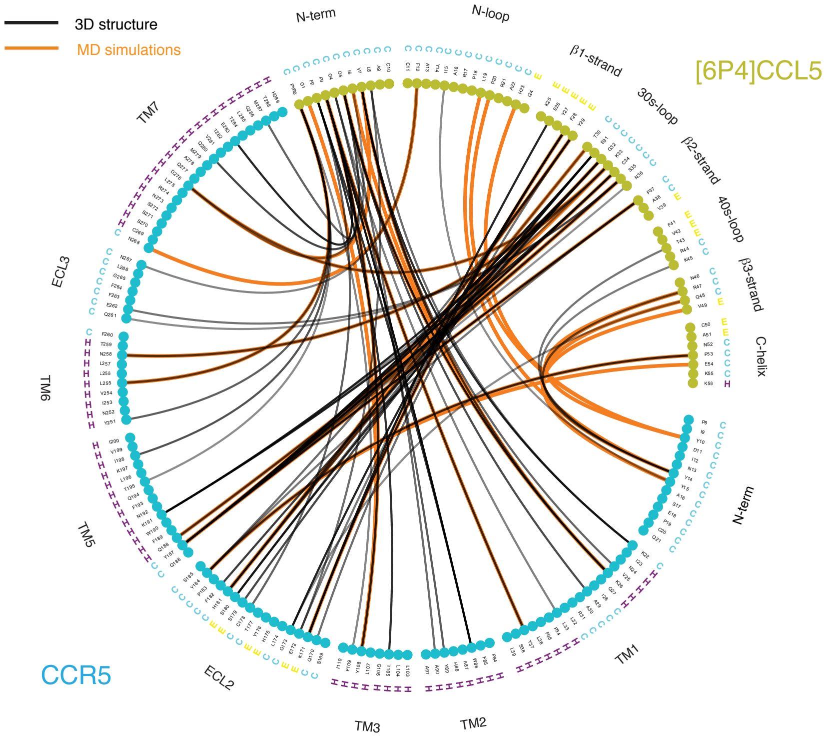


Figure S5. Conformational dynamics of [6P4]CCL5 in MD simulations and residue-residue contacts between [6P4]CCL5 and sulfated CCR5 in cryo-EM structure and MD simulations. (A) Representative conformational ensemble of [6P4]CCL5 – with sulfated Y10 and Y14 – in MD simulations. In each frame, the structure of the N-terminal residues (0-10) of the chemokine are superposed. Only one receptor structure is shown for clarity. The core domain of the chemokine (bound between the N-terminus and extracellular loop 2 (ECL2) of the receptor) is relatively flexible. (B) Residue-residue contacts between CCR5 (sulfated at Y10 and Y14; cyan section of the flare plot) and 6P4[CCL5] (yellow olive section) in the cryo-EM structure and during MD simulations. For clarity, in CCR5 only the regions that contact the chemokine are shown. Black lines denote the interactions observed in the starting frame (equilibrated model of the full-length receptor based on the cryo-EM data with the modeled N-terminus). Orange lines denote the interactions observed in the MD trajectories (for clarity, only the most prevalent interactions are shown).

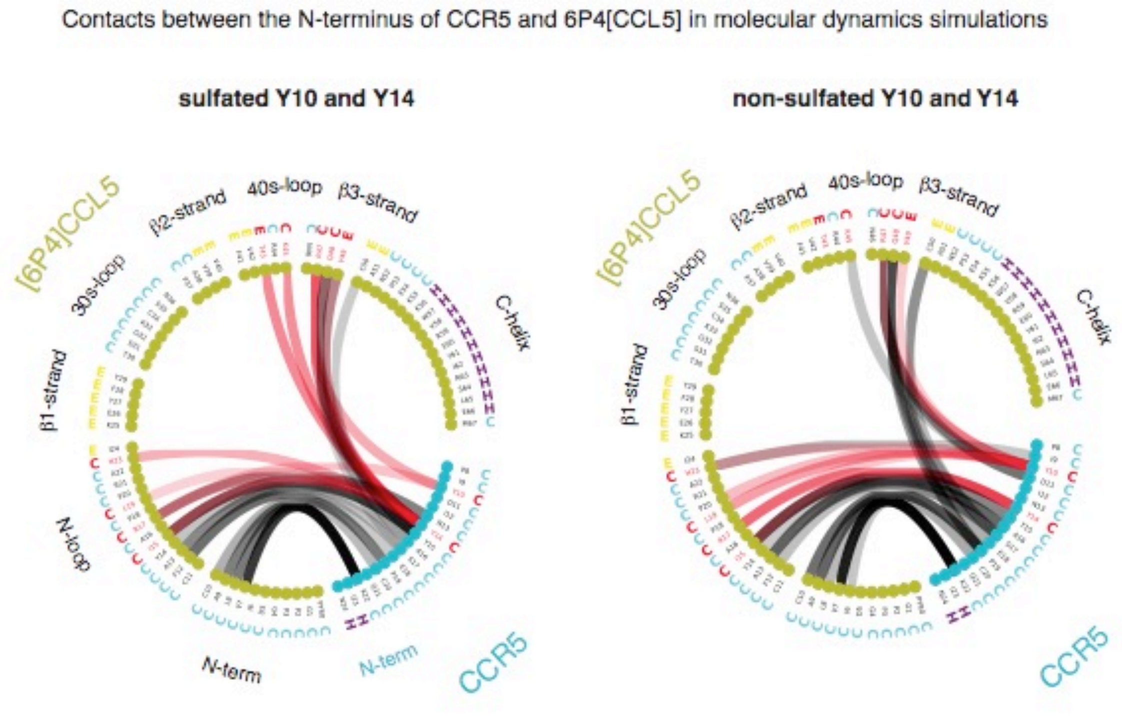


Figure S6. Effect of tyrosine sulfation on the interaction between [6P4]CCL5 and the CCR5 N-terminus. Residue-residue contacts between the N-terminus of CCR5 (residues 8-24, cyan) and 6P4[CCL5] (full sequence, yellow olive) in MD simulations in the presence (left) and absence (right) of sulfation at residues Tyr10 and Tyr14 of CCR5 N-terminus. Red lines denote the interactions mediated by Y10 and Y14. Sulfated Y10 and Y14 result in a higher number of contacts between the receptor N-terminus and the 40-s loop and the β3-strand regions of the chemokine.

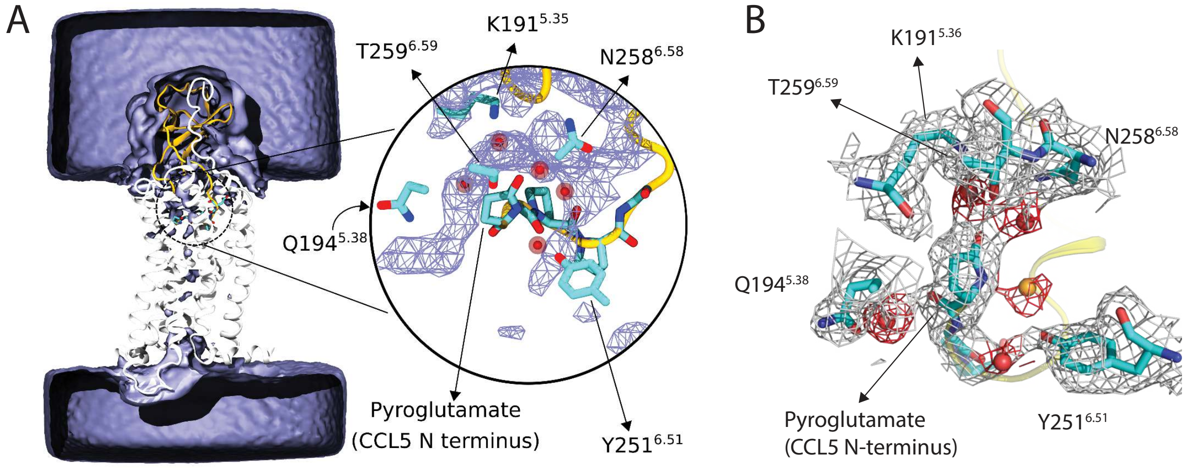


Figure S7. Water molecules in CRS2. (A) MD simulation of CCR5 bound to [6P4]CCL5. All replicas of the MD simulations (see Supplementary Table S4) were used to compute the average water density (blue surface). The inset shows a closer view of the average water density (blue mesh) solvating the PCA group of [6P4]CCL5, which includes the crystallographic water molecules observed in the [5P7]CCL5 complex. These data shows that the N-terminal PCA group interacts with nearby CCR5 residues through water-mediated hydrogen bonds both in [5P7]CCL5 and [6P4]CCL5. (B) Detail of the vicinity of the N-terminal pyroglutamate (PCA) in the structure of CCR5 bound to [5P7]CCL5 (PDB ID: 5UIW). The backbone of CCL5 is traced in yellow and the PCA and nearby residues of CCR5 are shown as sticks. Electron densities are shown as a grey mesh for the chemokine and the receptor and in red for water molecules. Re-refinement of the deposited X-ray data allowed us to model additional water molecules (red spheres) in addition of those present in the original structure (orange sphere).

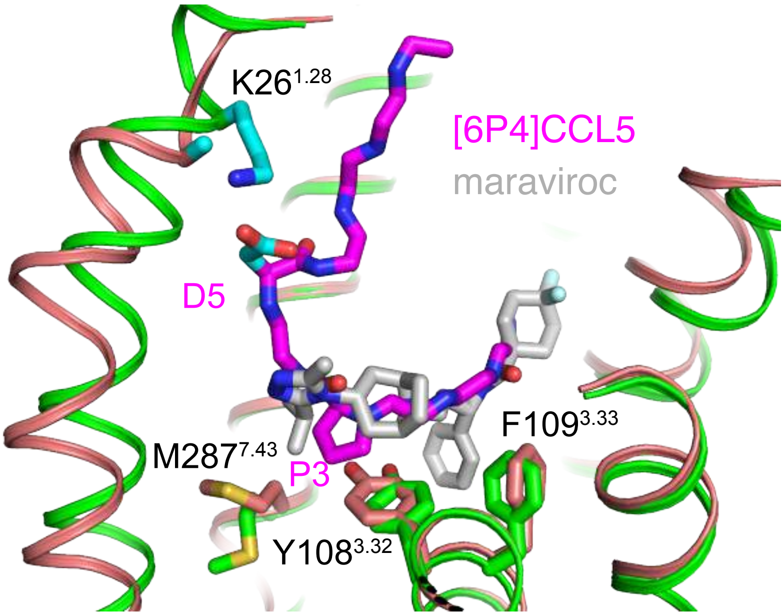


Figure S8. Comparison between the binding poses of the agonist [6P4]CCL5 N-terminus and the antagonist maraviroc in CRS2 of CCR5. Same view of CRS2 as Fig. 2C, showing the agonist [6P4]CCL5 (magenta) bound to active CCR5 (green; this structure) and the antagonist maraviroc (grey) bound to inactive CCR5 (light red; PDB ID: 4MBS). The salt bridge residues K26$^{1.28}$-CCR5 and D5-[6P4]CCL5 (both in cyan) as well P3-[6P4]CCL5 and Y108$^{3.32}$, F109$^{3.33}$, and M287$^{7.43}$ in the receptor (in active and inactive CCR5) are shown as sticks. The azabicyclo group of maraviroc does not insert as deeply as P3 of [6P4]CCL5 at the equivalent position. As a consequence, the sidechain of M287$^{7.43}$ remains in the inactive conformation (see Fig. 2C). The phenyl group of maraviroc, instead, reaches deeply into the receptor between the residues of the aromatic cluster, possibly blocking its function as a signal relay.

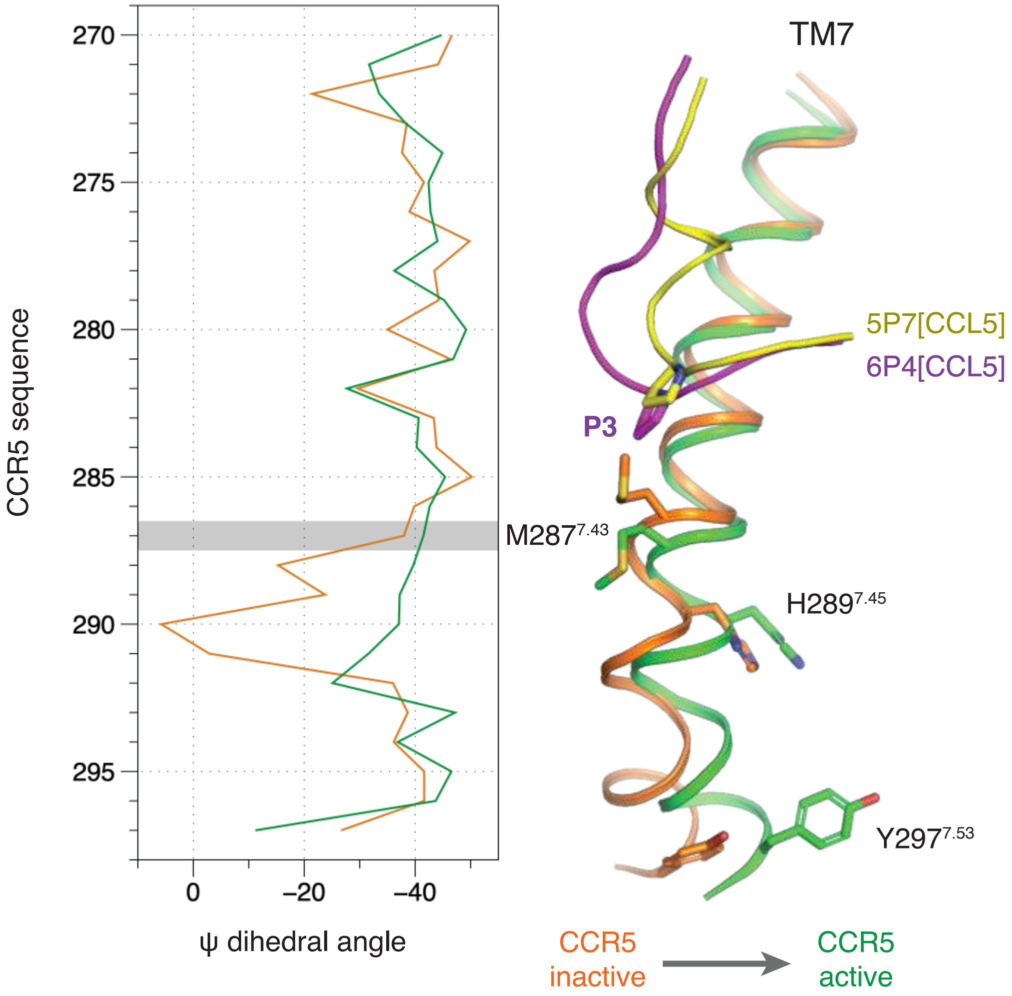


Figure S9. Comparison of the structure of TM7 between inactive [5P7]CCL5-bound and active [6P4]CCL5-bound CCR5. Left: ψ dihedral angles along TM7 reveals significant changes in the structure of the helical backbone at residues T288$^{7.44}$ – C291$^{7.47}$. Right: These changes occur together with the relocation of the M287$^{7.43}$ side chain (forced by the deeper insertion of the chemokine N-terminus at P3) and translate into the movement of the cytoplasmic side of TM7 (and the key residues H289$^{7.45}$ and Y297$^{7.53}$) towards the receptor core.

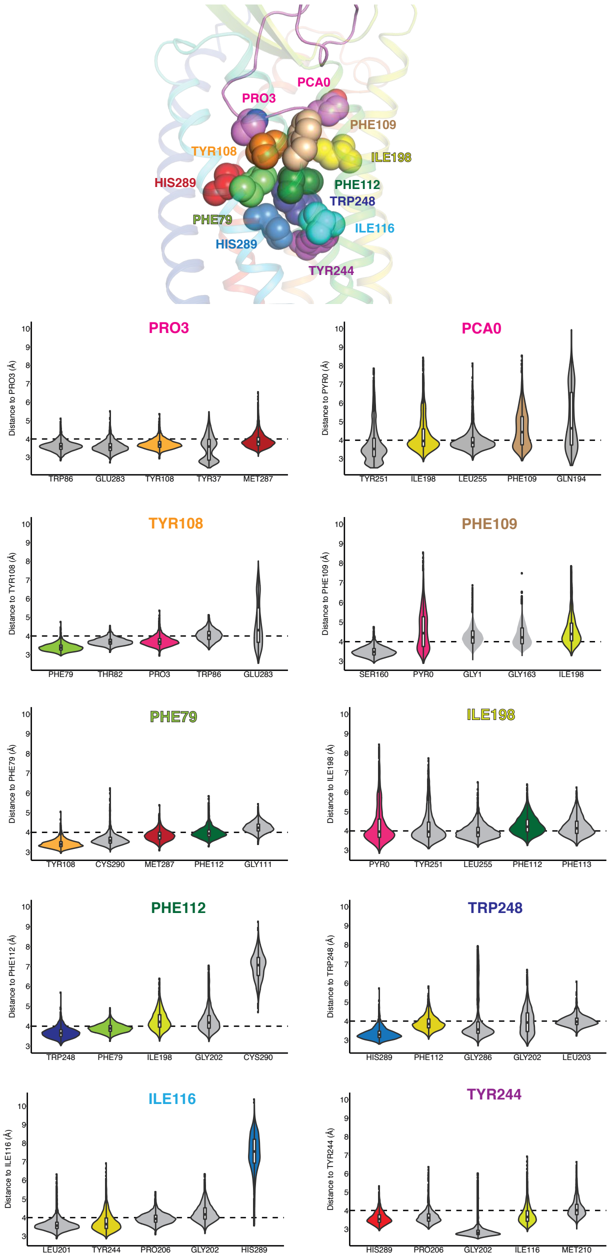


Figure S10. Residue contacts in the aromatic connector. (A) Residues PCA0 and P3 in [6P4]CCL5 and residues in the aromatic connector (Y108$^{3.32}$, F109$^{3.33}$ and F112$^{3.36}$) and its vicinity (F79$^{2.52}$, I198$^{5.43}$, H289$^{7.45}$) linking to W248$^{6.48}$ and Ile116$^{3.40}$ and Y244$^{6.44}$ in the PIF motif of CCR5 are shown as spheres.(B) The violin plots depict the populations of residue-residue distances between residues in panel A. Distances have been calculated from all replicas in our MD simulations. Coloring of the violin plots corresponds to that of panel A. The MD data shows that the contacts depicted in A are persistent, connecting the N-terminal residues of the chemokine to key activation switches of CCR5.

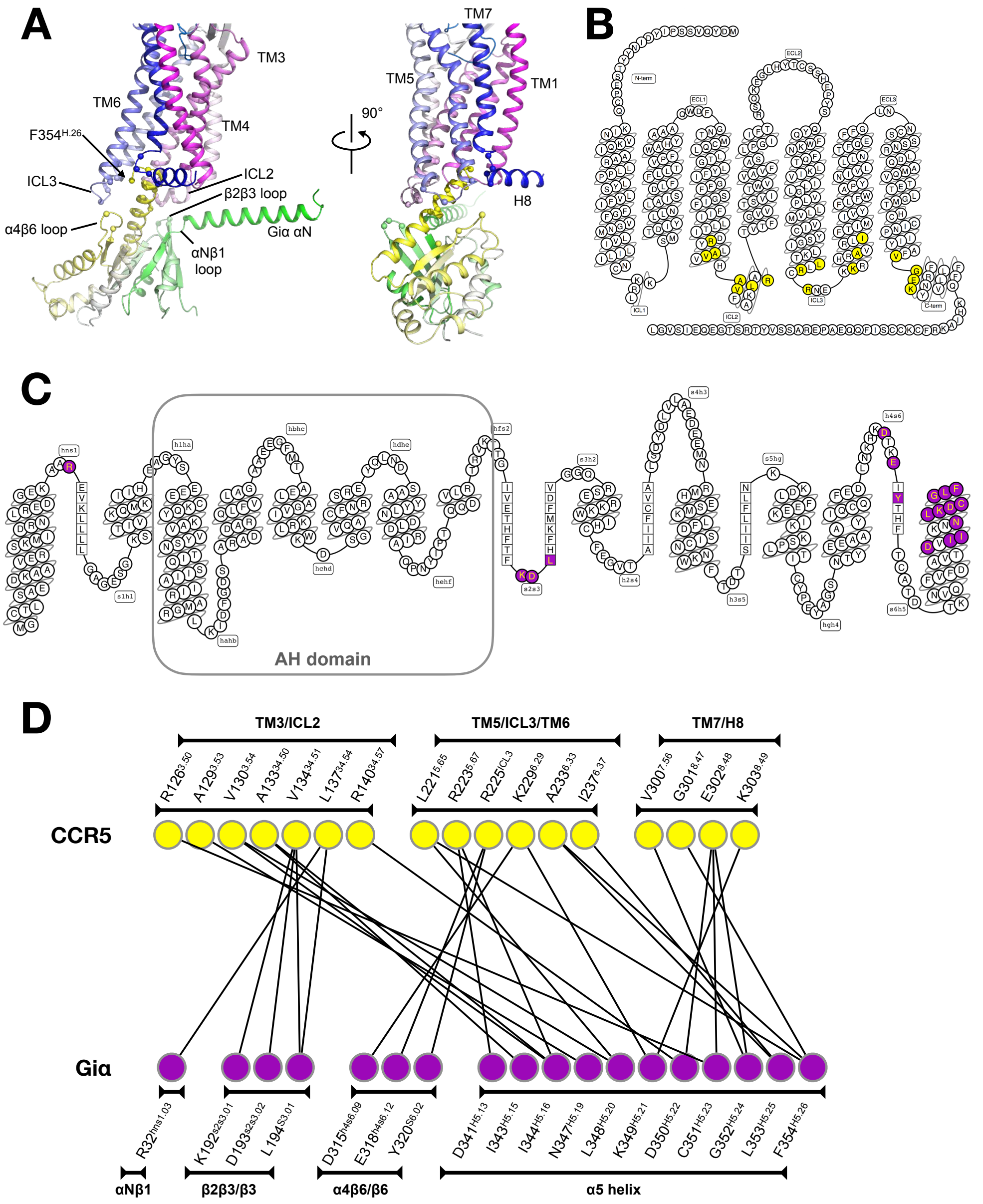


Figure S11. Contact interface between CCR5 and $G\alpha_i$. (A) Structure of CCR5 and $G\alpha_i$ in the complex colored in blue-white-magenta (receptor) and green-white-yellow (G protein) spectra. The α carbon atoms of contact residues (within 4 Å) are shown as spheres. (B) Snake plot of CCR5; contacts with $G\alpha_i$ are colored in yellow. (C) Snake plot of $G\alpha_{i1}$; contacts with CCR5 are colored in purple. (D) Table of residue-residue contacts between CCR5 and $G\alpha_i$.

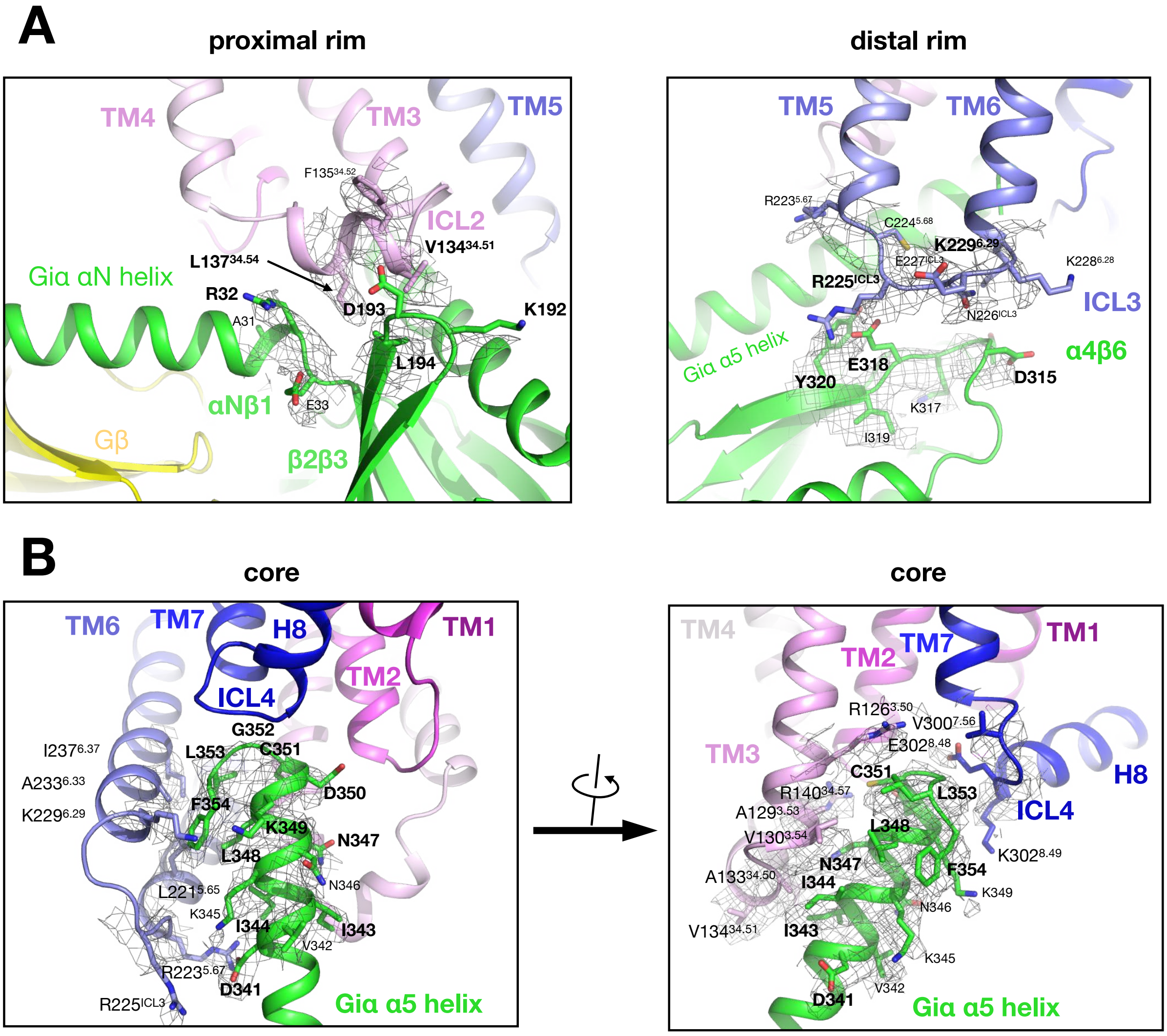


Figure S12. Electron density at the contact interface between CCR5 and Gα_i. (A) Left: contacts at the proximal rim of the interface, formed by ICL2 of CCR5 and the αNβ1/β2β3 loops of Gα_i. The mesh depicts the cryo-EM map contoured at a 7.5 σ cut-off for residues 137-139 of CCR5 and 31-33 of Gα_i. Right: contacts at the distal rim of the interface, formed by ICL3 of CCR5 and the α4β6 loop of Gα_i. The mesh depicts the cryo-EM map contoured at a 7.5 σ cut-off for residues 224-227 of CCR5 and 314-319 of Gα_i. (B) Contacts at the core of the interface, formed by ICL1 and ICL4 the cytoplasmic ends of TM3, TM5, and TM6 of CCR5 and the α5 helix of Gα_i. CCR5 residues within 4 Å of the Gα_i α5 helix and residues 341-354 of the Gα_i α5 helix are displayed as sticks with the grey mesh depicting the cryo-EM map at a 7.5 σ cut-off. Residues on the α5 helix residing within 4 Å to CCR5 are highlighted in bold.

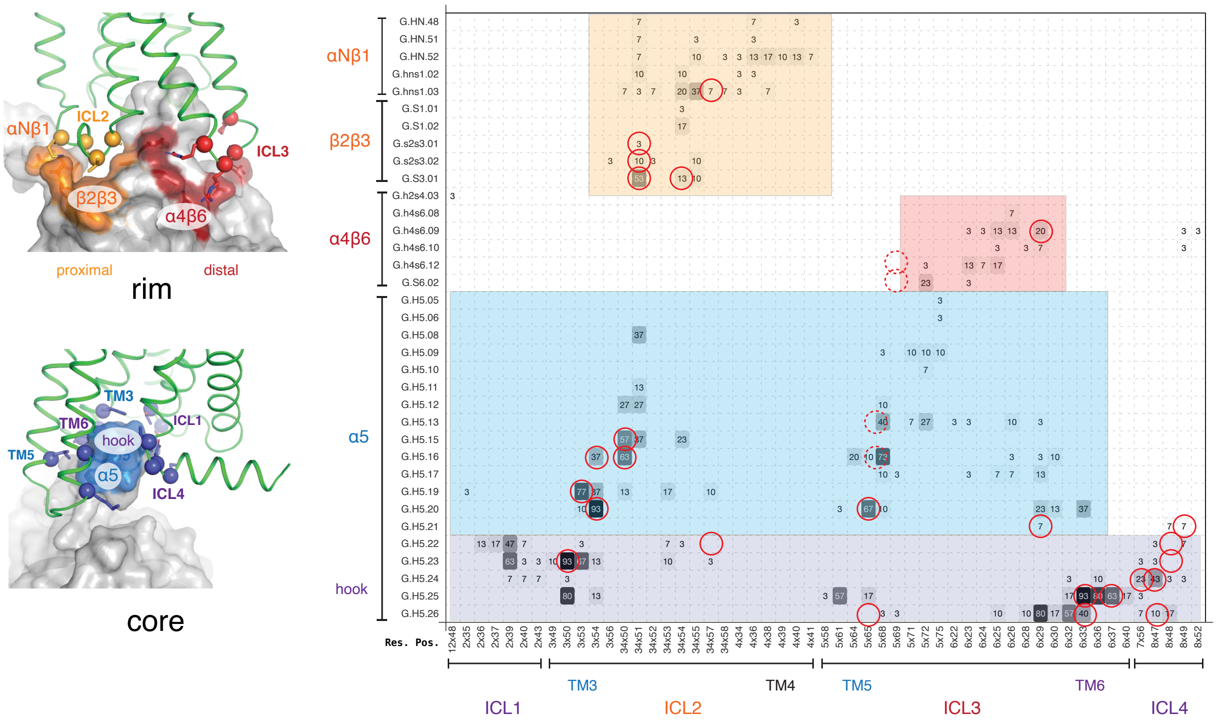


Figure S13. Receptor/Gα protein contact map of currently available GPCRs/G_i complexes. The map was obtained from www.gpcr.org. GPCR residues (in ICL1, ICL2, ICL3, and ICL4 and in the cytoplasmic ends of TM3, TM4, TM5, and TM6) are in the X-axis and $G\alpha_i$ residues (in the αNβ1, β2β3, α4β6 loops and in the α5 helix) are in the Y-axis. The number of contacts at each position (in all available complexes) is depicted in the map in grey scale (light grey: few contacts at this position across complexes; dark grey: many contacts). Red circles represent the contacts present in the CCR5/[6P4]CCL5 complex. Dashed red circles represent contacts present in our complex but shifted by one residue compared to other complexes. The proximal rim of the interface is highlighted in orange, the distal rim in red, and the core in light (base of α5) and dark (hook of α5) blue. These interfaces are shown on the 3D molecular structure in the left. At the proximal rim of the interface (orange), the CCR5/[6P4]CCL5 features several of the 'expected' contacts (red circles). On the other hand, the distal rim of the interface presents contacts shifted in other complexes (dashed red circles). At the core of the interface, many of the expected contacts are present (red circles), while there are several contacts unique for the CCR5/[6P4]CCL5 complex (red circles with white background). While there is certain variability in the precise composition of the interfaces, the CCR5/[6P4]CCL5 complex is overall similar to existing complexes, with some variability at the distal rim and the α5 hook.

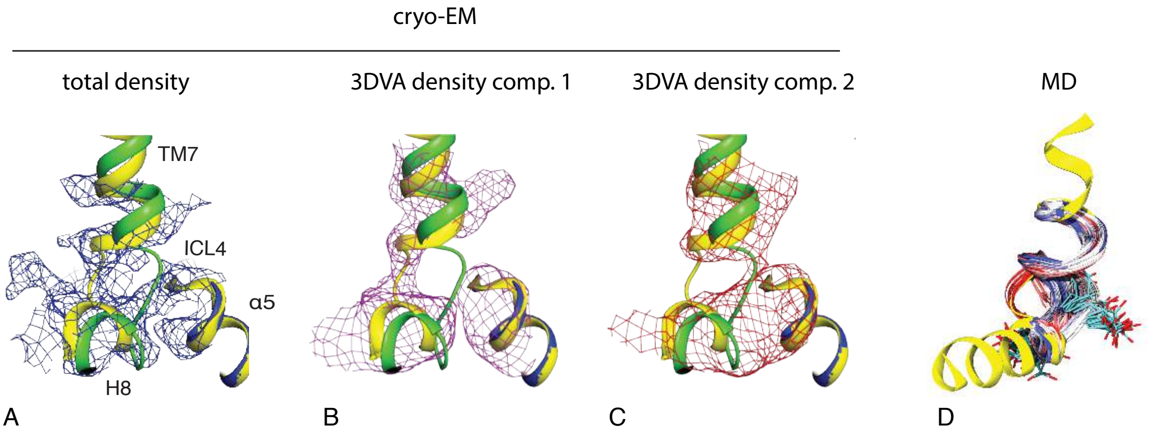
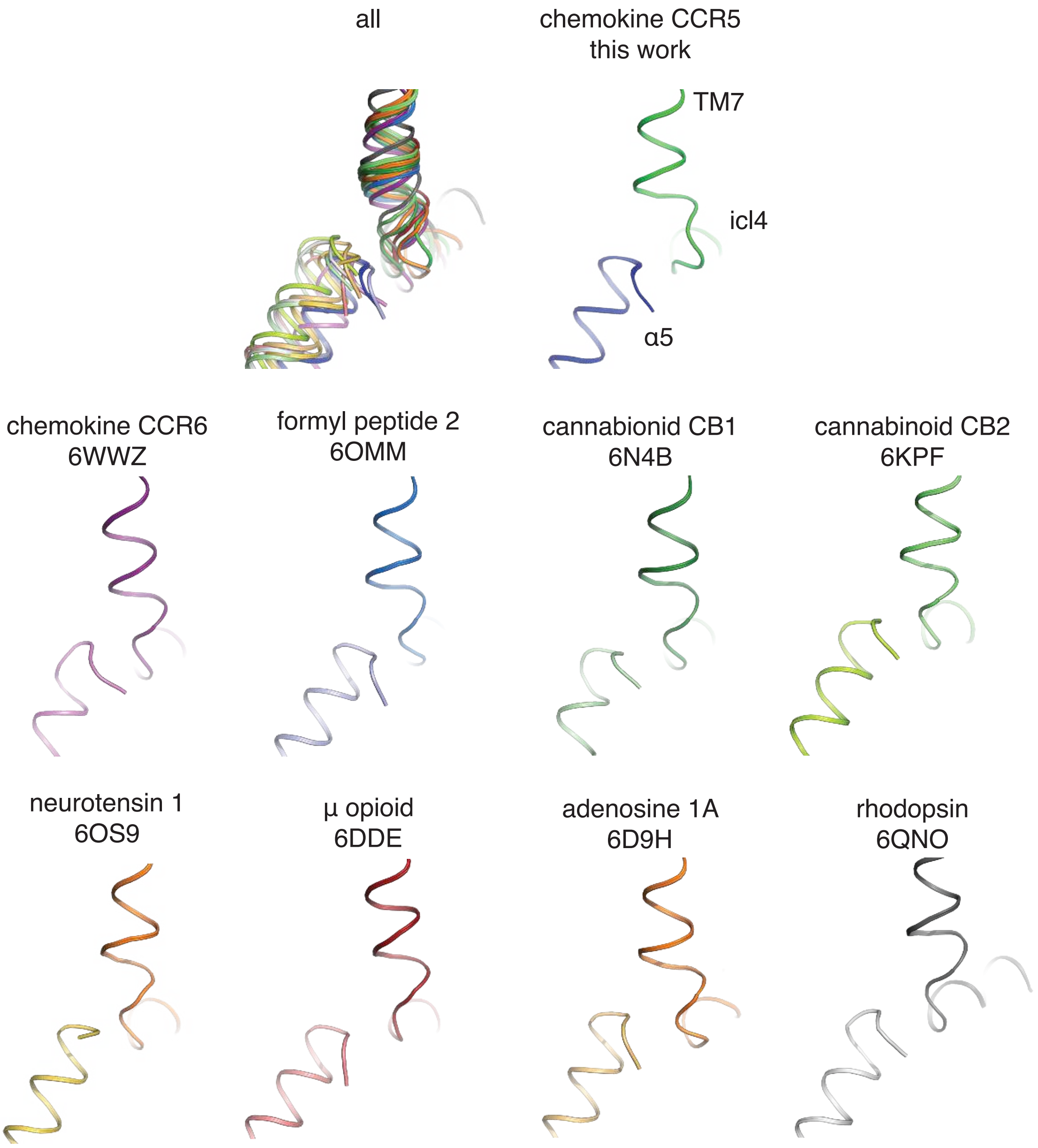


Figure S14. Structural plasticity of ICL4 in the cryo-EM density and in the MD simulations of the [6P4]CCL5•CCR5 complex. (A) Total cryo-EM density at the position of ICL4 (blue) together with the deposited structural model (yellow) and an alternative model obtained from 3D Variability Analysis (3DVA; green). (B) Cryo-EM density map of one 3DVA component together with the two models of ICL4. The deposited model (yellow) of ICL4 fits the generated density. (C) Cryo-EM density map of a further 3DVA component together with the two models of ICL4. The alternative model (green) of ICL4 fits the generated density. (D) MD simulation of [6P4]CCL5•CCR5 complex starting from the cryo-EM model. The CCR5 TM7-ICL4-HX8 junction in the cryo-EM structure described in this work is shown as a yellow cartoon. A representative conformational ensemble of CCR5 residues 296 to 304 from the simulations (in the absence of $G\alpha_i$) is depicted as colored cartoons. Time evolution during the simulation is represented in a red-to-blue color gradient (red – beginning of the simulation; blue – end of the simulation). The side chain of residue E302$^{8.48}$ is shown as sticks. The simulations indicate that the local structure of ICL4 is plastic and moves from the cryo-EM model conformation to the conformation observed in other structures, thus responding to the absence of the $G\alpha_i$. In contrast, in the presence of the bound $G\alpha_i$, ICL4 and the side chain of E302$^{8.48}$ adopt two conformations in the cryo-EM structure.



```
>ccr5       AMQVTETLGMTHCCINPIIYAFVGEKFRN
>ccr6       TKTVTEVLAFLHCCLNPVLYAFIGQKFRN
>fp2-6omm   LVNPTSSLAFFNSCLNPMLYVFVGQDFRE
>cb1-6n4b   VFAFCSMLCLLNSTVNPIIYALRSKDLRH
>cb2-6kpf   AFAFCSMLCLINSMVNPVIYALRSGEIRS
>nt1-6os9   FYMVTNALFYVSSTINPILYNLVSANFRH
>mop-6dde   SWHFCIALGYTNSCLNPVLYAFLDENFKR
>aa1r-6d9h  LTYIAIFLTHGNSAMNPIVYAFRIQKFRV
>rho-6qno   FMTIPAFFAKTSAVYNPVIYIMMNKQFRN
```

Figure S15. Structure and sequence of ICL4 in available GPCRs in complex with $G_{i/o}$. (A) A structural alignment (on the GPCR transmembrane bundle) of GPCR/$G_{i/o}$ structures reveals the plasticity of ICL4. (B) A sequence alignment of ICL4 in GPCRs bound to $G_{i/o}$ shows the sequence variability in this region.

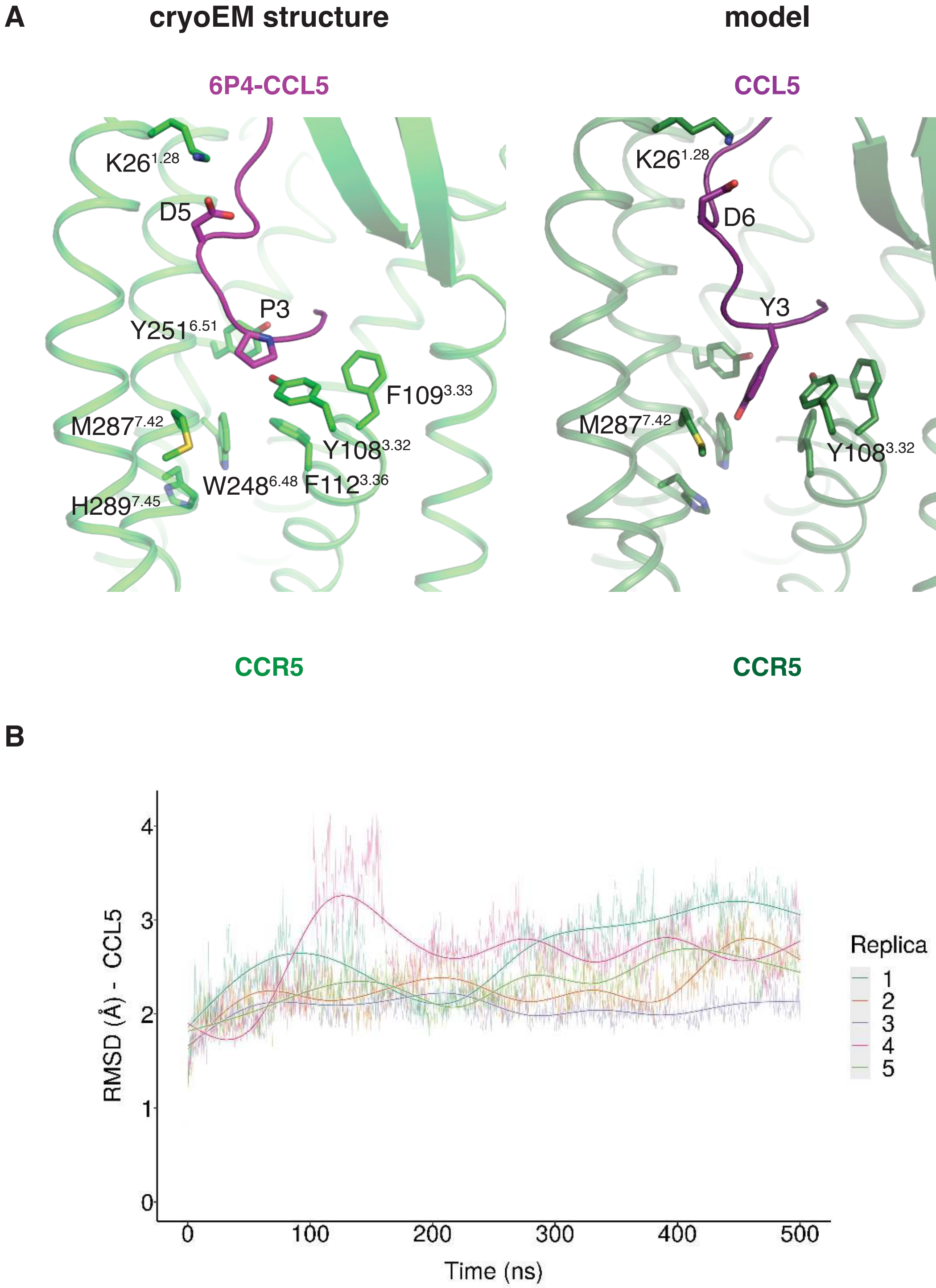


Figure S16. Model of wild type CCL5 bound to CCR5. (A) Left: cryo-EM structure of the [6P4]CCL5•CCR5 complex. D5 in the chemokine interacts with $K26^{1.28}$ in the receptor stabilizing the conformation of the CCL5 N-terminus. P3 in the chemokine pushes on $M287^{7.42}$ and $Y108^{3.32}$ in the aromatic connector. Right: model of wild type CCL5 in complex with CCR5. In this case, residues D6 and Y3 in the chemokine may play the same role as D5 and P3 in our structure. (B) Root mean square deviation (RMSD) of the modeled CCL5 in MD simulations of the CCR5•CCL5 complex. Each line represents an independent simulation replica, and the overlaid bold lines represent smoothed averages. The data show that the modeled pose of CCL5 is stable during the simulations.

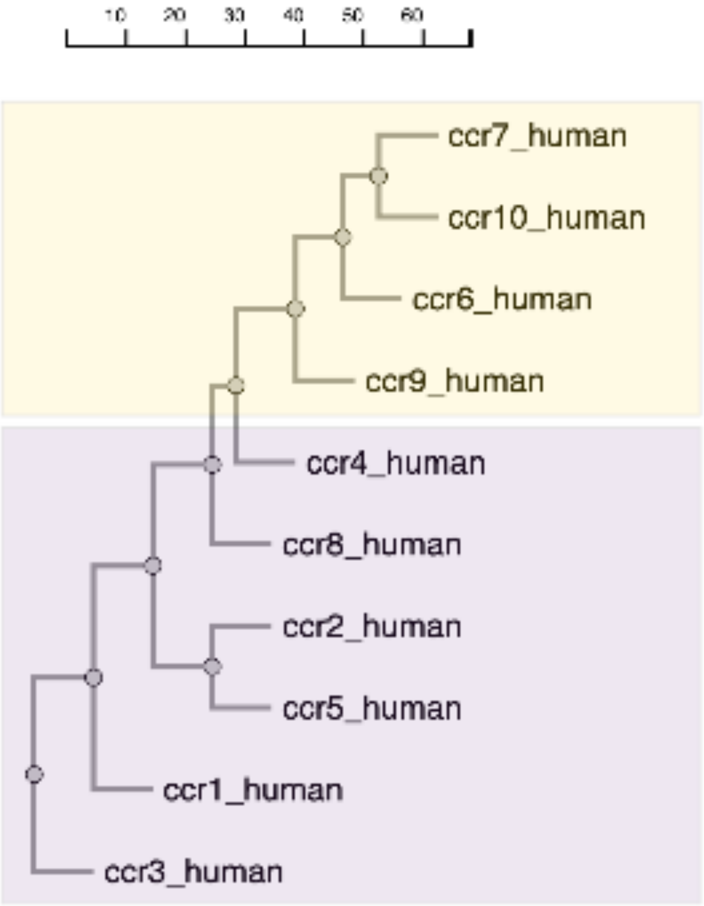


Figure S17. Phylogenetic tree of CC chemokine receptors. The tree was obtained at www.gpcrdb.org (see Methods), and shows how CC chemokine receptors containing $W^{6.48}$ (CCR1-5 and CCR8; magenta) and those containing $Q^{6.48}$ (CCR6, CCR7, CCR9, and CCR10; yellow) form distinct subgroups (see Figure 6 in the main text).

Supplementary Video Legends

Supplementary Video S1. 3D variability analysis (3DVA) of the [6P4]CCL5•CCR5•Gi•Fab16 cryo-EM map. The analysis resulted in variability components 0, 1 and 2 displayed in the videos S1a, S1b and S1c. The following motions can be perceived in the 3DVA: bending of the C-terminal helix of [6P4]CCL5 towards CCR5; variation of the insertion depth of the [6P4]CCL5 N-terminus in the CCR5 binding pocket; opening of the binding pocket of CCR5 by the rearrangement of the 7TM bundle; correlated motions of G_i and Fab16; twisting and contraction of the transmembrane helix bundle of CCR5; and elongation of TM3, which maintains the contact of ICL2 with $G\alpha_i$. The dynamic nature of GPCRs produces heterogeneous density maps in cryo-EM studies, which leads to locally blurred density maps. 3DVA allows to visualize this dynamic behavior of the complex and reveals several modes of possible functional significance which cannot be described by one density map and single structure.

Supplementary Video S2. Simulation of the interaction between the sulfated CCR5 N-terminus and [6P4]CCL5. CCR5 is displayed as white cartoons with its N-terminus in green. Sulfated tyrosines Y10 and Y14 are shown as sticks (cyan carbon atoms) and [6P4]CCL5 is shown as a pink surface.

Supplementary Video S3. Representative simulation of the CCL5•CCR5 model. CCR5 is represented as green cartoons, and CCL5 in magenta. Key residues are shown as cyan sticks. In the molecular dynamics trajectory, residue Y3 of CCL5 (at the center of the image) engages the aromatic connector of CCR5 (Y108$^{3.32}$, F109$^{3.33}$, F112$^{3.36}$ (center left), and Y251$_{6.51}$ (center right)). Residue E6 of CCL5 engages K26 of CCR5 forming a salt bridge (top). Both features are preserved during the MD simulations.

Table S1: Cryo-EM data collection, refinement and validation statistics.

Data collection and processing

Super-resolution pixel size (Å)	0.556
Voltage (kV)	200
Electron exposure (e–/Å^2)	49
Defocus range	-1.0 to -2.0
Pixel size (Å)	1.112
Symmetry imposed	C1
Initial particle images (no.)	2 639 161
Final particle images (no.)	345 458
Map resolution (Å)	3.13
FSC threshold	0.143

Refinement

Initial model used (PDB code)	5UIW, 5KDO, 6QNK
Model resolution (Å)	3.15
FCS threshold	0.143
Map sharpening B factor (Å^2)	-90
Model composition	
Non-hydrogen atoms	11172
Protein residues	1420
B factors (Å^2)	
Protein	43.4
R.m.s. deviations	
Bond length (Å)	0.006
Bond angles (°)	0.924
Validation	
MolProbability score	1.52
Clashscore	7.66
Rotamer outliers (%)	0.16
Ramachandran plot	
Favoured (%)	97.50
Allowed (%)	2.50
Disallowed (%)	0.00

Table S2: N-terminal sequences and CCR5 signaling activities of the panel of CCL5 analogs used for sequence activity analysis in Figure 2.

		CCL5 sequence position									
analog	**Ca2+ signaling activity**	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1P1-CCL5	BD	L	S	P	V	S	S	Q	S	S	A
1P2-CCL5	93.8	F	S	P	L	S	S	Q	S	S	A
1P3-CCL5	BD	L	S	P	M	S	S	Q	S	P	A
1P4-CCL5	BD	W	S	P	L	S	S	Q	S	P	A
1P5-CCL5	BD	L	S	P	Q	S	S	L	S	S	S
1P6-CCL5	26.5	Z	S	P	G	S	S	W	S	A	A
1P7-CCL5	93.5	M	S	P	L	S	S	Q	A	S	A
2P1-CCL5	BD	F	V	P	Q	S	G	Q	S	T	P
2P10-CCL5	BD	Z	G	P	P	M	M	Q	S	T	P
2P11-CCL5	38.6	T	G	P	P	G	G	Q	S	T	P
2P12-CCL5	93.7	V	G	P	L	S	Q	Q	A	T	P
2P13-CCL5	13.4	Z	F	P	P	G	G	Q	S	T	P
2P14-CCL5	52.7	F	A	P	M	S	Q	Q	S	T	P
2P15-CCL5	99.5	A	A	P	L	S	Q	Q	S	T	P
2P2-CCL5	BD	L	V	P	Q	P	G	Q	S	T	P
2P3-CCL5	BD	Z	G	P	P	L	M	Q	T	T	P
2P4-CCL5	BD	M	V	P	Q	S	G	Q	S	T	P
2P5-CCL5	BD	Z	G	P	P	M	M	Q	T	T	P
2P6-CCL5	36.7	Z	G	P	P	G	G	Q	T	T	P
2P7-CCL5	103.9	F	A	P	M	S	Q	Q	S	T	S
2P8-CCL5	95.6	Z	G	P	L	S	G	Q	S	T	P
2P9-CCL5	81.5	Z	G	P	P	G	G	Q	S	T	P
5P1-CCL5	BD	Z	G	P	P	L	M	W	L	Q	V
5P10-CCL5	BD	Z	G	P	P	L	M	W	L	Q	T
5P11-CCL5	BD	Z	G	P	P	L	M	W	R	G	S
5P12-CCL5	BD	Z	G	P	P	L	M	A	T	Q	S
5P13-CCL5	BD	Z	G	P	P	L	M	W	L	G	G
5P14-CCL5	BD	Z	G	P	P	L	M	S	L	Q	V
5P15-CCL5	42.7	Z	G	P	P	L	M	S	L	S	V
5P16-CCL5	BD	Z	G	P	P	L	M	G	L	S	V
5P2-CCL5	BD	Z	G	P	P	L	M	W	L	Q	S
5P3-CCL5	BD	Z	G	P	P	L	M	W	M	Q	V
5P4-CCL5	BD	Z	G	P	P	L	M	W	M	Q	S
5P5-CCL5	BD	Z	G	P	P	L	M	W	T	Q	V
5P6-CCL5	BD	Z	G	P	P	L	M	W	T	Q	S
5P7-CCL5	BD	Z	G	P	P	L	M	A	L	Q	S
5P8-CCL5	BD	Z	G	P	P	L	M	S	T	Q	S
5P9-CCL5	BD	Z	G	P	P	L	M	S	F	Q	S
6P1-CCL5	31.8	Z	G	P	P	G	G	G	G	L	G

6P10-CCL5	96.3	Z	G	P	P	G	D	T	V	L	W
6P11-CCL5	90.1	Z	G	P	P	G	S	Y	D	Y	S
6P12-CCL5	14.9	Z	G	P	P	L	G	A	G	S	S
6P13-CCL5	22.6	Z	G	P	P	L	G	S	M	G	P
6P14-CCL5	BD	Z	G	P	P	L	D	F	G	G	A
6P15-CCL5	26.0	Z	G	P	P	M	G	G	T	S	A
6P16-CCL5	BD	Z	G	P	P	M	Q	G	G	L	S
6P17-CCL5	BD	Z	G	P	P	M	M	A	G	L	S
6P18-CCL5	BD	Z	G	P	P	L	Q	A	S	V	T
6P19-CCL5	24.8	Z	G	P	P	M	S	G	H	S	T
6P2-CCL5	73.1	Z	G	P	P	G	D	G	G	Q	V
6P20-CCL5	BD	Z	G	P	P	M	S	A	Y	Q	V
6P3-CCL5	85.1	Z	G	P	P	G	D	G	G	S	V
6P4-CCL5	87.7	Z	G	P	P	G	D	I	V	L	A
6P5-CCL5	51.1	Z	G	P	P	G	G	G	G	Q	S
6P6-CCL5	51.7	Z	G	P	P	G	G	G	G	T	R
6P7-CCL5	45.0	Z	G	P	P	G	S	W	S	S	V
6P8-CCL5	34.1	Z	G	P	P	M	G	G	Q	V	T
6P9-CCL5	41.1	Z	G	P	P	G	D	T	Y	Q	A
7P1-CCL5	BD	Z	G	P	P	G	Q	W	Y	Q	S
7P10-CCL5	BD	Z	G	P	P	L	Q	W	R	Q	S
7P11-CCL5	12.7	Z	G	P	P	L	Q	W	M	Q	V
7P12-CCL5	BD	Z	G	P	P	L	M	W	S	Q	V
7P13-CCL5	BD	Z	G	P	P	G	Q	W	S	Q	V
7P14-CCL5	11.8	Z	G	P	P	L	Q	W	M	Q	A
7P15-CCL5	BD	Z	G	P	P	L	Q	W	F	Q	V
7P16-CCL5	BD	Z	G	P	P	L	Q	W	T	Q	V
7P17-CCL5	BD	Z	G	P	P	L	S	W	L	Q	S
7P18-CCL5	BD	Z	G	P	P	L	S	W	Q	Q	S
7P19-CCL5	BD	Z	G	P	P	L	S	W	L	Q	V
7P2-CCL5	BD	Z	G	P	P	L	S	W	S	Q	V
7P3-CCL5	31.1	Z	G	P	P	G	D	W	S	Q	V
7P4-CCL5	BD	Z	G	P	P	M	S	W	S	Q	S
7P5-CCL5	BD	Z	G	P	P	G	S	W	S	Q	V
7P6-CCL5	BD	Z	G	P	P	Q	G	W	S	Q	V
7P7-CCL5	BD	Z	G	P	P	Q	S	W	S	Q	A
7P8-CCL5	BD	Z	G	P	P	G	Q	W	G	Q	V
7P9-CCL5	BD	Z	G	P	P	G	M	W	S	Q	S
8P1-CCL5	88.5	Z	G	P	L	S	Q	G	S	Q	V
8P10-CCL5	97.2	Z	G	P	L	G	Q	Q	G	Q	V
8P2-CCL5	86.5	Z	G	P	L	S	Q	A	S	Q	V
8P3-CCL5	77.0	Z	G	P	L	S	Q	A	F	Q	V
8P4-CCL5	96.9	Z	G	P	L	S	Q	S	S	Q	V
8P5-CCL5	100.5	Z	G	P	L	S	S	Q	S	Q	V
8P6-CCL5	97.6	Z	G	P	L	S	G	W	A	Q	V

8P7-CCL5	88.6	Z	G	P	L	G	Q	W	A	Q	V
8P8-CCL5	73.2	Z	G	P	L	S	Q	W	Q	Q	V
8P9-CCL5	93.3	Z	G	P	L	S	Q	Q	G	Q	V
9P1-CCL5	10.3		Q	G	Q	R	I	S	D	T	T
9P10-CCL5	95.5	Z	W	V	M	G	S	D	T	T	P
9P12-CCL5	97.8	F	F	Q	T	P	S	D	T	T	P
9P14-CCL5	81.4	Y	Q	Q	F	G	S	D	T	T	P
9P3-CCL5	BD	M	S	K	M	Q	S	D	T	T	P
9P6-CCL5	17.4	M	S	Q	F	R	S	D	T	T	P
9P8-CCL5	33.4	L	G	Q	K	K	S	D	T	T	T
9P9-CCL5	BD	M	S	Q	R	I	S	D	T	T	P
M10-CCL5	BD	M	S	P	Y	S	M	Q	T	T	P
M11-CCL5	BD	Z	G	P	P	L	M	Q	S	A	A
M12-CCL5	BD	M	S	P	L	S	S	W	L	Q	V
M13-CCL5	84.3	M	S	P	L	S	S	Q	A	Q	V
M14-CCL5	48.4	M	S	P	L	L	M	W	L	Q	V
M15-CCL5	BD	Z	G	P	L	S	G	W	L	Q	V
M16-CCL5	BD	Z	G	P	L	L	M	W	L	Q	V
M17-CCL5	BD	Z	G	P	L	S	M	W	L	Q	V
M18-CCL5	30.9	Z	G	P	L	S	G	Q	L	Q	V
M19-CCL5	74.1	Z	G	P	L	S	G	Q	S	Q	V
M20-CCL5	14.2	Z	G	P	P	G	D	W	L	Q	V
M21-CCL5	14.6	Z	G	P	P	L	M	S	V	L	A
M22-CCL5	BD	Z	G	P	P	L	M	G	L	Q	V
M23-CCL5	BD	Z	G	P	P	L	M	A	L	Q	V
M24-CCL5	BD	Z	G	P	P	L	M	Q	L	Q	V
M25-CCL5	BD	Z	G	P	P	L	M	N	L	Q	V
M26-CCL5	BD	Z	G	P	P	L	M	D	L	Q	V
M27-CCL5	BD	Z	G	P	P	L	M	R	L	Q	V
M28-CCL5	BD	Z	G	P	P	L	M	T	L	Q	V
M29-CCL5	BD	Z	G	P	P	L	M	G	T	Q	S
M30-CCL5	29.6	Z	G	P	P	L	M	S	T	Q	V
M31-CCL5	BD	Z	G	P	P	L	M	A	T	Q	S
M33-CCL5	BD	Z	G	P	P	L	M	G	T	Q	S
M34-CCL5	18.9	Z	G	P	P	L	M	S	T	Q	S
M36-CCL5	BD	Z	G	P	P	L	M	V	T	Q	S
M37-CCL5	17.2	Z	G	P	P	L	M	S	L	Q	S
M38-CCL5	BD	Z	G	P	P	L	M	A	L	S	V
M39-CCL5	10.6	Z	G	P	P	L	M	S	G	Q	S
M40-CCL5	BD	Z	G	P	P	L	M	S	S	Q	S
M41-CCL5	10.3	Z	G	P	P	L	M	S	A	Q	S
M42-CCL5	33.4	Z	G	P	P	L	M	S	V	Q	S
M43-CCL5	11.5	Z	G	P	P	L	M	S	L	Q	A
M44-CCL5	77.4	Z	G	P	P	L	M	S	L	T	V
M45-CCL5	68.5	Z	G	P	P	L	M	S	L	A	V

M46-CCL5	20.1	Z	G	P	P	L	M	S	G	Q	V
M7-CCL5	BD	Z	G	P	Y	S	S	D	T	T	P
M8-CCL5	BD	M	S	P	P	L	M	Q	T	T	P
M9-CCL5	23.8	M	S	P	P	L	S	D	T	T	P
Met-CCL5	15.0	M	S	P	Y	S	S	D	T	T	P
NNY-CCL5	84.5	ø	ø	P	Y	S	S	D	T	T	P
PSC-CCL5	100.0	ø	ø	#	#	S	S	D	T	T	P
CCL5	54.2		S	P	Y	S	S	D	T	T	P

Signaling activities (% of PSC-CCL5 signal at 300 nM (E_{max}) concentration) are reported from [Gaertner et al (2008) PNAS 105:17706–17711] except analogs marked §, which were tested subsequently using the same methods used in [Gaertner et al (2008) PNAS 105:17706–17711]. Z: pyroglutmate residue, øø: nonanoyl, ##: thioprolyl-cyclohexylglycyl, BD: signal below assay detection limit.

Table S3: Results of cellular Ca2+ flux activation assays.

Activation by [6P4]CCL5 mutants of HEK cell clones stably expressing CCR5

	[6P4]CCL5 (batch 1)	[6P4]CCL5 (batch 2)	[6P4]CCL5-D5A	[6P4]CCL5-D5K	CCL5
R0	1.102	1.09	1.061	1.062	1.096
95% CI	1.055 to 1.148	1.055 to 1.124	1.049 to 1.073	1.052 to 1.072	1.066 to 1.126
Rmax	2.537	2.607	2.078	1.617	2.133
95% CI	2.456 to 2.622	2.544 to 2.672	2.027 to 2.134	1.545 to 1.717	2.046 to 2.231
EC50	9.601	10.28	55.21	121.2	31.17
95% CI	7.148 to 12.89	8.354 to 12.65	46.88 to 65.16	85.00 to 177.6	22.28 to 43.49

Activation by [6P4]CCL5 mutants of CHO cell clones stably expressing CCR5

	[6P4]CCL5 (batch 1)	[6P4]CCL5 (batch 2)	[6P4]CCL5-D5A	[6P4]CCL5-D5K	CCL5
R0	1.063	1.041	1.045	1.051	1.039
95% CI	1.036 to 1.090	1.019 to 1.063	1.033 to 1.057	1.036 to 1.065	0.9987 to 1.078
Rmax	2.583	2.531	2.024	1.367	2.228
95% CI	2.515 to 2.655	2.480 to 2.584	1.968 to 2.086	1.267 to 1.709	2.051 to 2.484
EC50	23.56	19.87	67.33	141.4	112.8
95% CI	19.68 to 28.22	17.22 to 22.94	56.52 to 80.54	53.94 to 555.0	74.21 to 179.7

Activation by [6P4]CCL5 of HEK cells transiently transfected with CCR5 mutants

	CCR5 WT	CCR5 E283A	CCR5 M287A	CCR5 Y108A
R0	1.004	1.013	1.026	0.9809
95% CI	0.9786 to 1.029	0.9911 to 1.034	1.006 to 1.045	0.9523 to 1.009
Rmax	2.138	1.729	1.328	1.694
95% CI	2.087 to 2.192	1.695 to 1.765	1.296 to 1.365	1.659 to 1.730
EC50	123.2	68.69	80.36	35.37
95% CI	102.0 to 149.0	54.52 to 86.68	48.43 to 134.5	27.44 to 45.59

The fluorescence response R was fitted to a 3-parameter model as R = R0 + [agonist]x(Rmax-R0)/(EC50+[agonist]), where Emax = Rmax-R0 and R0 presents the baseline.

Table S4: List of MD simulations of the CCL5/CCR5 complex.

System	**Simulation time (ns)**	**Number of Replicas**	**Details**
1	500	5	CCR5–[6P4]CCL5. Non-sulfated residues Y10 and Y14
2	500	5	CCR5–[6P4]CCL5. Sulfated residues Y10 and Y14
3	500	5	CCR5–CCL5. Sulfated residues Y10 and Y14

3 Preparation of a stable [6P4]CCL5•CCR5•G_i signaling complex for Cryo-EM analysis

3.1 Original Manuscript

Reference: Isaikina P.* Tsai C.-J., Petrovic I., Rogowski M., Meng Dürr A., Grzesiek S.* Preparation of a stable CCL5•CCR5•G_i signaling complex for Cryo-EM analysis. *Methods in Cell Biology, Elsevier*, **2022**
https://doi.org/10.1016/bs.mcb.2022.03.001
*Corresponding author

Short description: This publication describes the biochemical preparation of the [6P4]CCL5•CCR5•G_i signaling complex, as well as the biophysical characterization of the super-agonist [6P4]CCL5. In addition, the Appendix of this Chapter describes the NMR characterization of the partial agonist [5P14]CCL5 in comparison to the antagonist [5P12]CCL5.

3.1.1 Abstract

The numerous chemokines and their cognate G protein-coupled chemokine receptors on the surface of leukocytes form a complex signaling network, which regulates the immune response and also other key physiological processes. Currently only a very limited number of structures of chemokine•chemokine receptor complexes have been solved. More structures are needed for the understanding of their mechanism of action and the rational design of drugs against these highly relevant therapeutic targets. Recently, we have determined the cryo-EM structure of the human wild-type CCR5 chemokine receptor, which is also the HIV-1 coreceptor, in its active conformation bound to the chemokine super-agonist [6P4]CCL5 and the heterotrimeric G_i protein. The structure provides the rationale for the sequence-activity relation of agonist and antagonist CCR5 chemokine ligands.

In this chapter, we present a detailed protocol for the preparation of the active agonist chemokine•CCR5•G_i complex for cryo-EM studies including quality controls and caveats. As such the protocol may serve as starting point for structural and biophysical studies of other chemokine•chemokine receptor complexes.

3.1.2 Introduction

Chemokines and their cognate receptors form a complex signaling system comprising more than 50 chemokines and 24 receptors in humans (1). They play a crucial role in a wide range of physiological processes mainly related to chemotactic trafficking of leukocytes from the blood to the tissues and around the lymphatic system (2). As such, the chemokine system is strongly implicated in many pathologies including AIDS (3), COVID-19 (4) and cancer (5–7). The chemokine/chemokine receptor pairs involved in specific pathologies constitute highly relevant therapeutic targets (7,8). Despite extensive efforts, to date only three drugs targeting chemokine receptors have been approved for clinical use (8,9). The primary reason is apparently a still limited understanding of the underlying interactions within the chemokine system.

Chemokines are a subgroup of soluble signaling proteins of 8–12 kDa size within the large cytokine family. They are divided into 4 subfamilies (CC, CXC, CX3C, and XC) according to the spacing and presence of structurally important N-terminal cysteine residues, which form disulfide bonds to other conserved cysteines in the protein core (10). Despite their functional diversity and relatively low sequence homology, the three-dimensional structures of chemokines are highly similar (11). While many chemokines form dimers and often higher oligomers, they bind to their membrane receptors as monomers. In contrast, chemokines interact strongly in oligomeric form with a ‘second type of receptor’ on the cell surface (10,12), the glycosaminoglycans (GAGs). This interaction is critical for leukocyte recruitment and is thought to create localized pools of chemokines that can then bind to their cognate membrane receptors (12).

Dimerization of CC chemokines (CCLs) occurs via intermolecular β-sheet formation of their unstructured N-termini, which precludes receptor binding for steric reasons (12,13). Higher oligomerization of CC chemokines occurs via additional interactions involving their C-terminal helix and second β-strand (14). In contrast, CXC chemokines (CXCLs) dimerize via an intermolecular extension of their antiparallel β-sheet (11), which appears compatible with receptor binding (12,15).

The dissociation constant of chemokine dimers is typically in the low micromolar range and pH- as well as temperature-dependent (13,16–18). To achieve high homogeneity of chemokine-chemokine receptor complex preparations for structural studies, it is essential that chemokines bind predominantly as monomers to their receptors. This is best achieved under conditions of low concentrations, suitable pH and by a high affinity (nanomolar) of the chemokine to the receptor.

Chemokine receptors are integral membrane proteins belonging to the class A of G protein-coupled receptors. Most receptors bind chemokines exclusively from one family (CC, CXC, CX3C, XC) and are classified accordingly (CCR, CXCR, CX3CR, XCR). This selectivity apparently results from distinct orientations of the chemokine N-terminus induced by the N-terminal cysteine motif (19). Within one family, the receptor-chemokine specificity is often not very strong, as many chemokines bind multiple chemokine receptors and many receptors bind multiple chemokines (1). An additional degree of complexity stems from post-translational modifications of the receptor such as sulfation and O-glycosylation, which contribute to chemokine affinity (13,20,21). Moreover, chemokine receptors can form homo- and heterodimers, as well as higher-order oligomers (22). The functional implications of receptor and chemokine oligomer formation are currently not well understood. It is clear, however, that both post-translational modifications and oligomerization vastly expand the arsenal of possible chemokine-receptor interactions. This makes the chemokine system a challenging target for structural biology and drug development.

A number of structures of chemokine receptors in their inactive conformation in complexes with small-molecule ligands have been solved by X-ray crystallography (23–29). In contrast, due to the larger size and increased dynamics of receptor complexes with intact chemokines, until recently only five full chemokine•chemokine receptor complex structures have been available (Figure 1). Two of these complexes, vMIP-II•CXCR4 (30) and [5P7]CCL5•CCR5 (31), are in the inactive receptor conformation. The remaining three are all complexes of the viral chemokine receptor US28. US28 is a special case, since it is constitutively active, and its inverse-agonist-bound as well as its apo structures (32,33) are all in active conformations. The situation has changed in the last two years due to the advances in cryo-EM technology, which enabled the structure determination of five further human chemokine•receptor complexes in their active conformation bound to G proteins (15,34–36).

One of these is the complex of the wild-type human CC chemokine receptor 5 (CCR5) with the chemokine super-agonist [6P4]CCL5, the heterotrimeric G_i protein and the stabilizing antibody fragment Fab16 (34). Since its discovery in 1996 (37–39), CCR5 has been under extensive investigation due to its major role in HIV infection (40), inflammation (41), the pathology of cancer (42), and COVID-19 (4). The sequence of [6P4]CCL5 differs from wild-type CCR5 by several amino acid substitutions in its flexible N-terminus that increase its agonistic efficacy and the affinity to CCR5. The comparison of the active [6P4]CCL5•CCR5•G_i complex to the inactive, antagonist chemokine [5P7]CCL5•CCR5 complex revealed the rationale for the sequence-activity relation of agonist and antagonist chemokines (34). Two additional structures of CCR5•G_i bound to CCL3 and CCL5 have now been published (36). The latter complexes were stabilized by fusing the chemokine C-terminus to the N-terminus of CCR5 and further mutations.

We present here a step-by-step protocol for the preparation of the complex between [6P4]CCL5, wild-type CCR5, G_i, and the stabilizing Fab16, which led to its successful structure determination (34). We include pertinent quality control steps and point out difficulties in the hope that the protocol may serve as a model for the preparation of other active chemokine•chemokine receptor•G protein complexes for cryo-EM structure determination and other biophysical studies.

3.1.3 Methods

Overview

A schematic diagram of the [6P4]CCL5•CCR5•Gi•Fab16 complex preparation for cryo-EM analysis including quality control steps is shown in Figure 2. The individual components [6P4]CCL5, CCR5, $G\alpha_i$, $G\beta\gamma$, and Fab16 are produced separately and then allowed to form the complex. Details of the production of the engineered chemokine CCL5 super-agonist analog [6P4]CCL5 (43) and the wild-type receptor CCR5 are given below, whereas detailed protocols for the production of $G\alpha_i$ (44), $G\beta\gamma$ (45), and Fab16 (46) have been presented in the indicated references.

Chemokine production

This section describes the expression and purification of the engineered super-agonist chemokine [6P4]CCL5. Other chemokines may be obtained in an analogous manner. [6P4]CCL5 is expressed in *Escherichia coli (E. coli)* as an N-terminal thioredoxin/His tag fusion protein (Figure 3A). A summary of the main production steps is shown in Figure 3B.

In brief, the expressed fusion protein is unfolded by 6 M guanidine hydrochloride and then purified by nickel-affinity chromatography, followed by reduction of its two disulfide bonds by 2-mercaptoethanol. Refolding and simultaneous formation of disulfide bonds is induced by adding the protein to a refolding redox buffer containing a mixture of reduced and oxidized glutathione. After refolding, the fusion protein is cleaved by enterokinase, resulting in [6P4]CCL5 with a blunt N-terminus starting at residue glutamine-0 (Q0). The fusion partner is then separated by reversed-phase chromatography. The N-terminal glutamine spontaneously cyclizes to pyroglutamate to form the mature [6P4]CCL5 (43). Full cyclization of the N-terminus and overall quality of the protein is assayed by either mass spectrometry (MS; Figure 3C) or NMR spectroscopy (Figure 3D).

Equipment and materials

- Temperature-controlled incubator shaker (Infors HT, Multitron)
- 1-L Erlenmeyer flask
- 5-L Erlenmeyer flask
- French pressure cell and press (Spectronic Instruments, FA-031)
- Ultracentrifuge (Beckman Coulter, Optima XE)
- Membrane filter (Merck, MF-Millipore, 0.45-μm pore size)
- Tabletop centrifuge (Eppendorf, 5430 R)
- Centrifugal filter units (Merck, Amicon Ultra-4 and -15)
- Pressure-based stirred concentrator cell (Merck, Amicon)
- Ultrafiltration polyether sulfone-based membrane discs [Merck, Biomax, 10-kDa molecular weight cut-off (MWCO)]
- Dialysis tubes (Spectrum Chemical, Spectra/Por, 12–14 kDa MWCO)
- Sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE) system (Bio-Rad)
- Precast 4–20% gradient SDS-PAGE gels (Bio-Rad)
- Fast protein liquid chromatography (FPLC) system (Cytiva, ÄKTA pure)
- Nickel affinity column (Cytiva, HisTrap HP)

- Preparative HPLC system (Beckman Coulter, System Gold)
- C4 reversed-phase HPLC column 250 x 21.20 mm, 10 micron, 300 Å (Phenomenex)
- Freeze dryer (Thermo Fisher Scientific, Savant)
- Microvolume spectrophotometer (Thermo Fisher Scientific, NanoDrop)
- C4 micro spin column for MS analysis (The Nest Group, MicroSpin)
- Electrospray ionization time-of-flight (ESI-TOF) mass spectrometer (Bruker, microTOF$_{LC}$)
- High-field (≥500 MHz) NMR spectrometer (Bruker, AVANCE)

Reagents

- *E. coli* strain BL21 (DE3)
- Lysogeny broth (LB) medium (Miller)
- *E. coli* minimal medium (47) and ^{15}N-labeled ammonium chloride for production of ^{15}N-labeled chemokine and NMR quality control
- Isopropyl β-D-thiogalactopyranoside (IPTG)
- Guanidine hydrochloride (GuHCl)
- Sodium chloride
- Sodium acetate (NaOAc)
- Tris-(hydroxymethyl)-aminomethane (Tris) hydrochloride
- L-arginine hydrochloride
- Glutathione, reduced (GSH) and oxidized (GSSG)
- Enterokinase [either commercial (New England biolabs) or produced in-house - an optimized protocol is available upon request]
- Trifluoroacetic acid (TFA)
- Acetonitrile (ACN)
- 2-mercaptoethanol
- Ethylenediaminetetraacetic acid (EDTA)

Buffers

Note: all solutions are prepared using ultrapure water (Milli-Q, resistivity ≥18.2 MΩ/cm). Buffer pH values are indicated for the temperature at which they are used.

- Resuspension buffer: 6 M GuHCl, 200 mM NaCl, 50 mM Tris, pH 8.0

- HisTrap elution buffer: 6 M GuHCl, 200 mM NaCl, 60 mM NaOAc, pH 4.0
- Refolding buffer: 550 mM L-arginine hydrochloride, 200 mM NaCl, 1 mM EDTA, 1 mM GSH, 0.1 mM GSSG, 50 mM Tris, pH 8.0
- Cleavage buffer: 200 mM NaCl, 2 mM $CaCl_2$, 20 mM Tris, pH 8.0
- Reversed-phase buffer A: 0.1% TFA
- Reversed-phase buffer B: 0.085% TFA in 90% ACN
- MS buffer A: 10% ACN, 0.1% formic acid
- MS buffer B: 90% ACN, 0.1% formic acid

DNA construct

A DNA construct of [5P14]CCL5 harboring an N-terminal thioredoxin fusion partner and an enterokinase cleavage site cloned into a pET32a vector (Figure 3A) was a gift from Prof. P. LiWang (UC Merced).

The DNA sequence of the [5P14]CCL5 construct was mutated by Quik-Change™ (Stratagene) polymerase chain reaction to obtain the [6P4]CCL5 construct. Its final amino acid sequence including fusion partners (chemokine sequence in bold) is:

MSDKIIHLTDDSFDTDVLKADGAILVDFWAEWCGPCKMIAPILDEIADEYQGKLTVAKLNIDQNPGTAPKYGIRGIPTLLLFKNGEVAATKVGALSKGQLKEFLDANLAGSGSGHMHHHHHHSSGLVPRGSGMKETAAAKFERQHMDSPDLGTDDDDK**QGPPGDIVLACCFAYIARPLPRAHIKEYFYTSGKCSNPAVVFVTRKNRQVCANPEKKWVREYINSLEMS**

In our hands, this construct worked best for the production of various chemokine analogs. An alternative construct containing an N-terminal protein G followed by an enterokinase cleavage site worked well for the production of CCL5-E66S and [5P12]CCL5 (13,48). However, cleavage of the fusion partner from [6P4]CCL5 and [5P14]CCL5 (data not shown) was inefficient in the latter, presumably due to the burial of the cleavage site by the differing N-termini.

Expression

- Transform *E. coli* BL21(DE3) strain with the [6P4]CCL5-encoding plasmid and plate it on an LB agar plate containing 100 µg/mL ampicillin. Incubate the plate overnight at 37 °C.

- Inoculate a single colony from the LB agar plate in 100–120 mL LB broth containing 100 µl/mL ampicillin and grow overnight at 37 °C at a shaking speed of 150 rpm. Note that for producing ^{15}N-labeled protein, bacteria are grown in minimal medium supplemented with ^{15}N-labeled ammonium chloride.
- Use 50–60 mL of primary inoculum to inoculate 1.5 L LB broth containing 100 µL/mL ampicillin in a 5-L Erlenmeyer flask [initial optical density at 600 nm (OD_{600}) □ 0.1] at 37 °C and 150 rpm.
- Lower the temperature to 22 °C once OD_{600} reaches 0.5–0.6 (this typically takes ~2–2.5 h).
- When OD_{600} reaches 0.8, induce the protein expression with 1.0 mM IPTG and grow the cells for another 20 hours.
- Harvest the cells by centrifugation at 3,000 *g* for 25 min. Discard the supernatant and store the pellet at -80 °C or proceed directly to purification.

Purification and cyclization of chemokine N-terminus

All purification steps are performed at 4 °C or on ice, unless indicated otherwise.

- Resuspend the cell pellet (□10 g) in 100 mL resuspension buffer and apply the sample to a French press. Repeat this step twice. Note that the 6 M GuHCl of the resuspension buffer denatures the chemokine.
- Centrifuge the resulting lysate at 27,000 *g* for 1 h and collect the supernatant.
- Add another 100 mL of resuspension buffer and filter the suspension through a 0.45-µm filter. Apply the filtrate on the FPLC system to an equilibrated, pre-packed 5-mL Ni affinity column at a flow rate of 0.8 mL/min.
- Wash the column with 10 column volumes (CV) of resuspension buffer at a flow rate of 1 mL/min, then elute the protein with 3–4 CV of HisTrap elution buffer at a flow rate of 0.8 mL/min.
- Add 10 mM of 2-mercaptoethanol to the eluate and mix with end-over-end rotation for 1 hour at room temperature. This step reduces the two disulfide bonds of the chemokine.
- Add the reduced eluate dropwise to a 10–15-fold volume excess of refolding buffer at 4 ˚C and incubate overnight under gentle stirring. During this step, the

disulfide bonds are reformed under the control of the GSH/GSSG redox system in the buffer and the chemokine is refolded.

- Concentrate the refolded protein to 20–30 mL using a pressure-based concentrator with a polyether sulfone-based ultrafiltration membrane of 10-kDa MWCO.
- Dialyze the concentrated solution against the cleavage buffer in dialysis membrane tubes with a 12–14 kDa MWCO. The volume of the dialysis buffer should be at least 50-fold larger than the sample volume. Change the dialysis buffer twice. Remove any precipitation by centrifugation for 10 min at 25,000 *g*.
- Determine the protein concentration by measuring the absorbance at 280 nm (A_{280}), using the microvolume photometer, then dilute the folded protein to 1.5–2.5 mg/mL.
- Prior to large-scale cleavage of the N-terminal fusion, perform test cleavages with differing amounts of enterokinase:
    - Prepare four 100-µL aliquots of the fusion protein at 2.0 mg/mL.
    - Add enterokinase to the fusion protein aliquots at ratios 1:100, 1:50, 1:25 and 1:10 (w/w) and incubate at room temperature for 16 hours with end-over-end mixing.
    - Analyze the cleavage efficiency by SDS-PAGE to determine the optimal protease:protein ratio, i.e., the minimal amount of enterokinase needed to obtain close to 100% cleavage.
- For large-scale cleavage, add enterokinase to the fusion protein at the determined optimal ratio and incubate at room temperature for 16 hours with end-over-end rotation. Remove trace amounts of precipitate by centrifuging for 10 min at 25,000 *g*. Verify protein cleavage by SDS-PAGE.
- For the final HPLC purification step, adjust the pH of the protein solution to 2.2 using 10% TFA, and then add ACN to a final concentration of 10%.
- Apply the protein to a C4 reversed-phase chromatography column (250 x 21.20 mm, 10 micron, 300 Å, 88 mL, Jupiter) connected to a preparative HPLC system and equilibrated with reversed-phase buffer A. Separate the protein from the fusion tag using the following water/acetonitrile gradient program at a flow rate of 10 mL/min:

  - 0.1 CV: 10% reversed-phase buffer B
  - 0.2 CV: 10% → 20% reversed-phase buffer B
  - 5 CV: 20% → 70% reversed-phase buffer B
  - 0.2 CV: 70% → 100% reversed-phase buffer B
  - 2 CV: 100% reversed-phase buffer B
- Collect the separated protein peaks into different fractions. [6P4]CCL5 elutes at 35–45% of reversed-phase buffer B. Analyze the purity of the eluted fractions by SDS-PAGE.
- Lyophilize the pure chemokine fractions in a freeze dryer. This typically takes approximately 24 hours at a pressure of ~100 µbar.
- Solubilize the lyophilizate in 25 mM phosphate buffer (pH 6) to a concentration of 1.0–1.5 mg/mL.
- Perform the cyclization reaction by incubating the sample in a table shaker at 800 rpm and 37 °C for 48 hours.

Quality control

Mass spectrometry

The overall quality of [6P4]CCL5, disulfide bond formation as well as efficiency of N-terminal glutamine cyclization may be assessed by ESI-TOF MS involving the following steps:

- Wash a C4 MicroSpin column with 100 µL pure ACN for 1 min at 500 *g* in a tabletop centrifuge.
- Add 100 µL H_2O and wash the column by centrifuging for 1 min at 500 *g*. Repeat this step twice.
- Apply sample (2–200 µL) to the column, e.g., 50 µL 0.2 mg/mL protein in MS buffer A for 1 min at 500 *g*.
- Wash the column twice with 50 µL of MS buffer A for 1 min at 500 *g*.
- Elute with 50 µL of MS buffer B for 2 min at 500 *g*.
- Apply eluate to ESI-TOF mass spectrometer.

Figure 3C shows representative examples of the MS characterization of [6P4]CCL5 before (top) and after N-terminal cyclization (bottom). The corresponding masses confirm the protein integrity and completeness of the cyclization step.

NMR spectroscopy

Alternatively, the quality of the produced chemokine and the N-terminal cyclization can be assayed at high resolution by NMR spectroscopy. For this, the chemokine analog is expressed in minimal medium containing ^{15}N-labeled ammonium chloride as the sole nitrogen source. A typical NMR sample is then prepared as a 270-µL volume of 100–200 µM ^{15}N-labeled chemokine in suitable buffer and placed in a Shigemi microtube, followed by measurement of a ^{1}H-^{15}N HSQC spectrum.

An example of a [6P4]CCL5 ^{1}H-^{15}N HSQC is shown in Figure 4D. [6P4]CCL5 exhibits well-dispersed ^{1}H-^{15}N backbone amide resonances characteristic of a folded protein containing β-sheets. The ^{1}H-^{15}N resonances are labeled by the respective assigned amino acids. Cyclization of Q0 results in a shift of the glycine 1 (G1) resonance from its non-cyclized (n) to its cyclized (c) form. The extent of cyclization can be quantified from the ratio of cyclized to non-cyclized G1 resonance intensities.

Chemokine receptor expression and membrane preparation

A schematic overview of the CCR5 chemokine receptor expression and purification is shown on the left side of Figure 2. CCR5 is expressed in baculovirus-infected insect cells (49) using the Invitrogen Bac-to-Bac system for the generation of the recombinant virus and Sf9 cells for expression. A step-by-step protocol for the virus production and insect cell expression can be found in the manufacturer's manual (https://tools.thermofisher.com/content/sfs/manuals/bactobac_man.pdf). The expression is followed by membrane preparation, solubilization, purification and complex formation steps. As the procedures are lengthy, it is most practical to first prepare frozen receptor membranes, and then proceed to the following steps as time permits.

Equipment and materials

- Temperature-controlled incubator shaker (Infors, HT Multitron)
- Bioreactor tubes (TPP, TubeSpin)
- 5-L plastic bottle for insect cells (VWR, 215-1435)
- Breathable rayon film (VWR, 391-1262)
- Automated cell counter (Invitrogen, Countess)
- Cell counting slides (Eve)

- Flow cytometer (Orflo, Moxi Flow)
- Flow cytometer cassettes (Orflo, Moxi Flow Type MF-F)
- SDS-PAGE system (Bio-Rad)
- Precast 4–20% gradient SDS-PAGE gels (Bio-Rad)
- Blotting system for western blot (Bio-Rad, Trans-Blot turbo)
- Western blot transfer packs (Bio-Rad, Trans-Blot turbo mini)
- Electric dispenser (IKA, ULTRA-TURRAX T25)
- Tabletop centrifuge (Eppendorf, 5430 R)
- Ultracentrifuge (Beckman Coulter, Optima XE)

Reagents

- *Spodoptera frugiperda* Sf9 cells (Oxford Expression Technologies)
- Insect cell culture medium (BioConcept, SF-4 baculo express ICM)
- Trypan blue
- Sodium chloride
- Potassium chloride
- Magnesium chloride
- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)
- Glycerol
- EDTA-free protease inhibitor tablets (Roche, cOmplete)
- FuGENE transfection reagent (Promega)
- Phycoerythrin (PE)-conjugated 2D7 CCR5-specific conformation-dependent antibody (BD Biosciences, PE mouse anti-human CD195, clone 2D7/CCR5)
- Monoclonal anti-DYKDDDK (anti-FLAG)-peroxidase antibody (Sigma-Aldrich, ANTI-FLAG M2-Peroxidase)
- PE-conjugated anti-baculovirus envelope gp64 antibody (Invitrogen)

Buffers

- Lysis buffer: 10 mM HEPES, 10 mM $MgCl_2$, 20 mM KCl, pH 7.5
- High-salt buffer: 10 mM HEPES, 10 mM $MgCl_2$, 20 mM KCl, 1 M NaCl, pH 7.5
- Freezing buffer: 10 mM HEPES, 10 mM $MgCl_2$, 20 mM KCl, 30% glycerol, pH 7.5
- Phosphate-buffered saline pH 7.4 (PBS; Gibco)

DNA construct

To allow for the detection by western blot and purification by an M2 anti-FLAG resin, the full-length wild-type CCR5 sequence was extended by a C-terminal PreScission protease cleavage site and a FLAG tag. This construct was cloned into a pFastBac1 vector (Invitrogen), which is under the control of the polH promoter. The full protein sequence including the C-terminal extension (CCR5 in bold) is:

MDYQVSSPIYDINYYTSEPCQKINVKQIAARLLPPLYSLVFIFGFVGNMLVILILINCKRLKSMTDIYLLNLAISDLFFLLTVPFWAHYAAAQWDFGNTMCQLLTGLYFIGFFSGIFFIILLTIDRYLAVVHAVFALKARTVTFGVVTSVITWVVAVFASLPGIIFTRSQKEGLHYTCSSHFPYSQYQFWKNFQTLKIVILGLVLPLLVMVICYSGILKTLLRCRNEKKRHRAVRLIFTIMIVYFLFWAPYNIVLLLNTFQEFFGLNNCSSSNRLDQAMQVTETLGMTHCCINPIIYAFVGEKFRNYLLVFFQKHIAKRFCKCCSIFQQEAPERASSVYTRSTGEQEISVGLGVAGLEVLFQGPDYKDDDDK

Expression and membrane preparation

- Produce recombinant P2 baculovirus stock according to the Bac-to-Bac protocol and store the virus at 4 °C. Use relatively fresh P2 virus (not older than 1 month) for all following steps.

Monitoring of insect cell parameters

Many of the following steps include monitoring of the insect cell parameters size, viability and cell density. The procedure is as follows:

- Take an aliquot of 10 µL cell culture and stain with 10 µL of trypan blue.
- Apply 10 µL of the stained cells to the cell counting slide of the automated cell counter and determine the cell parameters.

Determination of optimal virus titer

The following describes small-scale test expressions at several virus dilutions to determine the optimal P2 virus titer for protein expression.

- Test the cell culture for the cell parameters.
- Prepare six 50-mL bioreactor test tubes containing each 20 mL of cell culture at a cell density of 2.5–3.0 10^6 cells/mL.

- Add virus to the test tubes to reach final concentrations of 0, 2, 5, 10, 20 and 40 mL virus solution per liter cell culture.
- At 6 hours post infection (hpi), take 100 µL from each test tube and determine the fraction of infected cells by flow cytometry as follows
    - Centrifuge the 100 µL in a tabletop centrifuge at 1500 *g* for 1 min.
    - Discard the supernatant, add 100 µL PBS and centrifuge again for 1 min (repeat this step twice).
    - Resuspend the cell pellet in 100 µL PBS and add 0.5 µL of 0.2 µg/µL PE-conjugated anti-baculovirus envelope gp64 antibody. Incubate for 15 min in the dark.
    - Centrifuge in a tabletop centrifuge at 1500 *g* for 1 min, discard the supernatant and resuspend the pellet by adding 100 µL PBS. Repeat this step twice.
    - Inject 75 µL into the fill port of the flow cytometer cassette. Adjust the gating of the flow cytometer to separate infected, non-infected and dead cells.
- The optimal virus titer is the minimal amount of P2 virus at which at least 90% of the cells are infected. In our hands, the optimal P2 titer was 10 mL virus per liter cell culture.
- Continue maintaining the remainder of the test culture, which has the optimal P2 virus titer, to determine the optimal harvest time in the subsequent steps.

Determination of optimal harvest time

To determine the optimal harvest time, monitor the continued culture by measuring cell parameters and taking samples for western blot at 24, 42, 44, 46 and 48 hpi.

- The cell size normally increases by 20-25% at 24 hpi, while the viability remains at about 98%. At 40–48 hpi, the viability usually drops to 85–90%.
- At 24 hpi and at about 40–46 hpi, i.e. ~85-90% viability, test the appearance level of correctly folded CCR5 at the insect surface by flow cytometry using the PE-conjugated 2D7 conformation-dependent CCR5 antibody. Include a negative control (e.g., uninfected cells) together with the samples.

- To determine the CCR5 expression levels during the time course, analyze the previously taken samples on SDS-PAGE followed by western blot with the anti-FLAG-peroxidase antibody.

The harvesting time is optimal when the viability drops to about 85% and the expression according to the western blot reaches the highest level. In our case this occurred at 46 hpi with more than 95% of the cells showing correctly folded CCR5 at their surface as indicated by the 2D7 antibody. We noted that longer expression times lead to unwanted receptor oligomerization.

Large-scale expression

- Inoculate 3 L of *Sf9* insect cells in a 5-L culture bottle with orbital shaking at 27 °C and 130 rpm (25 cm orbit).
- Infect the cells at a cell density of 2.5–3.0 million cells/mL by adding P2 virus at the determined optimal virus titer.
- At 42–46 hpi, confirm that more than 90% of the infected cells are expressing properly folded CCR5 by flow cytometry using the PE-conjugated 2D7 CCR5-specific conformation-dependent antibody.
- Measure the cell viability and cell concentration at the determined optimal harvesting time (the viability should be about 85–90%). Harvest the cells by centrifuging at 1500 *g* for 15 min.
- Freeze the cell pellets in liquid nitrogen and store at -80 °C, or continue directly to membrane preparation.

Membrane preparation

- Resuspend the cell pellet from 1 L of Sf-9 cells in 15 mL of lysis buffer.
- Add a protease inhibitor cocktail tablet and homogenize with an electric dispenser.
- Transfer the suspension to an ultracentrifuge tube. Centrifuge at 120,000 *g* for 1 hour, discard the supernatant and resuspend in ~15 mL high-salt buffer. Repeat the process three times with high-salt buffer and then once with lysis buffer.

- Resuspend the pellet in freezing buffer supplemented by a protease inhibitor cocktail tablet to a final volume of 15 mL. Freeze the suspension in liquid nitrogen and store at -80 ˚C until further use.

Chemokine receptor purification and complex formation

The following describes the purification of CCR5 from 1 L of Sf9 cell culture and the subsequent formation of the receptor complex. In our hands, this protocol yields 0.2 mg of pure receptor per liter of cell culture. For more extensive screening of grid freezing conditions (~27 grids, see below), the preparation is scaled up to processing at least 3 L cell culture. All purification steps are performed on ice or at 4 °C.

Equipment and materials

- Analytical protein liquid chromatography system equipped with autosampler, fraction collector, UV/VIS absorbance and fluorescence detectors (Thermo Fisher Scientific, Ultimate 3000)
- Gel filtration column (Cytiva, Superdex 200 Increase 10/300)
- SDS-PAGE system (Bio-Rad)
- Precast 4–20% gradient SDS-PAGE gels (Bio-Rad)
- Dounce homogenizer (Kontes)
- Ultracentrifuge (Beckman Coulter, Optima XE)
- Tabletop centrifuge (Eppendorf, 5430 R)
- Centrifugal filter unit with ultracel-100 membrane (Amicon Ultra-0.5 and Ultra-4)
- Microvolume spectrophotometer (Thermo Fisher Scientific, NanoDrop)

Reagents

- Iodoacetamide
- Sodium chloride
- Potassium chloride
- Magnesium chloride
- 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES)
- Glycerol
- Dithiothreitol
- Lauryl maltose neopentyl glycol (LMNG)
- Adenosine 5'-triphosphate (ATP)

- Guanosine 5′-*O*-(3′-thiotriphosphate) (GTPγS)
- EDTA-free protease inhibitor tablets (Roche cOmplete)
- Anti-FLAG M2 resin (Sigma-Aldrich)
- Apyrase (NEB)
- Purified human Gαi_1 subunit from *E. coli* expression (44) in Gα_i buffer
- Purified bovine transducin $\beta_1\gamma_1$ subunit (Gβγt) from bovine retina (45) in Gβγ buffer
- Purified Fab16 from hybridoma cell culture secreting IgG16 (46) in Fab16 buffer

Buffers

- Lysis buffer: 10 mM HEPES, 10 mM $MgCl_2$, 20 mM KCl, pH 7.5
- 2x solubilization buffer: 100 mM HEPES, 800 mM NaCl, 1% LMNG (prepared from 5% (w/v) stock), pH 7.5
- FLAG binding buffer: 25 mM HEPES, 400 mM NaCl, 10% glycerol, 0.01% LMNG, pH 7.5
- FLAG wash buffer: 25 mM HEPES, 400 mM NaCl, 10% glycerol, 0.01% LMNG, 5 mM ATP, 10 mM $MgCl_2$, pH 7.5
- FLAG elution buffer: 25 mM HEPES, 400 mM NaCl, 10% glycerol, 0.01% LMNG, 0.2 mg/mL FLAG peptide, pH 7.5
- Size-exclusion chromatography (SEC) buffer: 25 mM HEPES, 150 mM NaCl, 0.01% LMNG, pH 7.5
- Gα_i buffer: 25 mM HEPES, 2 mM dithiothreitol, 150 mM NaCl, 20% glycerol, pH 7.5
- Gβγ buffer: 10 mM HEPES, 4 mM $MgCl_2$, 2 mM dithiothreitol, 150 mM NaCl, 30% glycerol, pH 7.5
- Fab16 buffer: PBS, 1.5 mM NaN_3, 10% glycerol, pH 7.4

Receptor purification

- Thaw the membranes from 1 L of cell culture and add lysis buffer to a final volume of 25 mL, together with 1 tablet of EDTA-free protein inhibitor cocktail and 2 mg/mL iodoacetamide. Dounce homogenize with 20 strokes and incubate for 1 hour on a rotisserie.

- Add 25 mL of 2x solubilization buffer to the homogenized sample and incubate for 3 hours under gentle mixing on the rotisserie. Centrifuge the solubilized sample at 140,000 *g* for 1 hour.
- Transfer the supernatant to a new reaction tube and add 1 mL of equilibrated anti-FLAG M2 affinity resin. Incubate overnight under gentle mixing on the rotisserie.
- Transfer the resin to an open glass column to collect the resin. Wash the resin consecutively with 10 CV FLAG binding buffer, 10 CV FLAG wash buffer, and 6 CV FLAG binding buffer.
- Remove the flow-through and add 4-5 CV FLAG elution buffer to the resin. Incubate for 15 min and collect the eluate.
- Determine the receptor concentration from A_{280} using the microvolume photometer and analyze the receptor by SDS-PAGE to estimate its purity. Adjust the final receptor concentration to 0.2–0.3 mg/mL in FLAG binding buffer.

Receptor complex formation

The following describes the formation of the [6P4]CCL5•CCR5•Gi complex in the absence and presence of the antibody fragment Fab16, which had been shown to stabilize other GPCR/G protein complexes by recognizing an interface between Gα and Gβγ subunits in the heterotrimer, and to confer resistance to GTPγS-triggered dissociation (50). As it was initially unclear whether Fab16 would also stabilize the [6P4]CCL5•CCR5•Gi complex, a respective assay was carried out (see below), which indicated significant stabilization also for this complex. In consequence, we proceeded to the full cryo-EM analysis only with the [6P4]CCL5•CCR5•Gi•Fab16 complex.

- Form the G_i protein heterotrimer by incubating 50 µL 10 mg/mL $Gα_i$ in $Gα_i$ buffer with 250 µL 2.5 mg/mL Gβγ in Gβγ buffer for 30 min on ice.
- Add the formed G_i protein heterotrimer to 4 mL 0.2 mg/mL (~2–3 µM monomeric) CCR5 eluted from the FLAG resin and incubate for 30 min.
- Add 90 µL 1 mg/mL cyclized [6P4]CCL5 to the preformed complex to reach a 1:1 stoichiometric ratio with CCR5, i.e. a final concentration of ~2–3 µM. Supplement the mixture with 25 mU/mL apyrase which hydrolyzes nucleotides in order to

obtain the nucleotide-free complex. Incubate the complex for 2 hours under gentle rotation. Keep 500 µL of the formed [6P4]CCL5•CCR5•G_i complex for further use.

- Add 800 µL 0.8 mg/mL Fab16 to the remaining complex solution, and incubate for 1 hour under gentle rotation.
- Concentrate the [6P4]CCL5•CCR5•G_i•Fab16 complex to 400–600 µL using a concentrator with 100 kDa MWCO. Spin the sample at 20,000 *g* for 10 min to remove any precipitate.
- Estimate the concentrations of both complexes from A_{280} using the microvolume photometer and take aliquots for SDS-PAGE analysis.

Receptor complex purification

- Connect the Superdex 200 Increase 10/300 gel filtration column to a liquid chromatography (HPLC or FPLC) system equipped with a UV absorbance detector, autosampler and fraction collector. Equilibrate the column with 2 CV of SEC buffer at a flow rate of 0.5 mL/min.
- Perform a SEC program, which sequentially injects the receptor complex solution to the column, applies 1 CV of SEC buffer at 0.5 mL/min, monitors the absorbance at 280 nm, and collects 150-µL fractions in a 96-well plate fraction collector.
- Determine A_{280} of the fractions with a microvolume spectrophotometer to estimate the receptor complex concentration.
- Analyze the SEC fractions and the individual complex components as standards by SDS-PAGE.
- Combine the fractions showing good purity and complex integrity and concentrate them for further use.

Figures 4A, B (left) show representative SEC profiles of the complexes without and with Fab16, respectively. A significant part of the complex contains oligomers and higher molecular weight impurities. However, these impurities were successfully separated from the main complexes, which elute at 10.9 mL for the complex without Fab16 and at 10.5 mL with Fab16. The SDS-PAGE analyses of both complexes are shown on the right side of Figures 4A, B. The fractions which were taken for the cryo-EM grid preparation are delimited by orange dashed lines.

Quality control of receptor complex by GTPγS binding assay

While the cryo-EM grids should be prepared immediately after complex formation (see below), the following test can be performed to assess the specificity and stability of the formed [6P4]CCL5•CCR5•G_i complex and the effect of Fab16. The test is based on the GTPγS-triggered dissociation of the G_i heterotrimer, which is followed by analytical fluorescence-detection size-exclusion chromatography (FSEC) using the intrinsic tryptophan fluorescence. As standards also the individual complex components are analyzed by FSEC. All steps are performed on ice or at 4 °C

- Prepare the following standards at concentrations of ~0.1–0.3 mg/mL and volumes of 110 µL in SEC buffer for FSEC analysis:
    - $G\alpha_i$
    - Gβγ
    - Fab16
    - G_i heterotrimer formed from $G\alpha_i$, Gβγ mixture incubated for 1 hour
    - G_i heterotrimer•Fab16 complex formed from $G\alpha_i$, Gβγ, Fab16 mixture incubated for 1 hour
- Dilute the SEC-purified [6P4]CCL5•CCR5•G_i and [6P4]CCL5•CCR5•G_i•Fab16 complexes to ~0.1 mg/mL and 220-µL volumes. Split each complex solution into two equal portions, add 100 µM GTPγS (10 mM stock solution) to one part, and incubate for 1 hour.
- Transfer all samples to glass vials and place them in the autosampler connected to the HPLC system with the Superdex 200 Increase 10/300 gel filtration column.
- Set up a program for the SEC run, which sequentially injects 80 µL of the samples to the column, then applies 1 CV of SEC buffer at a flow rate of 0.4 mL/min, and monitors the protein tryptophan fluorescence (λ_{ex} = 280 nm; λ_{em} = 380 nm).

Figure 4C shows representative chromatograms of this stability analysis. The [6P4]CCL5•CCR5•G_i complex (red) dissociates in the presence of GTPγS (dark blue). Similar to other GPCR/G protein complexes (50), the addition of Fab16 (orange) was found to stabilize the CCR5/G_i complex in the presence of GTPγS (light blue).

Cryo-EM grid preparation

The grids for cryo-EM analysis should be prepared immediately from the freshly purified complex. The following gives a short overview of the preparation. More detailed protocols can be found elsewhere (51).

Equipment and materials

- Glow Discharge system for EM grids (Quorum, GloQube)
- EM grids (Quantifoil, R1.2/1.3 200-mesh copper)
- Plunge freezer (Thermo Fisher Scientific, Vitrobot Mark IV)
- Cryogen cup (Thermo Fisher Scientific, Vitrobot)
- Filter paper (Thermo Fisher Scientific, Standard Vitrobot filter paper, 55/20 mm, Grade 595)
- 60-mL syringe
- Tweezers for EM grid handling (e.g., Electron Microscopy Sciences, Dumont #7)
- Cryo grid boxes with lid (Agar Scientific)
- Cryo Dewar
- Liquid nitrogen
- Ethane gas

Cryo-EM grid preparation

- Prepare liquid ethane from ethane gas using the assembled cryogen cup cooled with liquid nitrogen.
- Glow discharge EM grids at 25 mA for 30 sec using the glow discharge system.
- Set the humidity to 100% and the temperature to 4 °C on the Vitrobot plunge freezer.
- Prepare purified complex samples at various concentrations ranging from 1 to 5 mg/mL (Table 1), either by concentrating in a centrifugal filter unit with a 100-kDa MWCO or diluting in SEC buffer.
- Load 3.5 µL of the purified complex at the various prepared concentrations to the carbon-coated side of the glow-discharged EM grid.
- Remove excess liquid on the grid by blotting with filter paper using various blot forces and time periods (Table 1) with triplicates for each condition to optimize ice thickness.

- Plunge the blot-dried grid into cooled liquid ethane (~ -180 °C) for rapid freezing in order to form a thin layer of amorphous ice.
- Transfer prepared grids to liquid nitrogen in a cryo Dewar for storage until use for cryo-EM analysis.

3.1.4 Discussion and Conclusion

The structure determination of chemokine-chemokine receptor complexes has long been challenging due to difficulties in obtaining highly purified chemokine receptors, the intrinsic tendencies of chemokines, receptors and their complexes to aggregate, heterogeneity and low stability of the complexes, and their intrinsic mobility. This chapter provides detailed protocols for the production of the superagonist [6P4]CCL5 chemokine variant and for its cognate receptor CCR5 in wild-type form, as well as for the formation of their complex with the G_i heterotrimer and the stabilizing Fab16 antibody fragment, which led to the successful cryo-EM structure determination of the [6P4]CCL5•CCR5•G_i complex.

In the current procedure, the individual complex components are prepared separately and then combined to form the complex. For proper interaction with CCR5, it is essential to obtain the CCL5 chemokine with a native N-terminus devoid of any further residues from cloning or *E. coli* expression. This is achieved by a DNA construct, which encodes CCL5 with a cleavable N-terminal thioredoxin His-tag fusion. After expression and an unfolding-reduction, reoxidation-refolding purification sequence, the fusion is efficiently cleaved by enterokinase to obtain the blunt-ended CCL5 N-terminus. In the specific case of [6P4]CCL5, the N-terminal glutamine then spontaneously cyclizes to form pyroglutamate. The protocol achieves a robust yield of ~10 mg pure and homogeneous [6P4]CCL5 from 1 liter of bacterial culture. The high quality of the purified chemokine can be assayed by mass spectrometry and NMR analysis. Other chemokines and their analogs may be obtained in a similar manner. Despite a certain tendency of [6P4]CCL5 to aggregate, a sufficient amount of monomeric [6P4]CCL5•CCR5•G_i for the cryo-EM analysis was obtained. However, other chemokines, such as native CCL5, may even be more prone to aggregation, which presents additional challenges.

The aggregation problem is compounded by the tendency of the chemokine receptors themselves to aggregate. In the present case, the expression of the full-length, wild-type CCR5 in insect cells yielded a mixture of CCR5 monomers, dimers and higher oligomers, which in our hands was considerably more heterogeneous than the expression product of a CCR5 construct optimized for crystallization of the inactive [5P7]CCL5•CCR5 complex (31). In this respect, we observed that longer CCR5 expression times in the insect cells lead to stronger oligomerization. Thus, it is very important to carefully optimize and control all insect cell expression parameters such as virus titer, harvesting time, level of receptor expression and receptor structural quality. For the latter, recognition by the structure-specific antibody 2D7 was a very useful indicator of properly folded CCR5 on the surface of the insect cells. In addition, the structural integrity and homogeneity of the receptor preparation depends critically on the appropriate detergent. Detailed protocols for detergent screening for GPCRs have been described elsewhere (52). In the present case, as in many others, LMNG proved best due to its high efficiency of CCR5 membrane extraction, its low critical micelle concentration (CMC), and consequently efficient removal by dialysis and lower background in the cryo-EM images.

Despite the heterogeneity of the CCR5 receptor preparation and of the formed [6P4]CCL5•CCR5•G_i complex solution, a reasonably homogeneous sample was obtained from this after complex formation with the chemokine and the G protein by SEC (Figure 4), which could then be further purified ‘in silico’ during the cryo-EM data analysis. Nevertheless, heterogeneity is clearly limiting all steps, and it was critical to form the complex at low (~2–3 μM) chemokine and receptor concentrations to avoid aggregation as much as possible. This is practical for [6P4]CCL5, which has a low nanomolar affinity for CCR5 (53). However, other chemokines such as wild-type CCL5 (53) often have weaker affinities for their receptors, which makes complex formation at low micromolar concentrations impossible. An alternative approach to reduce aggregation could be the co-expression of G protein and/or chemokine with the receptor. Co-expression with the G protein may increase the affinity of agonist chemokines and has been successful for the cryo-EM structure determination of a number of other active, non-chemokine GPCR•G protein complexes (54–56). Co-expression with the [5P7]CCL5 antagonist chemokine has been used for the X-ray structure determination of the [5P7]CCL5•CCR5 complex (31). Again, however, these approaches depend on a sufficient affinity of the chemokine for the receptor (57).

In summary, with the advances in cryo-EM technology and in the production of chemokine receptors, chemokines and G proteins, the structure determination and in-depth biophysical characterization of many further chemokine•receptor•G protein complexes seems in reach. This is a prerequisite for deciphering the complex interactions of the components of the chemokine signaling system and to move forward in the rational drug design against these important therapeutic targets.

3.1.5 Tables

Table 1. Cryo-EM grid plunge-freezing conditions

Receptor complex concentration [mg/mL]	**Blot force [Vitrobot units]**	**Blot time [s]**
1	20	2
	10	3
	20	3
2.5	20	2
	10	3
	20	3
5	20	2
	10	3
	20	3

3.1.6 Figures

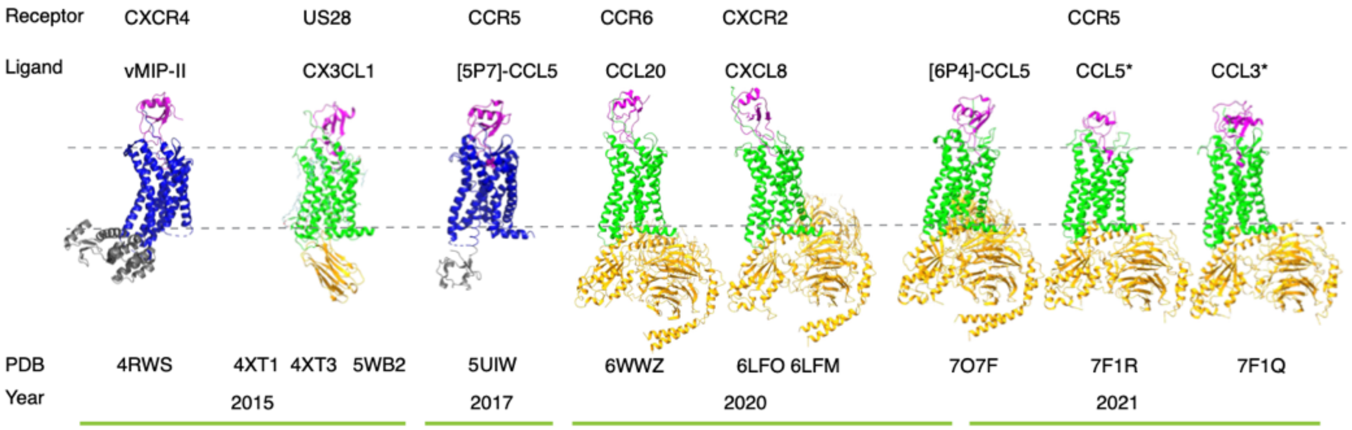


Figure 1. Overview of currently available chemokine•chemokine receptor complex structures. Receptors in their active and inactive conformation are colored in green and blue, respectively. Only CCR5 has been solved in both active, agonist- and inactive, antagonist-bound states. Chemokines are shown in magenta, intracellular binding partners in orange, and stabilizing intra-receptor fusion proteins in grey. CCR5 complexes marked with asterisks were expressed as chemokine-chemokine receptor fusion proteins.

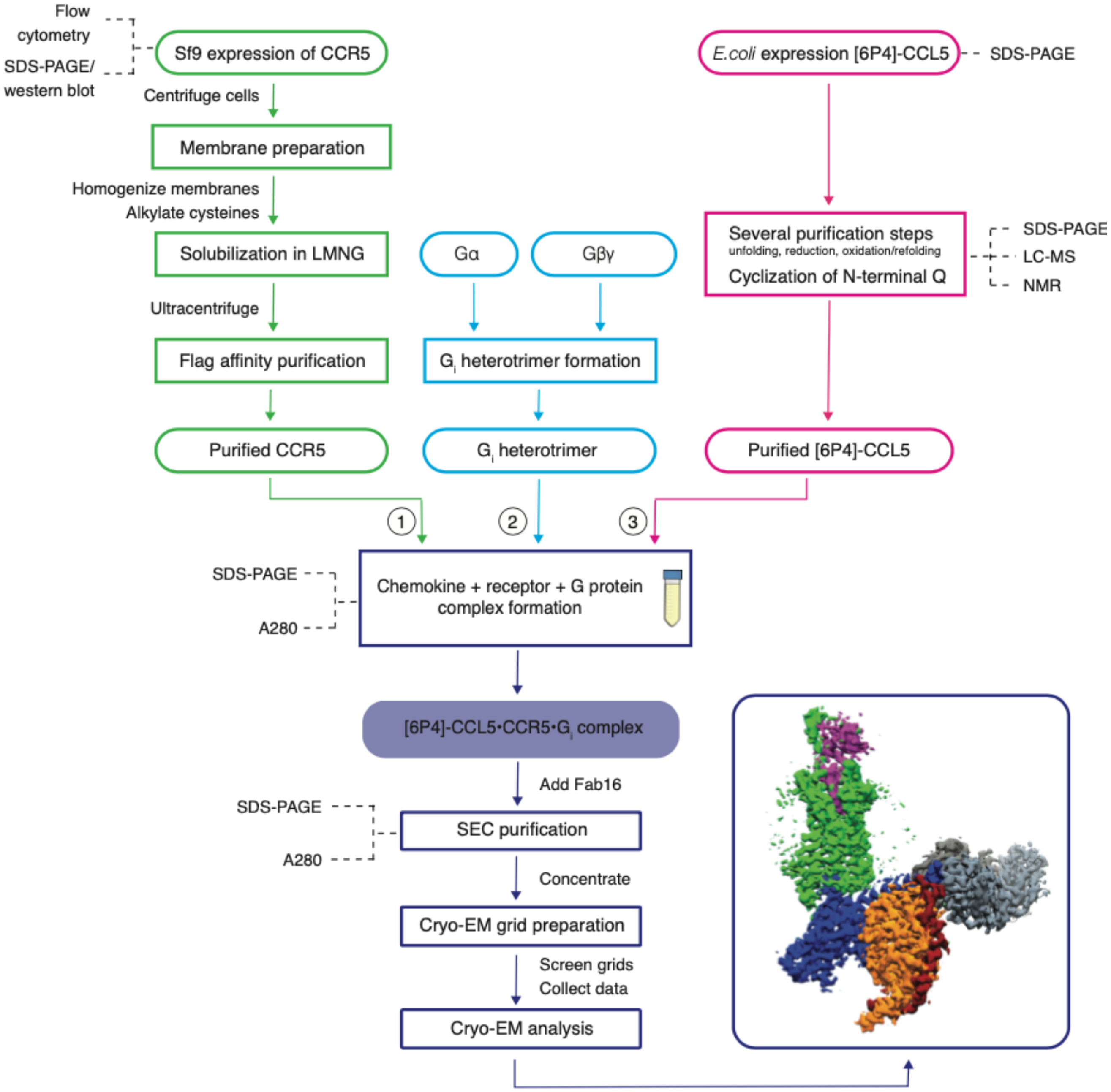


Figure 2. Overall scheme of the [6P4]CCL5•CCR5•Gi•Fab16 complex preparation for cryo-EM studies.

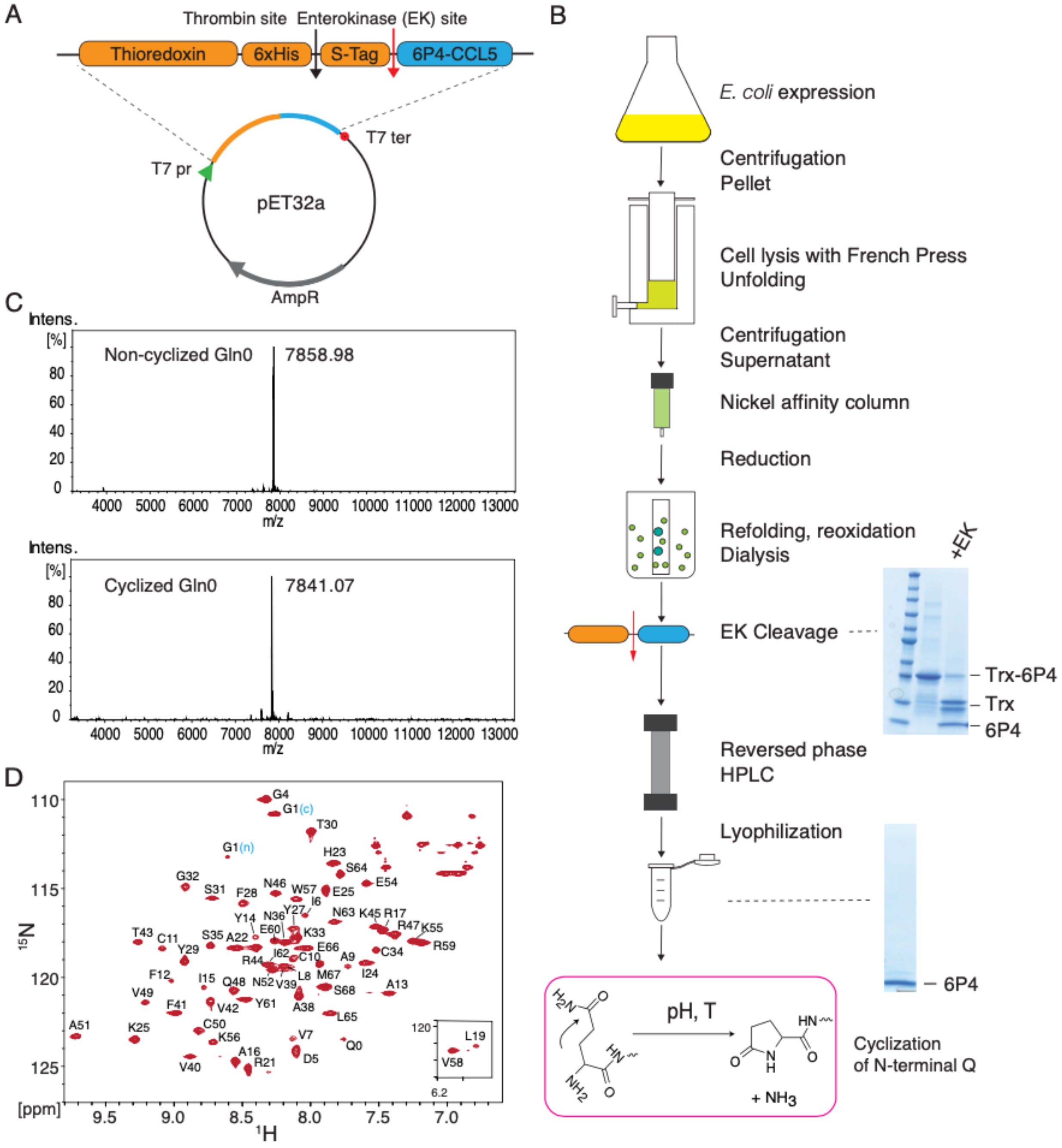


Figure 3. Chemokine preparation and characterization. A. [6P4]CCL5 expression construct. B. Production scheme for [6P4]CCL5 chemokine cyclized at its N-terminal glutamine Q0 to pyroglutamate and controls by SDS-PAGE analysis. C. Mass spectra of purified N-terminally non-cyclized and cyclized [6P4]CCL5. D. 1H-15N HSQC spectrum of [6P4]CCL5 in 25 mM phosphate, 5% D2O, 0.02% NaN3, pH 3.8 obtained with acquisition times of 86 ms (1H) and 56 ms (15N) at 600 MHz and 25 °C. The NMR data were processed with the NMRpipe suite of programs (58) and analyzed using the program NMRFAM-Sparky (59). As Q0 of [6P4]CCL5 was not yet fully cyclized, resonances of both non-cyclized and cyclized molecular species are visible for glycine-1 [G1(n) and G1(c), respectively].

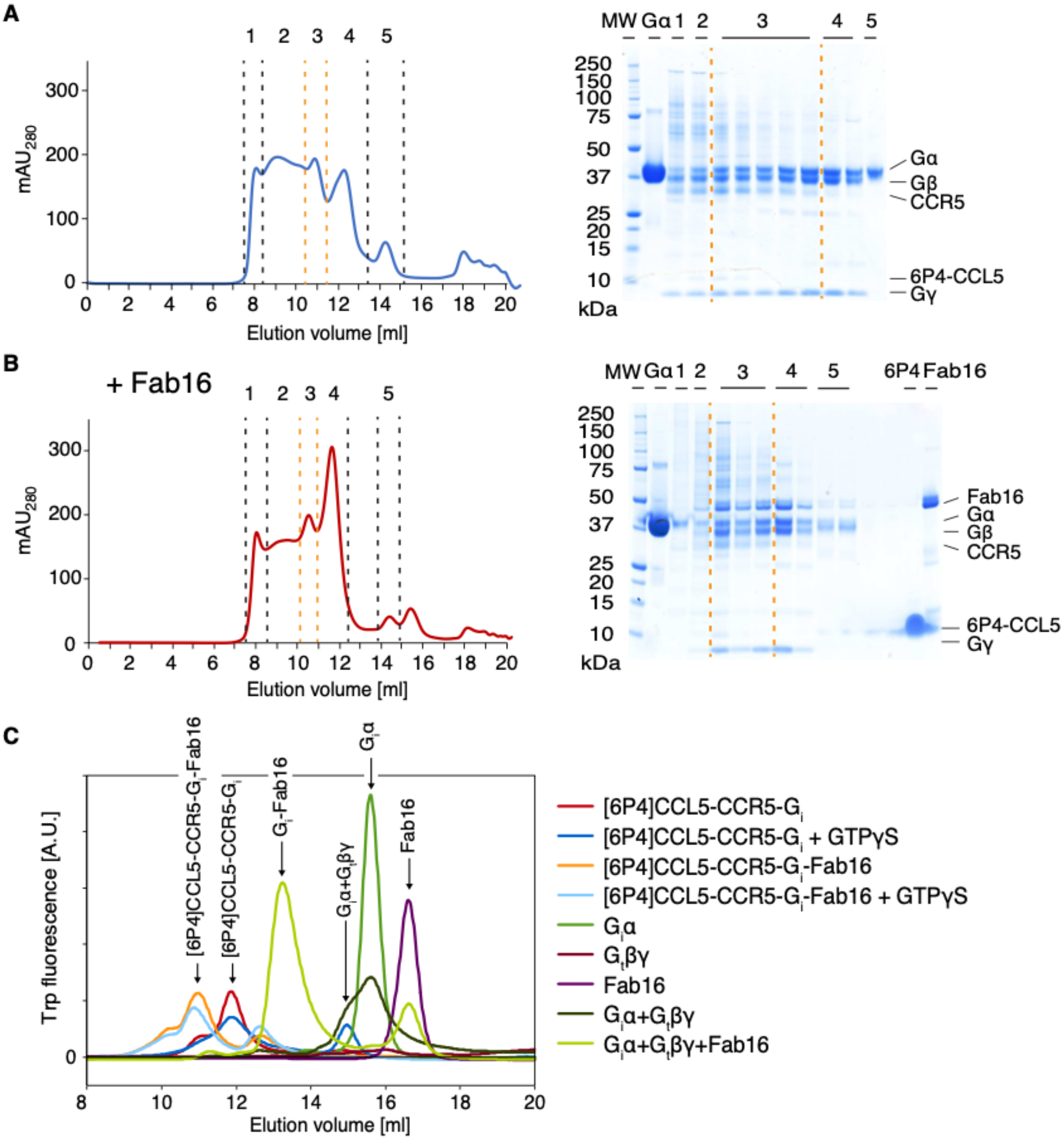


Figure 4. Characterization of the [6P4]CCL5•CCR5•G_i complex. A. Size-exclusion chromatography (SEC) profiles of [6P4]CCL5•CCR5•G_i complex (left) and SDS-PAGE analysis of the SEC fractions (right). The fractions are marked with numbers at the top and are separated by dashed lines in the chromatogram. Orange dashed lines delimit the fractions of the complex. B. Same analysis as A for the [6P4]CCL5•CCR5•G_i•Fab16 complex. C. Superimposed FSEC profiles of the [6P4]CCL5•CCR5•G_i complex with and without added Fab16 in presence of GTPγS. The FSEC profiles of the individual complex components are shown as controls. The [6P4]CCL5•CCR5•G_i complex (red) disintegrates partially in the presence of GTPγS (dark blue). The stability of the complex in the presence of GTPγS (light blue) is increased upon addition of Fab16 (orange), as the antibody fragment constrains the conformational flexibility of the G_i heterotrimer.

3.1.7 Acknowledgments

This work was supported by the Swiss National Science Foundation (Grants 31-201270 and IZLIZ3-200298 to S.G.). We thank Prof. P. LiWang (UC Merced) for the gift of the [5P14]CCL5 DNA construct, J. Mühle (Paul Scherrer Institute, Villigen) for preparation of the Fab16, Philip Pamula (Paul Scherrer Institute, Villigen) for preparation of the Gi protein, and Prof. G. Schertler (Paul Scherrer Institute, Villigen) and Prof. O. Hartley (U. Geneva) for helpful discussions.

3.1.8 Conflict of Interest

The authors declare no conflict of interest.

3.1.9 References

1. Bachelerie F, Ben-Baruch A, Burkhardt AM, Combadiere C, Farber JM, Graham GJ, et al. International Union of Basic and Clinical Pharmacology. LXXXIX. Update on the Extended Family of Chemokine Receptors and Introducing a New Nomenclature for Atypical Chemokine Receptors. Ohlstein EH, editor. Pharmacol Rev. 2014 Jan;66(1):1–79.

2. Scurci I, Martins E, Hartley O. CCR5: Established paradigms and new frontiers for a ‘celebrity’ chemokine receptor. Cytokine. 2018 Sep 1;109:81–93.

3. Brelot A, Chakrabarti LA. CCR5 Revisited: How Mechanisms of HIV Entry Govern AIDS Pathogenesis. Journal of Molecular Biology. 2018 Aug 17;430(17):2557–89.

4. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020 Aug;38(8):970–9.

5. Aldinucci D, Casagrande N. Inhibition of the CCL5/CCR5 Axis against the Progression of Gastric Cancer. International Journal of Molecular Sciences. 2018 May;19(5):1477.

6. Kraus S, Kolman T, Yeung A, Deming D. Chemokine Receptor Antagonists: Role in Oncology. Curr Oncol Rep. 2021 Sep 4;23(11):131.

7. Ortiz Zacarías NV, Bemelmans MP, Handel TM, de Visser KE, Heitman LH. Anticancer opportunities at every stage of chemokine function. Trends in Pharmacological Sciences. 2021 Nov 1;42(11):912–28.

8. Lai WY, Mueller A. Latest update on chemokine receptors as therapeutic targets. Biochemical Society Transactions. 2021 Jun 1;49(3):1385–95.

9. Zhao S, Wu B, Stevens RC. Advancing Chemokine GPCR Structure Based Drug Discovery. Structure. 2019 Mar 5;27(3):405–8.

10. Stone MJ, Hayward JA, Huang C, E. Huma Z, Sanchez J. Mechanisms of Regulation of the Chemokine-Receptor Network. International Journal of Molecular Sciences. 2017 Feb;18(2):342.

11. Miller MC, Mayo KH. Chemokines from a Structural Perspective. International Journal of Molecular Sciences. 2017 Oct;18(10):2088.

12. Handel TM, Dyer DP. Perspectives on the Biological Role of Chemokine:Glycosaminoglycan Interactions. J Histochem Cytochem. 2021 Feb 1;69(2):87–91.

13. Duma L, Häussinger D, Rogowski M, Lusso P, Grzesiek S. Recognition of RANTES by Extracellular Parts of the CCR5 Receptor. Journal of Molecular Biology. 2007 Jan 26;365(4):1063–75.

14. Wang X, Watson C, Sharp JS, Handel TM, Prestegard JH. Oligomeric Structure of the Chemokine CCL5/RANTES from NMR, MS, and SAXS Data. Structure. 2011 Aug 10;19(8):1138–48.

15. Liu K, Wu L, Yuan S, Wu M, Xu Y, Sun Q, et al. Structural basis of CXC chemokine receptor 2 activation and signalling. Nature. 2020 Sep 3;585(7823):135–40.

16. Chen YC, Chen SP, Li JY, Chen PC, Lee YZ, Li KM, et al. Integrative Model to Coordinate the Oligomerization and Aggregation Mechanisms of CCL5. Journal of Molecular Biology. 2020 Feb 14;432(4):1143–57.

17. Fernando H, Chin C, Rösgen J, Rajarathnam K. Dimer Dissociation Is Essential for Interleukin-8 (IL-8) Binding to CXCR1 Receptor*. Journal of Biological Chemistry. 2004 Aug 27;279(35):36175–8.

18. Ravindran A, Sawant KV, Sarmiento J, Navarro J, Rajarathnam K. Chemokine CXCL1 Dimer Is a Potent Agonist for the CXCR2 Receptor *. Journal of Biological Chemistry. 2013 Apr 26;288(17):12244–52.

19. Wedemeyer MJ, Mahn SA, Getschman AE, Crawford KS, Peterson FC, Marchese A, et al. The chemokine X-factor: structure-function analysis of the CXC motif at CXCR4 and ACKR3. J Biol Chem. 2020 Aug 11;jbc.RA120.014244.

20. Kessler N, Akabayov SR, Moseri A, Cohen LS, Sakhapov D, Bolton D, et al. Allovalency observed by transferred NOE: interactions of sulfated tyrosine residues in the N-terminal segment of CCR5 with the CCL5 chemokine. The FEBS Journal [Internet]. 2020 [cited 2021 Jan 13];doi: 10.1111/febs.15503. Online ahead of print.(n/a). Available from: https://febs.onlinelibrary.wiley.com/doi/abs/10.1111/febs.15503

21. Szpakowska M, Fievez V, Arumugan K, van Nuland N, Schmit JC, Chevigné A. Function, diversity and therapeutic potential of the N-terminal domain of human chemokine receptors. Biochemical Pharmacology. 2012 Nov 15;84(10):1366–80.

22. D'Agostino G, García-Cuesta EM, Gomariz RP, Rodríguez-Frade JM, Mellado M. The multilayered complexity of the chemokine receptor system. Biochemical and Biophysical Research Communications. 2020 Jul 23;528(2):347–58.

23. Apel AK, Cheng RKY, Tautermann CS, Brauchle M, Huang CY, Pautsch A, et al. Crystal Structure of CC Chemokine Receptor 2A in Complex with an Orthosteric Antagonist Provides Insights for the Design of Selective Antagonists. Structure. 2019 Mar 5;27(3):427-438.e5.

24. Jaeger K, Bruenle S, Weinert T, Guba W, Muehle J, Miyazaki T, et al. Structural Basis for Allosteric Ligand Recognition in the Human CC Chemokine Receptor 7. Cell. 2019 Aug 22;178(5):1222-1230.e10.

25. Oswald C, Rappas M, Kean J, Doré AS, Errey JC, Bennett K, et al. Intracellular allosteric antagonism of the CCR9 receptor. Nature. 2016 Dec 15;540(7633):462–5.

26. Peng P, Chen H, Zhu Y, Wang Z, Li J, Luo RH, et al. Structure-Based Design of 1-Heteroaryl-1,3-propanediamine Derivatives as a Novel Series of CC-Chemokine Receptor 5 Antagonists. J Med Chem. 2018 Nov 8;61(21):9621–36.

27. Tan Q, Zhu Y, Li J, Chen Z, Han GW, Kufareva I, et al. Structure of the CCR5 Chemokine Receptor-HIV Entry Inhibitor Maraviroc Complex. Science. 2013 Sep 20;341(6152):1387–90.

28. Wu B, Chien EYT, Mol CD, Fenalti G, Liu W, Katritch V, et al. Structures of the CXCR4 Chemokine GPCR with Small-Molecule and Cyclic Peptide Antagonists. Science. 2010 Nov 19;330(6007):1066–71.

29. Zheng Y, Qin L, Ortiz Zacarías NV, de Vries H, Han GW, Gustavsson M, et al. Structure of CC Chemokine Receptor 2 with Orthosteric and Allosteric Antagonists. Nature. 2016 Dec 15;540(7633):458–61.

30. Qin L, Kufareva I, Holden LG, Wang C, Zheng Y, Zhao C, et al. Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. Science (New York, NY). 2015 Mar 6;347(6226):1117–22.

31. Zheng Y, Han GW, Abagyan R, Wu B, Stevens RC, Cherezov V, 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 Jun;46(6):1005-1017.e5.

32. Burg JS, Ingram JR, Venkatakrishnan AJ, Jude KM, Dukkipati A, Feinberg EN, et al. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. Science. 2015 Mar 6;347(6226):1113–7.

33. Miles TF, Spiess K, Jude KM, Tsutsumi N, Burg JS, Ingram JR, et al. Viral GPCR US28 can signal in response to chemokine agonists of nearly unlimited structural degeneracy. eLife. 2018 Jun 8;7:e35850.

34. Isaikina P, Tsai CJ, Dietz N, Pamula F, Grahl A, Goldie KN, et al. Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. Science Advances. 2021 Jun 1;7(25):eabg8685.

35. Wasilko DJ, Johnson ZL, Ammirati M, Che Y, Griffor MC, Han S, et al. Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20. Nat Commun. 2020 Dec;11(1):3031.

36. Zhang H, Chen K, Tan Q, Shao Q, Han S, Zhang C, et al. Structural basis for chemokine recognition and receptor activation of chemokine receptor CCR5. Nat Commun. 2021 Jul 6;12(1):4151.

37. Combadiere C, Ahuja SK, Lee Tiffany H, Murphy PM. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1α, MIP-1β, and RANTES. Journal of Leukocyte Biology. 1996;60(1):147–52.

38. Raport C, Gosling J, Schweickart V, Gray P, Charo I. Molecular cloning and functional characterization of a novel human CC chemokine receptor (CCR5) for RANTES, MIP-1beta, and MIP-1alpha. The Journal of biological chemistry. 1996 Jul 19;271(29):17161–6.

39. Samson M, Labbe O, Mollereau C, Vassart G, Parmentier M. Molecular Cloning and Functional Expression of a New Human CC-Chemokine Receptor Gene. Biochemistry. 1996 Jan 1;35(11):3362–7.

40. Alkhatib G. The biology of CCR5 and CXCR4: Current Opinion in HIV and AIDS. 2009 Mar;4(2):96–103.

41. Martin-Blondel G, Brassat D, Bauer J, Lassmann H, Liblau RS. CCR5 blockade for neuroinflammatory diseases — beyond control of HIV. Nat Rev Neurol. 2016 Feb;12(2):95–105.

42. Aldinucci D, Borghese C, Casagrande N. The CCL5/CCR5 Axis in Cancer Progression. Cancers. 2020 Jul 2;12(7):1765.

43. Gaertner H, Cerini F, Escola JM, Kuenzi G, Melotti A, Offord R, et al. Highly potent, fully recombinant anti-HIV chemokines: Reengineering a low-cost microbicide. Proceedings Of The National Academy Of Sciences Of The United States Of America. 2008 Nov 17;105(46):17706–11.

44. Sun D, Flock T, Deupi X, Maeda S, Matkovic M, Mendieta S, et al. Probing Gαi1 protein activation at single–amino acid resolution. Nat Struct Mol Biol. 2015 Sep;22(9):686–94.

45. Maeda S, Sun D, Singhal A, Foggetta M, Schmid G, Standfuss J, et al. Crystallization Scale Preparation of a Stable GPCR Signaling Complex between Constitutively Active Rhodopsin and G-Protein. PLOS ONE. 2014 Jun 30;9(6):e98714.

46. Tsai CJ, Marino J, Adaixo R, Pamula F, Muehle J, Maeda S, et al. Cryo-EM structure of the rhodopsin-Gαi-βγ complex reveals binding of the rhodopsin C-terminal tail to the gβ subunit. eLife. 2019 Jun 28;8:e46041.

47. Cai M, Huang Y, Sakaguchi K, Clore GM, Gronenborn AM, Craigie R. An efficient and cost-effective isotope labeling protocol for proteins expressed in Escherichia coli. J Biomol NMR. 1998 Jan;11(1):97–102.

48. Wiktor M, Hartley O, Grzesiek S. Characterization of Structure, Dynamics, and Detergent Interactions of the Anti-HIV Chemokine Variant 5P12-RANTES. Biophysical Journal. 2013 Dec;105(11):2586–97.

49. Smith GE, Summers MD, Fraser MJ. Production of human beta interferon in insect cells infected with a baculovirus expression vector. Molecular and cellular biology. 1983 Dec;3(12):2156–65.

50. Maeda S, Koehl A, Matile H, Hu H, Hilger D, Schertler GFX, et al. Development of an antibody fragment that stabilizes GPCR/G-protein complexes. Nat Commun [Internet]. 2018 Sep 13 [cited 2020 Jul 16];9. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6137068/

51. Goldie KN, Abeyrathne P, Kebbel F, Chami M, Ringler P, Stahlberg H. Cryo-electron Microscopy of Membrane Proteins. In: Kuo J, editor. Electron Microscopy: Methods and Protocols [Internet]. Totowa, NJ: Humana Press; 2014 [cited 2021 Sep 8]. p. 325–41. (Methods in Molecular Biology). Available from: https://doi.org/10.1007/978-1-62703-776-1_15

52. Pamula F, Mühle J, Blanc A, Nehmé R, Edwards PC, Tate CG, et al. Strategic Screening and Characterization of the Visual GPCR-mini-G Protein Signaling Complex for Successful Crystallization. JoVE (Journal of Visualized Experiments). 2020 Mar 16;(157):e60747.

53. Rico CA, Berchiche YA, Horioka M, Peeler JC, Lorenzen E, Tian H, et al. High-Affinity Binding of Chemokine Analogs that Display Ligand Bias at the HIV-1 Coreceptor CCR5. Biophysical Journal. 2019 Sep;117(5):903–19.

54. Qiao A, Han S, Li X, Li Z, Zhao P, Dai A, et al. Structural basis of G $_s$ and G $_i$ recognition by the human glucagon receptor. Science. 2020 Mar 20;367(6484):1346–52.

55. Xu P, Huang S, Zhang H, Mao C, Zhou XE, Cheng X, et al. Structural insights into the lipid and ligand regulation of serotonin receptors. Nature. 2021 Apr 15;592(7854):469–73.

56. Zhuang Y, Xu P, Mao C, Wang L, Krumm B, Zhou XE, et al. Structural insights into the human D1 and D2 dopamine receptor signaling complexes. Cell. 2021 Feb 18;184(4):931-942.e18.

57. Gustavsson M, Zheng Y, Handel TM. Production of chemokine/chemokine receptor complexes for structural and biophysical studies. Methods Enzymol. 2016;570:233–60.

58. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995 Nov 1;6(3):277–93.

59. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2015 Apr 15;31(8):1325–7.

3.2 Appendix: NMR analysis of CCL5 analogs

3.2.1 Background

Wild-type CCL5 tends to form higher oligomers (216) via two distinct basic dimer building blocks [Figure 3.1, (247)]. One dimer is constituted via an intermolecular antiparallel β-sheet of the N-terminal amino acids comprising mostly hydrophobic interactions. The second dimer is formed by predominantly electrostatic interactions of the C-terminal part with the core of the structure, in particular a salt bridge between residues E66 and K25 (247). As such, one monomer from each dimer is accessible for further interactions leading to higher oligomerization.

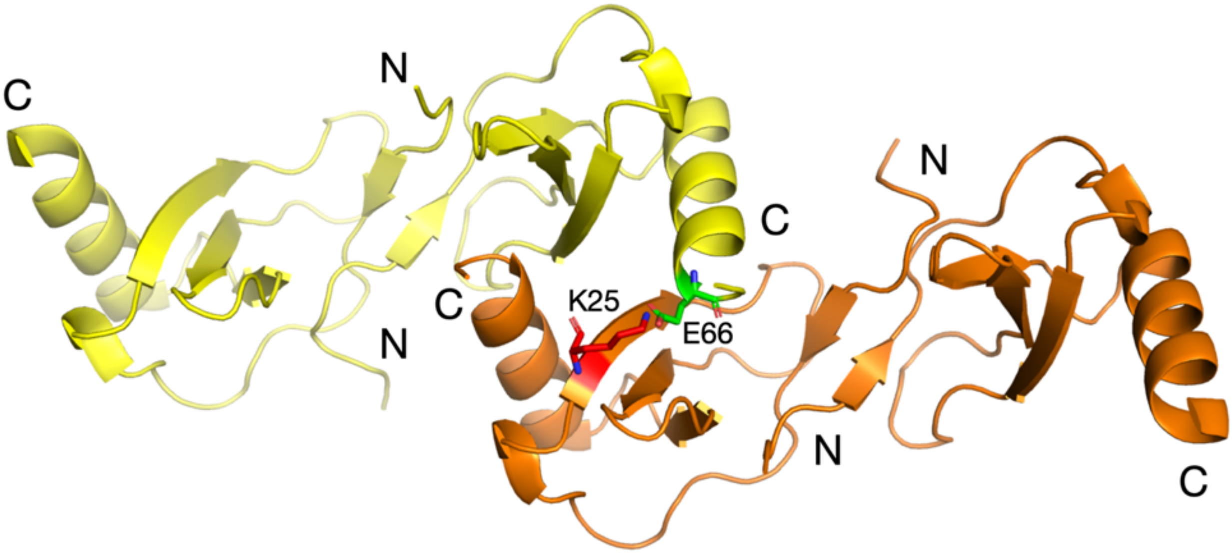


Figure 3.1. Structure of CCL5 tetramer (PDB: 2L9H) (247). The two N-terminal dimers are depicted in yellow and orange. A further C-terminal dimer is formed by electrostatic interactions involving a salt bridge between residues E66 (green) and K25 (red) of the two N-terminal dimers, lending itself to higher oligomerization.

The mutation of residue glutamate-66 to serine (E66S) prevents the formation of the intermolecular salt bridge and reduces the oligomerization tendency (248). Therefore, the E66S mutant of CCL5 was used to study its interactions with a CCR5 N-terminal peptide (216), and its monomer-dimer equilibrium has been characterized in detail by NMR titrations. In a further publication (249), the same mutation was introduced into the chemokine analog [5P12]CCL5 (Table 3.1) to prevent higher oligomerization. It was observed that unlike CCL5-E66S (216), [5P12]CCL5-E66S did not form the N-terminal dimer (249).

However, most biological studies of CCL5 and its analogs have been carried out without the E66S mutation (242,250–255). Furthermore, many engineered N-terminal CCL5 mutants without E66S mutation are not prone to higher oligomer formation, particularly at low pH values (~pH 4), which prevents C-terminal electrostatic dimer formation since the carboxylates of glutamic and aspartic acids are neutralized (256). One of these analogs, [P2]CCL5 (Table 3.1), has been described as a monomer in solution at pH of 5.5 (257).

For these reasons, ^1H-^{15}N HSQC spectra of [5P12]CCL5-E66S and [5P12]CCL5 were compared to probe their structural differences and dimerization tendencies. In addition, an NMR analysis of two further CCL5 analogs, [5P12]CCL5 (antagonist) and [5P14]CCL5 (partial agonist), was carried out to characterize any structural differences and thereby possibly explain their distinct pharmacology. Furthermore, the stability of [5P14]CCL5 at different pH values was probed by NMR.

Table 3.1. The N-terminal sequences of CCL5 analogs.

Name	N-terminal sequence	Pharmacology
CCL5	`.SPYSSDTTPCC-`	agonist
[6P4]CCL5	`QGPPGDIVLACC-`	super-agonist
[5P14]CCL5	`QGPPLMSLQVCC-`	partial agonist
[5P12]CCL5	`QGPPLMATQSCC-`	antagonist
[P2]CCL5	`FSPLSSQSSACC-`	agonist

3.2.2 Results

According to a previously published protocol (249), [5P12]CCL5-E66S was expressed in *E. coli* BL21(DE3) as a fusion protein with an N-terminal protein G followed by an enterokinase cleavage site for a blunt cut. However, using the same construct for other N-terminal CCL5 variants, such as [5P14]CCL5, resulted in inefficient cleavage of the fusion partner, presumably due to reduced accessibility of the cleavage site by the variation of the N-terminus. Therefore, the construct was changed to an N-terminal thioredoxin fusion containing a histidine tag and an enterokinase cleavage site. Using this construct, both [5P12] and [5P14]CCL5 could be expressed and purified from *E. coli* BL21(DE3) with high yield. A detailed protocol for this production was given in the preceding section of this Chapter (258).

Characterization of [5P12]CCL5

Both ^{15}N-labeled [5P12]CCL5 and [5P12]CCL5-E66S show almost identical, well-dispersed ^{1}H-^{15}N HSQC spectra (Figure 3.2). At the given chemokine concentration of 100 µM and pH of 3.8, only a single population corresponding to the monomeric form (249) is present in the spectra. Significant chemical shift differences of the ^{1}H-^{15}N resonances are observed only in the immediate vicinity of the mutation site (Figure 3.3). Thus, the E66S mutation does not affect the chemokine structure. Subsequent experiments on other CCL5 analogs were done without the E66S mutation.

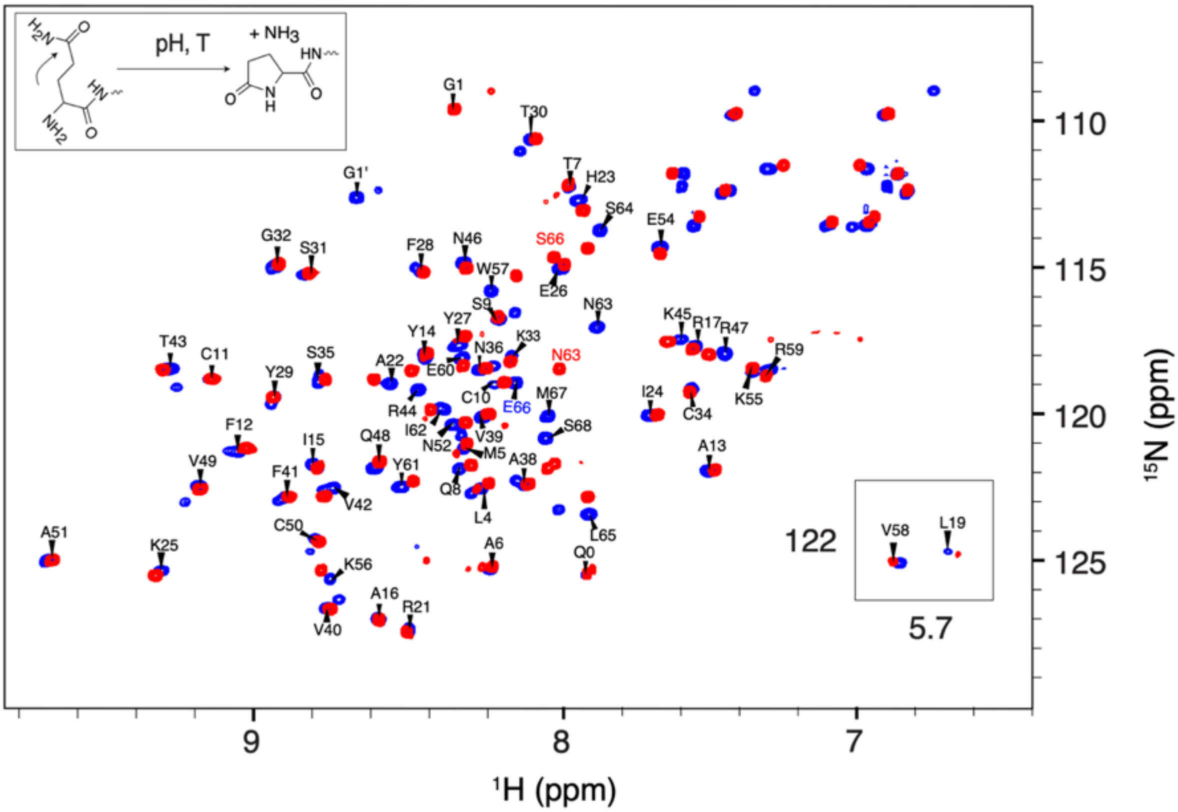


Figure 3.2. ^{1}H-^{15}N HSQCs of 100 µM ^{15}N-labled [5P12]CCL5 (blue) and 190 µM [5P12]CCL5-E66S (red) [25 mM sodium phosphate, pH 3.8, 5% D_2O, 0.01% sodium azide, 25°C, 600 MHz]. Resonances are labeled with assignment information. The N-terminal glutamine Q0 undergoes cyclization to pyroglutamate. The chemical reaction is shown in the insert on the top left. Resonances marked by G1 and G1' correspond to the non-cyclized and cyclized N-termini, respectively. In addition, the *cis-trans* isomerization of adjacent prolines causes a peak splitting.

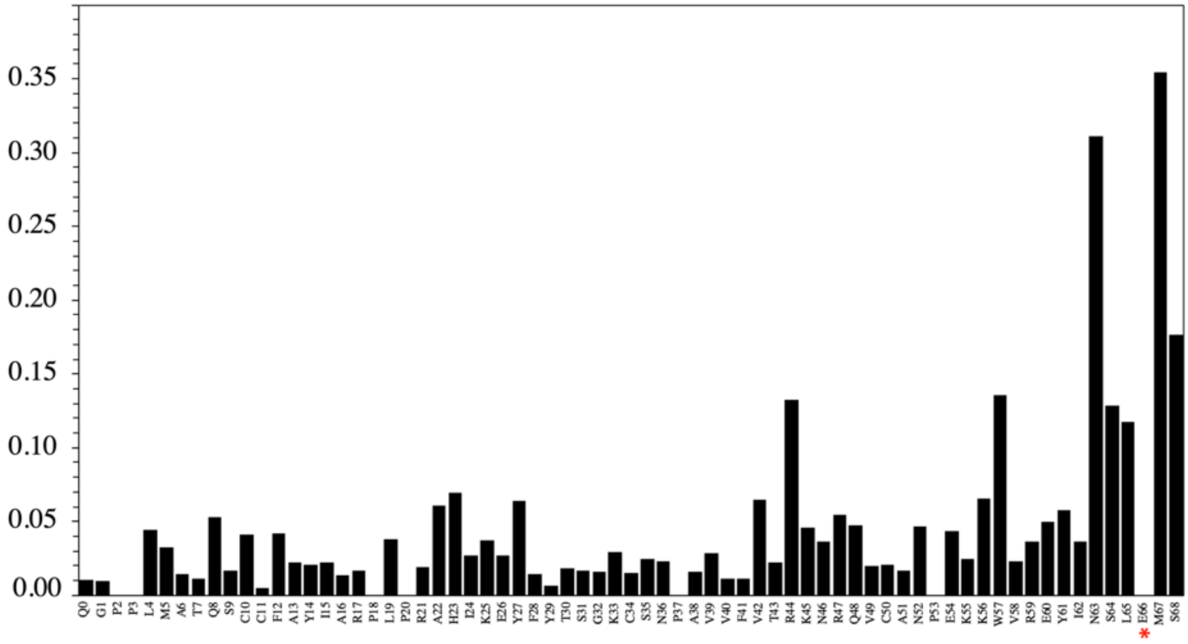


Figure 3.3. ^1H-^{15}N chemical shift differences between [5P12]CCL5 and [5P12]CCL5-E66S. The weighted average of ^1H-^{15}N chemical shift differences $\Delta\delta_{av} = [\Delta\delta_H{}^2 + (\Delta\delta_N/5)^2]^{1/2}$ is shown as a function of residue number. Prolines without observable ^1H-^{15}N resonance and E66S are left blank. Overall, the chemical shift differences are not significant, except for residues close to the E66S mutation site (marked by a red asterisk).

Comparison of [5P12]CCL5 and [5P14]CCL5

[5P14]CCL5 also gives a well-dispersed ^1H-^{15}N HSQC spectrum characteristic for folded proteins containing β-sheets (Figure 3.4). Similar to [5P12]CCL5, all observable 63 (total 69 residues with 6 prolines) ^1H-^{15}N backbone resonances are detected at pH 3.8. The resonances of the folded protein core are very similar between [5P12]CCL5 and [5P14]CCL5, whereas the differing N-terminal resonances of [5P14]CCL5 were assigned by 3D ^{15}N-NOESY-HSQC and ^{15}N-TOCSY-HSQC experiments.

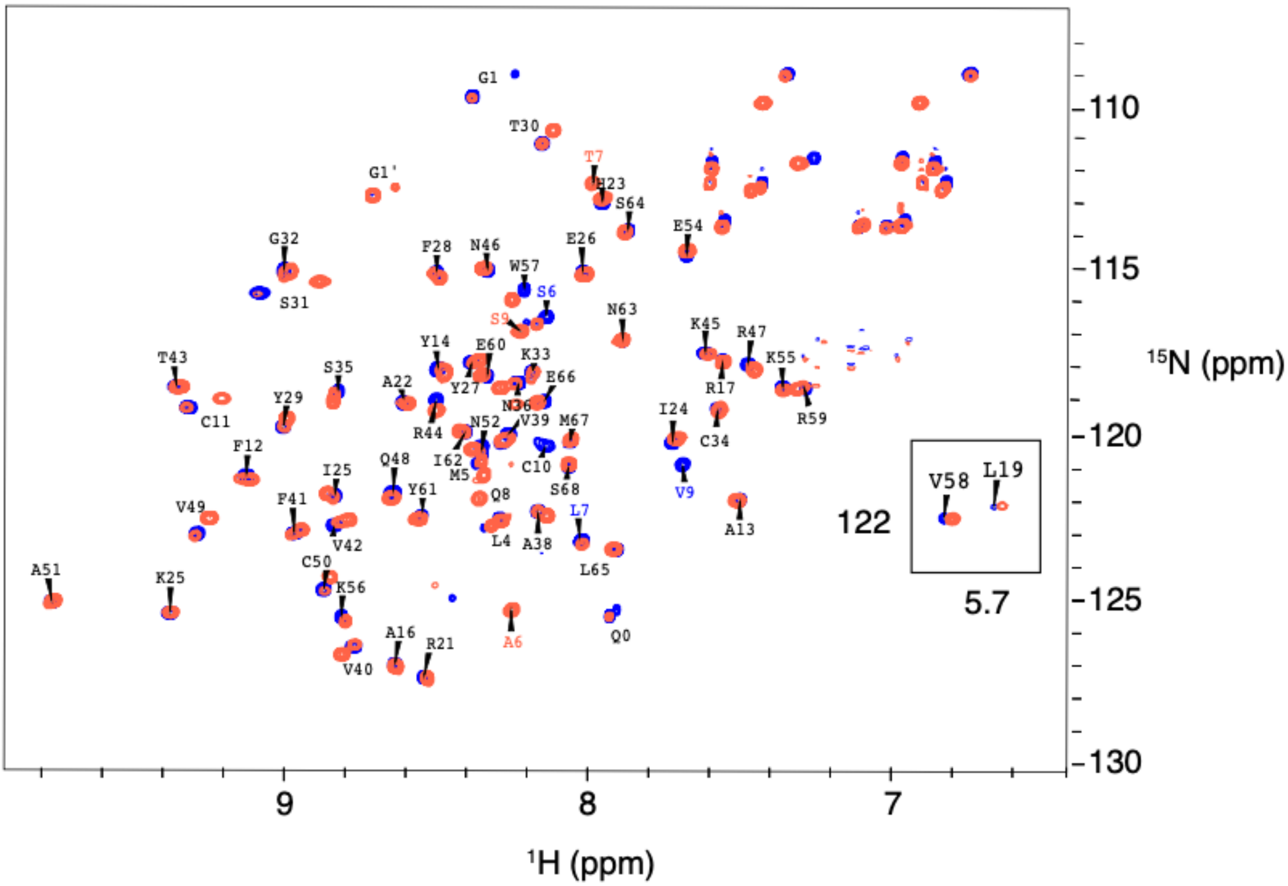


Figure 3.4. ^{1}H-^{15}N HSQC spectra of 100 μM [5P12]CCL5 (red) and 150 μM [5P14]CCL5 (blue) [25 mM sodium phosphate, pH 3.8, 5% D_2O, 0.01% sodium azide, 25°C, 600 MHz]. Resonances are labeled with assignment information.

All larger ^{1}H-^{15}N chemical shift differences between [5P14]CCL5 and [5P12]CCL5 are located in the immediate vicinity of the differing N-termini, whereas only smaller differences are observed in the core structure that can be linked to propagation via the two disulfide bonds C10-C34 and C11-50 (Figure 3.5 A, B). Previous work has shown that the N-terminal residues 0-9 of monomeric CCL5 are mobile on the nanosecond time scale, whereas the rest of the protein is well folded (249). The almost identical chemical shifts within the folded cores of [5P14]CCL5 and [5P12]CCL5 indicate that their core structures are essentially the same.

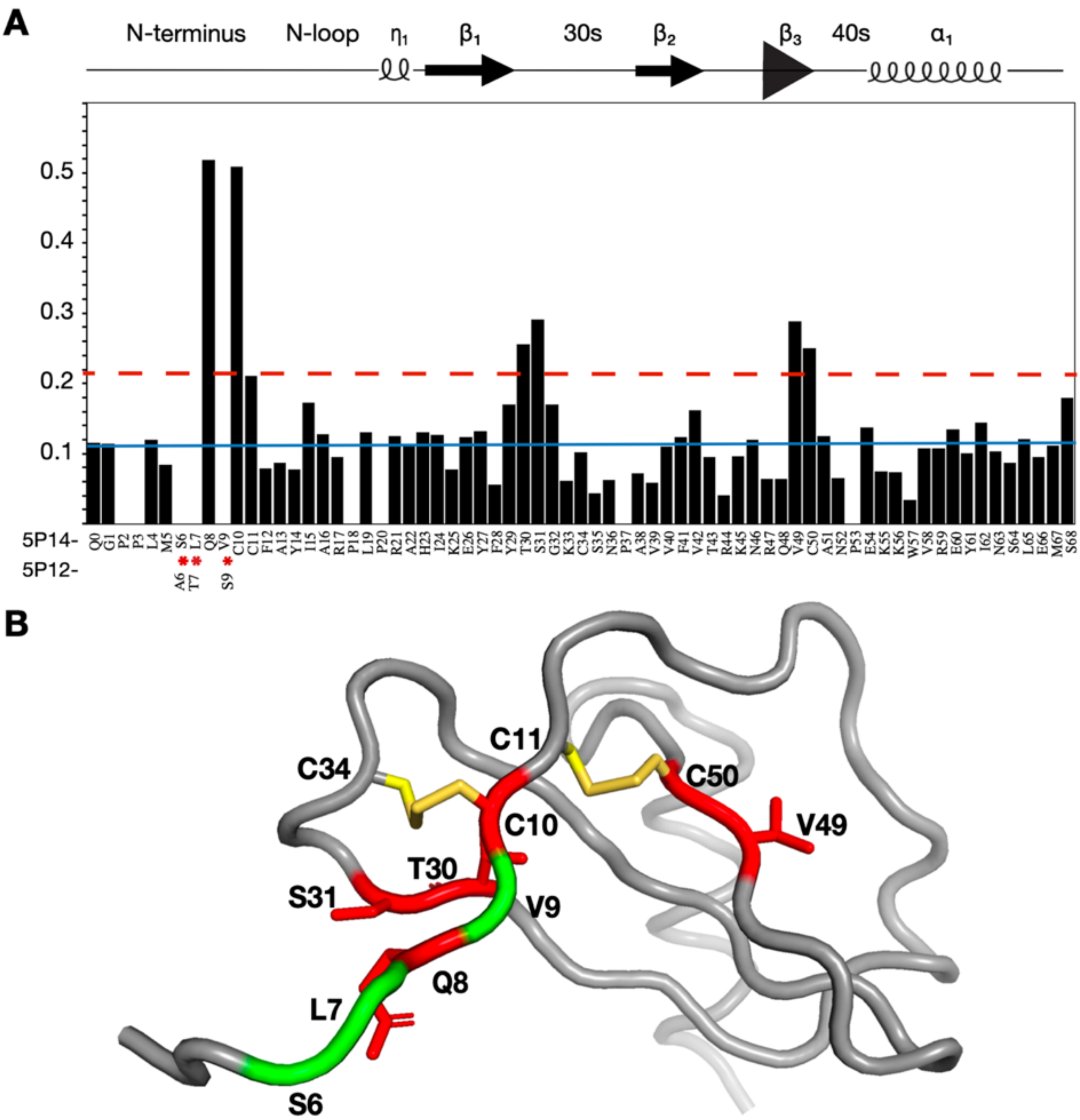


Figure 3.5. Analysis of [5P14]CCL5 and [5P12]CCL5 based on their NMR spectra. A. Weighted average of ^{1}H-^{15}N chemical shift differences $\Delta\delta_{av} = [\Delta\delta_{H}^{2} + (\Delta\delta_{N}/5)^{2}]^{1/2}$ between [5P14]CCL5 and [5P12]CCL5 as a function of residue number. The blue line represents the average $\Delta\delta_{av}$ value, whereas the red dotted line represents the average $\Delta\delta_{av}$ plus one standard deviation (SD). A schematic representation of the secondary structure elements is shown at the top. (B) The $\Delta\delta_{av}$ values are mapped to the crystal structure of CCL5 (PDB:1EQT). Chemical shift changes above the red line in panel A ($\Delta\delta_{av}$ plus standard deviation) are shown in red stick representation, the mutation sites in green. Disulfide-bonded cysteine residues are shown in yellow. N-terminal amino acids differing between [5P14]CCL5 and [5P12]CCL5 are indicated by a red asterisk.

Characterization of [5P14]CCL5

The measured ^{1}H^{N} T_2 times of 30–40 ms for 100–200 μM [5P14]CCL5 (partial agonist) at 25°C and pH 3.8 are comparable to the T_2 times of 35 ms observed for ubiquitin, which has a MWT of 8.6 kDa that is similar to the 8.0 kDa MWT of [5P14]CCL5. This, together with an SDS-PAGE analysis of [5P14]CCL5 under non-reducing conditions, indicates that [5P14]CCL5 is a monomer under the conditions used.

Since the interaction of CCL5 with CCR5 occurs at physiological pH, [5P14]CCL5 was titrated to higher pH values in an HSQC experiment (Figure 3.6). As indicated before, 63 ^1H-^{15}N backbone resonances of [5P14]CCL5 are detected at pH 3.8. In contrast, at pH 7.5, only 47 resonances are detected. Many resonances disappear due to fast amide hydrogen exchange with water of exposed amide groups, which are mainly located in flexible regions such as the N-terminus and loops. Relative to pH 3.8, the chemical shifts and intensities of the observable resonances are not significantly changed, indicating that the overall structure is preserved. The $^1H^N$ T_2 times measured at the higher pH values at 25°C and at concentrations below 100 μM were in the same range of 30-40 ms, suggesting that [5P14]CCL5 does not undergo strong aggregation at pH 7.5.

Similar to the mentioned CCL5 analogs, the spectrum of 100-200 μM [6P4]CCL5 (strong agonist) at pH 3.8 shows a single set of peaks corresponding to a monomer (see Chapter 3.1). Complex formation with the full CCR5 receptor for cryo-EM studies was then further pursued with the [6P4]CCL5 at physiological pH at concentrations below 10 μM.

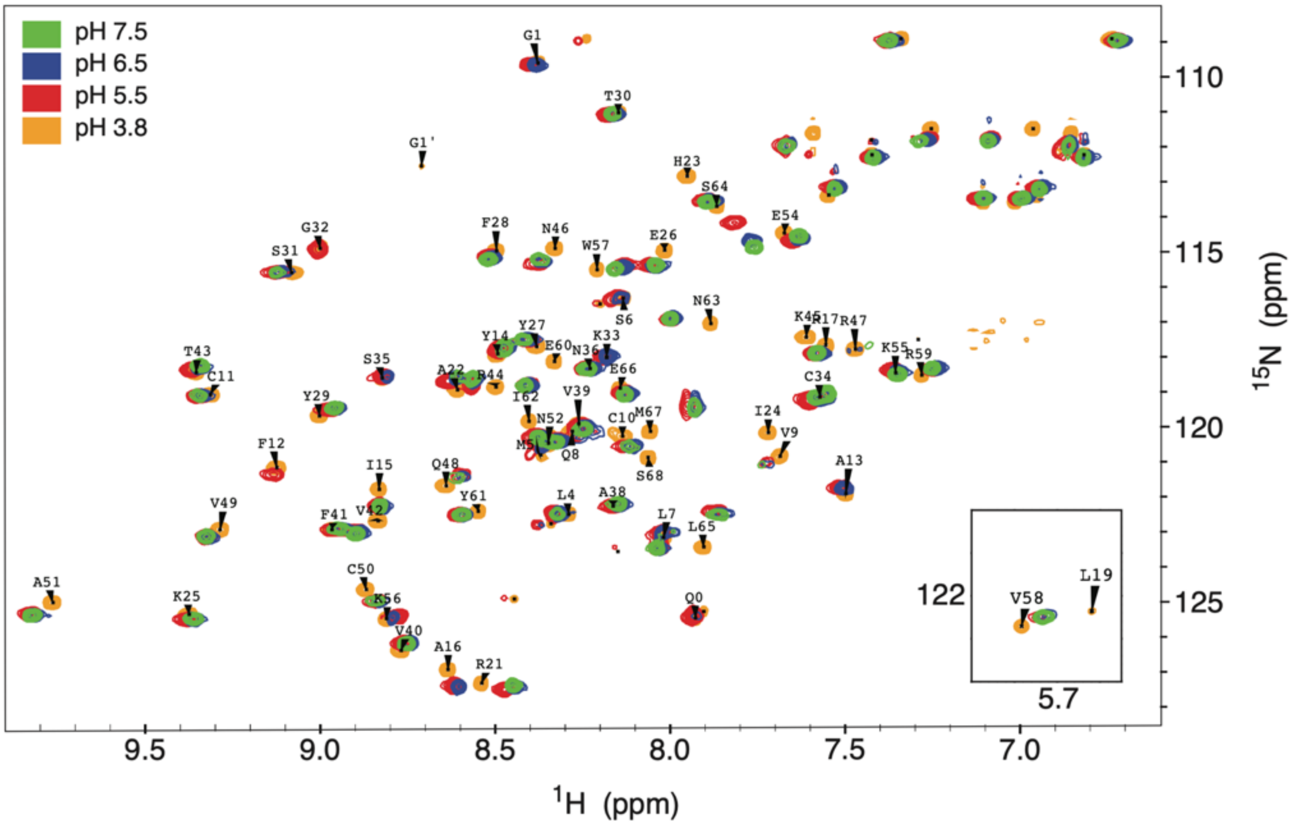


Figure 3.6. Comparison of ^{1}H-^{15}N HSQCs of [5P14]CCL5 within the pH range from 3.8 to 7.5 [25 mM sodium phosphate, 5% D_2O, 0.01% sodium azide, 25°C, 600 MHz]. The color code for each titration point is indicated on the top left.

3.2.3 Conclusion

The NMR analysis of the [5P12]CCL5 and its E66S mutant revealed no significant differences at pH 3.8 and concentrations of 100-200 μM. Therefore, all other chemokine analogs were produced without the E66S mutation. A comparison of the ^{1}H-^{15}N HSQC spectra of [5P12]CCL5 (antagonist) and [5P14]CCL5 (partial agonist) confirmed that their core residues (10-68) are folded into an identical three-dimensional structure. In contrast, their N-termini that contain the sequence differences are flexible. Hence, as expected, the distinct functional behavior of these CCL5 analogs must result from the specific interactions of their differing N-termini with CCR5. An NMR pH titration showed that [5P14]CCL5 is stable and well-folded at physiological pH, similar to [5P12]CCL5 and [6P4]CCL5 at the measured concentrations (<100 μM). This suggested that complexes of CCL5 analogs and CCR5 could be formed at physiological pH at low concentrations, which was indeed the case (see Chapter 2).

4 A key GPCR phosphorylation motif discovered in arrestin2·CCR5 phosphopeptide complexes

4.1 Original Manuscript

Reference: Isaikina P.[+,*], Petrovic I.[+], Jacob R.P., Sarma P., Ranjan A., Baruah M., Panwalkar V., Maier T., Shukla A.K.*, Grzesiek S.* A key GPCR phosphorylation motif discovered in arrestin2•CCR5 phosphopeptide complexes. *bioRxiv,* **2022**
https://doi.org/10.1101/2022.10.10.511578
[+]Equal contribution
*Corresponding author

Short description: This manuscript describes a structural and functional analysis of arrestin2 in apo form and complexes with several CCR5 phosphopeptides. Together with a sequence analysis of GPCR intracellular loops 3 and C-terminal tails, these data reveal key phosphorylation sites responsible for stable GPCR•arrestin interactions and their contributions to the CCR5•arrestin2 function.

4.1.1 Summary

The two non-visual arrestin isoforms, arrestin2 and arrestin3 recognize and bind hundreds of G protein-coupled receptors (GPCRs) with different phosphorylation patterns leading to distinct functional outcomes. The impact of phosphorylation on arrestin interactions has been well studied only for very few GPCRs. Here we have characterized the interactions between the phosphorylated CC chemokine receptor 5 (CCR5) and arrestin2. We detected several new CCR5 phosphorylation sites, which are necessary for stable complex formation with arrestin2. Crystal structures of arrestin2 in apo form and in complexes with CCR5 C-terminal phosphopeptides together with NMR spectroscopy, biochemical and functional assays revealed three phosphoresidues in a pXpp motif that are essential for the arrestin2 interactions and activation. The same phosphoresidue cluster is present in other receptors, which form stable complexes with arrestin2. We propose that the identified pXpp motif is responsible for robust arrestin2 recruitment in many GPCRs. An analysis of available sequences, structural and functional information on other GPCR•arrestin interactions suggests that a particular arrangement of phosphoresidues within the GPCR intracellular loop 3 and C-terminal tail determines arrestin2 and 3 isoform specificity. Taken together, our findings demonstrate how multi-site phosphorylation controls GPCR•arrestin interactions and provide a framework to probe the intricate details of arrestin activation and signaling.

4.1.2 Introduction

GPCRs represent a large family of cell-surface receptors mediating signaling events via G proteins and arrestins (1). The agonist-induced G protein signaling is terminated by the phosphorylation of the receptor C-terminal tail and/or intracellular loops (2) primarily via GPCR kinases (GRKs), which subsequently leads to the binding of arrestins to the receptor C-terminal tail and core (3–6). The receptor•arrestin complex acts as a scaffold for various further signaling proteins thereby activating e.g. ERK1/2 and MAP or inducing receptor internalization (2,7). The many hundreds non-visual GPCRs in the human body are regulated by two genetically and structurally conserved non-visual arrestin subtypes, arrestin2 and arrestin3 (also known as β-arrestin1 and β-arrestin2, respectively). While both isoforms can bind to the same receptor, their interaction may activate different signaling partners (8). The arrestin-mediated signaling depends on the phosphorylation pattern of the GPCR C-terminal tail, which is modulated by the interactions of the GPCR with various GRKs (9,10). These findings have led to the hypothesis of a phosphorylation 'barcode' for the receptor C-terminal tail (11).

The molecular details of the specificity of GPCR-arrestin interactions and the causes for the different functional outcomes of varying phosphorylation patterns are still poorly understood. Previously solved full-length GPCR•arrestin complex structures comprise a rhodopsin-arrestin1 fusion complex (12,13), a fusion complex of the engineered constitutively active arrestin2 with the truncated 5HT2B serotonin receptor (14), an arrestin2 complex with a chimera of the M2 muscarinic receptor (M2R) and the vasopressin 2 receptor C-terminal phosphopeptide (V2Rpp) (15), a β1AR-V2Rpp chimera•arrestin2 complex (16), as well as native V2R•arrestin2 (17) and NTR1•arrestin2 (18,19) complexes. Of these, the rhodopsin fusion, the M2R-V2Rpp and β1AR-V2Rpp chimera, as well as the truncated 5HT2B serotonin receptor fusion complexes show similar orientations of the arrestin. However, the arrestin2 orientation differs significantly in the NTR1 and V2R complexes, which have native receptor C-terminal tails. The phosphate groups in the C-terminal tails have only been resolved in the M2R-V2Rpp, β1AR-V2Rpp, and V2R complexes. Higher resolution has been obtained in complex structures of arrestins with phosphorylated GPCR C-terminal tail peptides, which have revealed the position and coordination of several phosphates in addition to the hallmarks of arrestin activation, i.e. the replacement of arrestin strand β20 by the phosphorylated receptor peptide, conformational changes within the arrestin loops and a twist between its N- and C-terminal domains. However, these studies have been limited to V2Rpps binding arrestin2 (20,21), an ACKR3 (formerly known as CXCR7) C-terminal phosphopeptide binding arrestin3 (22), and a rhodopsin C-terminal phosphopeptide binding arrestin1 (visual arrestin) (23).

We have recently determined the structure of the human chemokine receptor 5 (CCR5) in an active complex with the chemokine super-agonist [6P4]CCL5 and the heterotrimeric G_i protein by cryo electron microscopy (24). CCR5 plays a major role in inflammation by recruiting and activating leukocytes (25). It is also the principal HIV coreceptor (26) and is involved in the pathology of cancer (27,28), neuroinflammation (25), and COVID-19 (29). The structure of the active [6P4]CCL5•CCR5•G_i complex has provided detailed insights into the mechanism of CCR5 activation via the [6P4]CCL5 N-terminus, which differs from other solved chemokine•GPCR complexes.

Much less is known about the molecular basis of chemokine receptor interactions with arrestins, which play a similar critical role as G protein interactions in the immune response and inflammatory signaling pathways (30). Here we have characterized the CCR5 phosphorylation induced by the super-agonist [6P4]CCL5 and GRK2. We discovered several new phosphorylation sites within the CCR5 C-terminal tail and used them together with previously reported sites to design a set of phosphopeptides for analyzing their effects on arrestin interactions. We determined high-resolution crystal structures of arrestin2 in apo form and in complex with two of these phosphopeptides with resolved positions of the phosphate groups in their electron density. The phosphate groups form a distinct pXpp phosphopeptide sequence motif that induces the active arrestin2 conformation by dominant electrostatic and beta-sheet interactions. An identical pXpp motif exists in the V2Rpp and forms the same activating interactions with arrestin2. A combination of NMR, biochemical and cellular assays confirmed the importance of this motif and quantified the contributions of individual phosphoresidues to arrestin binding and activation. A comparison of the new structures to other arrestin complexes together with an analysis of GPCR sequences provides hints on the molecular basis of arrestin2/arrestin3 isoform specificity.

4.1.3 Results

Revisiting CCR5 phosphorylation by a G protein kinase

Agonist-driven phosphorylation of serine and threonine residues in GPCRs by GRKs is generally required for arrestin binding. Although CCR5 has seven such potential phosphorylation sites in its C-terminal tail, phosphorylation of only four C-terminal serine residues has been reported (31,32). To clarify this situation, we characterized CCR5 phosphorylation by GRK2 that plays a prominent role in the regulation of many chemokine receptors, including CCR5 (33–36). For this, GRK2 was co-expressed with CCR5 in insect cells and phosphorylation induced by the addition of the super-agonist chemokine [6P4]CCL5. Phosphoproteomics of the purified CCR5 clearly revealed phosphorylation of S349 and S325, which had not previously been reported, in addition to a double phosphorylation of S336 and S337, as well as a not well-defined single phosphorylation site in the region T340 to T343 (Figures 1A, S1). Similar assignment ambiguities have been related previously to the heterogeneity of phosphorylation by GRKs (19). A further western blot analysis with CCR5 phosphosite-specific antibodies confirmed the previously not reported phosphorylation of T340 as well as that of the serine residues S336/337, S342, and S349 (Figure 1B).

To assess the impact of these various CCR5 phosphoresidues on arrestin2 binding and conformation, we designed seven distinct CCR5 synthetic phosphopeptides in addition to the V2Rpp phosphopeptide mimicking V2R phosphorylation (21), which served as a control (Figure 1C). These designed CCR5 phosphopeptides did not include pS325, which is unlikely to play a major role in the arrestin interaction due to its location at the beginning of the CCR5 C-terminal tail directly following the palmitoylated cysteine residues (Figure 1A). This is confirmed by functional assays (see below).

CCR5 phosphorylation levels govern arrestin2 binding

The interaction of the synthesized phosphopeptides with arrestin2 was first investigated by solution NMR spectroscopy using a truncated arrestin construct (arrestin2$^{1\text{-}393}$, residues 1-393), which comprises the C-terminal strand β20, but lacks the disordered C-terminal tail (15). Upon addition of the phosphopeptides, the ^{1}H-^{15}N

HSQC-TROSY spectra of arrestin2$^{1\text{-}393}$ [^{15}N-labeled, ~80% deuterated] show continuous shifts in several resonance positions (Figure 2A, S2) indicating fast to intermediate chemical exchange on the microsecond time scale. A non-linear fit of the chemical shift changes to binding isotherms provided dissociation constants K_D in the ten to hundred micromolar range (Figures 2A, S2), which is similar to previous observations on phosphopeptide•arrestin1 complexes (23).

Functional binding of the phosphopeptides to arrestin is expected to induce a conformational change, where arrestin's C-terminal strand β20 is released from the β-sheet with its N-terminal β-strand and replaced by the phosphopeptide (see below). This exposes previously inaccessible cleavage sites in the arrestin C-terminus (R393 for arrestin2), which can be probed by trypsin digestion (37). Such a trypsin proteolysis carried out on full-length arrestin2 (arrestin2$^{1\text{-}418}$, Figure 2B) confirmed the functional conformational changes induced by the binding of the synthetic phosphopeptides.

Figure 2C summarizes the quantitative results of the NMR titrations and the trypsin proteolysis assays. The tightest binding was observed for the six-fold phosphorylated CCR5 6P peptide ($K_D = 45 \pm 6$ μM) and the eight-fold phosphorylated V2Rpp ($K_D = 19 \pm 3$ μM), whereas the three- and four-fold phosphorylated peptides have weaker affinities in the hundred micromolar K_D range, agreeing with the expectation that the phosphorylation level dominates the arrestin•peptide interaction. However, individual phosphorylation sites contribute differently to the overall affinity. Hence, 5P1 (lacking pT343) shows weaker affinity ($K_D = 147 \pm 7$ μM) whereas 5P2 (comprising pT343) has an affinity ($K_D = 54 \pm 6$ μM) similar to 6P suggesting that phosphorylation of T343 is essential for tight arrestin binding.

The trypsin proteolysis rates of the arrestin•phosphopeptide complexes correlate with the determined affinities. Whereas the low-affinity ($K_D > 100$ μM) peptides (5P1 and all three- or four-fold phosphorylated peptides) had a trypsin digestion midpoint at around 30 minutes, for both 6P and V2Rpp the digested band became dominant already after only 5 minutes (Figure 2B). Of note, 5P2 comprising pT343 also had an accelerated digestion midpoint at about 20 minutes.

To probe active arrestin2 conformation induced by the synthetic CCR5 phosphopeptides, we incubated full-length arrestin2 with the peptides and synthetic antibody fragment Fab30 that selectively recognizes active arrestin2 (21). The mixtures were analyzed by size exclusion chromatography (SEC) (Figure 2D). Both CCR5 phosphopeptide 6P (red) or V2Rpp (green) induced a shift of the elution volume to ~2.3 ml from the ~2.6 ml observed for the inactive apo form (black), which indicates the formation of stable phosphopeptide•arrestin2•Fab30 complexes. An identical result was obtained for the peptide 3P2 (blue), which is only phosphorylated at the central T340, S342, and T343 residues, despite its more than four-fold lower affinity. In contrast, all other CCR5 phosphopeptides with similar low affinities as 3P2 resulted in mixtures between active and inactive arrestin populations. These results suggest that the 3P2 phosphorylation pattern pXpp is specific, necessary, and sufficient for inducing the active arrestin2 conformation as assayed by Fab30 binding, while the other 6P phosphosites only contribute to the overall affinity.

Crystal structures of human arrestin2 in apo state and in complex with distinct CCR5 phosphopeptides

To obtain insights into structural changes of arrestin2 induced by phosphopeptide binding, two complexes of human arrestin2 (arrestin2$^{1\text{-}359}$, lacking the C-terminal strand β20, see below) with the CCR5 4P and 6P phosphopeptides and the stabilizing Fab30 were prepared and their crystal structures determined at resolutions of 3.2 and 3.5 Å, respectively (Figure 3A, Table S1). Furthermore, also the crystal structure of full-length human arrestin2 was determined in its apo state at a resolution of 2.3 Å (Figure 3A). The latter is highly similar to the structure of bovine arrestin2 (PDB 1G4M) with the conserved fold composed of the two N- and C-terminal β-sandwich domains, the characteristic parallel β-sheet between the N- and C-terminal strands β1/β20, as well as a disordered C-terminal tail beyond residue 396 (38).

While only parts of the phosphopeptides have well-defined electron densities in both phosphopeptide complex structures (Figure S3A), the phosphopeptide electron density is better defined in the 6P complex structure, indicating lower structural disorder in this region. For 6P, eight peptide residues (V338–E345) including the three phosphosites (pT340, pS342, pT343) have clearly defined density. In contrast, the electron density for 4P is reasonably defined only for four residues (A335–V338) including only two phosphosites (pS336, pS337) (Figure 3B). Both 6P and 4P coordinate as extended β-strands with the arrestin2 N-terminal strand β1 in an antiparallel manner with their observable phosphate groups interacting with the same arrestin2 residues (see below). Strikingly, however, there is a register shift (Figure 3C) such that the two consecutive phosphosites pS336, pS337 of the 4P complex take the positions of the two consecutive phosphosites pS342, pT343 in the 6P complex.

Both phosphopeptide complexes exhibit the hallmarks of arrestin2 activation, i.e., the formation of the anti-parallel intermolecular β-sheet between the arrestin strand β1 and the phosphopeptide, which replaces strand β20 of the intramolecular parallel β-sheet in the inactive apo arrestin2 structure, a twist of the C-domain relative to the N-domain of approximately 21°, and significant changes in the lariat, finger and middle loops (Figure 3D). The two phosphopeptides induced very minor differences in the orientations of Fab30 relative to arrestin2 (Figure S3B) and consequently also very minor differences in unit cell dimensions and crystal packing of their crystallized complexes despite identical space groups (Table S1). As compared to the 4P complex, the electron density of the 6P complex is less well-defined for arrestin2 residues 64-70 (center of finger loop) and 308-313 (at the end of the lariat loop). This may be caused by the variations in the crystal packing, but also indicates structural plasticity of these loops, which participate in receptor (finger loop) and clathrin (lariat loop) binding (39).

Key CCR5 phosphorylation sites responsible for arrestin2 activation

The polar core of apo arrestin harbors a network of highly conserved ionic interactions formed by D26 and R169 in the N-domain, D290 and D297 on the lariat loop of the C-domain and R393 located in an extended stretch after the C-terminal strand β20 (Figure 4A). Together with a salt bridge between R25 and E389 and the hydrogen bonds forming the parallel β-sheet between strands β1 and β20, these strong interactions connect the two arrestin domains and stabilize the inactive conformation.

The central element of arrestin activation is the disruption of this polar core by the β-strand exchange of the arrestin strand β20 with the receptor phosphopeptide and the subsequent relocation of the gate loop (21). Arrestin21-359, which was used for solving the crystal structures of the active complexes with the CCR5 phosphopeptides, lacks strand β20, thereby facilitating the formation of the intermolecular anti-parallel β-sheet to arrestin2 strand β1. In the 6P complex, the intermolecular β-sheet comprises CCR5 residues 341 to 344 (Figure 4A). The first visible phosphoresidue pT340 forms intermolecular salt bridges to R25, K11, and K294 (lariat loop), thereby apparently disrupting the salt bridges between R169 and D290/D297, pulling the lariat loop towards the N-domain, and forcing a twist of the C-domain. The 6P complex is further stabilized by extensive electrostatic interactions of CCR5 residues pS342, pT343, and E345 with arrestin2 residues R7, K10 on strand β1 and R103, K107 on helix α1. In addition, the phosphate of pS342 forms a salt bridge to R67 on Fab30 (Figure S3B).

While the 4P peptide also establishes an intermolecular anti-parallel β-sheet with strand β1 of arrestin2, there are fewer stable intermolecular contacts with no engagement of the arrestin2 lariat loop (Figure 4A). The two visible phosphoresidues, pS336 and pS337, form intermolecular salt bridges with arrestin2 residues R7, K10, and K107 in a similar way as the 6P phosphoresidues pS342 and pT343.

<u>*Cellular assays*</u>

The effect of individual CCR5 phosphorylation sites on arrestin2 recruitment and conformation was tested in a cellular context using several previously developed luciferase (NanoBiT) complementation assays (40) on seven S/A and T/A C-terminal point mutations of full-length CCR5.

For testing the direct recruitment of arrestin2 by CCR5 in response to [6P4]CCL5 stimulation, CCR5 fused to a small fragment (SmBiT) and arrestin2 fused to a large fragment (LgBiT) of luciferase (NanoBiT) were co-expressed in HEK293 cells and the complementation-induced luminescence was measured upon stimulation with the super-agonist [6P4]CCL5 (Figure 4B, S4A). With the exception of the strongly attenuated T343A, all other CCR5 mutants recruited arrestin2 with similar efficiency as wild-type CCR5. We attribute the reduced recruitment by T343A to the abolishment of the direct ionic interaction of the pT343 phosphate group with the side chain of the arrestin N-terminal residue R7 (Figure 4A), which also leads to a three-fold reduction in the phosphopeptide affinity as assayed by NMR (see above, Figure 2).

The intrabody30 (Ib30, a single chain derivative of Fab30) specifically recognizes the activated conformation of arrestin2 (41). For testing the formation of the activated conformation, Ib30 fused to LgBiT and arrestin2 fused to SmBiT, and the respective

CCR5 mutant were co-expressed and the cells stimulated by [6P4]CCL5 (Figure 4B, S4B). Whereas most mutations had only a moderate effect, the CCR5-T340A and S342A mutants almost completely abolished arrestin recognition by Ib30. For S342A this is expected, since pS342 directly contacts Fab30 in the 6P•arrestin2•Fab30 complex (Figure 3A). In contrast, the strong effect of T340A mutation apparently indicates a genuine change of the arrestin2 conformation, which strongly impedes recognition by Ib30. This agrees with the notion that pT340 is directly involved in the arrestin2 activating motion by pulling the lariat loop towards the N-domain (see above and below).

Finally, we also assayed endocytosis for the three CCR5 mutants (T340A, S342A and T343A) by co-expression with arrestin2 fused to SmBiT and the endofin FYVE domain, which targets early endosomes, fused to LgBiT (Figure S4C). In agreement with the arrestin2 recruitment assays, only CCR5-T343A significantly decreased endocytosis levels.

The phosphorylation motif pXpp is responsible for stable arrestin recruitment

CCR5 vs V2R

The phosphorylation patterns of the V2R and its interactions with arrestin2 have been studied extensively by a range of biophysical and computational methods (9,20,21,42,43). Little is known about the interactions of other receptors with arrestin2. The present structures of the CCR5 C-terminal peptides in complex with arrestin2 provide an opportunity for a detailed comparison.

Both CCR5 6P•arrestin2•Fab30 and V2Rpp•arrestin2•Fab30 (PDB: 4JQI) complexes have an overall very similar organization (Figure S5). Despite V2Rpp containing five more phosphorylation sites than CCR5 6P, both peptides form a completely analogous antiparallel β-sheet between their central residues and strand β1 of arrestin2 (Figure 5A). This β-sheet is stabilized by an identical set of contacts between positively charged residues of arrestin2 and a central pXpp motif on the phosphopeptides (6P: pT340, pS342, pT343; V2Rpp: pT360, pS362, pS363). Similar to pT360 of V2Rpp, pT340 of CCR5 6P is oriented towards the connector region between arrestin's N- and C-domains and forms salt bridges to K11 and R25 in the

N-domain as well as K294 on the lariat loop. The latter interaction may constitute a main component of the force that reorients the arrestin domains upon activation. This notion is corroborated by Ib30-based functional studies, which show that both pT340 in CCR5 (Figure 4B) and pT360 in V2R (42) strongly impact the conformation of arrestin2. Thus, this first phosphoresidue of the pXpp motif appears crucial to drive arrestin activation. Of note, the interaction of this first phosphoresidue is not present in the 4P complex where only the consecutive pS336 and pS337 take the role of the last two phosphoresidues in the pXpp motif (Figure 4B). However, the presence of Fab30 and crystal packing may stabilize the active arrestin domain orientation even for this incomplete pXpp motif.

Besides these main interactions, further phosphosite interactions are observable in the V2Rpp complex, which are not present in the CCR5 6P complex: pT359 is involved in crystal contacts and pS357 as well as pS364 interact with the same arrestin residues as their respective adjacent pT360 and pT363. Remarkably, the N-terminal V2Rpp residues pT347 and pS350 form contacts with observable electron density to the arrestin2 finger loop and nearby residues (Figure 5A). However, these two phosphoresidues do not form visible interactions with arrestin2 in full-length receptor complexes such as the M2R-V2Rpp chimera•arrestin2 (PDB: 6U1N), the β1AR-V2Rpp chimera•arrestin2 (PDB: 6TKO) (Figure 5A) and the recently solved V2R•arrestin2 structures (17). Rather in these complexes, only the phosphosites beyond V2R residue 356 coordinate with the arrestin2 N-domain in an identical manner as in the V2Rpp and CCR5 6P complex structures. This is presumably due to steric constraints imposed by the direct binding of the receptor core to the arrestin2 finger loop (15–17). An inspection of the β1AR-V2Rpp chimera•arrestin2 structure (Figure 5A) shows that a flexible linker of about 15 residues connects the receptor helix 8 and the phosphorylation motif recognized by the arrestin2 N-domain.

To prove that wild-type CCR5 is able to engage arrestin2 robustly, we reconstituted a complex between arrestin2$^{1\text{-}393}$, [6P4]CCL5, and GRK2-phosphorylated CCR5, for which the phosphorylation of each individual phosphosite had been verified by western blot analysis (see Methods). The complex was assembled on FLAG beads and stabilized by Fab30, washed extensively to remove excess arrestin and Fab30, and then further purified with SEC. Negative-stain EM of the complex and subsequent 2D classification revealed CCR5•arrestin2•Fab30 particles with arrestin in core- and tail-engaged as well tail-only-engaged arrangements. The latter may be caused by the dynamics of binding and/or heterogeneous phosphorylation of CCR5 (Figure 5B). These results prove that similar to NTR1 and V2R also GRK-phosphorylated CCR5 forms a stable complex with arrestin2.

4.1.4 Discussion

Generalization to other receptors

While our NMR data indicate that cumulative phosphorylation of the CCR5 C-terminus is important for arrestin2 affinity, the SEC data and Ib30 cellular assay show that a specific arrangement of phosphoresidues in a pXpp motif is needed for robust activation of arrestin2 and stable complex formation (Figures 2, 4). A comparison of the C-terminal sequences of structurally and functionally characterized GPCRs in the context of arrestin2 recruitment (Table S2) reveals that this pXpp motif is commonly found at a distance of 15 to 30 residues downstream of helix 8 (defined by the GPCRdb numbering scheme (44)) (Figure 5C). This is the case for V2R, NTR1, and the chemokine receptors CXCR3 and CXCR4 (Figure 5C). A similar pattern occurs in many other chemokine receptors, such as CCR3 and CCR4 (Figures 5C, 6B). In agreement with these observations, the indicated phosphorylation sites have been suggested as key arrestin binding motifs in V2R (42), rhodopsin (23), and NTR1 (19). In summary (Figure 5D), this suggests that the pXpp cluster which follows a flexible linker of at least 15 amino acids after helix 8 is crucial for full arrestin2 engagement and stable arrestin2 complex formation. As judged from the high similarity of the solved CCR5 6P and V2Rpp complexes and the functional data, the first phosphoresidue of the pXpp cluster engages the lariat loop thereby triggering arrestin activation, whereas the last residue seems more important for overall arrestin2 recruitment (Figure 4B).

Structural insights into arrestin isoform recognition

The arrestin2 and arrestin3 isoforms of non-visual arrestins have highly conserved sequences and similar three-dimensional structures. Both arrestins can desensitize GPCRs, however, their localization in cells differs to some extent. Whereas arrestin3 is localized in the cytoplasm, arrestin2 is found in both the cytoplasm and nucleus (45,46). Despite having similar interactions with client proteins, in many cases they play distinct roles in downstream outcomes (8).

Based on the characteristics of their agonist-dependent arrestin interaction, GPCRs have been separated into ‘class A’ and ‘class B’ subcategories (7,47). Class A receptors such as adrenergic, muscarinic, dopamine, μ-opioid, and 5-hydroxytryptamine receptors bind arrestin3 with higher affinity than arrestin2, but the respective complexes are transient and dissociate at or near the plasma membrane (Table S2). In contrast, class B receptors such as the angiotensin II type 1 receptor (AT1aR), NTR1, V2R rhodopsin, the complement C5a receptor, CCR5 and several others (Table S2) bind both arrestins with approximately equal, but higher affinity than class A receptors, and form long-lived arrestin complexes that traffic into endosomes.

Arrestin2 vs arrestin3 phosphopeptide complexes

To obtain insights into the specific recognition of both arrestin isoforms by GPCRs, we compared the structure of the arrestin2•CCR5 6P complex with that of arrestin3 in complex with the C-terminal ACKR3 phosphopeptide [ACKR3pp, PDB 6K3F (22), Figure 6A]. Similar to the CCR5 6P peptide, ACKR3pp binds in an extended conformation to the groove formed by the N-terminal arrestin β-sheet. However, its position is shifted towards the arrestin finger loop and it does not form an antiparallel β-sheet with the arrestin3 N-terminal strand β1. The arrestin3 binding groove also has a higher positive charge density than arrestin2, which extends towards the finger loop thereby explaining the engagement of ACKR3pp residue pS335 with this region. Both peptide positions overlap at the sites of pT340 (6P) and pT342 (ACKR3pp), interacting with a similar set of residues in arrestin2 and arrestin3, respectively. Interestingly, pS337 of CCR5, although not resolved in our structure and manually placed for comparison in Figure 6A would fit well into the arrestin3 binding interface and would form the same set of charge interactions as pT338 of ACKR3pp. In agreement with these observations, functional data on CCR5 indicate that an S337A mutant reduces both arrestin3 and arrestin2 binding to a similar extent, whereas arrestin3 binding is less reduced than arrestin2 binding for S342A and S349A mutants (31). This hints at a weaker role of CCR5 phosphoresidues beyond S337 in arrestin3 recruitment.

Receptor ICL3 and C-terminal tail sequence analysis

Besides these structurally well-detected interactions between the GPCR C-terminal tails and arrestins, also intracellular loops (ILs) interact with arrestins (12,13,15,19), albeit they mostly have not been resolved. In particular, ICL3 is in very close proximity to arrestin (see e.g., β1AR-V2Rpp•arrestin2 complex, Figure 5A). To relate these structural findings to arrestin specificity we aligned the ICL3 and C-terminal tail sequences of GPCRs that are well-characterized with respect to arrestin interactions (Table S2) together with receptors from similar families (Figure 6B).

A visual inspection of the location of potential phosphorylation sites (S/T) and lengths of the IL3s and the C-terminal tails clearly puts the analyzed receptors into two distinct groups. The first group (Figure 6B, top) comprises receptors with very short IL3s (< ~5aa) containing few possible phosphorylation sites and short (~20–50 aa) C-terminal tails with dense clusters of serine and threonine residues often harboring the pXpp motif. Receptors of this group are almost all peptide-binding GPCRs including chemokine receptors with a number of them known to form stable arrestin complexes and characterized as class B receptors (Table S2). As indicated, both CCR5 and V2R belong to this group. The high density of potential phosphorylation sites and the specific recognition of the pXpp motif in this class of receptors may explain the higher affinity for arrestins and the long lifetime of the arrestin complexes.

The second group of receptors (Figure 6B, bottom) have longer to extremely long (>100 aa) IL3s and C-terminal tails with diverse lengths ranging from very short (< 10 aa, muscarinic, dopamine, and some 5-hydroxytryptamine receptors) to very long (~150 aa, α-adrenergic receptors). The density of potential phosphorylation sites within ICL3 and the C-terminal tail is lower for this second group than for the first group. The receptors in this second group, which have been characterized for arrestin binding (Table S2), interact with arrestins in a transient manner and in part have a preference towards arrestin3 as would be expected for class A receptors. Interestingly, the pXpp motif can also be found in the long ICL3 of some of these receptors such as M2R, which may enable specific arrestin2 binding also via ICL3 as evident from the Ib30 recognition of such complexes (40).

This diversity of C-terminal tail lengths and the low density of phosphorylation sites in the class A receptor group suggest that arrestin3 recruitment is less dependent on the exact position of phosphoresidues within the C-terminal tail than arrestin2 recruitment. The more extended positive surface of the binding groove may accommodate this diversity of phosphosites and lead to lower affinity/transient arrestin binding. Moreover, the structure of active arrestin3 in complex with ACKR3pp (Figure 6A) shows that the peptide interaction surface is very close to the finger and middle loops, which are engaged with the receptor core. Considering this close proximity of the arrestin3 binding interface to the receptor core, it is likely that both the C-terminal tail and ICL3 participate in the arrestin3 complex formation. Of note, D2R, which has a long ICL3, but no C-terminal tail after helix 8, recruits arrestin3 even in absence of GRK phosphorylation suggesting that arrestin3 recruitment might be less dependent on phosphorylation (48).

4.1.5 Conclusion

In conclusion our structural and biophysical analysis of the interactions between the CCR5 phosphorylated C-terminus and arrestin2 has identified a key pXpp motif for arrestin2 recruitment and activation, which appears to be conserved across many class B GPCRs such as V2R, NTR1, rhodopsin, CXCR3 and CXCR4. An analysis of GPCR ICL3 and C-terminal phosphorylation sites and their class A or B arrestin interaction behavior together with a structural comparison of arrestin2 and arrestin3 phosphopeptide binding modes revealed salient sequence features for arrestin2 and arrestin3 isoform specificity. This may provide a framework for a more rigorous characterization of other GPCR arrestin systems, which are needed to obtain a comprehensive picture of the diverse arrestin binding modes and their relation to GPCR phosphorylation.

4.1.6 Methods

Peptide synthesis

Phosphorylated peptides corresponding to the last 22 residues of human CCR5 receptor (3P1, 3P2, 4P, 5P1, 5P2, 6P) or the last 29 residues of the V2 receptor (V2Rpp) were obtained from the Tufts University Core Facility for peptide synthesis. The non-phosphorylated CCR5 peptide (0P) was obtained from GenScript and contained a biotinylated N-terminus.

Constructs

The genes encoding wild-type, full-length human CCR5 with a C-terminal 3C cleavage site followed by a FLAG tag for expression in the baculovirus Sf9 insect cell system and [6P4]CCL5 for expression in *E. coli* have been described before (24,49).

The plasmid encoding GPCR kinase subtype 2 used for the receptor phosphorylation in insect cells (GRK2-CAAX) was a gift from Robert Lefkowitz (Addgene plasmid #166224 (50)).

Full-length arrestin2$^{1\text{-}418}$ (C150L, C242V, C251V, C269S) in vector pET-28a (+) was obtained from GenScript. The complete construct contained an N-terminal hexahistidine tag followed by a TEV cleavage site (ENLYFQG) and the arrestin2 sequence (bold):

MGSSHHHHHHSSGENLYFQG**MGDKGTRVFKKASPNGKLTVYLGKRDFVDHIDLVDPVDGVVLVDPEYLKERR**
VYVTLTCAFRYGREDLDVLGLTFRKDLFVANVQSFPPAPEDKKPLTRLQERLIKKLGEHAYPFTFEIPPNLPCSV
TLQPGPEDTGKACGVDYEVKAFLAENLEEKIHKRNSVRLVIRKVQYAPERPGPQPTAETTRQFLMSDKPLHLEAS
LDKEIYYHGEPISVNVHVTNNTNKTVKKIKISVRQYADIVLFNTAQYKVPVAMEEADDTVAPSSTFSKVYTLTPF
LANNREKRGLALDGKLKHEDTNLASSTLLREGANREILGIIVSYKVKVKLVVSRGGLLGDLASSDVAVELPFTLM
HPKPKEEPPHREVPENETPVDTNLIELDTNDDDIVFEDFARQRLKGMKDDKEEEEDGTGSPQLNNR

The truncated arrestin2$^{1\text{-}359}$ and arrestin2$^{1\text{-}393}$ constructs were obtained by introducing stop codons (TAA) into arrestin2$^{1\text{-}418}$ at position 360 and 394, respectively, via standard QuickChange polymerase chain reactions. For crystallization of the apo form, arrestin2$^{1\text{-}418}$ was obtained from a DNA construct cloned into a pGEX vector harboring an N-terminal GST tag followed by an HRV-3C cleavage site.

Protein expression and purification

CCR5

Non-phosphorylated wild-type full-length CCR5 and super-agonist chemokine [6P4]CCL5 were expressed in Sf9 cells and *E. coli*, respectively, and purified according to previous protocols (24,49).

Phosphorylated CCR5 was obtained by co-expression with untagged GRK2-CAAX in Sf9 insect cells following a protocol described for the phosphorylation of the β2-adrenergic receptor (6) and optimized as follows. Once the viability dropped to ~85–90% [44 hour post-infection (hpi)], cells were stimulated by addition of 500 nM [6P4]CCL5. After 2 h incubation at 37 °C, cells were harvested and kept at -80 °C until further use. The phosphorylation level was assayed using western blot analysis with phospho-specific CCR5 antibodies. Membrane preparation and receptor purification were carried out as previously described (49).

Arrestin

Arrestin$2^{1\text{-}418}$, arrestin$2^{1\text{-}393}$ and arrestin$2^{1\text{-}359}$ were expressed in *E. coli* BL21 (DE3) strain cultured in Lysogeny broth (LB). For the preparation of the deuterated ^{15}N-labeled arrestin$2^{1\text{-}393}$ NMR sample, cells were grown in D_2O/$^{15}NH_4Cl$ M9 minimal medium. Cells were grown at 37 °C until the optical density at 600 nm reached 0.7–0.8. Thereafter protein expression was induced by the addition of 25 μM isopropyl β-D-thiogalactopyranoside and the temperature lowered to 18 °C for an overnight expression, after which the cells were harvested by centrifugation. Proteins were purified on a Ni-NTA HiTrap HP column (GE Life Sciences) and the His tag was removed by overnight cleavage with TEV protease (homemade). The cleaved protein was further separated from impurities by a reverse IMAC step on a Ni-NTA HiTrap HP column, followed by concentration in a Vivaspin 20 concentrator [10-kDa MWCO (molecular weight cutoff)] and final gel filtration step on a HiLoad 16/600 Superdex 200 pg gel filtration column (GE Healthcare) equilibrated with 20 mM HEPES, 150 mM NaCl, pH 7.4 (SEC buffer I). The protein purity was confirmed by SDS-PAGE.

The GST-tagged arrestin$2^{1\text{-}418}$ was expressed under the same conditions, purified using a GST HiTrap column, followed by removal of the GST tag by homemade

PreScission protease and a final gel filtration step on a HiLoad 16/600 Superdex 200 pg gel filtration column.

Preparation of Fab30

Fab30 was purified as described previously (21), with slight modifications. Briefly, overnight a primary culture of Fab30-transformed *E. coli* M55244 strain was inoculated into 1 L 2x YT media (Himedia, Cat. no. G034) and allowed to grow for 8 hours at 30 °C. After 8 hours the culture was pelleted down and redissolved in 1 L CRAP media [7 mM $(NH_4)_2SO_4$, 14 mM KCl, 2.4 mM sodium citrate, 5.4 g/L yeast extract, 5.4 g/L casein hydrolyzates, 0.11 M MOPS buffer pH 7.3, 0.55% (w/v) glucose, 7 mM $MgSO_4$] grown for 16-18 h at 30 °C. Harvested cells were lysed in 20 mM HEPES (SRL, Cat. no. 63732), 100 mM NaCl, 0.5 mM $MgSO_4$, 0.5% Triton-X 100, pH 8.0. Cell debris was separated by high-speed centrifugation. After loading the cell lysate onto a Protein L beads (Capto™ L, GE Healthcare, Cat. no. 17-5478-02) column, nonspecific proteins were removed through extensive washing [20 mM HEPES, 100 mM NaCl, pH 8.0]. Bound protein was eluted in 0.1 M acetic acid, pH 3.0, and neutralized with 1 M HEPES, pH 8.0. Eluted protein was desalted using a PD10 column (GE Healthcare, Cat. no. 17085101) in 20 mM HEPES, 150 mM NaCl, pH 8.0. The purified protein was stored at -80 °C in 10% glycerol until further use.

Western blot analysis of CCR5 phosphorylation by phospho-specific CCR5 antibodies

All the following steps were performed at room temperature. 10 µL of non-phosphorylated or phosphorylated FLAG-purified CCR5 were separated by SDS-PAGE and transferred to a nitrocellulose membrane. Blots were then blocked with 1% BSA in Tris-Buffered Saline-Tween (TBST) for 1 h and then incubated with different primary (rabbit polyclonal) phospho-CCR5 antibodies [pS336/pS337-, pS342-, pT340-CCR5 (all 7TM antibodies) and pS349-CCR5 (Thermo Fisher)] at 1:2000 dilution for 1 h. Blots were washed twice for 5 min with TBST and incubated with HRP-coupled anti-rabbit secondary antibody (Thermo Fisher) at 1:5000 dilution for 1 h in the dark. Blots were then washed three times for 5 min with TBST and developed using western blotting substrate chemiluminescent detection.

Mass spectrometry

Sample preparation

The samples of non-phosphorylated and phosphorylated FLAG-purified CCR5 were prepared as technical triplicates. For this, 5 µg of the receptor was reduced and alkylated for 10 min at 95 °C in 50 µl of 1% sodium deoxycholate, 0.1 M ammonium bicarbonate, 10 mM TCEP, 15 mM chloroacetamide, pH 8.3. The sample was split into two, and the two halves were digested with either Sequencing Grade Modified Trypsin or endoproteinase Glu-C (both Promega, Madison, Wisconsin, enzyme:receptor=1:50 w/w) for 12 h at 37 °C. The samples were then acidified by the addition of 50 mM HCl (from 2 M HCl stock) and incubated for 15 min at 37 °C. Subsequently, the precipitated detergent was removed by centrifugation at 10,000g for 15 min and the peptides were separated from the reaction mixture using a C18 spin column (BioPureSPN MINI, The Nest Group, Inc.) according to the manufacturer's instructions. Samples were then dried under vacuum and stored at -80 °C until further use.

LC-MS analysis

For LC-MS analysis, the dried peptide samples were solubilized at a concentration of 1 pmol/µl in 98% water, 2% acetonitrile, 0.15% formic acid. 5 µl of each sample was then subjected to LC-MS analysis by a Q Exactive Plus mass spectrometer fitted with an EASY-nLC 1000 liquid chromatography system (both Thermo Fisher Scientific). Peptides were resolved using an EasySpray RP-HPLC column (75 µm × 25 cm) at a flow rate of 0.2 µL/min and a pre-column setup under a linear gradient ranging from 5% buffer B (80% acetonitrile, 0.1% formic acid in water) in 95% buffer A (0.1% formic acid in water) to 45% buffer B over 60 minutes. The mass spectrometer was operated in DDA (data-dependent acquisition) mode with a total cycle time of approximately 1 s. Each MS1 scan was followed by high-collision-dissociation (HCD) of the 20 most abundant precursor ions with the dynamic exclusion set to 5 seconds. For MS1, 3e6 ions were accumulated in the Orbitrap over a maximum time of 25 ms and scanned at a resolution of 70,000 FWHM (at 200 m/z). MS2 scans were acquired at a target setting of 1e5 ions, maximum accumulation time of 110 ms and resolution of 17,500 FWHM (at 200 m/z). Singly charged ions,

ions with charge state ≥ 6 and ions with unassigned charge state were excluded from triggering MS2 events. The normalized collision energy was set to 27%, the mass isolation window to 1.4 m/z, and one microscan was acquired for each spectrum.

In addition to the DDA LC-MS analysis, a targeted MS analysis was carried out focusing on the peptides with the phosphorylation sites of interest. For this, the sequence of CCR5_HUMAN was downloaded from uniprot.org (download 2022/10/22), imported into the Skyline software (https://skyline.ms/project/home/software/Skyline/begin.view), and the corresponding peptides manually phosphorylated at the expected sites of phosphorylation. The peptide ion masses containing 2^+ and 3^+ ions were exported as a mass isolation list from Skyline v21.2 and imported to the MS acquisition software. Targeted LC-MS analysis was carried out using the same settings as for DDA LC-MS with the following changes: the resolution of MS1 scans was reduced to 35,000 FWHM (at 200 m/z), the AGC target for MS2 scans was set to 3e6, the maximal fill time to 50 ms, and the mass isolation window to 0.4 m/z.

MS data analysis

The acquired raw files were converted to the mascot generic file (mgf) format using the msconvert tool [part of ProteoWizard, version 3.0.4624 (2013-6-3)]. Using the MASCOT algorithm (Matrix Science, Version 2.4.1), the mgf files were searched against a decoy database containing normal and reverse sequences of the predicted UniProt entries of *Spodoptera frugiperda* (www.ebi.ac.uk, release date 2020/10/22), the protein CCR5_HUMAN and commonly observed contaminants (in total 56,642 sequences) generated using the SequenceReverser tool from the MaxQuant software (version 1.0.13.13). The precursor ion tolerance was set to 10 ppm and fragment ion tolerance was set to 0.02 Da. The search criteria were set to requiring full trypsin specificity (cleavage after lysine or arginine unless followed by proline) and GluC specificity (cleavage after aspartate or glutamate unless followed by proline). At most 3 miscleavages were allowed, carbamidomethylation (C) was set as fixed modification, and phosphorylation (STY), oxidation (M) and acetylation (protein N-terminus) were set as variable modifications. The database search results were imported into Scaffold (version 5.1.0) and filtered to 1% FDR (false discovery rate) on the protein and peptide level using the built-in LFDR algorithm.

Trypsin proteolysis assay

Arrestin2 trypsin proteolysis assays were carried out by first incubating 100 µL of 20 µM arrestin2$^{1\text{-}418}$ with a 3-molar excess of the respective phosphopeptide at room temperature for 5-10 min in SEC buffer I. Thereafter, 1 ng of Trypsin Gold (Promega) was added, and the mixture incubated at 35 ℃ under gentle shaking (500 rpm). Samples for SDS-PAGE (10 µL sample mixed with 10 µL of 4x SDS loading buffer) were taken at 0, 5, 10, 20, 30, 60 and 90 min. The reaction was quenched by boiling the samples for 10 min at 95 ℃. Samples were then loaded on 4-20% precast gradient gels and visualized with Instant Blue protein stain (Abcam). Control reactions were run with apo arrestin2$^{1\text{-}418}$ and arrestin2$^{1\text{-}418}$ incubated with the phosphopeptide 0P.

NMR titrations

^{15}N-, ^{2}H-labeled arrestin2$^{1\text{-}393}$ (50 µM) NMR samples were prepared in SEC buffer I supplemented with 5% D_2O and 0.03% NaN_3 as 270-µl volumes in Shigemi tubes. The phosphopeptides were titrated into these samples to concentrations of 0–500 µM and the interactions monitored by ^{1}H-^{15}N HSQC-TROSY spectra recorded on a Bruker AVANCE 14.1 T (600 MHz) spectrometer equipped with a TCI cryoprobe at 303 K.

NMR data were processed with NMRPipe (51) and analyzed with NMRFAM-SPARKY (52). K_D values were obtained by nonlinear least-squares fitting using Matlab (Matlab_R2021b, MathWorks, Inc.) and the following equation:

$$\Delta\delta = \Delta\delta_{\max}\frac{[\mathrm{L}] + [\mathrm{P}] + \mathrm{K_D} - \sqrt{([\mathrm{L}] + [\mathrm{P}] + \mathrm{K_D})^2 - 4[\mathrm{P}][\mathrm{L}]}}{2[\mathrm{P}]} \qquad (1)$$

where $\Delta\delta = \delta_{\text{apo}}$-$\delta_{\text{bound}}$ is the difference in the ^{1}H^{N} or ^{15}N arrestin2 chemical shift, $\delta_{\max}$ is the difference between apo and the fully ligand-bound state, and [P] and [L] are the total protein (arrestin2) and ligand (phosphopeptide) concentrations, respectively.

CCR5 phosphopeptide•arrestin2•Fab30 complex formation and purification

Arrestin2$^{1\text{-}359}$ (30 µM) was incubated with a 5-molar excess of CCR5 phosphopeptides for 1 h on ice in SEC buffer II (20 mM HEPES, 150 mM NaCl, pH 6.9). Then, a 1.2-molar excess of Fab30 was added, followed by incubation for 1.5 h in the cold room. Samples were then concentrated in an Amicon concentrator (MWCO 50 kDa) and separated by SEC using a self-packed 4-ml S200 10/300 SEC column (length 25 mm, diameter 4.6 mm) and monitoring protein absorbance at 280 nm. The complex quality of each SEC fraction was evaluated by SDS-PAGE and visualized with Instant Blue protein stain. Fractions showing a fully formed complex were pooled.

X-ray crystallography

Both apo arrestin2 and its phosphopeptide/Fab30 complexes were crystallized by sitting drop vapor diffusion at room temperature from 1:1 mixtures of protein in SEC buffer II and crystallization buffer. Respective protein concentrations before mixing and crystallization buffers were: (i) 10 mg/ml apo arrestin2$^{1-418}$, 100 mM magnesium formate dihydrate, 100 mM bis-tris, 15% PEG3350, pH 7.0 and (ii) ~4-6 mg/ml arrestin2$^{1-359}$•Fab30•CCR5 phosphopeptide, 100 mM magnesium formate dihydrate, 100 mM bis-tris, 15% PEG3350, pH 7.0. In both cases, crystals formed and reached their final size within 24–48 h. Thereafter, they were quickly soaked in the crystallization buffer mixed with 20–25% ethylene glycol, and then flash-frozen in liquid nitrogen.

Diffraction data were collected at the Swiss Light Source, Paul Scherrer Institute, Villigen, Switzerland at beamline X06DA and X06SA, processed with XDS (53) and scaled using Aimless. For the 6P•arrestin2$^{1-359}$•Fab30 complex, a single data set was collected. For the 4P•arrestin2$^{1-359}$•Fab30 complex two data sets from the same crystal were merged, and for the apo arrestin2$^{1-418}$ four data sets of three different crystals were integrated. All structures were determined by molecular replacement with PHASER contained in the CCP4 software package (54) , using 4JQI (21) for complexes and 1G4M (38) for apo arrestin2 as search models. Models building was performed with COOT (Emsley and Cowtan 2004) and refinement with BUSTER-TNT (55) and PHENIX (56). The final models were evaluated with MolProbity and visualized with PyMOL. Data and refinement statistics are summarized in Table S1.

[6P4]CCL5•CCR5•arrestin2 complex formation and negative-stain EM analysis

Membranes containing phosphorylated CCR5 from 1 L Sf9 cell culture were solubilized in 50 mM HEPES, 400 mM NaCl, 0.5% LMNG, pH 7.4 and incubated with 1 mL M2 anti-FLAG resin in the cold room overnight. To form the [6P4]CCL5•CCR5•arrestin2 complex, the resin was then incubated with ~3-5 μM arrestin2$^{1\text{-}393}$ and 5 μM [6P4]CCL5 for 1 h. The complex was then stabilized by the addition of a 1.2-fold excess of Fab30, followed by a further 1-hour incubation. Thereafter, the resin was packed into a column and washed with 10 column volumes (CV) wash buffer (25 mM HEPES, 150 mM NaCl, 10 % glycerol, 0.01% LMNG, pH 7.4). Then the complex was eluted with 3 CV elution buffer (25 mM HEPES, 150 mM NaCl, 0.01% LMNG, 0.2 mg/ml FLAG peptide, pH 7.4) and further purified by SEC using a Superdex 200 Increase 10/300 GL column preequilibrated with SEC buffer III (10 mM HEPES, 150 mM NaCl, 0.01% LMNG, pH 7.4). The quality of each SEC fraction was assessed by SDS-PAGE. Fractions showing good complex integrity and purity were combined.

For negative-stain EM, the complex was diluted to 0.05 mg/ml in SEC buffer III. 5 μl of this solution was then applied onto freshly glow-discharged carbon-coated 300-mesh copper grids (produced in-house) and blotted with filter paper. The grids were stained with 2% (w/v) uranyl acetate for 30 s and imaged at a magnification of x135,000 on a FEI Tecnai G2 Spirit TEM operated at 80 kV and equipped with an EMSIS Veleta camera. 2D classification of single particles was carried out with CryoSPARC v.3.1.

NanoBiT assays

The receptor constructs including wild-type CCR5 and all phosphosite mutants were synthesized from GenScript and subcloned in pcDNA3.1(+) vector with an N-terminal FLAG tag. For receptor-based NanoBiT assay, receptor constructs bearing a carboxyl-terminus SmBiT spaced with a flexible linker were cloned in the lab using enzymes *Kpn*I and *Sma*I in pCAGGS vector.

For assessing agonist-induced arrestin recruitment, trafficking, and conformational variability of CCR5^{WT} and phosphosite mutants, a luciferase enzyme-linked complementation-based assay (NanoBiT assay) was used following the protocol described earlier (57). For arrestin2$^{1-418}$ recruitment, receptor-SmBiT (1.5 µg) and LgBiT-arrestin2$^{1-418}$ (1.5 µg) constructs were used (Figure 4B) to transfect HEK293 cells using polyethylenimine (PEI) with DNA:PEI ratio as 1:3. Similarly, for arrestin2$^{1-418}$ endosomal trafficking, receptor constructs (3 µg), SmBiT-arrestin2$^{1-418}$ (3.5 µg) and LgBiT-FYVE (3.5 µg) were used (Figure S3). Assessing the conformational diversity of arrestin2$^{1-418}$ bound to receptors was done by transfecting cells with receptors (3 µg), SmBiT-arrestin2$^{1-418}$ (2 µg) and LgBiT-Ib30 (5 µg) (Figure 4B). After transfection, all the NanoBiT-based assays follow a common set of steps. Briefly, 16 h post-transfection cells were trypsinized and harvested, followed by resuspension in assay buffer containing 1x HBSS (Gibco, Cat. no. 14065-056), 0.01% bovine serum albumin (BSA, SRL, Cat. no. 83803), 5 mM 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES), pH 7.4 with 10 µM coelenterazine (GoldBio, Cat. no. CZ05). Afterward, cells were seeded at a density of $1x10^5$ cells well^{-1} in a white 96-well plate and incubated for 1.5 h at 37 °C followed by 30 min at room temperature. After incubation baseline luminescence was measured using a multimode plate reader and then cells were stimulated with varying doses of [6P4]CCL5 followed by measurement of luminescence signal for 10-20 cycles and average data from 5th to 10th cycle were analyzed and presented using GraphPad Prism 9 (v3) software.

Receptor surface expression assay

In order to measure the surface expression of the receptors in different assays, we used a previously described whole cell-based surface ELISA assay (58). Post 24 h of transfection, cells were seeded into a 0.01% poly-D-Lysine precoated 24-well plate at a density of $2x10^5$ cells per well and allowed to adhere and grow for 24 h. The next day, cells were washed with ice-cold TBS, fixed with 4% PFA (w/v in TBS) on ice for 20 min. After fixing, cells were washed again three times with TBS and incubated in 1% BSA prepared in TBS at room temperature for 1.5 h. Thereafter, the cells were incubated with anti-FLAG M2-HRP antibody (Sigma, Cat. no. A8592) (1:5000, 1 h at room temperature) followed by three washes with 1% BSA in TBS. The plates were developed with TMB-ELISA substrate (Thermo Fisher Scientific, Cat. no. 34028) until the light blue color appeared. The reaction was quenched by transferring 100 µl of the colored solution to another 96-well plate already having 100 µl of 1 M H_2SO_4, and the absorbance was recorded at 450 nm. For normalization of the ELISA reading with the total cell content of each well, cells were incubated with 0.2% (w/v) Janus Green (Sigma, Cat. no. 201677) for 15 min at room temperature after washing twice with TBS. For removing excess stain cells were washed thoroughly with water followed by developing the satin with the addition of 800 µl of 0.5 N HCl in each well, of which 200 µl of this solution was transferred to a 96-well plate for measuring the absorbance at 595 nm. The ELISA signal was normalized from the A450/A595 ratio and the values were plotted using the GraphPad Prism 9 (v3).

Data availability

The following structural models for human arrestin2 have been deposited in the Protein Data Bank: 4P•arrestin2•Fab30 complex (PDB 8AS2), 6P•arrestin2•Fab30 complex (PDB 8AS3) and apo arrestin2 (PDB 8AS4). The CCR5 phosphoproteomics data have been deposited in the ProteomeXchange repository (PXD036220).

Acknowledgments

This work was supported by the Swiss National Science Foundation (grants 201270 and IZLIZ3-200298 to S.G., 179323 to T.M.), the Indo-Swiss grant from the Department of Biotechnology to A.K.S (IC-12044(11)/4/2021-ICD-D), and by a Fellowship for Excellence by the Biozentrum Basel International PhD Program to I.P. We gratefully acknowledge Dr. A. Schmidt and U.K. Lanner (Biozentrum Proteomics Core facility) for recording and analyzing proteomics data on phosphorylated CCR5, C. Alampi and Dr. M. Chami (Biozentrum, BioEM facility) for their support with electron microscopy, the beamline staff at Swiss Light Source (PSI) for crystallographic data collection and support at beamlines X06SA and X06DA, Dr. H. Dwivedi-Agnihotri and Dr. M. Baidya (IIT) for their help with cellular assays, M. Rogowski and I. Hertel (Biozentrum) for expression and purification of [6P4]CCL5 and proteases.

Author contributions

P.I., I.P., V.P., A.K.S., and S.G. conceived the study. I.P., P.I., A.R., V.P. and M.B. designed, expressed and purified proteins. P.I. prepared phosphorylated CCR5 samples and analyzed respective data. I.P. recorded NMR experiments and analyzed NMR data. R.P.J. and V.P. crystallized apo arrestin2. R.P.J. and P.I. crystallized arrestin2 in complex with 4P and 6P CCR5 phosphopeptides. P.I., I.P., S.G., and A.K.S. analyzed the structures with input from R.P.J. and T.M.. P.I. prepared the sample for negative stain EM, recorded and processed the EM data. A.R. designed the cellular experiments with guidance from A.K.S., P.S., and A.R. generated all the constructs and mutants for the cellular assays and performed their validation. P.I., I.P. and S.G. wrote the manuscript with input from A.K.S..

Declaration of interests

The authors have no competing interests.

4.1.7 Figures

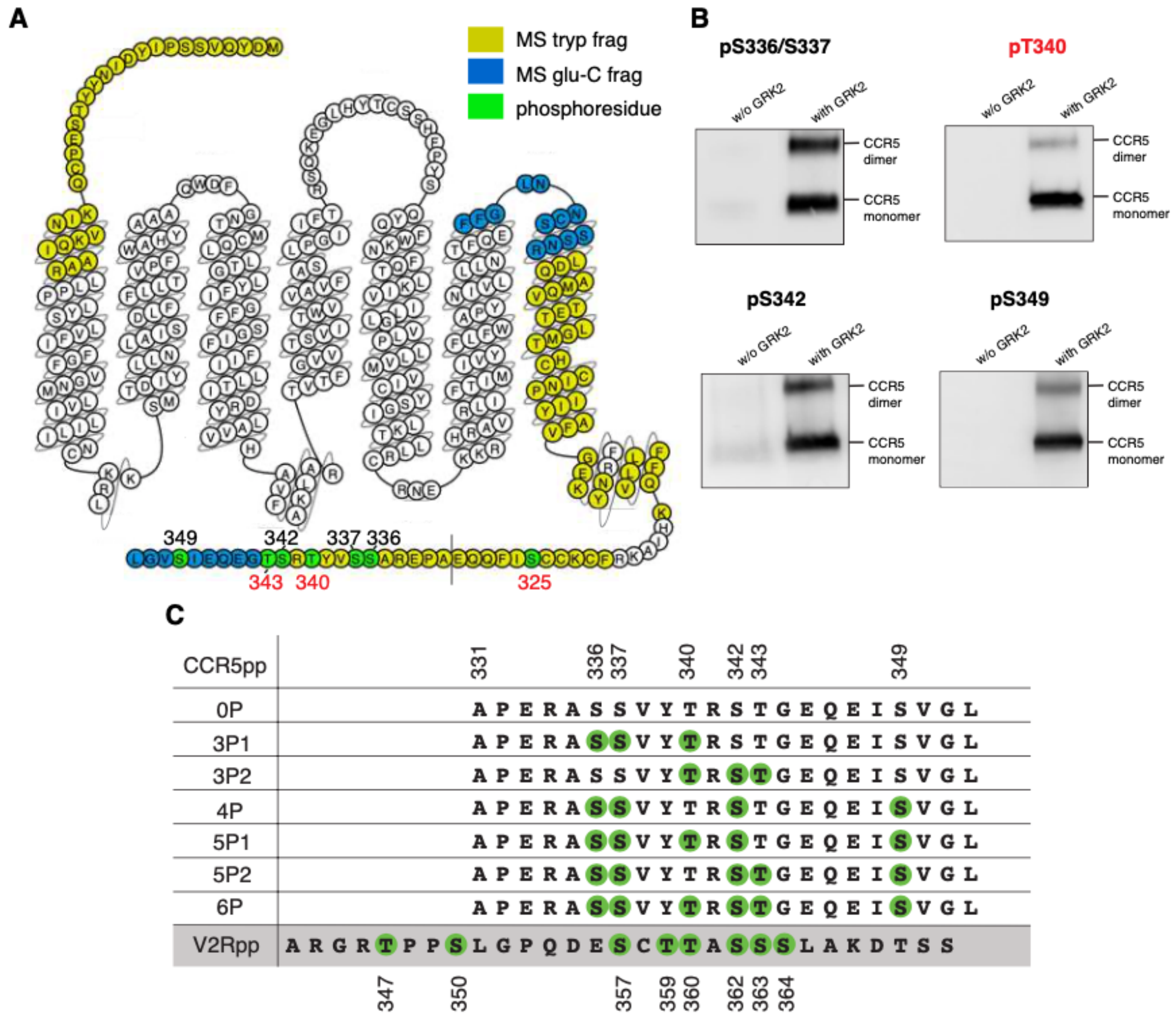
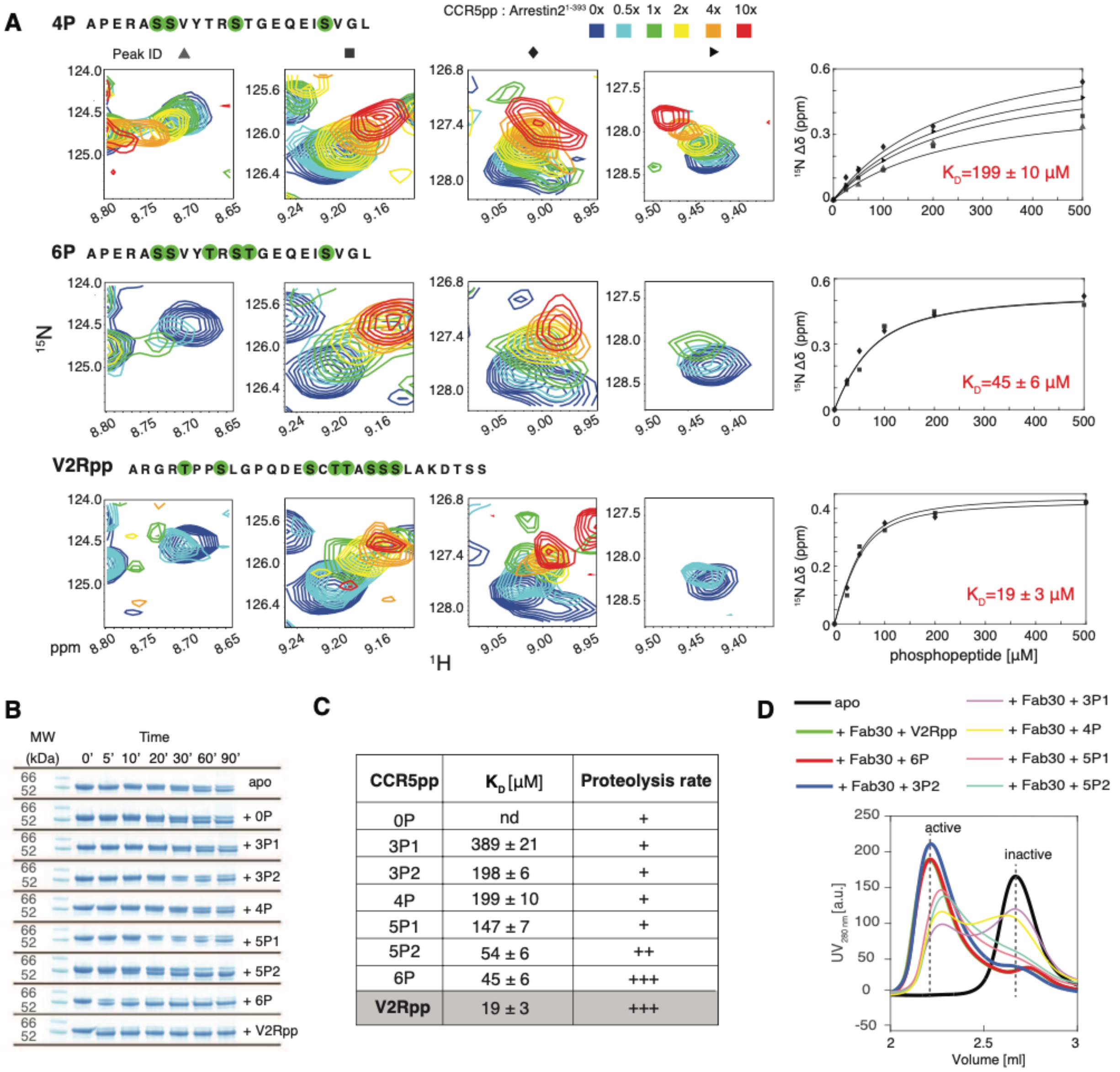


Figure 1. Detection of CCR5 phosphorylation and synthetic phosphopeptide design. (A) Mass spectrometric sequence coverage of GRK2-phosphorylated CCR5 using trypsin (yellow) and glu-C digestion (blue). Phosphosites detected by mass spectrometry and western blot analysis are indicated in green. (B) Western blot images of phosphorylated CCR5 detected with phosphosite-specific antibodies. (C) Designed synthetic phosphopeptides (0P–6P) corresponding to the last 22 residues of CCR5 based on the detected CCR5 phosphorylation. A further V2R C-terminal phosphopeptide (V2Rpp) was used as a control. Serine or threonine phosphorylation are indicated by a green circle.



CCR5pp	K$_D$ [µM]	Proteolysis rate
0P	nd	+
3P1	389 ± 21	+
3P2	198 ± 6	+
4P	199 ± 10	+
5P1	147 ± 7	+
5P2	54 ± 6	++
6P	45 ± 6	+++
V2Rpp	19 ± 3	+++

Figure 2. Arrestin2 interaction with CCR5 phosphopeptides. (A) Left: small regions of 1H-15N TROSY spectra showing resonance shifts of four selected arrestin21-393 residues upon CCR5 4P, 6P and V2Rpp binding. Right: detected chemical shift changes as a function of phosphopeptide concentration. Solid lines depict global non-linear least-square fits to the data points with respective dissociation constants (see Methods). Analogous fits for the other phosphopeptide titrations are shown in Figure S1. (B) Trypsin proteolysis assay of arrestin21-418 in apo form and in complexes with various phosphopeptides visualized by SDS-PAGE. (C) Summarized NMR titration and trypsin proteolysis results. Errors in KD represent one standard deviation. Tighter peptide binding is correlated with faster arrestin2 digestion. (D) SEC profiles showing arrestin2 activation in the presence of phosphopeptides and Fab30. The active arrestin2•phosphopeptide complexes, which are recognized by Fab30, elute at lower volumes than the inactive apo form. Arrestin2 in presence of CCR5 6P, 3P2 and V2Rpp elutes only as active complexes, whereas the other peptides lead to mixtures of inactive and active arrestin2 forms.

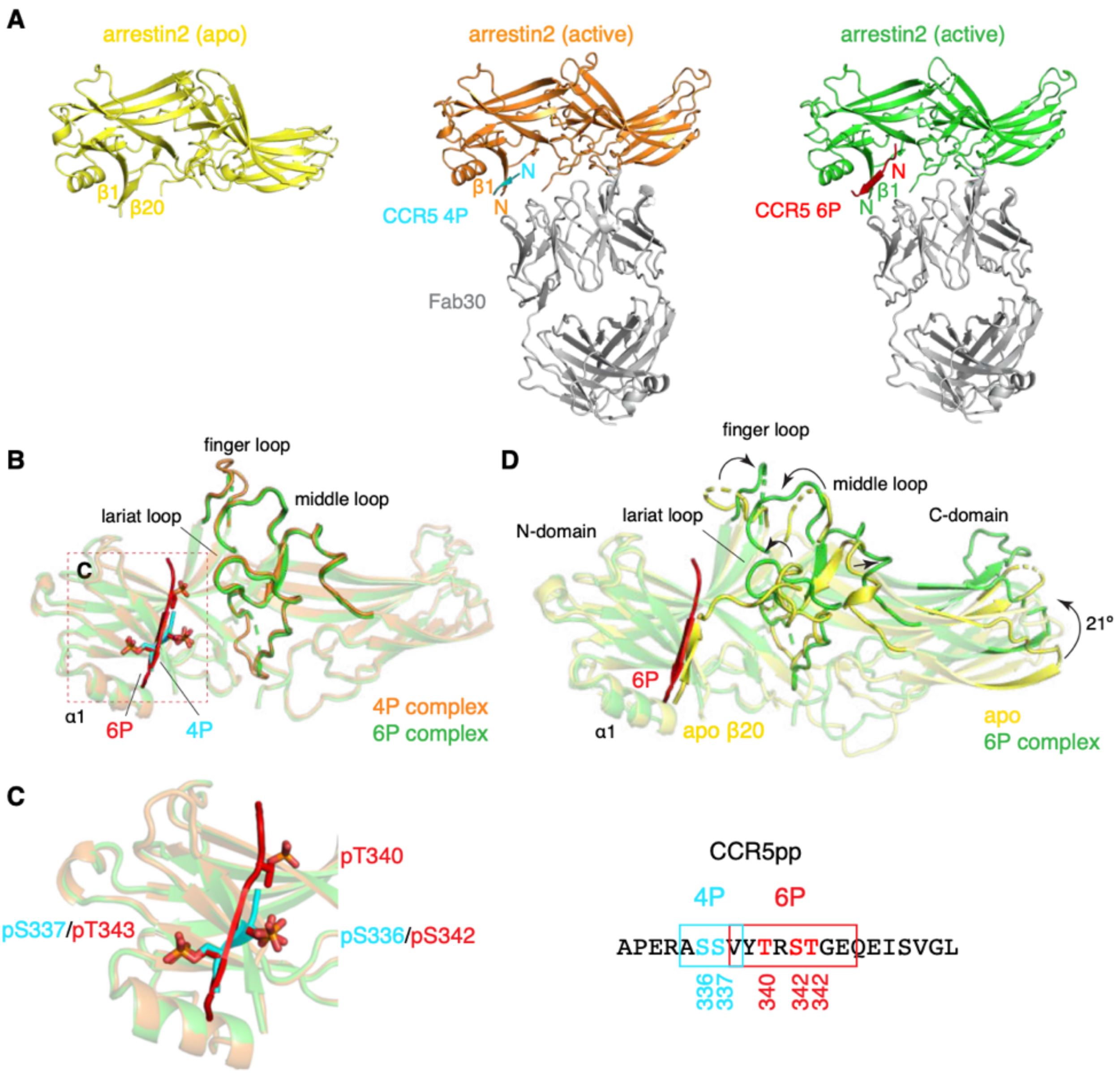


Figure 3. X-ray crystallographic structures of arrestin2 in apo form and in complexes with two CCR5 phosphopeptides. (A) Solved structures of apo arrestin2 (yellow), 4P•arrestin2 (arrestin2: orange, 4P: cyan), and 6P•arrestin2 (arrestin2: green, 6P: red) complexes. Both arrestin2 phosphopeptide complexes were stabilized with Fab30 (gray) (B) Overlay of the two arrestin2 complexes with the 4P (cyan) and 6P (red) phosphopeptides. (C) Detailed view of the peptide binding interfaces of both complexes (left) and alignment of the respective CCR5 C-terminal residues (right) showing a stable interaction in the electron density of the arrestin2 complexes. (D) Structural overlay of inactive, apo arrestin2 (yellow) and active arrestin2 (green) in complex with 6P (red). Salient arrestin2 conformational changes upon activation are indicated by black arrows.

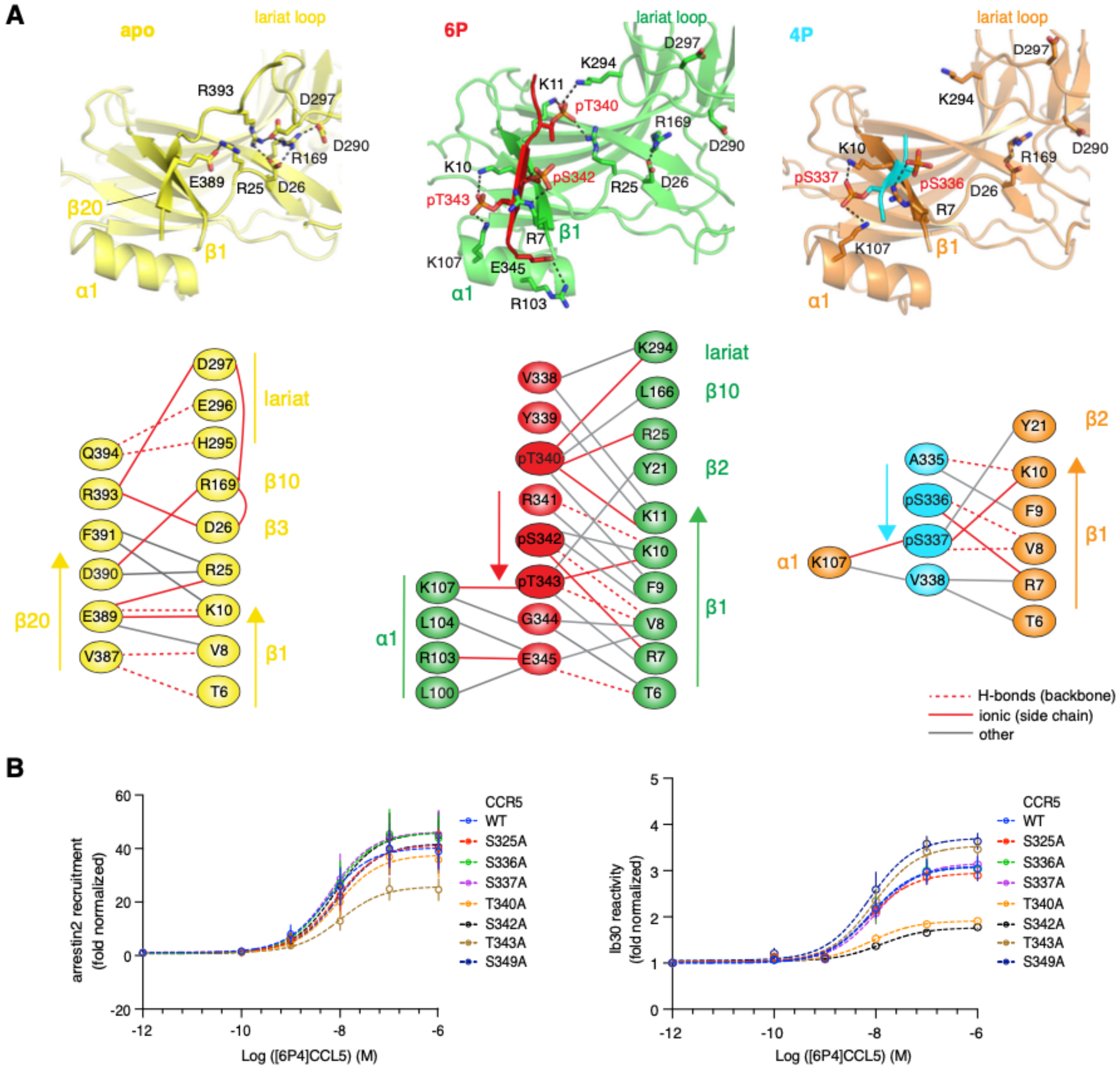


Figure 4. Detailed view of arrestin2 receptor C-terminal tail interactions. (A) Structural details of the arrestin2 N-domain in apo form (left) and in complexes with the CCR5 6P (center) and 4P phosphopeptides (right). Important residues stabilizing the inactive and active conformations are depicted in stick representation with key interactions as dashed lines. Upon arrestin2 activation by the peptides, the polar core is perturbed. Schematic diagrams of the key residue interactions are shown below each structural panel. The color coding follows Figure 3. (B) Super-agonist [6P4]CCL5-induced arrestin2 activation by CCR5 monitored with NanoBiT assays in HEK293 cells. Left: [6P4]CCL5-induced arrestin2 recruitment by wild-type CCR5 and seven S/A or T/A CCR5 C-terminal point mutants. Right: [6P4]CCL5-induced arrestin2 conformational changes monitored by the Ib30 assay for the same CCR5 constructs.

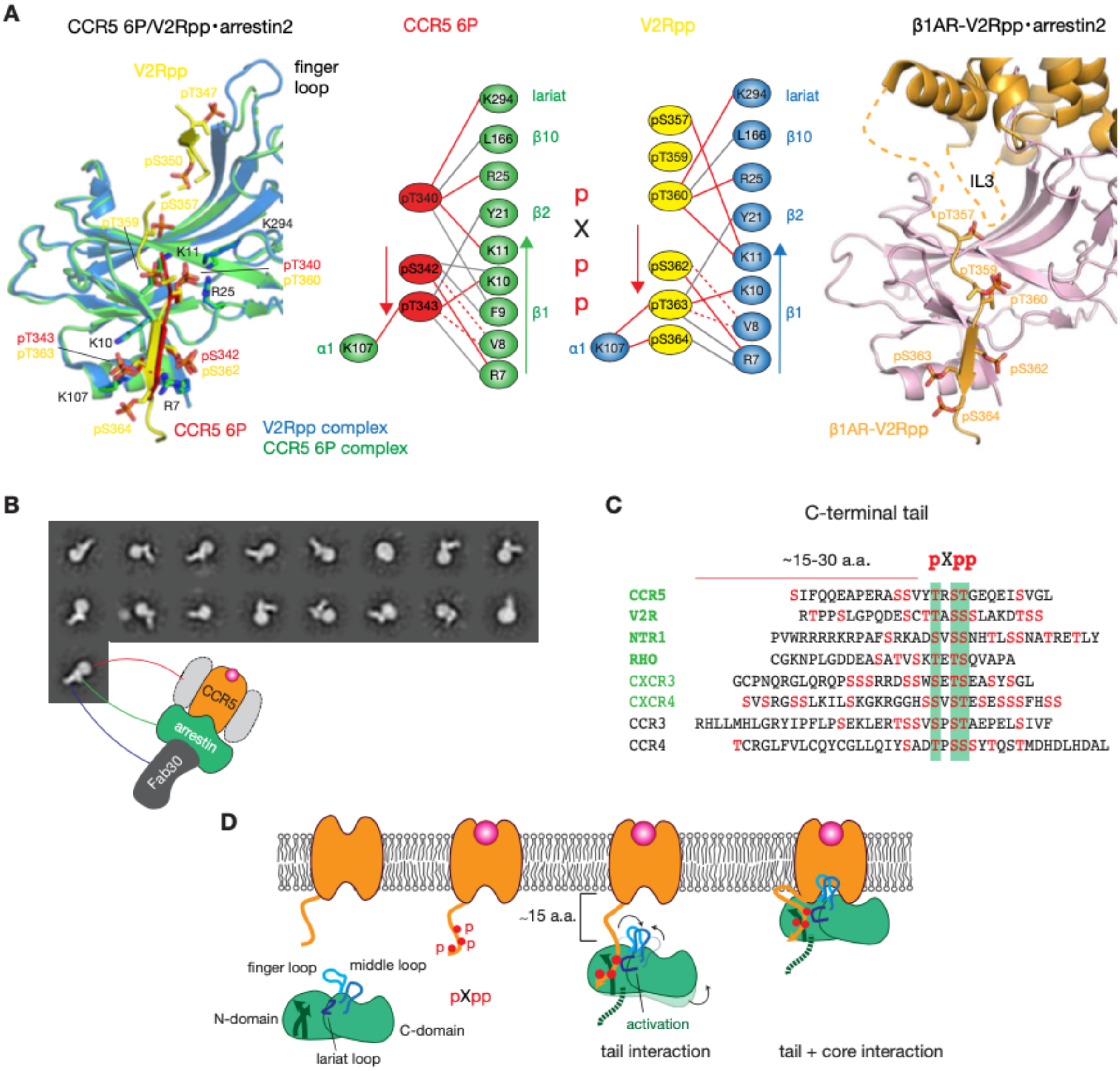


Figure 5. The pXpp motif is responsible for stable arrestin2 recruitment. (A) Left: Overlay of binding interfaces of arrestin2 with CCR5 6P (6P: red, arrestin2: green) and V2Rpp (V2Rpp: yellow, arrestin2: blue, PDB 4JQI). Phosphoresidues are shown in stick representation. Center: schematic diagrams of key 6P/V2R•arrestin2 interactions. Solid red lines represent polar interactions, dashed red lines H-bonds, gray lines other interactions. Right: V2Rpp•arrestin2 binding interface in the β1AR-V2Rpp•arrestin2 cryo-EM structure (PDB 6TKO). Unresolved parts of the ICL3 and C-terminal tail are depicted as dashed lines. (B) Single-particle analysis of negative-stain EM images of the [6P4]CCL5•CCR5•arrestin2•Fab30 complex. (C) Sequence alignment of C-terminal tails of several GPCRs (green), which are known to form a stable complex with arrestin2. Bold green indicates solved complex structures, light green functionally well-characterized receptors. Receptors indicated in black are examples of other not well characterized chemokine receptors harboring a pXpp motif. The GPCR C-terminal tail sequences were downloaded from the GPCRdb (44). The common pXpp motif (highlighted in green) is typically located about 15–30 residues downstream from receptor helix 8. (D) Schematic mechanism of arrestin2 (green) activation upon binding to an agonist(magenta)-stimulated class B GPCR (orange) containing a pXpp C-terminal cluster (red dots).

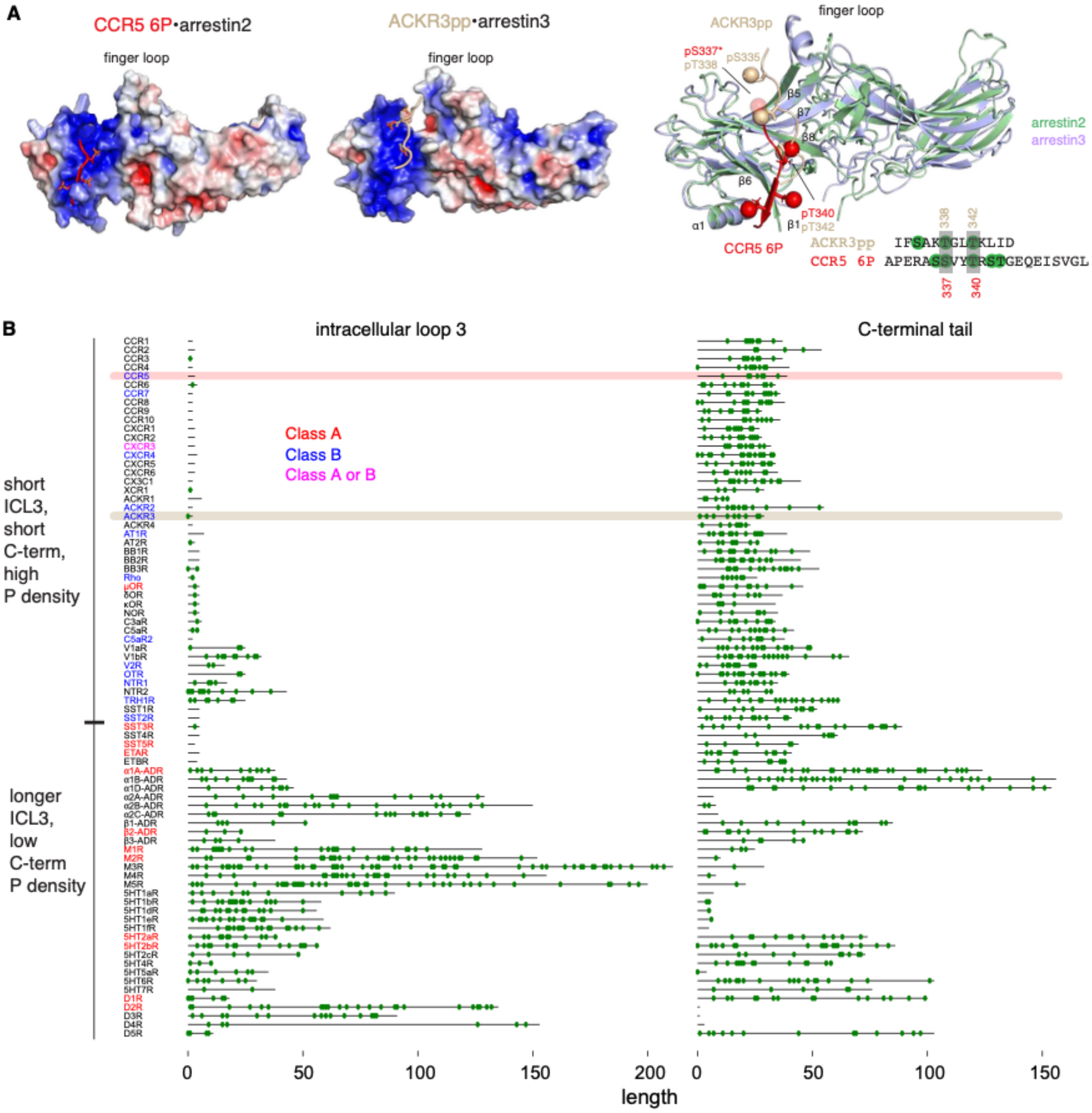


Figure 6. Structural differences in phosphopeptide recognition by arrestin2 and arrestin3. (A) Structural comparison of arrestin2 bound to the CCR5 6P phosphopeptide and arrestin3 bound to the ACKR3pp (PDB 6K3F). Left and center: arrestin electric surface charge densities (red: positive, blue: negative) of the 6P•arrestin2 and ACKR3pp•arrestin3 complexes. Right: superposition of the two complex structures (arrestin2: pale green, CCR5 6P: red, arrestin3: blue, ACKR3pp: wheat). Phosphates are shown as spheres (the position of the CCR5 pS337* phosphate is modeled). A sequence alignment is shown at the bottom with the overlapping phosphoresidue recognition sites indicated by gray boxes. (B) Sequence alignments of intracellular loops 3 (ICL3) and C-terminal tails of representative GPCRs. The amino acid sequences of GPCR ICL3 and C-terminal tails were downloaded from the GPCRdb (44). Potential phosphorylation sites (S and T) are shown as green circles. Receptors, which have been characterized for the type of arrestin binding (Table S2) are colored in red for class A, blue for class B, and purple for receptors, which exhibit isoform-dependent class A or B behavior. Sequences for CCR5 and ACKR3, for which the structures of arrestin complexes are shown in panel A, are highlighted.

4.1.8 References

1. Lefkowitz RJ, Shenoy SK. Transduction of Receptor Signals by ß-Arrestins. Science. 2005 Apr 22;308(5721):512–7.

2. Gurevich VV, Gurevich EV. GPCR Signaling Regulation: The Role of GRKs and Arrestins. Frontiers in Pharmacology [Internet]. 2019 [cited 2022 May 31];10. Available from: https://www.frontiersin.org/article/10.3389/fphar.2019.00125

3. Cahill TJ, Thomsen ARB, Tarrasch JT, Plouffe B, Nguyen AH, Yang F, et al. Distinct conformations of GPCR–β-arrestin complexes mediate desensitization, signaling, and endocytosis. Proceedings of the National Academy of Sciences. 2017 Mar 7;114(10):2562–7.

4. Kumari P, Srivastava A, Banerjee R, Ghosh E, Gupta P, Ranjan R, et al. Functional competence of a partially engaged GPCR–β-arrestin complex. Nat Commun. 2016 Nov 9;7(1):1–16.

5. Lohse MJ, Hoffmann C. Arrestin Interactions with G Protein-Coupled Receptors. In: Gurevich VV, editor. Arrestins - Pharmacology and Therapeutic Potential [Internet]. Berlin, Heidelberg: Springer; 2014 [cited 2022 May 31]. p. 15–56. (Handbook of Experimental Pharmacology). Available from: https://doi.org/10.1007/978-3-642-41199-1_2

6. Shukla AK, Westfield GH, Xiao K, Reis RI, Huang LY, Tripathi-Shukla P, et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature. 2014 Aug;512(7513):218–22.

7. Smith JS, Rajagopal S. The β-Arrestins: Multifunctional Regulators of G Protein-coupled Receptors. J Biol Chem. 2016 Apr 22;291(17):8969–77.

8. Srivastava A, Gupta B, Gupta C, Shukla AK. Emerging Functional Divergence of β-Arrestin Isoforms in GPCR Function. Trends in Endocrinology & Metabolism. 2015 Nov 1;26(11):628–42.

9. Latorraca NR, Masureel M, Hollingsworth SA, Heydenreich FM, Suomivuori CM, Brinton C, et al. How GPCR Phosphorylation Patterns Orchestrate Arrestin-Mediated Signaling. Cell. 2020 Dec 23;183(7):1813-1825.e18.

10. Tobin AB. G-protein-coupled receptor phosphorylation: where, when and by whom. British Journal of Pharmacology. 2008;153(S1):S167–76.

11. Nobles KN, Xiao K, Ahn S, Shukla AK, Lam CM, Rajagopal S, et al. Distinct Phosphorylation Sites on the β2-Adrenergic Receptor Establish a Barcode That Encodes Differential Functions of β-Arrestin. Science Signaling [Internet]. 2011 Aug 9 [cited 2022 Sep 16]; Available from: https://www.science.org/doi/10.1126/scisignal.2001707

12. Kang Y, Zhou XE, Gao X, He Y, Liu W, Ishchenko A, et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. Nature. 2015 Jul;523(7562):561–7.

13. Zhou XE, He Y, de Waal PW, Gao X, Kang Y, Van Eps N, et al. Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. Cell. 2017 Jul 27;170(3):457-469.e13.

14. Cao C, Barros-Álvarez X, Zhang S, Kim K, Dämgen MA, Panova O, et al. Signaling snapshots of a serotonin receptor activated by the prototypical psychedelic LSD. Neuron. 2022 Oct 5;110(19):3154-3167.e7.

15. Staus DP, Hu H, Robertson MJ, Kleinhenz ALW, Wingler LM, Capel WD, et al. Structure of the M2 muscarinic receptor–β-arrestin complex in a lipid nanodisc. Nature. 2020 Mar;579(7798):297–302.

16. Lee Y, Warne T, Nehmé R, Pandey S, Dwivedi-Agnihotri H, Chaturvedi M, et al. Molecular basis of β-arrestin coupling to formoterol-bound β1-adrenoceptor. Nature. 2020 Jul 30;583(7818):862–6.

17. Bous J, Fouillen A, Orcel H, Trapani S, Cong X, Fontanel S, et al. Structure of the vasopressin hormone–V2 receptor–β-arrestin1 ternary complex. Science Advances [Internet]. 2022 Sep [cited 2022 Sep 13]; Available from: https://www.science.org/doi/10.1126/sciadv.abo7761

18. Yin W, Li Z, Jin M, Yin YL, de Waal PW, Pal K, et al. A complex structure of arrestin-2 bound to a G protein-coupled receptor. Cell Res. 2019 Dec;29(12):971–83.

19. Huang W, Masureel M, Qu Q, Janetzko J, Inoue A, Kato HE, et al. Structure of the neurotensin receptor 1 in complex with β-arrestin 1. Nature. 2020 Mar;579(7798):303–8.

20. He QT, Xiao P, Huang SM, Jia YL, Zhu ZL, Lin JY, et al. Structural studies of phosphorylation-dependent interactions between the V2R receptor and arrestin-2. Nat Commun. 2021 Apr 22;12(1):2396.

21. Shukla AK, Manglik A, Kruse AC, Xiao K, Reis RI, Tseng WC, et al. Structure of active β-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. Nature. 2013 May;497(7447):137–41.

22. Min K, Yoon HJ, Park JY, Baidya M, Dwivedi-Agnihotri H, Maharana J, et al. Crystal Structure of β-Arrestin 2 in Complex with CXCR7 Phosphopeptide. Structure. 2020 Sep 1;28(9):1014-1023.e4.

23. Mayer D, Damberger FF, Samarasimhareddy M, Feldmueller M, Vuckovic Z, Flock T, et al. Distinct G protein-coupled receptor phosphorylation motifs modulate arrestin affinity and activation and global conformation. Nat Commun. 2019 Dec;10(1):1261.

24. Isaikina P, Tsai CJ, Dietz N, Pamula F, Grahl A, Goldie KN, et al. Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. Science Advances. 2021 Jun 1;7(25):eabg8685.

25. Scurci I, Martins E, Hartley O. CCR5: Established paradigms and new frontiers for a 'celebrity' chemokine receptor. Cytokine. 2018 Sep;109:81–93.

26. Alkhatib G. The biology of CCR5 and CXCR4: Current Opinion in HIV and AIDS. 2009 Mar;4(2):96–103.

27. Jiao X, Nawab O, Patel T, Kossenkov AV, Halama N, Jaeger D, et al. Recent Advances Targeting CCR5 for Cancer and Its Role in Immuno-Oncology. Cancer Research. 2019 Oct 1;79(19):4801–7.

28. Kraus S, Kolman T, Yeung A, Deming D. Chemokine Receptor Antagonists: Role in Oncology. Curr Oncol Rep. 2021 Sep 4;23(11):131.

29. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020 Aug;38(8):970–9.

30. Sharma D, Parameswaran N. Multifaceted role of β-arrestins in inflammation and disease. Genes Immun. 2015 Dec;16(8):499–513.

31. Martins E, Brodier H, Rossitto-Borlat I, Ilgaz I, Villard M, Hartley O. Arrestin Recruitment to C-C Chemokine Receptor 5: Potent C-C Chemokine Ligand 5 Analogs Reveal Differences in Dependence on Receptor Phosphorylation and Isoform-Specific Recruitment Bias. Mol Pharmacol. 2020 Nov 1;98(5):599–611.

32. Oppermann M, Mack M, Proudfoot AEI, Olbrich H. Differential Effects of CC Chemokines on CC Chemokine Receptor 5 (CCR5) Phosphorylation and Identification of Phosphorylation Sites on the CCR5 Carboxyl Terminus*. Journal of Biological Chemistry. 1999 Mar 26;274(13):8875–85.

33. Aramori I, Ferguson SS, Bieniasz PD, Zhang J, Cullen B, Cullen MG. Molecular mechanism of desensitization of the chemokine receptor CCR-5: receptor signaling and internalization are dissociable from its role as an HIV-1 co-receptor. EMBO J. 1997 Aug 1;16(15):4606–16.

34. Kleibeuker W, Jurado-Pueyo M, Murga C, Eijkelkamp N, Mayor Jr F, Heijnen CJ, et al. Physiological changes in GRK2 regulate CCL2-induced signaling to ERK1/2 and Akt but not to MEK1/2 and calcium. Journal of Neurochemistry. 2008;104(4):979–92.

35. Olbrich H, Proudfoot AEI, Oppermann M. Chemokine-induced phosphorylation of CC chemokine receptor 5 (CCR5). Journal of Leukocyte Biology. 1999;65(3):281–5.

36. Vroon A, Heijnen CJ, Kavelaars A. GRKs and arrestins: regulators of migration and inflammation. Journal of Leukocyte Biology. 2006;80(6):1214–21.

37. Xiao K, Shenoy SK, Nobles K, Lefkowitz RJ. Activation-dependent Conformational Changes in β-Arrestin 2*. Journal of Biological Chemistry. 2004 Dec 31;279(53):55744–53.

38. Han M, Gurevich VV, Vishnivetskiy SA, Sigler PB, Schubert C. Crystal Structure of β-Arrestin at 1.9 Å: Possible Mechanism of Receptor Binding and Membrane Translocation. Structure. 2001 Sep 1;9(9):869–80.

39. Kang DS, Kern RC, Puthenveedu MA, von Zastrow M, Williams JC, Benovic JL. Structure of an Arrestin2-Clathrin Complex Reveals a Novel Clathrin Binding Domain That Modulates Receptor Trafficking. J Biol Chem. 2009 Oct 23;284(43):29860–72.

40. Baidya M, Kumari P, Dwivedi-Agnihotri H, Pandey S, Sokrat B, Sposini S, et al. Genetically encoded intrabody sensors report the interaction and trafficking of β-arrestin 1 upon activation of G-protein–coupled receptors. Journal of Biological Chemistry. 2020 Jul 24;295(30):10153–67.

41. Baidya M, Kumari P, Dwivedi-Agnihotri H, Pandey S, Chaturvedi M, Stepniewski TM, et al. Key phosphorylation sites in GPCR s orchestrate the contribution of β-Arrestin 1 in ERK 1/2 activation. EMBO Rep [Internet]. 2020 Sep 3 [cited 2022 Jul 20];21(9). Available from: https://onlinelibrary.wiley.com/doi/10.15252/embr.201949886

42. Dwivedi-Agnihotri H, Chaturvedi M, Baidya M, Stepniewski TM, Pandey S, Maharana J, et al. Distinct phosphorylation sites in a prototypical GPCR differently orchestrate β-arrestin interaction, trafficking, and signaling. Sci Adv. 2020 Sep 11;6(37):eabb8368.

43. Shiraishi Y, Kofuku Y, Ueda T, Pandey S, Dwivedi-Agnihotri H, Shukla AK, et al. Biphasic activation of β-arrestin 1 upon interaction with a GPCR revealed by methyl-TROSY NMR. Nat Commun. 2021 Dec 9;12(1):7158.

44. Isberg V, de Graaf C, Bortolato A, Cherezov V, Katritch V, Marshall FH, et al. Generic GPCR residue numbers – aligning topology maps while minding the gaps. Trends in Pharmacological Sciences. 2015 Jan 1;36(1):22–31.

45. Scott MGH, Le Rouzic E, Périanin A, Pierotti V, Enslen H, Benichou S, et al. Differential Nucleocytoplasmic Shuttling of β-Arrestins: CHARACTERIZATION OF A LEUCINE-RICH NUCLEAR EXPORT SIGNAL IN β-ARRESTIN2*. Journal of Biological Chemistry. 2002 Oct 4;277(40):37693–701.

46. Wang P, Wu Y, Ge X, Ma L, Pei G. Subcellular Localization of β-Arrestins Is Determined by Their Intact N Domain and the Nuclear Export Signal at the C Terminus*. Journal of Biological Chemistry. 2003 Mar 28;278(13):11648–53.

47. Oakley RH, Laporte SA, Holt JA, Caron MG, Barak LS. Differential Affinities of Visual Arrestin, βArrestin1, and βArrestin2 for G Protein-coupled Receptors Delineate Two Major Classes of Receptors*. Journal of Biological Chemistry. 2000 Jun 2;275(22):17201–10.

48. Asher WB, Terry DS, Gregorio GGA, Kahsai AW, Borgia A, Xie B, et al. GPCR-mediated β-arrestin activation deconvoluted with single-molecule precision. Cell. 2022 May 12;185(10):1661-1675.e16.

49. Isaikina P, Tsai CJ, Petrovic I, Rogowski M, Dürr AM, Grzesiek S. Preparation of a stable CCL5·CCR5·Gi signaling complex for Cryo-EM analysis. Methods Cell Biol. 2022;169:115–41.

50. Inglese J, Koch WJ, Caron MG, Lefkowitz RJ. Isoprenylation in regulation of signal transduction by G-protein-coupled receptor kinases. Nature. 1992 Sep;359(6391):147–50.

51. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A. NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR. 1995 Nov 1;6(3):277–93.

52. Lee W, Tonelli M, Markley JL. NMRFAM-SPARKY: enhanced software for biomolecular NMR spectroscopy. Bioinformatics. 2015 Apr 15;31(8):1325–7.

53. Kabsch W. XDS. Acta Crystallogr D Biol Crystallogr. 2010 Feb 1;66(Pt 2):125–32.

54. Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, et al. Overview of the CCP4 suite and current developments. Acta Cryst D. 2011 Apr 1;67(4):235–42.

55. Bricogne G, Blanc E, Brandl M, Flensburg C, Keller P, Paciorek W, et al. BUSTER version 2.10. 0. Global Phasing Ltd, Cambridge, UK. 2011;

56. Adams PD, Grosse-Kunstleve RW, Hung LW, Ioerger TR, McCoy AJ, Moriarty NW, et al. PHENIX: building new software for automated crystallographic structure determination. Acta Cryst D. 2002 Nov 1;58(11):1948–54.

57. Inoue A, Raimondi F, Kadji FMN, Singh G, Kishi T, Uwamizu A, et al. Illuminating G-Protein-Coupling Selectivity of GPCRs. Cell. 2019 Jun 13;177(7):1933-1947.e25.

58. Pandey S, Roy D, Shukla AK. Chapter 8 - Measuring surface expression and endocytosis of GPCRs using whole-cell ELISA. In: Shukla AK, editor. Methods in Cell Biology [Internet]. Academic Press; 2019 [cited 2022 Jul 11]. p. 131–40. (G Protein-Coupled Receptors, Part B; vol. 149). Available from: https://www.sciencedirect.com/science/article/pii/S0091679X18301377

4.2 Supplementary Material

Supplementary Tables

Table S1. Statistics on diffraction data and refinement of human arrestin2 (apo) and CCR5pp•arrestin2•Fab30 complexes.

	4P•arrestin2•Fab	6P•arrestin2•Fab	human arrestin2
PDB Identifier	8AS2	8AS3	8AS4
Wavelength (Å)	1.00000	0.99999	0.999999
Resolution range	46.8 – 3.2	46.5- 3.50	44.75 – 2.30
Space group	I 21 21 21	I 21 21 21	P 1 21 1
Unit cell	116.8 122.6	116.3 121.1	62.0 71.4 115.4
α, β, γ (°)	90.0 90 90.0	90.0 90.0 90.0	90.0 97.7 90.0
Total reflections	461,526 (42,112)	178,306 (17,607)	1,216,192
Unique reflections	17,524 (1,702)	13,152 (1,295)	44,486 (4,409)
Multiplicity	26.3 (26.8)	13.6 (13.6)	27.3 (27.0)
Completeness (%)	99.9 (100)	98.3 (90.3)	99.8 (99.0)
Mean I/sigma(I)	11.6 (1.1)	12.2 (0.7)	12.1 (1.6)
Wilson B-factor	120.1	171.8	48.6
R-merge (%)	22.0 (399.2)	14.4 (525.4)	24.0 (246.1)
CC1/2	0.998 (0.662)	0.999 (0.317)	0.999 (0.790)
Reflections used in	17,439 (1,691)	13,005 (1,171)	44,388 (4,397)
R-work	0.281 (0.428)	0.283 (0.442)	0.224 (0.361)
R-free	0.317 (0.488)	0.335 (0.454)	0.252 (0.407)
Number of atoms	6,096	6,067	5,892
Protein residues	788	783	325
RMSD(bonds) (Å)	0.003	0.004	0.011
RMSD(angles) (°)	0.68	0.75	1.66
Ramachandran	94.0	96.0	96.5
Ramachandran	0	0	0
Rotamer	0	0	0.5
Clash score	2.30	4.87	3.78
Average B-factor	197.4	199.9	80.3

*The values in parentheses correspond to the highest resolution shell.

Table S2. GPCR classification based of the characteristics of agonist-dependent arrestin interactions.

Receptor name	**Class**	**Method**	**Reference**
α1B-ADR	A	Functional assays	(1)
β2-ADR	A	Functional assays	(1)
μOR	A	Functional assays	(1)
ETAR	A	Functional assays	(1)
D1R	A	Functional assays	(1)
D2R	A	Functional assays	(2)
5HT2aR	A	Functional assays	(3)
5HT2bR	A	Functional and structural data	(4)
M1R	A	FRET	(5)
M2R	A	Functional and structural data on M2R-V2Rpp (indirect evidence*)*	(6,7)
SST3R	A	Functional assays	(8)
SST5R	A	Functional assays	(8)
CXCR3B	A	Functional assays	(9)
CXCR3A	B	Functional assays	(9)
CXCR4	B	Functional assays	(10)
CCR5	B	Functional assays	(10)
CCR7	B	Functional assays	(11)
ACKR2	B	Functional assays	(12)
ACKR3	B	Functional assays and structural data	(13,14)
V2R	B	Functional assays and structural data	(1)
NTR1	B	Functional assays and structural data	(1)
C5aR2	B	Functional assays	(12)
Rho	B	Biochemical and biophysical data	(15)
AT1R	B	Functional assays	(1)
OTR	B	Functional assays	(1)
SST2R	B	Functional assays	(8)
TRH1R	B	Functional assays	(1)

Supplementary Figures

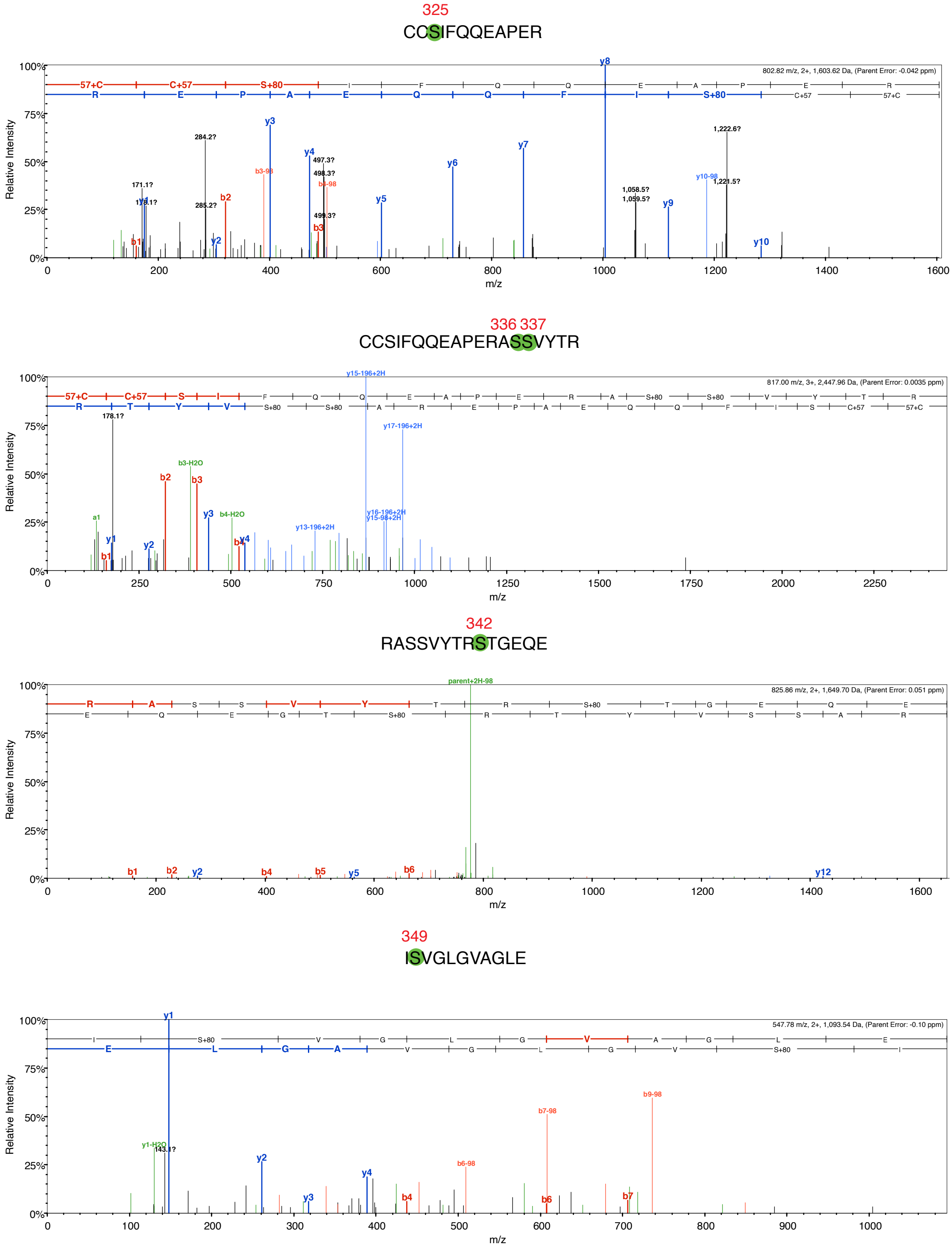


Figure S1. Mass spectra used for detection of CCR5 phosphorylation induced by GRK2. Representative spectra of the identified peptides bearing phosphorylated residues. Residues with higher occupancies are highlighted by green circles.

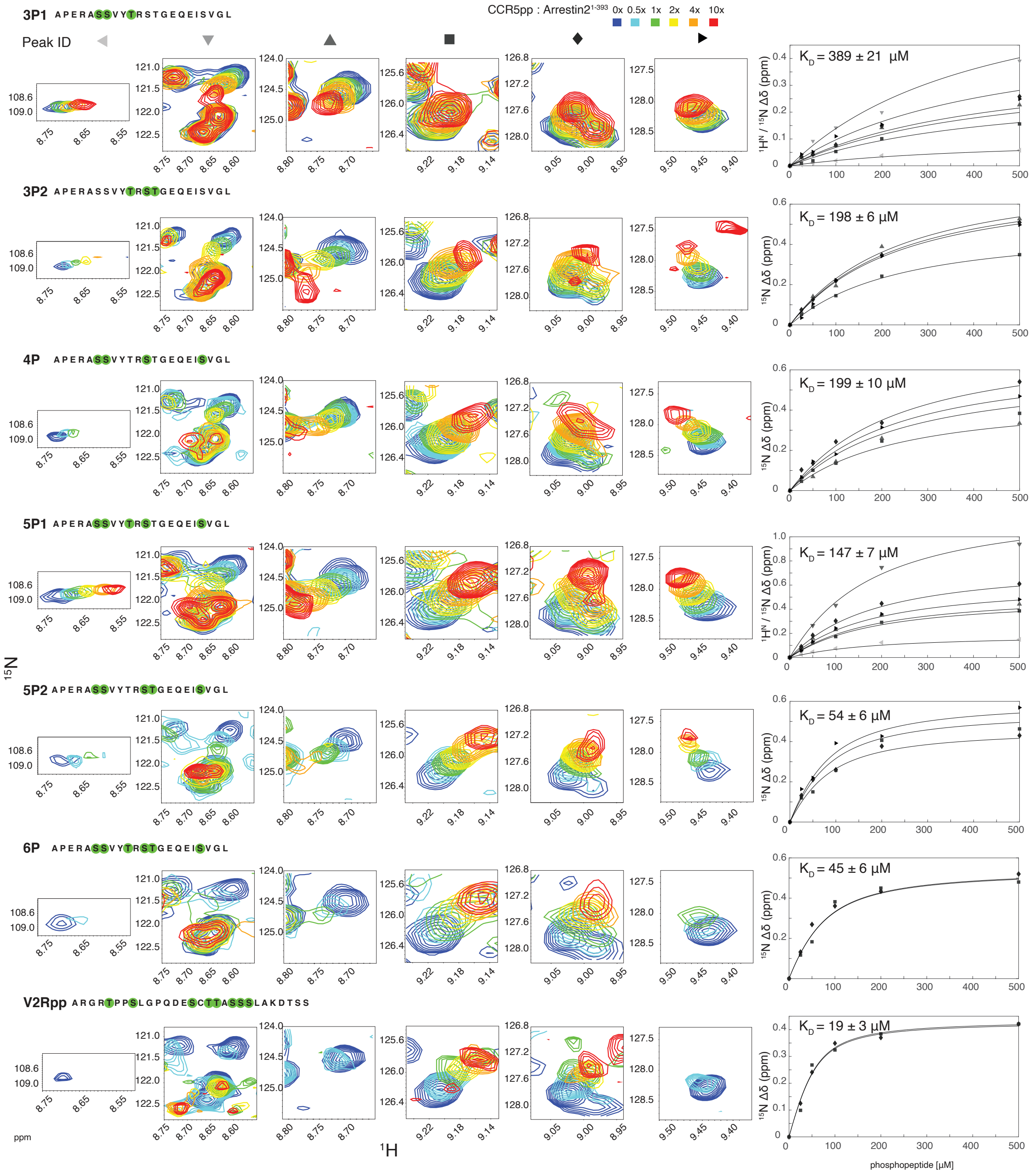


Figure S2. NMR titration of arrestin21-393 with CCR5 phosphopeptides. Left: small regions of 1H-15N TROSY spectra showing resonance shifts of four selected arrestin21-393 residues upon CCR5 phosphopeptide binding. The individual peptides are indicated at the top of each row. Right: detected chemical shift changes as a function of phosphopeptide concentration. Solid lines depict global non-linear least-square fits to the data points with respective dissociation constants (see Methods).

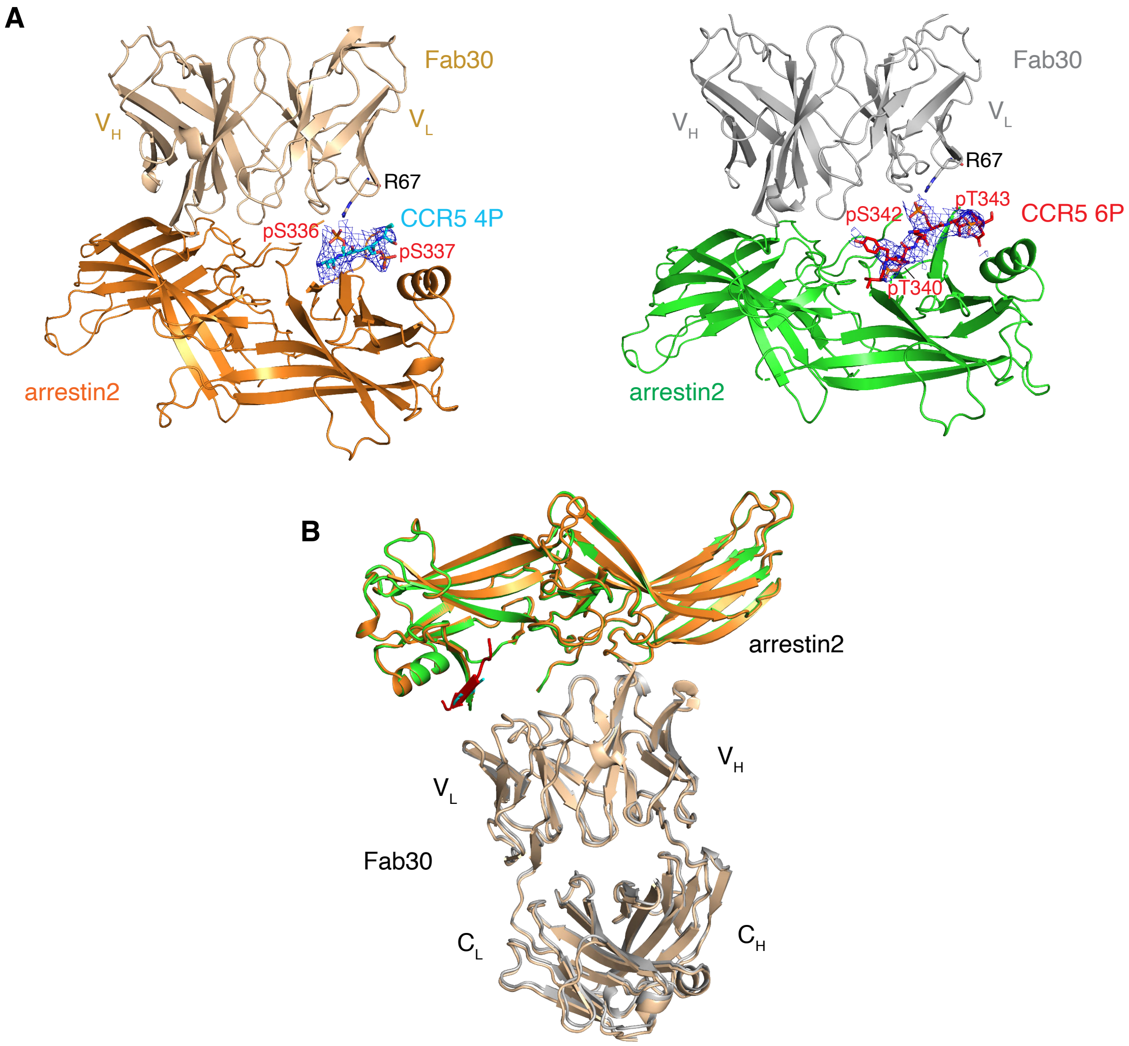


Figure S3. Phosphopeptide electron density and Fab30 coordination in arrestin2•CCR5pp crystal structures. (A) 2Fo-Fc electron density maps of the CCR5 4P (left) and 6P (right) phosphopeptides and contacts between the phosphopeptides and Fab30. The electron densities are contoured at 1.0 σ cut-off within 6 Å and displayed as a blue mesh. (B) Overall coordination of the 6P•arrestin2 and 4P•arrestin2 complexes with Fab30. Fab30 is colored in beige in the 4P complex and in light gray in the 6P complex. Only very minor conformation changes of Fab30 upon binding to the two CCR5 phosphopeptides are detectable.

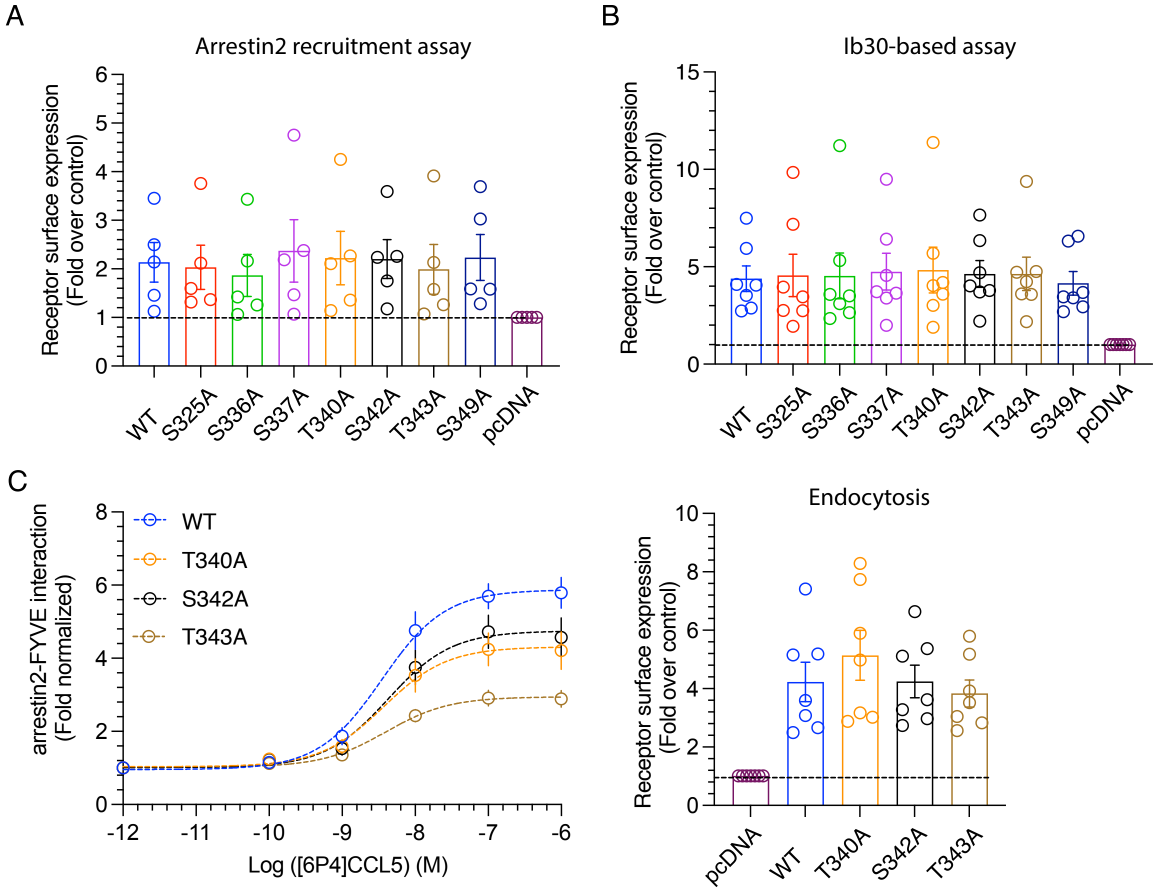


Figure S4. Arrestin2 trafficking assay and surface expression of the various CCR5 phospho-deficient mutants in different cellular assays. (A) Surface expression levels of wild-type CCR5 and all S/A and T/A CCR5 mutants used in the arrestin2 recruitment NanoBiT assay. (B) Surface expression levels of wild-type CCR5 and all S/A and T/A CCR5 mutants used in the Ib30-based NanoBiT assay. (C) Left: arrestin2 trafficking assay on wild-type CCR5 and selected CCR5 mutants (T340A, S342A and T343A) in response to [6P4]CCL5 stimulation. Right: surface expression of the CCR5 constructs used in this assay.

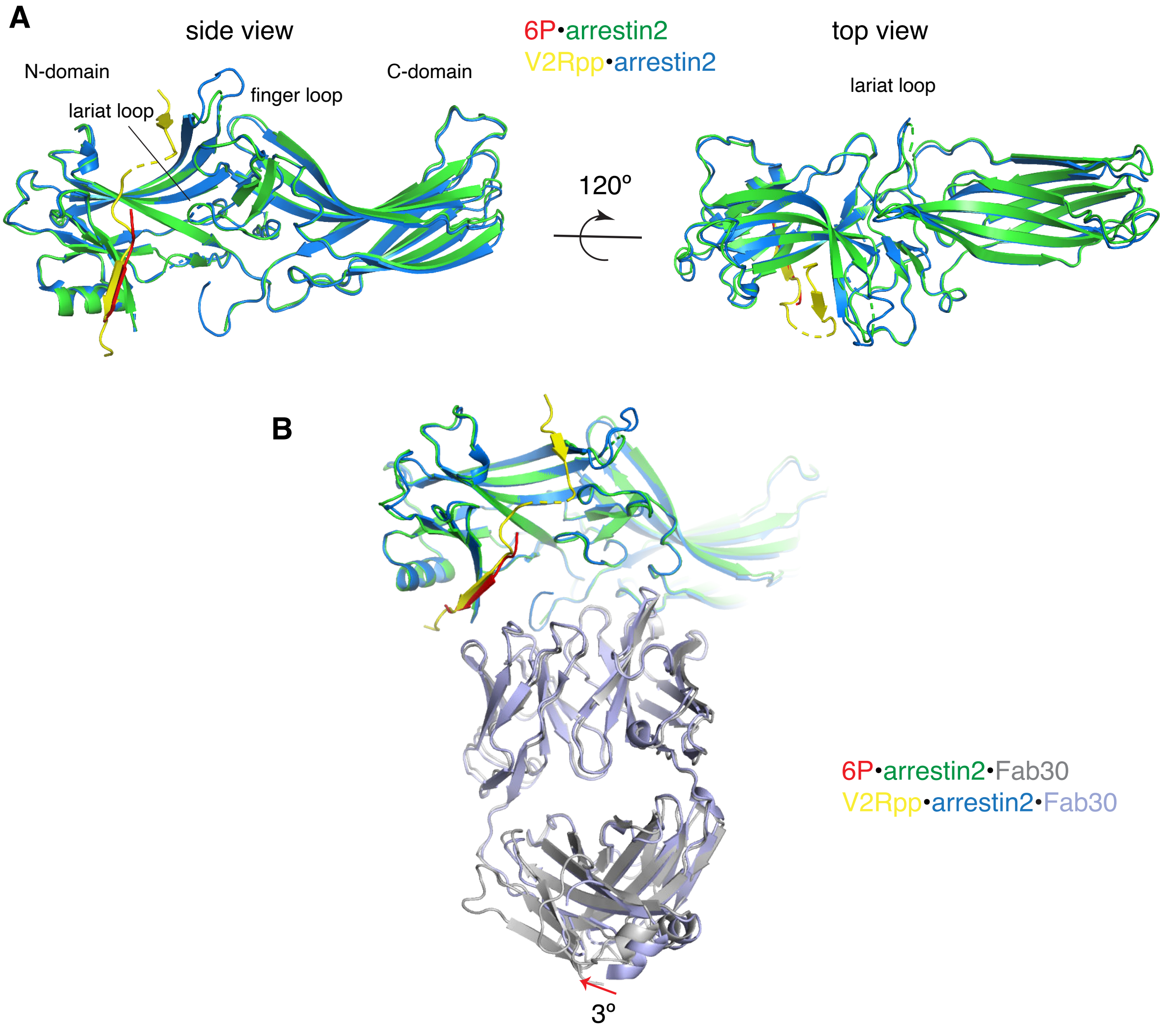


Figure S5. Overall structural comparison between 6P•arrestin2•Fab30 and V2Rpp•arrestin2•Fab30. (A) Superposition of 6P•arrestin2•Fab30 and V2Rpp•arrestin2•Fab30 structures in side and top views. The color code is indicated and identical to Figure 5A. (B) Overall Fab30 binding in both structures. The color code is indicated and follows panel A.

Supplementary References

1. Oakley RH, Laporte SA, Holt JA, Caron MG, Barak LS. Differential Affinities of Visual Arrestin, βArrestin1, and βArrestin2 for G Protein-coupled Receptors Delineate Two Major Classes of Receptors*. Journal of Biological Chemistry. 2000 Jun 2;275(22):17201–10.

2. Asher WB, Terry DS, Gregorio GGA, Kahsai AW, Borgia A, Xie B, et al. GPCR-mediated β-arrestin activation deconvoluted with single-molecule precision. Cell. 2022 May 12;185(10):1661-1675.e16.

3. Schmid CL, Raehal KM, Bohn LM. Agonist-directed signaling of the serotonin 2A receptor depends on β-arrestin-2 interactions in vivo. Proceedings of the National Academy of Sciences. 2008 Jan 22;105(3):1079–84.

4. Cao C, Barros-Álvarez X, Zhang S, Kim K, Dämgen MA, Panova O, et al. Signaling snapshots of a serotonin receptor activated by the prototypical psychedelic LSD. Neuron [Internet]. 2022 Sep 9 [cited 2022 Sep 20]; Available from: https://www.sciencedirect.com/science/article/pii/S0896627322007528

5. Jung SR, Kushmerick C, Seo JB, Koh DS, Hille B. Muscarinic receptor regulates extracellular signal regulated kinase by two modes of arrestin binding. Proceedings of the National Academy of Sciences. 2017 Jul 11;114(28):E5579–88.

6. Staus DP, Wingler LM, Choi M, Pani B, Manglik A, Kruse AC, et al. Sortase ligation enables homogeneous GPCR phosphorylation to reveal diversity in β-arrestin coupling. Proceedings of the National Academy of Sciences. 2018 Apr 10;115(15):3834–9.

7. Staus DP, Hu H, Robertson MJ, Kleinhenz ALW, Wingler LM, Capel WD, et al. Structure of the M2 muscarinic receptor–β-arrestin complex in a lipid nanodisc. Nature. 2020 Mar;579(7798):297–302.

8. Tulipano G, Stumm R, Pfeiffer M, Kreienkamp HJ, Höllt V, Schulz S. Differential β-Arrestin Trafficking and Endosomal Sorting of Somatostatin Receptor Subtypes*. Journal of Biological Chemistry. 2004 May 14;279(20):21374–82.

9. Smith JS, Alagesan P, Desai NK, Pack TF, Wu JH, Inoue A, et al. C-X-C Motif Chemokine Receptor 3 Splice Variants Differentially Activate Beta-Arrestins to Regulate Downstream Signaling Pathways. Mol Pharmacol. 2017 Aug 1;92(2):136–50.

10. Liebick M, Henze S, Vogt V, Oppermann M. Functional consequences of chemically-induced β-arrestin binding to chemokine receptors CXCR4 and CCR5 in the absence of ligand stimulation. Cellular Signalling. 2017 Oct;38:201–11.

11. Zidar DA, Violin JD, Whalen EJ, Lefkowitz RJ. Selective engagement of G protein coupled receptor kinases (GRKs) encodes distinct functions of biased ligands.

Proceedings of the National Academy of Sciences. 2009 Jun 16;106(24):9649–54.

12. Pandey S, Kumari P, Baidya M, Kise R, Cao Y, Dwivedi-Agnihotri H, et al. Intrinsic bias at non-canonical, β-arrestin-coupled seven transmembrane receptors. Molecular Cell. 2021 Nov 18;81(22):4605-4621.e11.

13. Rajagopal S, Kim J, Ahn S, Craig S, Lam CM, Gerard NP, et al. β-arrestin- but not G protein-mediated signaling by the "decoy" receptor CXCR7. Proceedings of the National Academy of Sciences. 2010 Jan 12;107(2):628–32.

14. Sarma P, Yoon HJ, Carino CMC, S D, Banerjee R, Yun Y, et al. Molecular insights into intrinsic transducer-coupling bias in the CXCR4-CXCR7 system. bioRxiv. 2022 Jun 7;2022.06.06.494935.

15. Mayer D, Damberger FF, Samarasimhareddy M, Feldmueller M, Vuckovic Z, Flock T, et al. Distinct G protein-coupled receptor phosphorylation motifs modulate arrestin affinity and activation and global conformation. Nat Commun. 2019 Dec;10(1):1261.

5 CCR5•arrestin2 complex assembly and its optimization for single-particle cryo-EM analysis

5.1.1 Introduction

The previous chapter has described the molecular basis of arrestin2 coupling to the phosphorylated C-terminal tail of CCR5, suggesting key elements needed for stable arrestin2 recruitment specific for arrestin-class B GPCRs. Once bound to the phosphorylated receptor tail, arrestin may further engage with the intracellular core of the receptor, which is required for GPCR desensitization. The structural details of the core engagement of arrestin are yet unknown for CCR5 and many other GPCRs.

So far, structures of fully engaged GPCR•arrestin complexes have been published for only five receptors (Figure 5.1). The structure determination of arrestin complexes is significantly more challenging compared to G protein complexes, often requiring considerable protein engineering and other optimization approaches. The reason for this is a relatively weak affinity of the arrestins to the receptor (259), intrinsic flexibility of the complexes, and structural plasticity of arrestin binding dictated by the varying receptor phosphorylation. Indeed, dissociation was reported for the β2AR-V2Rpp•arrestin2•G_s megaplex in cryo-EM data along with heterogeneous binding poses and blurred arrestin density, presumably due to this flexibility (260). Cross-linking the GPCR•arrestin complexes, as in the case of NTR1 (70), or expressing the receptor and arrestin as a single fusion protein (HTR2B and rhodopsin complexes), helped to overcome some of the mentioned problems (71,72,152). Further strategies, which have contributed to solving structures of M2R and β2AR complexes, are the replacement of the native receptor C-terminal tail with that of V2R and stabilizing the complexes with antibody fragments (150,153).

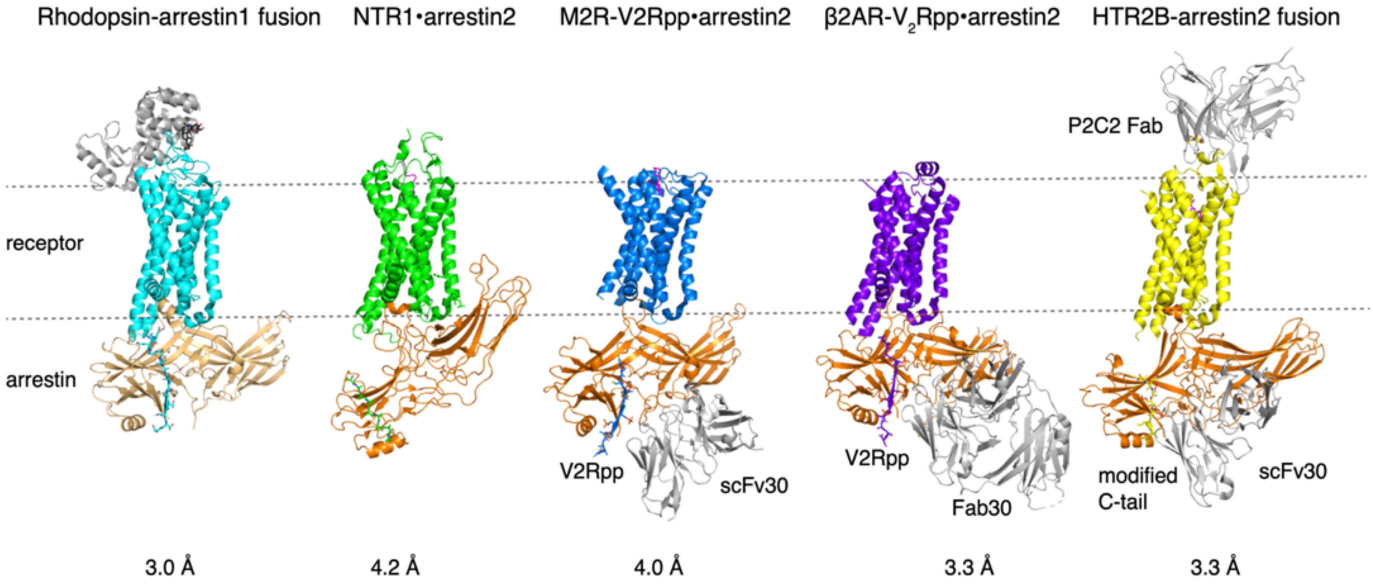


Figure 5.1. Overview of solved GPCR•arrestin structures at the time of writing (December 2022, PDB codes: 5W0P, 6UP7, 6U1N, 6TKO, 7SRS). The structures are aligned on the receptor chain (arrestin2: orange, arrestin1: wheat, other stabilizing proteins: grey).

The solved structures exhibit significant variations in the arrestin binding pose relative to the receptor core (Figure 5.1) (70,149–153). The structures and knowledge of their dynamics should explain the functional diversity of arrestins. However, specific elements, which define the way arrestin engages particular receptors, have not been described yet. Thus, additional arrestin complex structures and data on their conformational diversity are needed to identify the basic molecular principles and the variations of arrestin receptor interactions and functional outcomes.

The following parts of this thesis describe attempts to assemble a stable wild-type CCR5•arrestin2 complex and obtain structural insights on this fully engaged complex triggered by the chemokine ligand.

5.1.2 Results and Discussion

[6P4]CCL5•CCR5 arrestin2•Fab30 complex assembly

To form a stable CCR5•arrestin2 complex, wild-type CCR5 with a C-terminal Flag tag was co-expressed in *Sf9* insect cells together with GRK2 bearing the CAAX motif (261) as a membrane anchor. The phosphorylation was triggered by the addition of the super-agonist [6P4]CCL5 and characterized by western blot with phosphospecific antibodies as well as phosphoproteomics as described in Chapter 4.

Subsequently, the detergent-solubilized phosphorylated receptor was purified using M2 Flag beads. For the complex formation, a minimal cysteine variant of truncated arrestin2 (arrestin2$^{1\text{-}393}$), which comprises the C-terminal strand β20 (see Chapter 4), was used to enhance its stability. Incubation of the purified CCR5 with arrestin2 and the stabilizing antibody fragment Fab30 resulted in visible aggregation. For this reason, the complex was prepared on the M2 Flag beads which had captured CCR5 (Figure 5.2A). This approach allowed to co-elute the agonist-bound CCR5•arrestin2•Fab30 complex without precipitation. Subsequently, arrestin2 and Fab30 excess was removed by washing the resin. The eluted complex was then concentrated and further purified by size-exclusion chromatography (Figure 5.2B). The purity and complex integrity during purification were followed by SDS-PAGE, which confirmed complex formation on the M2 Flag beads and its presence on the respective SEC fractions (Figure 5.2B).

The purified complex was then subjected to negative stain EM, which showed intact complex particles (Figure 5.2C), despite a dilution of the complex to the tens of nanomolar range. This suggests that the affinity of the arrestin2 to the receptor is at least in this range. The 2D classes of the complex show that arrestin is mainly engaged in the receptor core.

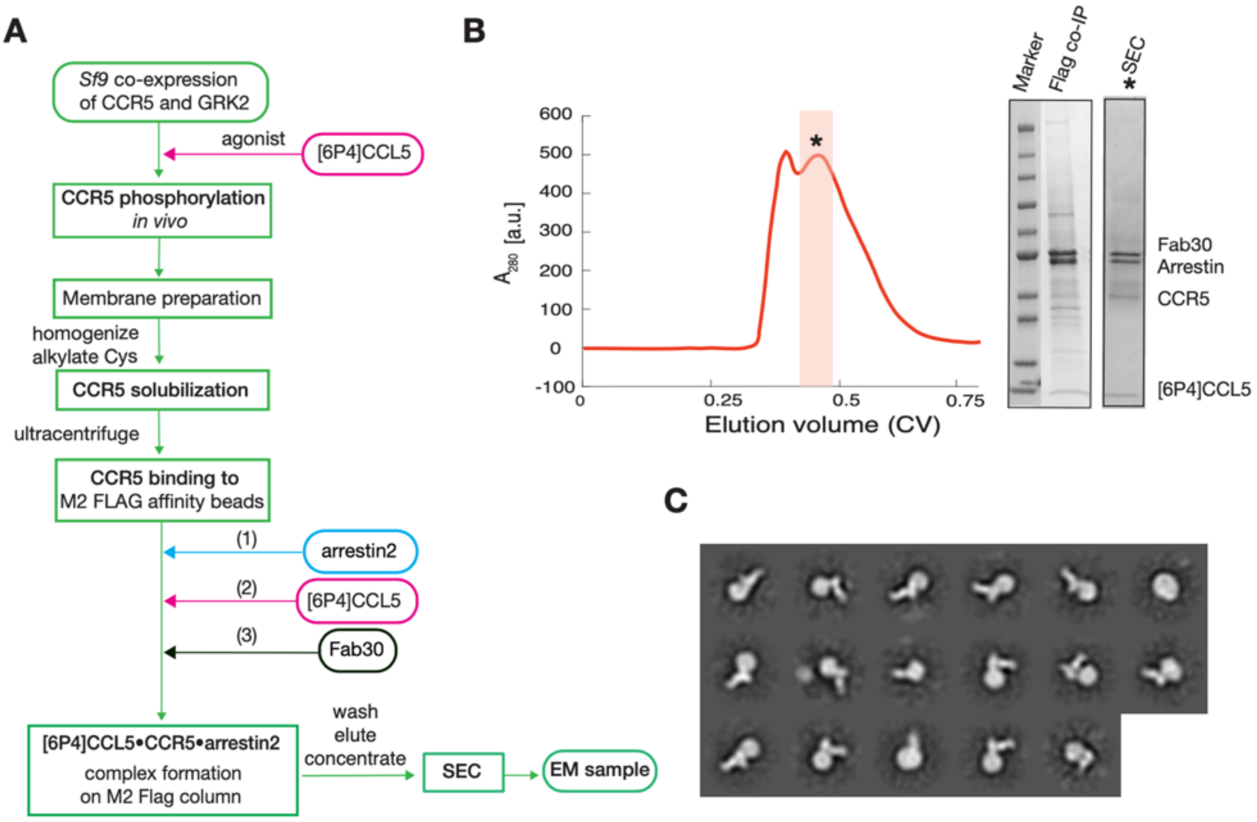


Figure 5.2. Overview of the CCR5•arrestin2 complex preparation for EM analysis. A. Overall scheme of the complex assembly. [6P4]CCL5, arrestin2, Fab30 were purified separately. The complex was assembled on M2 Flag beads bound to CCR5, followed by SEC. B. Representative SEC profile of the complex and SDS-PAGE analysis of the M2 Flag co-elution and the combined SEC fractions of the complex. C. Representative 2D classes from negative stain EM micrographs of the complex.

Initial cryo-EM analysis of the[6P4]CCL5•CCR5•arrestin2•Fab30 complex

While negative stain EM provides valuable information on the sample homogeneity and the overall architecture of the complex, the resolution is limited to about 20 Å. High-resolution details of the complex should be obtainable by single-particle cryo-EM analysis. Structure determination by cryo-EM requires that macromolecular particles be distributed in random orientations in a thin layer of vitreous ice (50–100 nm) on the grid. However, the ice thickness and particle distribution are often challenging to control and depend on many factors. Thus, cryo-EM sample preparation is still a bottleneck in the cryo-EM workflow and usually requires iterative optimization cycles (125,262).

The classical paper blotting approach is the most commonly used to create a thin vitreous ice layer with commercially available devices, such as the Vitrobot

(ThermoFisher Scientific) or Leica EM GP (Leica) plungers. Both instruments have humidity and temperature-controlled chambers and are operated in semiautomated mode.

The Vitrobot was used to prepare grids of the freshly purified [6P4]CCL5•CCR5 arrestin2•Fab30 complex. Before data collection, multiple conditions, including blot force and time, as well as sample concentration were screened using a 200 kV Talos instrument equipped with a Ceta 16M pixel camera (Table 5.1). A blot force of 15 instrument units and a blot time of 2.5 seconds with sample concentrations between 1.5 to 2 mg/ml provided optimal ice thickness and particle distributions on the cryo-EM grid. These conditions were close to the ones used for the [6P4]CCL5•CCR5•G_i complex. However, unlike the G protein complex, the arrestin complex particles were mainly aggregated on the grid (Figure 5.3A). Only a few areas with suboptimal particle distribution could be seen. The 2D class averages obtained from the deconvoluted micrograph movies collected on such grid areas using a 200 kV Glacios instrument equipped with a K3 camera show weak, blurry densities without any defined features (not shown).

Table 5.1. Summary of grid preparation conditions of the [6P4]CCL5•CCR5•arrestin2• Fab30 complex.

Vitrification device	**Device-specific Parameters**		**Sample concentration** [mg/ml]	**Additives**
	Blot force [device unit]	Blot time [s]		
Vitrobot Mark IV	10	1.5	1.5–2	-
	10	2.5	1.5–2	-
	10	3.5	1.5–2	-
	15	1.5	1.5–2	-
	15	2.5	1.5–2	-
	15	3.5	1.5–2	-
	20	1.5	1.5–2	-
	20	2.5	1.5–2	-
	20	3.5	1.5–2	-
	15	2.5	2.5	-
	15	2.5	1.0	-
	15	2.5	1.5–2	1 µM Amphipol A8-35
	15	3.5	1.5–2	1 µM Amphipol A8-35
	15	2.5	1.5–2	4 mM Fos-Choline-8
	15	3.5	1.5–2	4 mM Fos-Choline-8
	15	2.5	1.5–2	pH 6.9
	15	2.5	1.5–2	pH 7.9
Leica EM CP	Blot time, [s]			
	2		1.5–2	-
	3		1.5–2	-
CryoWriter	Withdrawal flow, µl/min			-
	1.5		1.5–2	-
	2		1.5–2	-
	2.5		1.5–2	
	2		1.5–2	1 mM Amphipol A8-35
	2.5		1.5–2	1 mM Amphipol A8-35
	2		1.5–2	4 mM Fos-Choline-8

The Leica plunger method with one-side blotting, which is milder, produced grids with more uniform ice thickness and slightly better particle distribution for the same sample (Figure 5.3B). Nevertheless, data collection on a 200 kV Glacios instrument equipped with a K3 camera and subsequent analysis resulted in similar low-quality 2D class averages (Figure 5.3C) as those obtained from the grids produced with the Vitrobot. Only a few particles display features of a complex, indicating that the particles are very heterogeneous and that in addition to aggregation, complex dissociation may have occurred. Since no significant dissociation of the complex was observed during negative stain EM analysis (Figure 5.2C), the complex damage must happen during the cryo-EM grid preparation.

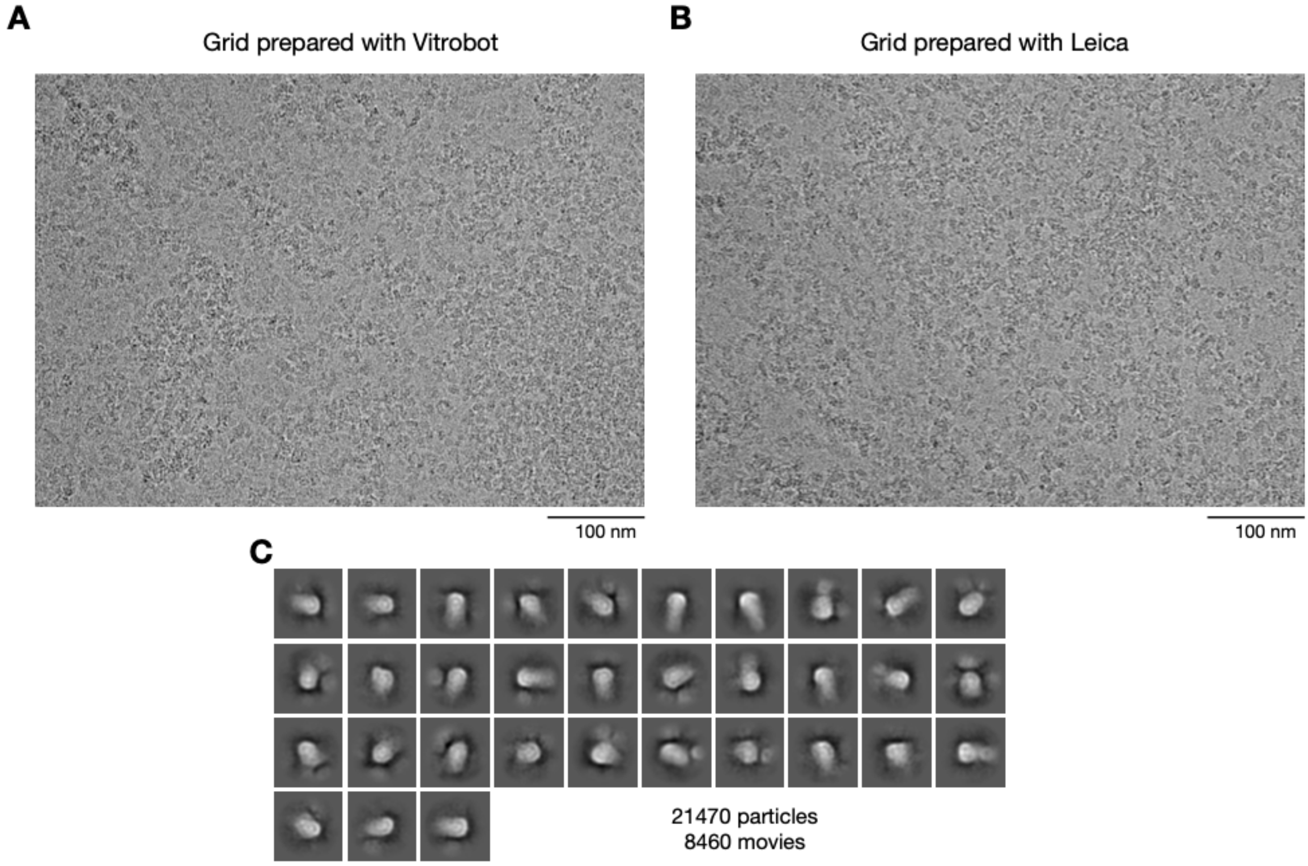


Figure 5.3. Examples of [6P4]CCL5•CCR5•arrestin2•Fab30 complex cryo-EM micrographs and particle distributions using different vitrification methods. Representative micrographs of the complex prepared using the Vitrobot plunger (A) and using the Leica EM CP plunger (B). 2D class averages of the complex prepared with the Leica plunger (C).

Cryo-EM sample optimization approaches

Aggregation or denaturation of biomacromolecules upon vitrification has been a persistent problem in cryo-EM, particularly for membrane proteins (125,263,264). The main reason for the sample damage has been attributed to interaction of the particles with the air-water interface. The denaturation effect at the air-water interface has been mentioned in early publications (265). It may happen by exposure of the macromolecule to the hydrophobic interface and/or by particle collision. A collision rate of 1000 or more per second has been estimated at the air interface for sample thicknesses of ≤100 nm (266). Recently, cryoelectron tomographic studies have shown that indeed the majority of protein particles are adsorbed to the air-water interface (267–269), leading to preferential orientation, local unfolding or complete denaturation. Some protein complexes, especially those with exposed hydrophobic regions or with certain surface charge distributions, have been shown to be more susceptible to air-water interface damage (269,270).

So far, there is no general solution to prevent damage of the macromolecules upon vitrification. Typically, sample optimization requires rigorous screening of conditions since the effects strongly depend on the sample. The problem may be approached systematically using the following strategies: (i) altering the conditions of the sample (e.g., salt concentration, pH, detergents, incorporation into nanodiscs); (ii) testing different grid types (e.g., graphene support, gold grids) or using blotting-free grid preparation methods, and (iii) modification of the chemical structure of the complex components (e.g., cross-linking, protein engineering).

Modifications of the chemical structure, such as protein engineering (104,150,152,153) and cross-linking (70), have been used for the majority of GPCR•arrestin complexes. However, they may significantly affect the observed interactions. Therefore, initially alternative approaches of cryo-EM sample optimizations were tested for the [6P4]CCL5•CCR5•arrestin2•Fab30 complex. All the tested conditions are listed in Table 5.1.

Effect of sample conditions

Initially, minor variations in the pH (from 6.9 to 7.9) were tested but had only a moderate effect on the particle distributions (not shown). The interaction between CCR5 and arrestin2 is dominated by electrostatics, with the components being of opposite charge. This means that a wide pH variation is not possible since it can unpredictably impact the complex. Therefore, other buffer components were varied.

Addition of secondary detergents or other surfactants has been reported to improve the particle distribution of other membrane proteins within the vitrified ice layer (267). Effects of amphipol A8-35 (151), which was used for vitrification of the V2R•arrestin2 complex, and of Fos-choline-8, one of the most commonly used detergent additives (271,272), were also tested for the CCR5•arrestin2 complex. The addition of amphipol A8-35 did not have any effect on the complex particle distribution. In contrast, Fos-Choline-8 improved the cryo-EM grid preparations by preventing aggregation to a large extent (Figure 5.4A). However, some areas of the grid still had particle clusters (Figure 5.4A). Furthermore, the particle density on the grid is low under the addition of Fos-Choline-8. Thus, it requires a significantly higher sample concentration to achieve reasonable particle coverage. A cryo-EM data set was collected on the sample with the Fos-Cholin-8 additive using a 200 kV Glacios instrument equipped with a K3 camera to estimate the complex integrity. Several rounds of 2D classification identified only very few particles that possibly could be identified as the complex of interest (Figure 5.4A). Moreover, the 2D class averages did not display high-resolution features even after extensive 2D classification, similar to the previous datasets collected on the grid without any additives (Figure 5.3C). Reasons for this behavior may be the heterogeneity of the complex, its partial dissociation, as well as the very low number of particles on the grid.

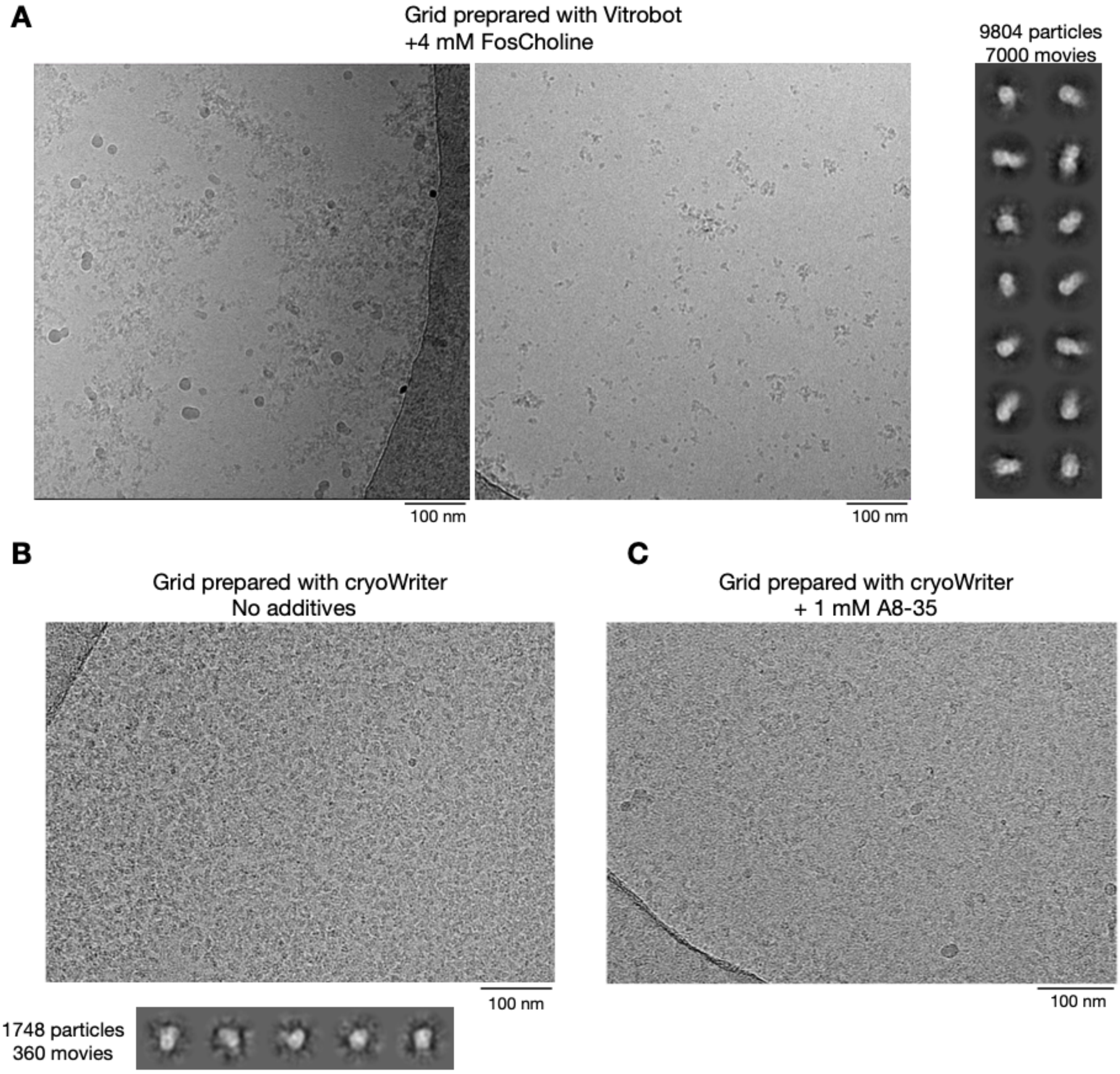


Figure 5.4. Different grid preparation approaches. Representative micrographs of the grid prepared with addition of FosCholine-12 using the Vitrobot (A), with the cryoWriter without additives (B), and with the cryoWriter and addition of amphipol A8-35 (C).

Blotting-free cryo-EM grid preparation

Classical cryo-EM grid preparation might harm macromolecules in two main ways: shear forces induced by paper blotting and long exposures to the air-water interface. Alternative cryo-EM grid preparation approaches, which avoid the paper blotting step or/and have faster freezing times, have been developed in recent years (206,273–278). One of them is a blotting-free microfluidic-based microcapillary writing device built in-house by the Braun group (Biozentrum, University of Basel) called the cryoWriter (276–278). The system can be operated in a fully automated regime, providing highly reproducible results and requiring only a few nanoliters of sample for cryo-EM grid preparation. Freshly prepared [6P4]CCL5•CCR5 arrestin2•Fab30 complex was plunge-frozen using the cryoWriter, and several conditions were tested (Table 5.1), including different withdrawal flows and the addition of protective molecules. Despite these improvements and reduced grid preparation times, the sample may still be affected by the air-water interface. Therefore, also protective additives were tested. However, representative movies collected on the 200 kV Glacios instrument with the K3 camera showed that the particle distribution with (Figure 5.4C) and without additives (Figure 5.4B) was not improved significantly compared with the Vitrobot preparations (Figure 5.3A). The 2D classification of a small dataset indicated that most of the complex particles were damaged despite using the blotting-free approach.

Modification of the complex components

Thus, more extensive changes to the arrestin complex itself seem to be required to obtain a high-resolution structure. An attempt to cross-link the complex using glutaraldehyde at concentrations varying from 1% to 0.1% resulted in complex precipitation and therefore was not continued. A further approach was made by exchanging the stabilizing antibody fragment from the highly charged Fab30 (pI of 8.7) to the more neutral single-chain fragment scFv30 (pI of 7.9). The complex was assembled in an identical manner as described for Fab30 (Figure 5.2A), and cryo-EM grids were prepared with the Vitrobot using a blot force of 10 instrument units and a blot time of 2 seconds. A dataset of about 5000 movies was recorded on the 200 kV Glacios instrument equipped with the K3 camera. The 2D classification analysis of the data resulted in higher-quality 2D classes of the [6P4]CCL5•CCR5•arrestin2•scFv30 complex (Figure 5.5). Despite these improvements, no high-resolution features, such as secondary structure elements, could be seen yet. Moreover, only about 3% (about 35,000) of the total particles had features of a complex with a very limited coverage of orientations. Thus, also this approach did not yield a high-resolution 3D map reconstruction.

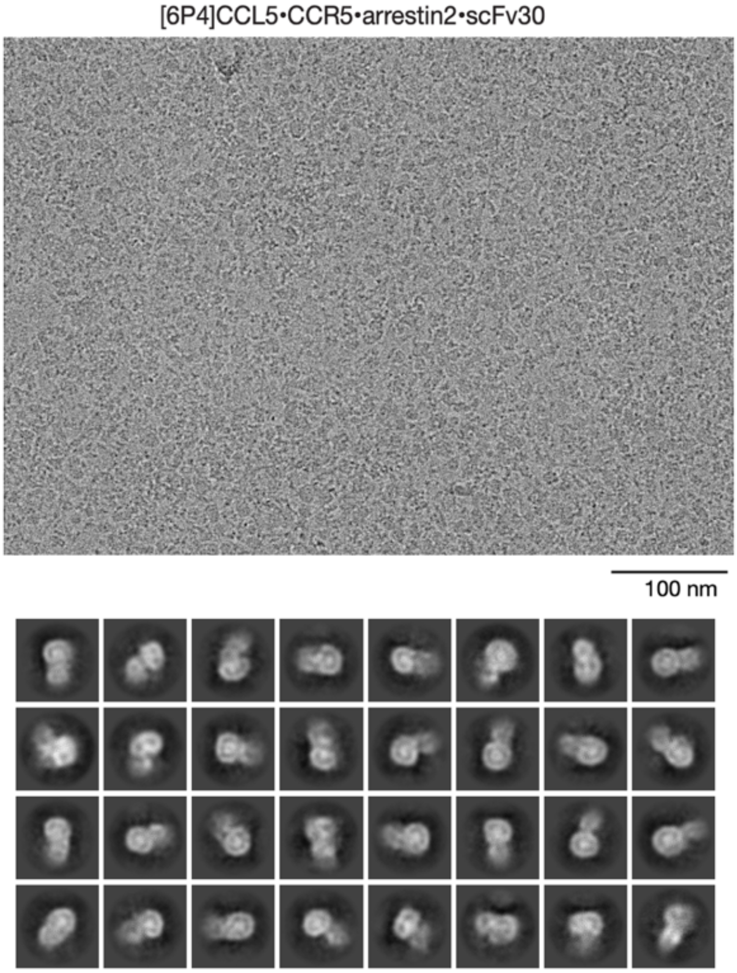


Figure 5.5. Representative micrograph of [6P4]CCL5•CCR5•arrestin2•scFv30 complex and obtained 2D class averages.

5.1.3 Conclusion and perspective

This chapter described the assembly of a stable [6P4]CCL5•CCR5•arrestin2•Fab30 complex *in vitro* using full-length wild-type CCR5. Initial cryo-EM experiments with conventional sample preparation approaches demonstrated that the complex is not stable under the grid preparation conditions, due to e.g., shear forces by the blotting paper and exposure to the air-water interface. It was attempted to overcome these problems by altering the buffer conditions of the sample and using blotting-free grid preparation methods. However, no significant improvements were obtained, indicating that the complex is intrinsically unstable at the air-water interface. The next logical step was the modification of the sample. Exchanging Fab30 with its derivative scFv30 improved the particle distribution and quality of the 2D classes. However, the majority of the sample seemed to be still affected by the air-water interface, interfering with the high-resolution reconstruction. Additionally, the initial cross-linking of [6P4]CCL5•CCR5•arrestin2•Fab30 complex trial did not provide usable results due to complex precipitation. Thus, further screening of chemical modifications and composition of the complex, in particular cross-linking with different reagents without antibody fragments, will be required for this highly dynamic and sensitive complex to obtain cryo-EM preparations amenable to high-resolution structure determination.

5.1.4 Material and Methods

Protein expression and purification

The production of [6P4]CCL5 from *E. coli* was described in Chapters 1 and 2. Wild-type full-length human CCR5 was modified with an N-terminal hemagglutinin signal peptide tag, a C-terminal 3C cleavage site followed by a Flag tag and a hexahistidine tag and cloned in pVL1393 vector. CCR5 was co-expressed with full-length GRK2 bearing a C-terminal CAAX motif in *Spodoptera frugiperda* (*Sf9)* insect cells. Detailed protocols of expression and purification are described in Chapters 3 and 4.

A minimal-cysteine variant (C150L, C242V, C251V, C269S) of arrestin2 (arrestin2$^{1-393}$) truncated after residue 393 with an N-terminal hexahistidine tag followed by a TEV cleavage site was expressed and purified from *E. coli* as described in Chapter 4.

Fab30 was a generous gift from Prof. Arun Shukla and its preparation is described in Chapter 4.

[6P4]CCL5•CCR5•arrestin2 complex formation

M2 Flag affinity beads bound to CCR5 were washed with 5 CV of buffer 1 [25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol, and 0.01% LMNG], followed by 3 CV of buffer 2 [25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol, 0.01% LMNG, 3 mM ATP, and 10 mM $MgCl_2$] and subsequently washed with 5 CV of buffer 3 [25 mM HEPES pH 7.4, 150 mM NaCl, 10% glycerol, 20 µM diC8-PIP2 and 0.01% LMNG]. Arrestin2 was added to the washed M2 Flag beads equilibrated in buffer 3 at a concentration of 3 µM and incubated for 15 minutes under gentle mixing at 4°C. The mixture was then supplemented with 5 µM [6P4]CCL5 and further incubated for 1 hour. Subsequently, a molar excess (1:1.2) of Fab30 or scFv30 was added and the mixture incubated for another hour. Thereafter unbound components were washed out from the resin with 10 CV of buffer 3. The formed complex was then eluted from the column with 3 CV of buffer 3 supplemented with 0.2 mg/ml of Flag peptide. The eluted mixture was concentrated with a 50-kDa molecular weight cut-off (MWCO) Amicon concentrator to about 250 µl and purified by size-exclusion chromatography (SEC) using a self-packed 4.2-ml S200 10/300 SEC column (length 25 mm, diameter 4.6 mm). The column was equilibrated with buffer 4 [25 mM HEPES, 150 mM NaCl, 0.01% LMNG, 0.5 µM [6P4]CCL5, pH 7.4]. Fractions containing the pure complex as judged from SDS-PAGE analysis were combined and concentrated with a 100-kDa MWCO Amicon concentrator to a final concentration of 1.5-2 mg/ml for cryo-EM analysis.

Negative stain EM

Shortly before grid preparation, the complex was diluted with a buffer containing 25 mM HEPES, 150 mM NaCl, and 0.01% LMNG to 0.05-0.1 mg/ml. The sample preparation was done according to a standard protocol (279). In short, 5 µl of the diluted complex was applied onto a glow-discharged carbon-coated 300-mesh copper grid, which was produced in-house. The sample was then blotted with filter paper. The grid was stained with 2% (w/v) uranyl acetate for 30 seconds. The obtained grids were imaged using a FEI Tecnai G2 Spirit TEM operated at 80 kV, or a Talos L120C TEM operated at 120 kV under low-dose of 20 e/Å^2 at a nominal magnification of x135,000 and with a defocus ranging from -0.7 to -2 μm. About 90 recorded micrographs were subjected to 2D classification analysis of single particles performed with CryoSPARC v3.1 (280).

Cryo-EM sample preparation

The complex samples were plunge-frozen freshy after purification. The additives (if used) were supplied immediately before grid preparation. The grids were glow-discharged for 30-45 seconds at 20 mA (unless specified otherwise) in the air plasma immediately before use. All grids were plunge-frozen into cooled liquid ethane. Blotting and plunge-freezing were performed using semiautomated or completely automated procedures. The following conditions are specific to the device used for the specimen preparation.

Grid preparation with Vitrobot plunger

The FEI Vitrobot Mark IV plunger (ThermoFisher Scientific) was operated at a humidity of 95% at 4°C. An aliquot of 3.5 μl of the complex was applied to a glow-discharged 200-mesh carbon-coated copper grid (Quantifoil R1.2/R1.3) and plunge-frozen using an equilibrated plunger. Different blotting forces and times were tested to optimize the ice thickness. Typically, the optimal blot force was about 10–15 instrument units and the optimal blotting time was 2.5 seconds with no waiting time. Several sample concentrations were tested ranging from 1 mg/ml to 2.5 mg/ml. The optimal complex concentration was 1.5–2 mg/ml.

Grid preparation with Leica plunger

The Leica CP plunger (Leica Microsystems) was operated at 15°C and a humidity of 70%. Grids were blotted single-sided with a blot time of 3 seconds.

CryoWriter

The principles of cryo-EM grid preparations with the in-house built cryoWriter machine have been described in detail elsewhere (276,281). The instrument was operated at a humidity of 80% at 5°C. The microcapillary for dispensing the sample on the grid was flashed with tens of nanoliters prior to use. Based on the written macro script, the sample was plunge-frozen in a user-free manner. In short, 17 nl of the complex at a concentration of 1.8 mg/ml was dispensed with a capillary on a 300-mesh carbon-coated copper grid (Quantifoil Cu300 R1.2/1.3). The liquid excess was withdrawn at a flow rate ranging from 1.5 to 2.5 µl/min. The additive molecules were injected into the gas stream simultaneously with the withdrawal of the sample excess.

Cryo-EM image acquisition

Cryo-EM grids of the complex were screened using a 200 kV Glacios TEM equipped with a Gatan K3 Summit direct electron detector or a 200 kV Talos F200C TEM equipped with a Ceta 16M Pixel CMOS camera. Automated data collection was carried out on a 200 kV Glacios Cryo-TEM equipped with a Gatan K3 Summit direct electron detector using the SerialEM software. Movie stacks of 40-50 frames were obtained with a defocus range of –1.9 to –2.5 µm at a nominal magnification of x45,000 (0.878 Å/pixel) with a total accumulated dose exposure of 50 to 60 e/Å^2.

Cryo-EM data processing

The data were processed using standard workflows based on algorithms implemented in cryoSPARC v3.1 (280).

6 Conclusion and Outlook

CCR5 is a chemokine receptor, which has been under extensive investigation not only due to its crucial role in HIV (233) but also in inflammation (234), the pathology of cancer (235,236), and COVID-19 (237). This thesis describes the structural and functional elucidation of this pivotal receptor and its complexes. The variable pharmacology of the CCL5 analogs and the activation mechanism of the CCR5 had not been understood for many years, largely due to difficulties in obtaining homogeneous and stable chemokine receptor complexes for structural studies. We developed a protocol for stable complex formation of full-length wild-type human CCR5 bound to the super-agonist [6P4]CCL5 and the heterotrimeric G_i protein (258) and solved the cryo-EM structure of this CCR5 complex in an active conformation (138). Comparison to the previously solved antagonist-bound CCR5 structure (246) allowed us to decipher the molecular details of the CCR5 activation mechanism, by which the N-terminus of the chemokine analog pushes onto specific structural motifs at the bottom of the orthosteric pocket thereby activating the canonical GPCR microswitch network. The structure also allowed to identify chemokine sequence signatures that drive either agonist or antagonist activity.

Based on the structure, we also modeled the interaction of the CCR5 with its native ligand, CCL5. Subsequently several further cryo-EM structures of CCR5 in complex with its N-terminally fused endogenous ligands (CCL3 and CCL5) were published. However, despite of this fusion, the density of the chemokine N-termini is not well-defined, indicating a dynamic nature of these interactions and impeding a detailed understanding of the mechanism (228). A dynamical description of these interactions at the atomic level will be required to obtain the precise activation mechanism of CCR5 by its endogenous ligands.

A comprehensive comparison of all available structures and sequences of CC chemokines and their receptors revealed that the CCR5 activation mechanism differs significantly from other CC chemokine receptors, which bind chemokines with shorter N-termini and involve a non-canonical, yet unclear activation mechanism (229). Interestingly, some chemokines with long N-termini, including CCL5 can undergo post-translational proteolytic processing leading to shorter N-termini. This probably constitutes an additional layer of the chemokine system regulation.

How the varying lengths of the chemokine N-termini influence receptor activation is an open question of considerable biological and medical interest. This could be investigated by measuring the response to and affinity of chemokines with varying N-terminal lengths in cellular assays on structure-based receptor mutants. While such data may provide a primary characterization, a detailed and comprehensive understanding of the chemokine N-terminal length effects will require further high-resolution structural and dynamical data.

While significant progress has been achieved in the structural elucidation of GPCR•G protein complexes, explaining the activation mechanisms of many receptors, much less is known about GPCR•arrestin interactions. In particular, no structural data are available on arrestin•chemokine receptor complexes. To address some of these questions, we characterized the CCR5 phosphorylation induced by the super-agonist [6P4]CCL5 and GRK2. Based on these findings, we designed a set of phosphopeptides to analyze their effects on arrestin interactions by a combination of NMR and biochemical assays. We then solved the high-resolution crystal structures of arrestin2 in apo form and in complex with two phosphopeptides (44). These data revealed a distinct pXpp phosphopeptide sequence motif that induces the active arrestin2 conformation by dominant electrostatic interactions. An identical pXpp motif exists in the vasopressin hormone-V2 receptor C-terminal phosphopeptide (V2Rpp) and in other GPCRs (70,282,283), which are known to form stable complexes with arrestin2. This suggests that this motif is crucially involved in stable arrestin2 recruitment. Additional structural and functional comparisons of GPCR•arrestin interactions together with an analysis of GPCR sequences provided hints on the molecular basis of the arrestin2/arrestin3 isoform specificity.

Although the latter part of the thesis has provided preliminary information on how multi-site GPCR phosphorylation controls receptor•arrestin interactions, the detailed molecular basis of arrestin's functional multiplicity and biased agonism towards arrestin has not been elucidated. The diverse functional outcomes of arrestin•GPCR interactions appear to be driven by distinct phosphorylation patterns, which induce distinct arrestin conformations. How these conformations are linked with specific functional outcomes, e.g., desensitization or endocytosis, is also poorly understood. Additional full-length, wild-type receptor•arrestin structures in combination with high-throughput methods for quantifying receptor phosphorylation in their native cellular context are expected to provide mechanistic insights on the GPCR signal transmission towards arrestin. Further high-resolution dynamics data on the various arrestin conformations induced by the different phosphorylation motifs and their subsequent interactions with signaling partners may then provide a more complete molecular understanding of arrestin downstream signaling.

It is hoped that the described findings and these suggested approaches will enhance the understanding of the highly complex chemokine-chemokine receptor network and other GPCR systems, which may serve to improve existing drugs and aid the discovery of novel effective compounds.

7 References

1. Takeda S, Kadowaki S, Haga T, Takaesu H, Mitaku S. Identification of G protein-coupled receptor genes from the human genome sequence. FEBS Letters. 2002;520(1–3):97–101.

2. S. Odoemelam C, Percival B, Wallis H, Chang MW, Ahmad Z, Scholey D, et al. G-Protein coupled receptors: structure and function in drug discovery. RSC Advances. 2020;10(60):36337–48.

3. Hauser AS, Attwood MM, Rask-Andersen M, Schiöth HB, Gloriam DE. Trends in GPCR drug discovery: new agents, targets and indications. Nat Rev Drug Discov. 2017 Dec 1;16(12):829–42.

4. Alexander SPH, Christopoulos A, Davenport AP, Kelly E, Mathie A, Peters JA, et al. THE CONCISE GUIDE TO PHARMACOLOGY 2019/20: G protein-coupled receptors. British Journal of Pharmacology. 2019;176(S1):S21–141.

5. Schertler GFX, Villa C, Henderson R. Projection structure of rhodopsin. Nature. 1993 Apr;362(6422):770–2.

6. Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature. 2013 Feb;494(7436):185–94.

7. Deupi X, Kobilka B. Activation of G Protein–Coupled Receptors. In: Advances in Protein Chemistry [Internet]. Academic Press; 2007 [cited 2022 Nov 1]. p. 137–66. (Mechanisms and Pathways of Heterotrimeric G Protein Signaling; vol. 74). Available from: https://www.sciencedirect.com/science/article/pii/S0065323307740044

8. Ballesteros JA, Weinstein H. [19] Integrated methods for the construction of three-dimensional models and computational probing of structure-function relations in G protein-coupled receptors. In: Sealfon SC, editor. Methods in Neurosciences [Internet]. Academic Press; 1995 [cited 2022 Nov 11]. p. 366–428. (Receptor Molecular Biology; vol. 25). Available from: https://www.sciencedirect.com/science/article/pii/S1043947105800497

9. Kooistra AJ, Mordalski S, Pándy-Szekeres G, Esguerra M, Mamyrbekov A, Munk C, et al. GPCRdb in 2021: integrating GPCR sequence, structure and function. Nucleic Acids Research. 2021 Jan 8;49(D1):D335–43.

10. Brelot A, Chakrabarti LA. CCR5 Revisited: How Mechanisms of HIV Entry Govern AIDS Pathogenesis. Journal of Molecular Biology. 2018 Aug 17;430(17):2557–89.

11. Lai WY, Mueller A. Latest update on chemokine receptors as therapeutic targets. Biochem Soc Trans. 2021 Jun 30;49(3):1385–95.

12. Pavlos NJ, Friedman PA. GPCR Signaling and Trafficking: The Long and Short of It. Trends in Endocrinology & Metabolism. 2017 Mar 1;28(3):213–26.

13. Lefkowitz RJ, Shenoy SK. Transduction of Receptor Signals by ß-Arrestins. Science. 2005 Apr 22;308(5721):512–7.

14. Marchese A, Paing MM, Temple BRS, Trejo J. G Protein–Coupled Receptor Sorting to Endosomes and Lysosomes. Annual Review of Pharmacology and Toxicology. 2008;48(1):601–29.

15. Smith JS, Pack TF, Inoue A, Lee C, Zheng K, Choi I, et al. Noncanonical scaffolding of Gαi and β-arrestin by G protein–coupled receptors. Science. 2021 Mar 12;371(6534):eaay1833.

16. Nguyen AH, Thomsen ARB, Cahill TJ, Huang R, Huang LY, Marcink T, et al. Structure of an endosomal signaling GPCR–G protein–β-arrestin megacomplex. Nat Struct Mol Biol. 2019 Dec;26(12):1123–31.

17. Thomsen ARB, Plouffe B, Cahill TJ, Shukla AK, Tarrasch JT, Dosey AM, et al. GPCR-G Protein-β-Arrestin Super-Complex Mediates Sustained G Protein Signaling. Cell. 2016 Aug 11;166(4):907–19.

18. Pandey S, Kumari P, Baidya M, Kise R, Cao Y, Dwivedi-Agnihotri H, et al. Intrinsic bias at non-canonical, β-arrestin-coupled seven transmembrane receptors. Molecular Cell. 2021 Nov 18;81(22):4605-4621.e11.

19. Calebiro D, Koszegi Z, Lanoiselée Y, Miljus T, O'Brien S. G protein-coupled receptor-G protein interactions: a single-molecule perspective. Physiological Reviews. 2021 Jul;101(3):857–906.

20. Syrovatkina V, Alegre KO, Dey R, Huang XY. Regulation, Signaling, and Physiological Functions of G-Proteins. Journal of Molecular Biology. 2016 Sep;428(19):3850–68.

21. Flock T, Hauser AS, Lund N, Gloriam DE, Balaji S, Babu MM. Selectivity determinants of GPCR–G-protein binding. Nature. 2017 May;545(7654):317–22.

22. Hauser AS, Avet C, Normand C, Mancini A, Inoue A, Bouvier M, et al. Common coupling map advances GPCR-G protein selectivity. Weis WI, Aldrich RW, Sunahara R, editors. eLife. 2022 Mar 18;11:e74107.

23. Bos JL, Rehmann H, Wittinghofer A. GEFs and GAPs: Critical Elements in the Control of Small G Proteins. Cell. 2007 Jun 1;129(5):865–77.

24. Masuho I, Skamangas NK, Muntean BS, Martemyanov KA. Diversity of the Gβγ complexes defines spatial and temporal bias of GPCR signaling. Cell Systems. 2021 Apr;12(4):324-337.e5.

25. Gurevich VV, Gurevich EV. GPCR Signaling Regulation: The Role of GRKs and Arrestins. Frontiers in Pharmacology [Internet]. 2019 [cited 2022 May 31];10. Available from: https://www.frontiersin.org/article/10.3389/fphar.2019.00125

26. Sato PY, Chuprun JK, Schwartz M, Koch WJ. The Evolving Impact of G Protein-Coupled Receptor Kinases in Cardiac Health and Disease. Physiological Reviews. 2015 Apr;95(2):377–404.

27. Komolov KE, Benovic JL. G protein-coupled receptor kinases: Past, present and future. Cellular Signalling. 2018 Jan 1;41:17–24.

28. Inglese J, Koch WJ, Caron MG, Lefkowitz RJ. Isoprenylation in regulation of signal transduction by G-protein-coupled receptor kinases. Nature. 1992 Sep;359(6391):147–50.

29. Carman CV, Barak LS, Chen C, Liu-Chen LY, Onorato JJ, Kennedy SP, et al. Mutational Analysis of Gβγ and Phospholipid Interaction with G Protein-coupled Receptor Kinase 2*. Journal of Biological Chemistry. 2000 Apr 7;275(14):10443–52.

30. Wolters V, Krasel C, Brockmann J, Bünemann M. Influence of Gαq on the Dynamics of M3-Acetylcholine Receptor–G-Protein–Coupled Receptor Kinase 2 Interaction. Mol Pharmacol. 2015 Jan 1;87(1):9–17.

31. Day PW, Carman CV, Sterne-Marr R, Benovic JL, Wedegaertner PB. Differential Interaction of GRK2 with Members of the Gαq Family. Biochemistry. 2003 Aug 1;42(30):9176–84.

32. Duan J, Liu H, Ji Y, Yuan Q, Li X, Wu K, et al. Structure of a G protein-coupled receptor with GRK2 and a biased ligand [Internet]. bioRxiv; 2022 [cited 2022 Nov 2]. p. 2022.10.19.512855. Available from: https://www.biorxiv.org/content/10.1101/2022.10.19.512855v1

33. Pronin AN, Carman CV, Benovic JL. Structure-Function Analysis of G Protein-coupled Receptor Kinase-5: ROLE OF THE CARBOXYL TERMINUS IN KINASE REGULATION*. Journal of Biological Chemistry. 1998 Nov 20;273(47):31510–8.

34. Jiang X, Benovic JL, Wedegaertner PB. Plasma Membrane and Nuclear Localization of G Protein–coupled Receptor Kinase 6A. MBoC. 2007 Aug;18(8):2960–9.

35. Sulon SM, Benovic JL. Targeting G protein–coupled receptor kinases to G protein–coupled receptors. Current Opinion in Endocrine and Metabolic Research. 2021 Feb 1;16:56–65.

36. Bradley SJ, Wiegman CH, Iglesias MM, Kong KC, Butcher AJ, Plouffe B, et al. Mapping physiological G protein-coupled receptor signaling pathways reveals a role for receptor phosphorylation in airway contraction. Proceedings of the National Academy of Sciences. 2016 Apr 19;113(16):4524–9.

37. Choi M, Staus DP, Wingler LM, Ahn S, Pani B, Capel WD, et al. G protein–coupled receptor kinases (GRKs) orchestrate biased agonism at the β2-adrenergic receptor. Science Signaling. 2018 Aug 21;11(544):eaar7084.

38. Lohse MJ, Hoffmann C. Arrestin Interactions with G Protein-Coupled Receptors. In: Gurevich VV, editor. Arrestins - Pharmacology and Therapeutic Potential [Internet]. Berlin, Heidelberg: Springer; 2014 [cited 2022 May 31]. p. 15–56. (Handbook of Experimental Pharmacology). Available from: https://doi.org/10.1007/978-3-642-41199-1_2

39. Gurevich EV, Gurevich VV. Arrestins: ubiquitous regulators of cellular signaling pathways. Genome Biology. 2006 Oct 2;7(9):236.

40. Srivastava A, Gupta B, Gupta C, Shukla AK. Emerging Functional Divergence of β-Arrestin Isoforms in GPCR Function. Trends in Endocrinology & Metabolism. 2015 Nov 1;26(11):628–42.

41. Kaya AI, Perry NA, Gurevich VV, Iverson TM. Phosphorylation barcode-dependent signal bias of the dopamine D1 receptor. Proceedings of the National Academy of Sciences. 2020 Jun 23;117(25):14139–49.

42. Sensoy O, Moreira IS, Morra G. Understanding the Differential Selectivity of Arrestins toward the Phosphorylation State of the Receptor. ACS Chem Neurosci. 2016 Sep 21;7(9):1212–24.

43. Vishnivetskiy SA, Raman D, Wei J, Kennedy MJ, Hurley JB, Gurevich VV. Regulation of Arrestin Binding by Rhodopsin Phosphorylation Level*. Journal of Biological Chemistry. 2007 Nov 2;282(44):32075–83.

44. Isaikina P, Petrovic I, Jakob RP, Sarma P, Ranjan A, Baruah M, et al. A key GPCR phosphorylation motif discovered in arrestin2•CCR5 phosphopeptide complexes. bioRxiv. 2022 Oct 12;2022.10.10.511578.

45. Oakley RH, Laporte SA, Holt JA, Caron MG, Barak LS. Differential Affinities of Visual Arrestin, βArrestin1, and βArrestin2 for G Protein-coupled Receptors Delineate Two Major Classes of Receptors*. Journal of Biological Chemistry. 2000 Jun 2;275(22):17201–10.

46. van Gastel J, Hendrickx JO, Leysen H, Santos-Otte P, Luttrell LM, Martin B, et al. β-Arrestin Based Receptor Signaling Paradigms: Potential Therapeutic Targets for Complex Age-Related Disorders. Frontiers in Pharmacology [Internet]. 2018 [cited 2022 Nov 3];9. Available from: https://www.frontiersin.org/articles/10.3389/fphar.2018.01369

47. Leonard AP, Appleton KM, Luttrell LM, Peterson YK. A High-Content, Live-Cell, and Real-Time Approach to the Quantitation of Ligand-Induced β-Arrestin2 and Class A/Class B GPCR Mobilization. Microscopy and Microanalysis. 2013 Feb;19(1):150–70.

48. Sente A, Peer R, Srivastava A, Baidya M, Lesk AM, Balaji S, et al. Molecular mechanism of modulating arrestin conformation by GPCR phosphorylation. Nat Struct Mol Biol. 2018 Jun;25(6):538–45.

49. Shukla AK, Westfield GH, Xiao K, Reis RI, Huang LY, Tripathi-Shukla P, et al. Visualization of arrestin recruitment by a G-protein-coupled receptor. Nature. 2014 Aug;512(7513):218–22.

50. Kumari P, Srivastava A, Ghosh E, Ranjan R, Dogra S, Yadav PN, et al. Core engagement with β-arrestin is dispensable for agonist-induced vasopressin receptor endocytosis and ERK activation. MBoC. 2017 Apr 15;28(8):1003–10.

51. Cahill TJ, Thomsen ARB, Tarrasch JT, Plouffe B, Nguyen AH, Yang F, et al. Distinct conformations of GPCR–β-arrestin complexes mediate desensitization, signaling, and endocytosis. Proceedings of the National Academy of Sciences. 2017 Mar 7;114(10):2562–7.

52. Janetzko J, Kise R, Barsi-Rhyne B, Siepe DH, Heydenreich FM, Kawakami K, et al. Membrane phosphoinositides regulate GPCR-β-arrestin complex assembly and dynamics. Cell [Internet]. 2022 Nov 10 [cited 2022 Nov 11]; Available from: https://www.sciencedirect.com/science/article/pii/S0092867422013605

53. Kang DS, Kern RC, Puthenveedu MA, von Zastrow M, Williams JC, Benovic JL. Structure of an Arrestin2-Clathrin Complex Reveals a Novel Clathrin Binding Domain That Modulates Receptor Trafficking. J Biol Chem. 2009 Oct 23;284(43):29860–72.

54. Thomsen ARB, Hahn H, Bunnett NW. Chapter 2 - Arrestin-mediated trafficking and compartmentalized biology of GPCRs. In: Gurevich VV, editor. Arrestins [Internet]. Academic Press; 2022 [cited 2022 Nov 1]. p. 9–24. Available from: https://www.sciencedirect.com/science/article/pii/B9780323857567000042

55. Gu Y jing, Sun W yi, Zhang S, Wu J jing, Wei W. The emerging roles of β-arrestins in fibrotic diseases. Acta Pharmacol Sin. 2015 Nov;36(11):1277–87.

56. Sharma D, Parameswaran N. Multifaceted role of β-arrestins in inflammation and disease. Genes Immun. 2015 Dec;16(8):499–513.

57. Kenakin T. Efficacy at g-protein-coupled receptors. Nat Rev Drug Discov. 2002 Feb;1(2):103–10.

58. Wacker D, Stevens RC, Roth BL. How Ligands Illuminate GPCR Molecular Pharmacology. Cell. 2017 Jul 27;170(3):414–27.

59. Katritch V, Fenalti G, Abola EE, Roth BL, Cherezov V, Stevens RC. Allosteric sodium in class A GPCR signaling. Trends in Biochemical Sciences. 2014 May 1;39(5):233–44.

60. Abiko LA, Dias Teixeira R, Engilberge S, Grahl A, Mühlethaler T, Sharpe T, et al. Filling of a water-free void explains the allosteric regulation of the β1-adrenergic receptor by cholesterol. Nat Chem. 2022 Oct;14(10):1133–41.

61. Hanson MA, Cherezov V, Griffith MT, Roth CB, Jaakola VP, Chien EYT, et al. A Specific Cholesterol Binding Site Is Established by the 2.8 Å Structure of the Human β2-Adrenergic Receptor. Structure. 2008 Jun 11;16(6):897–905.

62. Kolb P, Kenakin T, Alexander SPH, Bermudez M, Bohn LM, Breinholt CS, et al. Community guidelines for GPCR ligand bias: IUPHAR review 32. British Journal of Pharmacology. 2022;179(14):3651–74.

63. Bock A, Bermudez M. Allosteric coupling and biased agonism in G protein-coupled receptors. The FEBS Journal. 2021;288(8):2513–28.

64. Gundry J, Glenn R, Alagesan P, Rajagopal S. A Practical Guide to Approaching Biased Agonism at G Protein Coupled Receptors. Frontiers in Neuroscience [Internet]. 2017 [cited 2022 Nov 8];11. Available from: https://www.frontiersin.org/articles/10.3389/fnins.2017.00017

65. Congreve M, de Graaf C, Swain NA, Tate CG. Impact of GPCR Structures on Drug Discovery. Cell. 2020 Apr 2;181(1):81–91.

66. Kato HE, Zhang Y, Hu H, Suomivuori CM, Kadji FMN, Aoki J, et al. Conformational transitions of a neurotensin receptor 1–Gi1 complex. Nature. 2019 Aug;572(7767):80–5.

67. Zhang M, Gui M, Wang ZF, Gorgulla C, Yu JJ, Wu H, et al. Cryo-EM structure of an activated GPCR–G protein complex in lipid nanodiscs. Nat Struct Mol Biol. 2021 Mar;28(3):258–67.

68. Gao Y, Hu H, Ramachandran S, Erickson JW, Cerione RA, Skiniotis G. Structures of the Rhodopsin-Transducin Complex: Insights into G-Protein Activation. Molecular Cell. 2019 Aug 22;75(4):781-790.e3.

69. Kang Y, Kuybeda O, de Waal PW, Mukherjee S, Van Eps N, Dutka P, et al. Cryo-EM structure of human rhodopsin bound to an inhibitory G protein. Nature. 2018 Jun;558(7711):553–8.

70. Huang W, Masureel M, Qu Q, Janetzko J, Inoue A, Kato HE, et al. Structure of the neurotensin receptor 1 in complex with β-arrestin 1. Nature. 2020 Mar;579(7798):303–8.

71. Kang Y, Zhou XE, Gao X, He Y, Liu W, Ishchenko A, et al. Crystal structure of rhodopsin bound to arrestin by femtosecond X-ray laser. Nature. 2015 Jul;523(7562):561–7.

72. Zhou XE, He Y, de Waal PW, Gao X, Kang Y, Van Eps N, et al. Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. Cell. 2017 Jul 27;170(3):457-469.e13.

73. Chen Q, Plasencia M, Li Z, Mukherjee S, Patra D, Chen CL, et al. Structures of rhodopsin in complex with G-protein-coupled receptor kinase 1. Nature. 2021 Jul 22;595(7868):600–5.

74. Jumper J, Evans R, Pritzel A, Green T, Figurnov M, Ronneberger O, et al. Highly accurate protein structure prediction with AlphaFold. Nature. 2021 Aug;596(7873):583–9.

75. Baek M, DiMaio F, Anishchenko I, Dauparas J, Ovchinnikov S, Lee GR, et al. Accurate prediction of protein structures and interactions using a 3-track neural network. Science. 2021 Aug 20;373(6557):871–6.

76. Kinch LN, Schaeffer RD, Kryshtafovych A, Grishin NV. Target classification in the 14th round of the critical assessment of protein structure prediction (CASP14). Proteins: Structure, Function, and Bioinformatics. 2021;89(12):1618–32.

77. Lee C, Su BH, Tseng YJ. Comparative studies of AlphaFold, RoseTTAFold and Modeller: a case study involving the use of G-protein-coupled receptors. Briefings in Bioinformatics. 2022 Sep 1;23(5):bbac308.

78. Mullard A. What does AlphaFold mean for drug discovery? Nature Reviews Drug Discovery. 2021 Sep 14;20(10):725–7.

79. Palczewski K, Kumasaka T, Hori T, Behnke CA, Motoshima H, Fox BA, et al. Crystal Structure of Rhodopsin: A G Protein-Coupled Receptor. Science [Internet]. 2000 Aug 4 [cited 2022 Nov 9]; Available from: https://www.science.org/doi/10.1126/science.289.5480.739

80. Rosenbaum DM, Cherezov V, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. GPCR Engineering Yields High-Resolution Structural Insights into β2-Adrenergic Receptor Function. Science. 2007 Nov 23;318(5854):1266–73.

81. Rasmussen SGF, Choi HJ, Rosenbaum DM, Kobilka TS, Thian FS, Edwards PC, et al. Crystal structure of the human β2 adrenergic G-protein-coupled receptor. Nature. 2007 Nov;450(7168):383–7.

82. Cherezov V, Rosenbaum DM, Hanson MA, Rasmussen SGF, Thian FS, Kobilka TS, et al. High-Resolution Crystal Structure of an Engineered Human β2-Adrenergic G Protein–Coupled Receptor. Science [Internet]. 2007 Nov 23 [cited 2022 Nov 9]; Available from: https://www.science.org/doi/10.1126/science.1150577

83. Riekel C, Burghammer M, Schertler G. Protein crystallography microdiffraction. Current Opinion in Structural Biology. 2005 Oct 1;15(5):556–62.

84. Caffrey M, Cherezov V. Crystallizing membrane proteins using lipidic mesophases. Nat Protoc. 2009 May;4(5):706–31.

85. Munk C, Mutt E, Isberg V, Nikolajsen LF, Bibbe JM, Flock T, et al. An online resource for GPCR structure determination and analysis. Nat Methods. 2019 Feb;16(2):151–62.

86. Rosenbaum DM, Zhang C, Lyons JA, Holl R, Aragao D, Arlow DH, et al. Structure and function of an irreversible agonist-β2 adrenoceptor complex. Nature. 2011 Jan;469(7329):236–40.

87. Ghanouni P, Gryczynski Z, Steenhuis JJ, Lee TW, Farrens DL, Lakowicz JR, et al. Functionally Different Agonists Induce Distinct Conformations in the G Protein Coupling Domain of the β2Adrenergic Receptor*. Journal of Biological Chemistry. 2001 Jan 1;276(27):24433–6.

88. Scheerer P, Park JH, Hildebrand PW, Kim YJ, Krauß N, Choe HW, et al. Crystal structure of opsin in its G-protein-interacting conformation. Nature. 2008 Sep;455(7212):497–502.

89. Rasmussen SGF, DeVree BT, Zou Y, Kruse AC, Chung KY, Kobilka TS, et al. Crystal structure of the β2 adrenergic receptor–Gs protein complex. Nature. 2011 Sep;477(7366):549–55.

90. Isogai S, Deupi X, Opitz C, Heydenreich FM, Tsai CJ, Brueckner F, et al. Backbone NMR reveals allosteric signal transduction networks in the β1-adrenergic receptor. Nature. 2016 Feb;530(7589):237–41.

91. Abiko LA, Dias Teixeira R, Engilberge S, Grahl A, Mühlethaler T, Sharpe T, et al. Filling of a water-free void explains the allosteric regulation of the β1-adrenergic receptor by cholesterol. Nat Chem. 2022 Oct;14(10):1133–41.

92. Abiko LA, Grahl A, Grzesiek S. High Pressure Shifts the β1-Adrenergic Receptor to the Active Conformation in the Absence of G Protein. J Am Chem Soc. 2019 Oct 23;141(42):16663–70.

93. Sun B, Feng D, Chu MLH, Fish I, Lovera S, Sands ZA, et al. Crystal structure of dopamine D1 receptor in complex with G protein and a non-catechol agonist. Nat Commun. 2021 Jun 3;12(1):3305.

94. Kobilka B. The Structural Basis of G-Protein-Coupled Receptor Signaling (Nobel Lecture). Angewandte Chemie International Edition. 2013;52(25):6380–8.

95. Lefkowitz RJ. A Brief History of G-Protein Coupled Receptors (Nobel Lecture). Angewandte Chemie International Edition. 2013;52(25):6366–78.

96. Rasmussen SGF, Choi HJ, Fung JJ, Pardon E, Casarosa P, Chae PS, et al. Structure of a nanobody-stabilized active state of the β2 adrenoceptor. Nature. 2011 Jan;469(7329):175–80.

97. Kruse AC, Ring AM, Manglik A, Hu J, Hu K, Eitel K, et al. Activation and allosteric modulation of a muscarinic acetylcholine receptor. Nature. 2013 Dec;504(7478):101–6.

98. Huang W, Manglik A, Venkatakrishnan AJ, Laeremans T, Feinberg EN, Sanborn AL, et al. Structural insights into µ-opioid receptor activation. Nature. 2015 Aug;524(7565):315–21.

99. Carpenter B, Nehmé R, Warne T, Leslie AGW, Tate CG. Structure of the adenosine A2A receptor bound to an engineered G protein. Nature. 2016 Aug;536(7614):104–7.

100. Tsai CJ, Pamula F, Nehmé R, Mühle J, Weinert T, Flock T, et al. Crystal structure of rhodopsin in complex with a mini-Go sheds light on the principles of G protein selectivity. Science Advances. 2018 Sep 19;4(9):eaat7052.

101. Weis WI, Kobilka BK. The Molecular Basis of G Protein–Coupled Receptor Activation. Annual Review of Biochemistry [Internet]. 2014 Jun 12 [cited 2021 Aug 6]; Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev-biochem-060614-033910

102. Manglik A, Kruse AC. Structural Basis for G Protein-Coupled Receptor Activation. Biochemistry. 2017 Oct 24;56(42):5628–34.

103. Grahl A, Abiko LA, Isogai S, Sharpe T, Grzesiek S. A high-resolution description of β1-adrenergic receptor functional dynamics and allosteric coupling from backbone NMR. Nat Commun. 2020 May 5;11(1):2216.

104. Zhou XE, Gao X, Barty A, Kang Y, He Y, Liu W, et al. X-ray laser diffraction for structure determination of the rhodopsin-arrestin complex. Sci Data. 2016 Apr 12;3(1):160021.

105. Stauch B, Cherezov V. Serial Femtosecond Crystallography of G Protein–Coupled Receptors. Annual Review of Biophysics [Internet]. 2018 May 24 [cited 2022 Nov 13]; Available from: https://www.annualreviews.org/doi/abs/10.1146/annurev-biophys-070317-033239

106. Zhang H, Qiao A, Yang D, Yang L, Dai A, de Graaf C, et al. Structure of the full-length glucagon class B G-protein-coupled receptor. Nature. 2017 Jun;546(7657):259–64.

107. García-Nafría J, Tate CG. Structure determination of GPCRs: cryo-EM compared with X-ray crystallography. Biochemical Society Transactions. 2021 Nov 1;49(5):2345–55.

108. Piper SJ, Johnson RM, Wootten D, Sexton PM. Membranes under the Magnetic Lens: A Dive into the Diverse World of Membrane Protein Structures Using Cryo-EM. Chem Rev. 2022 Sep 14;122(17):13989–4017.

109. Xing C, Zhuang Y, Xu TH, Feng Z, Zhou XE, Chen M, et al. Cryo-EM Structure of the Human Cannabinoid Receptor CB2-Gi Signaling Complex. Cell. 2020 Feb;180(4):645-654.e13.

110. Zhao P, Liang YL, Belousoff MJ, Deganutti G, Fletcher MM, Willard FS, et al. Activation of the GLP-1 receptor by a non-peptidic agonist. Nature. 2020 Jan;577(7790):432–6.

111. Ping YQ, Mao C, Xiao P, Zhao RJ, Jiang Y, Yang Z, et al. Structures of the glucocorticoid-bound adhesion receptor GPR97–Go complex. Nature. 2021 Jan 28;589(7843):620–6.

112. Zhang M, Gui M, Wang ZF, Gorgulla C, Yu JJ, Wu H, et al. Cryo-EM structure of an activated GPCR–G protein complex in lipid nanodiscs. Nature Structural & Molecular Biology. 2021 Feb 25;1–10.

113. Zhuang Y, Xu P, Mao C, Wang L, Krumm B, Zhou XE, et al. Structural insights into the human D1 and D2 dopamine receptor signaling complexes. Cell. 2021 Feb;184(4):931-942.e18.

114. Su M, Paknejad N, Zhu L, Wang J, Do HN, Miao Y, et al. Structures of β1-adrenergic receptor in complex with Gs and ligands of different efficacies. Nat Commun. 2022 Jul 14;13(1):4095.

115. Zhang X, Johnson RM, Drulyte I, Yu L, Kotecha A, Danev R, et al. Evolving cryo-EM structural approaches for GPCR drug discovery. Structure. 2021 Sep 2;29(9):963-974.e6.

116. Bammes BE, Rochat RH, Jakana J, Chen DH, Chiu W. Direct electron detection yields cryo-EM reconstructions at resolutions beyond 3/4 Nyquist frequency. Journal of Structural Biology. 2012 Mar 1;177(3):589–601.

117. Yip KM, Fischer N, Paknia E, Chari A, Stark H. Atomic-resolution protein structure determination by cryo-EM. Nature. 2020 Nov;587(7832):157–61.

118. Nakane T, Kotecha A, Sente A, McMullan G, Masiulis S, Brown PMGE, et al. Single-particle cryo-EM at atomic resolution. Nature. 2020 Nov;587(7832):152–6.

119. Thompson RF, Iadanza MG, Hesketh EL, Rawson S, Ranson NA. Collection, pre-processing and on-the-fly analysis of data for high-resolution, single-particle cryo-electron microscopy. Nat Protoc. 2019 Jan;14(1):100–18.

120. Punjani A, Zhang H, Fleet DJ. Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nat Methods. 2020 Dec;17(12):1214–21.

121. Punjani A, Fleet DJ. 3D Variability Analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM [Internet]. Biophysics; 2020 Apr [cited 2020 Dec 10]. Available from: http://biorxiv.org/lookup/doi/10.1101/2020.04.08.032466

122. Zivanov J, Nakane T, Forsberg BO, Kimanius D, Hagen WJ, Lindahl E, et al. New tools for automated high-resolution cryo-EM structure determination in RELION-3. Egelman EH, Kuriyan J, editors. eLife. 2018 Nov 9;7:e42166.

123. Scheres SHW. A Bayesian View on Cryo-EM Structure Determination. Journal of Molecular Biology. 2012 Jan 13;415(2):406–18.

124. Weissenberger G, Henderikx RJM, Peters PJ. Understanding the invisible hands of sample preparation for cryo-EM. Nat Methods. 2021 May;18(5):463–71.

125. Kampjut D, Steiner J, Sazanov LA. Cryo-EM grid optimization for membrane proteins. iScience. 2021 Mar;24(3):102139.

126. Grisshammer R. The quest for high-resolution G protein-coupled receptor–G protein structures. Proc Natl Acad Sci USA. 2020 Mar 16;202002665.

127. Campbell MG, Cheng A, Brilot AF, Moeller A, Lyumkis D, Veesler D, et al. Movies of Ice-Embedded Particles Enhance Resolution in Electron Cryo-Microscopy. Structure. 2012 Nov 7;20(11):1823–8.

128. Scheres SH. Beam-induced motion correction for sub-megadalton cryo-EM particles. Kühlbrandt W, editor. eLife. 2014 Aug 13;3:e03665.

129. Nogales E, Scheres SHW. Cryo-EM: A Unique Tool for the Visualization of Macromolecular Complexity. Molecular Cell. 2015 May 21;58(4):677–89.

130. Danev R, Belousoff M, Liang YL, Zhang X, Eisenstein F, Wootten D, et al. Routine sub-2.5 Å cryo-EM structure determination of GPCRs. Nat Commun. 2021 Jul 15;12(1):4333.

131. Kato HE, Zhang Y, Hu H, Suomivuori CM, Kadji FMN, Aoki J, et al. Conformational transitions of a neurotensin receptor 1–Gi1 complex. Nature. 2019 Aug;572(7767):80–5.

132. Seven AB, Barros-Álvarez X, de Lapeyrière M, Papasergi-Scott MM, Robertson MJ, Zhang C, et al. G-protein activation by a metabotropic glutamate receptor. Nature. 2021 Jul;595(7867):450–4.

133. Punjani A, Fleet DJ. 3D Variability Analysis: Resolving continuous flexibility and discrete heterogeneity from single particle cryo-EM [Internet]. Biophysics; 2020 Apr [cited 2020 Dec 10]. Available from: http://biorxiv.org/lookup/doi/10.1101/2020.04.08.032466

134. Marino J, Schertler GFX. A set of common movements within GPCR-G-protein complexes from variability analysis of cryo-EM datasets. Journal of Structural Biology. 2021 Jun 1;213(2):107699.

135. Majerle A, Hadži S, Aupič J, Satler T, Lapenta F, Strmšek Ž, et al. A nanobody toolbox targeting dimeric coiled-coil modules for functionalization of designed

protein origami structures. Proc Natl Acad Sci USA. 2021 Apr 27;118(17):e2021899118.

136. Robertson MJ, Papasergi-Scott MM, He F, Seven AB, Meyerowitz JG, Panova O, et al. Structure determination of inactive-state GPCRs with a universal nanobody. Nat Struct Mol Biol. 2022 Nov 17;1–8.

137. Wu M, Lander GC, Herzik MA. Sub-2 Angstrom resolution structure determination using single-particle cryo-EM at 200 keV. Journal of Structural Biology: X. 2020 Jan 1;4:100020.

138. Isaikina P, Tsai CJ, Dietz N, Pamula F, Grahl A, Goldie KN, et al. Structural basis of the activation of the CC chemokine receptor 5 by a chemokine agonist. Science Advances. 2021 Jun 1;7(25):eabg8685.

139. Piper SJ, Johnson RM, Wootten D, Sexton PM. Membranes under the Magnetic Lens: A Dive into the Diverse World of Membrane Protein Structures Using Cryo-EM. Chem Rev. 2022 Sep 14;122(17):13989–4017.

140. Chen H, Chen K, Huang W, Staudt LM, Cyster JG, Li X. Structure of S1PR2–heterotrimeric G13 signaling complex. Science Advances. 2022 Mar 30;8(13):eabn0067.

141. Inoue A, Raimondi F, Kadji FMN, Singh G, Kishi T, Uwamizu A, et al. Illuminating G-Protein-Coupling Selectivity of GPCRs. Cell. 2019 Jun 13;177(7):1933-1947.e25.

142. Huang S, Xu P, Shen DD, Simon IA, Mao C, Tan Y, et al. GPCRs steer Gi and Gs selectivity via TM5-TM6 switches as revealed by structures of serotonin receptors. Molecular Cell. 2022 Jul 21;82(14):2681-2695.e6.

143. Suno R, Sugita Y, Morimoto K, Takazaki H, Tsujimoto H, Hirose M, et al. Structural insights into the G protein selectivity revealed by the human EP3-Gi signaling complex. Cell Reports. 2022 Sep;40(11):111323.

144. Duan J, Shen DD, Zhao T, Guo S, He X, Yin W, et al. Molecular basis for allosteric agonism and G protein subtype selectivity of galanin receptors. Nat Commun. 2022 Mar 15;13(1):1364.

145. Harris JA, Faust B, Gondin AB, Dämgen MA, Suomivuori CM, Veldhuis NA, et al. Selective G protein signaling driven by substance P–neurokinin receptor dynamics. Nat Chem Biol. 2022 Jan;18(1):109–15.

146. Liu Q, Yang D, Zhuang Y, Croll TI, Cai X, Dai A, et al. Ligand recognition and G-protein coupling selectivity of cholecystokinin A receptor. Nat Chem Biol. 2021 Dec;17(12):1238–44.

147. Qian Y, Ma Z, Liu C, Li X, Zhu X, Wang N, et al. Structural insights into adhesion GPCR ADGRL3 activation and Gq, Gs, Gi, and G12 coupling. Molecular Cell. 2022 Oct;S1097276522009662.

148. Duan J, Liu H, Ji Y, Yuan Q, Li X, Wu K, et al. Structure of a G protein-coupled receptor with GRK2 and a biased ligand. bioRxiv. 2022 Oct 19;2022.10.19.512855.

149. Zhou XE, He Y, de Waal PW, Gao X, Kang Y, Van Eps N, et al. Identification of Phosphorylation Codes for Arrestin Recruitment by G Protein-Coupled Receptors. Cell. 2017 Jul 27;170(3):457-469.e13.

150. Lee Y, Warne T, Nehmé R, Pandey S, Dwivedi-Agnihotri H, Chaturvedi M, et al. Molecular basis of β-arrestin coupling to formoterol-bound β1-adrenoceptor. Nature. 2020 Jul 30;583(7818):862–6.

151. Bous J, Fouillen A, Orcel H, Trapani S, Cong X, Fontanel S, et al. Structure of the vasopressin hormone–V2 receptor–β-arrestin1 ternary complex. Science Advances [Internet]. 2022 Sep [cited 2022 Sep 13]; Available from: https://www.science.org/doi/10.1126/sciadv.abo7761

152. Cao C, Barros-Álvarez X, Zhang S, Kim K, Dämgen MA, Panova O, et al. Signaling snapshots of a serotonin receptor activated by the prototypical psychedelic LSD. Neuron [Internet]. 2022 Sep 9 [cited 2022 Sep 20]; Available from: https://www.sciencedirect.com/science/article/pii/S0896627322007528

153. Staus DP, Hu H, Robertson MJ, Kleinhenz ALW, Wingler LM, Capel WD, et al. Structure of the M2 muscarinic receptor–β-arrestin complex in a lipid nanodisc. Nature. 2020 Mar;579(7798):297–302.

154. Nguyen AH, Lefkowitz RJ. Signaling at the endosome: cryo-EM structure of a GPCR–G protein–beta-arrestin megacomplex. FEBS J. 2021 Mar 8;febs.15773.

155. Cao C, Kang HJ, Singh I, Chen H, Zhang C, Ye W, et al. Structure, function and pharmacology of human itch GPCRs. Nature. 2021 Dec;600(7887):170–5.

156. Zhou Y, Daver H, Trapkov B, Wu L, Wu M, Harpsøe K, et al. Molecular insights into ligand recognition and G protein coupling of the neuromodulatory orphan receptor GPR139. Cell Res. 2022 Feb;32(2):210–3.

157. Lin X, Li M, Wang N, Wu Y, Luo Z, Guo S, et al. Structural basis of ligand recognition and self-activation of orphan GPR52. Nature. 2020 Mar;579(7797):152–7.

158. Liang YL, Khoshouei M, Radjainia M, Zhang Y, Glukhova A, Tarrasch J, et al. Phase-plate cryo-EM structure of a class B GPCR–G-protein complex. Nature. 2017 Jun;546(7656):118–23.

159. Zhang Y, Sun B, Feng D, Hu H, Chu M, Qu Q, et al. Cryo-EM structure of the activated GLP-1 receptor in complex with a G protein. Nature. 2017 Jun;546(7657):248–53.

160. Josephs TM, Belousoff MJ, Liang YL, Piper SJ, Cao J, Garama DJ, et al. Structure and dynamics of the CGRP receptor in apo and peptide-bound forms. Science. 2021 Apr 9;372(6538):eabf7258.

161. Cao J, Belousoff MJ, Liang YL, Johnson RM, Josephs TM, Fletcher MM, et al. A structural basis for amylin receptor phenotype. Science. 2022 Mar 25;375(6587):eabm9609.

162. Qiao A, Han S, Li X, Li Z, Zhao P, Dai A, et al. Structural basis of G $_s$ and G $_i$ recognition by the human glucagon receptor. Science. 2020 Mar 20;367(6484):1346–52.

163. Lin H, Xiao P, Bu RQ, Guo S, Yang Z, Yuan D, et al. Structures of the ADGRG2–Gs complex in apo and ligand-bound forms. Nat Chem Biol. 2022 Nov;18(11):1196–203.

164. Xiao P, Guo S, Wen X, He QT, Lin H, Huang SM, et al. Tethered peptide activation mechanism of the adhesion GPCRs ADGRG2 and ADGRG4. Nature. 2022 Apr;604(7907):771–8.

165. Lala T, Hall RA. Adhesion G protein-coupled receptors: structure, signaling, physiology, and pathophysiology. Physiological Reviews. 2022 Oct;102(4):1587–624.

166. Ping YQ, Mao C, Xiao P, Zhao RJ, Jiang Y, Yang Z, et al. Structures of the glucocorticoid-bound adhesion receptor GPR97–Go complex. Nature. 2021 Jan;589(7843):620–6.

167. Ping YQ, Xiao P, Yang F, Zhao RJ, Guo SC, Yan X, et al. Structural basis for the tethered peptide activation of adhesion GPCRs. Nature. 2022 Apr;604(7907):763–70.

168. Barros-Álvarez X, Nwokonko RM, Vizurraga A, Matzov D, He F, Papasergi-Scott MM, et al. The tethered peptide activation mechanism of adhesion GPCRs. Nature. 2022 Apr;604(7907):757–62.

169. Patil DN, Singh S, Laboute T, Strutzenberg TS, Qiu X, Wu D, et al. Cryo-EM structure of human GPR158 receptor coupled to the RGS7-Gβ5 signaling complex. Science. 2022 Jan 7;375(6576):86–91.

170. Lin S, Han S, Cai X, Tan Q, Zhou K, Wang D, et al. Structures of Gi-bound metabotropic glutamate receptors mGlu2 and mGlu4. Nature. 2021 Jun;594(7864):583–8.

171. Shen C, Mao C, Xu C, Jin N, Zhang H, Shen DD, et al. Structural basis of GABAB receptor–Gi protein coupling. Nature. 2021 Jun;594(7864):594–8.

172. Velazhahan V, Ma N, Pándy-Szekeres G, Kooistra AJ, Lee Y, Gloriam DE, et al. Structure of the class D GPCR Ste2 dimer coupled to two G proteins. Nature

[Internet]. 2020 Dec 2 [cited 2020 Dec 5]; Available from: http://www.nature.com/articles/s41586-020-2994-1

173. Qi X, Friedberg L, De Bose-Boyd R, Long T, Li X. Sterols in an intramolecular channel of Smoothened mediate Hedgehog signaling. Nat Chem Biol. 2020 Dec;16(12):1368–75.

174. Tsutsumi N, Mukherjee S, Waghray D, Janda CY, Jude KM, Miao Y, et al. Structure of human Frizzled5 by fiducial-assisted cryo-EM supports a heterodimeric mechanism of canonical Wnt signaling. Zeng YA, Boudker O, Hannoush R, editors. eLife. 2020 Aug 7;9:e58464.

175. Nygaard R, Zou Y, Dror RO, Mildorf TJ, Arlow DH, Manglik A, et al. The Dynamic Process of β2-Adrenergic Receptor Activation. Cell. 2013 Jan 31;152(3):532–42.

176. Grahl A, Isogai S, Sharpe T, Grzesiek S. NMR backbone dynamics reveals mechanism of ligand to effector site allosteric coupling in the β1-adrenergic receptor. :26.

177. Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, et al. Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. Cell. 2015 May 21;161(5):1101–11.

178. Gregorio GG, Masureel M, Hilger D, Terry DS, Juette M, Zhao H, et al. Single-molecule analysis of ligand efficacy in β2AR–G-protein activation. Nature. 2017 Jul;547(7661):68–73.

179. Dror RO, Arlow DH, Maragakis P, Mildorf TJ, Pan AC, Xu H, et al. Activation mechanism of the β2-adrenergic receptor. Proceedings of the National Academy of Sciences. 2011 Nov 15;108(46):18684–9.

180. Latorraca NR, Venkatakrishnan AJ, Dror RO. GPCR Dynamics: Structures in Motion. Chem Rev. 2017 Jan 11;117(1):139–55.

181. Altenbach C, Kusnetzow AK, Ernst OP, Hofmann KP, Hubbell WL. High-resolution distance mapping in rhodopsin reveals the pattern of helix movement due to activation. Proceedings of the National Academy of Sciences. 2008 May 27;105(21):7439–44.

182. Hoffmann C, Zürn A, Bünemann M, Lohse MJ. Conformational changes in G-protein-coupled receptors—the quest for functionally selective conformations is open. British Journal of Pharmacology. 2008;153(S1):S358–66.

183. Janetopoulos C, Jin T, Devreotes P. Receptor-Mediated Activation of Heterotrimeric G-Proteins in Living Cells. Science. 2001 Mar 23;291(5512):2408–11.

184. Marullo S, Bouvier M. Resonance energy transfer approaches in molecular pharmacology and beyond. Trends in Pharmacological Sciences. 2007 Aug 1;28(8):362–5.

185. Lohse MJ, Vilardaga JP, Bünemann M. Direct optical recording of intrinsic efficacy at a G protein-coupled receptor. Life Sciences. 2003 Dec 5;74(2):397–404.

186. Opitz C, Isogai S, Grzesiek S. An economic approach to efficient isotope labeling in insect cells using homemade 15N-, 13C- and 2H-labeled yeast extracts. J Biomol NMR. 2015 Jul 1;62(3):373–85.

187. Franke B, Opitz C, Isogai S, Grahl A, Delgado L, Gossert AD, et al. Production of isotope-labeled proteins in insect cells for NMR. J Biomol NMR. 2018 Jul 1;71(3):173–84.

188. Pervushin K, Riek R, Wider G, Wüthrich K. Transverse Relaxation-Optimized Spectroscopy (TROSY) for NMR Studies of Aromatic Spin Systems in 13C-Labeled Proteins. J Am Chem Soc. 1998 Jul 1;120(25):6394–400.

189. Shimada I, Ueda T, Kofuku Y, Eddy MT, Wüthrich K. GPCR drug discovery: integrating solution NMR data with crystal and cryo-EM structures. Nat Rev Drug Discov. 2019 Jan;18(1):59–82.

190. Wu FJ, Bumbak F, Tanipour MH, Asadollahi K, Vaid TM, Sethi A, et al. Chapter 13:Characterizing Conformational Diversity of G Protein-coupled Receptors by Solution NMR Spectroscopy. In: NMR Spectroscopy for Probing Functional Dynamics at Biological Interfaces [Internet]. 2022 [cited 2022 Dec 6]. p. 346–82. Available from: https://pubs.rsc.org/en/content/chapter/bk9781839162091-00346/978-1-83916-209-1

191. Wu FJ, Rieder PS, Abiko LA, Rößler P, Gossert AD, Häussinger D, et al. Nanobody GPS by PCS: An Efficient New NMR Analysis Method for G Protein Coupled Receptors and Other Large Proteins. J Am Chem Soc. 2022 Nov 30;144(47):21728–40.

192. Solt AS, Bostock MJ, Shrestha B, Kumar P, Warne T, Tate CG, et al. Insight into partial agonism by observing multiple equilibria for ligand-bound and Gs-mimetic nanobody-bound β1-adrenergic receptor. Nat Commun. 2017 Nov 27;8(1):1795.

193. Wu FJ, Williams LM, Abdul-Ridha A, Gunatilaka A, Vaid TM, Kocan M, et al. Probing the correlation between ligand efficacy and conformational diversity at the α1A-adrenoreceptor reveals allosteric coupling of its microswitches. Journal of Biological Chemistry. 2020 May 22;295(21):7404–17.

194. Bokoch MP, Zou Y, Rasmussen SGF, Liu CW, Nygaard R, Rosenbaum DM, et al. Ligand-specific regulation of the extracellular surface of a G-protein-coupled receptor. Nature. 2010 Jan;463(7277):108–12.

195. Frei JN, Broadhurst RW, Bostock MJ, Solt A, Jones AJY, Gabriel F, et al. Conformational plasticity of ligand-bound and ternary GPCR complexes studied by 19F NMR of the β1-adrenergic receptor. Nat Commun. 2020 Feb 3;11(1):669.

196. Huang SK, Pandey A, Tran DP, Villanueva NL, Kitao A, Sunahara RK, et al. Delineating the conformational landscape of the adenosine A2A receptor during G protein coupling. Cell. 2021 Apr 1;184(7):1884-1894.e14.

197. Ye L, Van Eps N, Zimmer M, Ernst OP, Scott Prosser R. Activation of the A2A adenosine G-protein-coupled receptor by conformational selection. Nature. 2016 May;533(7602):265–8.

198. Manglik A, Kim TH, Masureel M, Altenbach C, Yang Z, Hilger D, et al. Structural Insights into the Dynamic Process of β2-Adrenergic Receptor Signaling. Cell. 2015 May 21;161(5):1101–11.

199. Liu JJ, Horst R, Katritch V, Stevens RC, Wüthrich K. Biased Signaling Pathways in β2-Adrenergic Receptor Characterized by 19F-NMR. Science. 2012 Mar 2;335(6072):1106–10.

200. Eddy MT, Lee MY, Gao ZG, White KL, Didenko T, Horst R, et al. Allosteric Coupling of Drug Binding and Intracellular Signaling in the A2A Adenosine Receptor. Cell. 2018 Jan 11;172(1):68-80.e12.

201. Imai S, Yokomizo T, Kofuku Y, Shiraishi Y, Ueda T, Shimada I. Structural equilibrium underlying ligand-dependent activation of β2-adrenoreceptor. Nat Chem Biol. 2020 Apr;16(4):430–9.

202. Martens C, Shekhar M, Lau AM, Tajkhorshid E, Politis A. Integrating hydrogen–deuterium exchange mass spectrometry with molecular dynamics simulations to probe lipid-modulated conformational changes in membrane proteins. Nat Protoc. 2019 Nov;14(11):3183–204.

203. Gavriilidou AFM, Hunziker H, Mayer D, Vuckovic Z, Veprintsev DB, Zenobi R. Insights into the Basal Activity and Activation Mechanism of the β1 Adrenergic Receptor Using Native Mass Spectrometry. J Am Soc Mass Spectrom. 2019 Mar 1;30(3):529–37.

204. Chung KY, Rasmussen SGF, Liu T, Li S, DeVree BT, Chae PS, et al. Conformational changes in the G protein Gs induced by the β2 adrenergic receptor. Nature. 2011 Sep;477(7366):611–5.

205. Yen HY, Liko I, Song W, Kapoor P, Almeida F, Toporowska J, et al. Mass spectrometry captures biased signalling and allosteric modulation of a G-protein-coupled receptor. Nat Chem. 2022 Nov 10;1–8.

206. Dandey VP, Budell WC, Wei H, Bobe D, Maruthi K, Kopylov M, et al. Time-resolved cryo-EM using Spotiton. Nat Methods. 2020 Sep;17(9):897–900.

207. Orville AM. Recent results in time resolved serial femtosecond crystallography at XFELs. Current Opinion in Structural Biology. 2020 Dec 1;65:193–208.

208. Gruhl T, Weinert T, Rodrigues M, Milne CJ, Ortolani G, Nass K, et al. ULTRAFAST STRUCTURAL CHANGES DIRECT THE FIRST MOLECULAR EVENTS OF

VISION [Internet]. Biochemistry; 2022 Oct [cited 2022 Nov 9]. Available from: http://biorxiv.org/lookup/doi/10.1101/2022.10.14.511948

209. Hughes CE, Nibbs RJB. A guide to chemokines and their receptors. FEBS J. 2018 Aug;285(16):2944–71.

210. Mollica Poeta V, Massara M, Capucetti A, Bonecchi R. Chemokines and Chemokine Receptors: New Targets for Cancer Immunotherapy. Frontiers in Immunology [Internet]. 2019 [cited 2022 Dec 7];10. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2019.00379

211. Steen A, Larsen O, Thiele S, Rosenkilde MM. Biased and G Protein-Independent Signaling of Chemokine Receptors. Frontiers in Immunology [Internet]. 2014 [cited 2022 Dec 8];5. Available from: https://www.frontiersin.org/articles/10.3389/fimmu.2014.00277

212. Zweemer AJM, Toraskar J, Heitman LH, IJzerman AP. Bias in chemokine receptor signalling. Trends in Immunology. 2014 Jun 1;35(6):243–52.

213. Ravindran A, Sawant KV, Sarmiento J, Navarro J, Rajarathnam K. Chemokine CXCL1 Dimer Is a Potent Agonist for the CXCR2 Receptor *. Journal of Biological Chemistry. 2013 Apr 26;288(17):12244–52.

214. Liu K, Wu L, Yuan S, Wu M, Xu Y, Sun Q, et al. Structural basis of CXC chemokine receptor 2 activation and signalling. Nature. 2020 Sep 3;585(7823):135–40.

215. Miller MC, Mayo KH. Chemokines from a Structural Perspective. International Journal of Molecular Sciences. 2017 Oct;18(10):2088.

216. Duma L, Häussinger D, Rogowski M, Lusso P, Grzesiek S. Recognition of RANTES by Extracellular Parts of the CCR5 Receptor. Journal of Molecular Biology. 2007 Jan 26;365(4):1063–75.

217. Handel TM, Dyer DP. Perspectives on the Biological Role of Chemokine:Glycosaminoglycan Interactions. J Histochem Cytochem. 2021 Feb 1;69(2):87–91.

218. Mellado M, Rodríguez-Frade JM, Vila-Coro AJ, Fernández S, Martín de Ana A, Jones DR, et al. Chemokine receptor homo- or heterodimerization activates distinct signaling pathways. The EMBO Journal. 2001 May 15;20(10):2497–507.

219. Kobayashi D, Endo M, Ochi H, Hojo H, Miyasaka M, Hayasaka H. Regulation of CCR7-dependent cell migration through CCR7 homodimer formation. Sci Rep. 2017 Aug 17;7(1):8536.

220. Martínez-Muñoz L, Villares R, Rodríguez-Fernández JL, Rodríguez-Frade JM, Mellado M. Remodeling our concept of chemokine receptor function: From monomers to oligomers. Journal of Leukocyte Biology. 2018;104(2):323–31.

221. Ge B, Lao J, Li J, Chen Y, Song Y, Huang F. Single-molecule imaging reveals dimerization/oligomerization of CXCR4 on plasma membrane closely related to its function. Sci Rep. 2017 Dec 4;7(1):16873.

222. Goth CK, Petäjä-Repo UE, Rosenkilde MM. G Protein-Coupled Receptors in the Sweet Spot: Glycosylation and other Post-translational Modifications. ACS Pharmacol Transl Sci. 2020 Apr 10;3(2):237–45.

223. Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC, Basnet H, et al. Structural Basis of CXCR4 Sulfotyrosine Recognition by the Chemokine SDF-1/CXCL12. Science Signaling. 2008 Sep 16;1(37):ra4–ra4.

224. Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, et al. Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. Science. 2016 Jan 8;351(6269):186–90.

225. Scurci I, Akondi KB, Pinheiro I, Paolini-Bertrand M, Borgeat A, Cerini F, et al. CCR5 tyrosine sulfation heterogeneity generates cell surface receptor subpopulations with different ligand binding properties. Biochimica et Biophysica Acta (BBA) - General Subjects. 2021 Jan;1865(1):129753.

226. Shao Z, Shen Q, Yao B, Mao C, Chen LN, Zhang H, et al. Identification and mechanism of G protein-biased ligands for chemokine receptor CCR1. Nat Chem Biol. 2022 Mar;18(3):264–71.

227. Shao Z, Tan Y, Shen Q, Hou L, Yao B, Qin J, et al. Molecular insights into ligand recognition and activation of chemokine receptors CCR2 and CCR3. Cell Discov. 2022 May 15;8(1):1–11.

228. Zhang H, Chen K, Tan Q, Shao Q, Han S, Zhang C, et al. Structural basis for chemokine recognition and receptor activation of chemokine receptor CCR5. Nat Commun. 2021 Jul 6;12(1):4151.

229. Wasilko DJ, Johnson ZL, Ammirati M, Che Y, Griffor MC, Han S, et al. Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20. Nat Commun. 2020 Dec;11(1):3031.

230. Samson M, Libert F, Doranz BJ, Rucker J, Liesnard C, Farber CM, et al. Resistance to HIV-1 infection in Caucasian individuals bearing mutant alleles of the CCR-5 chemokine receptor gene. Nature. 1996 Aug;382(6593):722–5.

231. Raport CJ, Gosling J, Schweickart VL, Gray PW, Charo IF. Molecular Cloning and Functional Characterization of a Novel Human CC Chemokine Receptor (CCR5) for RANTES, MIP-1β, and MIP-1α*. Journal of Biological Chemistry. 1996 Jul 19;271(29):17161–6.

232. Combadiere C, Ahuja SK, Lee Tiffany H, Murphy PM. Cloning and functional expression of CC CKR5, a human monocyte CC chemokine receptor selective for MIP-1α, MIP-1β, and RANTES. Journal of Leukocyte Biology. 1996;60(1):147–52.

233. Alkhatib G. The biology of CCR5 and CXCR4: Current Opinion in HIV and AIDS. 2009 Mar;4(2):96–103.

234. Gerard C, Rollins BJ. Chemokines and disease. Nat Immunol. 2001 Feb;2(2):108–15.

235. Aldinucci D, Borghese C, Casagrande N. The CCL5/CCR5 Axis in Cancer Progression. Cancers. 2020 Jul 2;12(7):1765.

236. Halama N, Zoernig I, Berthel A, Kahlert C, Klupp F, Suarez-Carmona M, et al. Tumoral Immune Cell Exploitation in Colorectal Cancer Metastases Can Be Targeted Effectively by Anti-CCR5 Therapy in Cancer Patients. Cancer Cell. 2016 Apr 11;29(4):587–601.

237. Chua RL, Lukassen S, Trump S, Hennig BP, Wendisch D, Pott F, et al. COVID-19 severity correlates with airway epithelium–immune cell interactions identified by single-cell analysis. Nat Biotechnol. 2020 Aug;38(8):970–9.

238. Cocchi F, DeVico AL, Garzino-Demo A, Arya SK, Gallo RC, Lusso P. Identification of RANTES, MIP-1α, and MIP-1β as the Major HIV-Suppressive Factors Produced by CD8+ T Cells. Science. 1995 Dec 15;270(5243):1811–5.

239. Schols D, Proost P, Struyf S, Wuyts A, De Meester I, Scharpé S, et al. CD26-processed RANTES(3–68), but not intact RANTES, has potent anti-HIV-1 activity. Antiviral Research. 1998 Oct 1;39(3):175–87.

240. Hartley O, Dorgham K, Perez-Bercoff D, Cerini F, Heimann A, Gaertner H, et al. Human Immunodeficiency Virus Type 1 Entry Inhibitors Selected on Living Cells from a Library of Phage Chemokines. Journal of Virology. 2003 Jun 15;77(12):6637–44.

241. Hartley O, Gaertner H, Wilken J, Thompson D, Fish R, Ramos A, et al. Medicinal chemistry applied to a synthetic protein: Development of highly potent HIV entry inhibitors. PNAS. 2004 Nov 23;101(47):16460–5.

242. Cerini F, Offord R, McGowan I, Hartley O. Stability of 5P12-RANTES, A Candidate Rectal Microbicide, in Human Rectal Lavage. AIDS Research and Human Retroviruses. 2017 Aug 1;33(8):768–77.

243. Gaertner H, Cerini F, Escola JM, Kuenzi G, Melotti A, Offord R, et al. Highly potent, fully recombinant anti-HIV chemokines: Reengineering a low-cost microbicide. PNAS. 2008 Nov 18;105(46):17706–11.

244. McGowan IM, Tzakis N, Kosak B, Korczak B, Engstrom J, Tomaszewska-Kiecana M, et al. Evaluation of the Safety, Acceptability, and Pharmacokinetic Profile of a Gel Formulation of OB-002 in Healthy Volunteers. AIDS Research and Human Retroviruses. 2021 Jun;37(6):453–60.

245. Bönsch C, Munteanu M, Rossitto-Borlat I, Fürstenberg A, Hartley O. Potent Anti-HIV Chemokine Analogs Direct Post-Endocytic Sorting of CCR5. PLOS ONE. 2015 Apr 29;10(4):e0125396.

246. Zheng Y, Han GW, Abagyan R, Wu B, Stevens RC, Cherezov V, 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 Jun;46(6):1005-1017.e5.

247. Wang X, Watson C, Sharp JS, Handel TM, Prestegard JH. Oligomeric Structure of the Chemokine CCL5/RANTES from NMR, MS, and SAXS Data. Structure (London, England : 1993). 2011 Aug 10;19(8):1138.

248. Czaplewski LG, McKeating J, Craven CJ, Higgins LD, Appay V, Brown A, et al. Identification of Amino Acid Residues Critical for Aggregation of Human CC Chemokines Macrophage Inflammatory Protein (MIP)-1α, MIP-1β, and RANTES: CHARACTERIZATION OF ACTIVE DISAGGREGATED CHEMOKINE VARIANTS*. Journal of Biological Chemistry. 1999 Jun 4;274(23):16077–84.

249. Dorgham K, Cerini F, Gaertner H, Melotti A, Rossitto-Borlat I, Gorochov G, et al. Chapter Three - Generating Chemokine Analogs with Enhanced Pharmacological Properties Using Phage Display. In: Handel TM, editor. Methods in Enzymology [Internet]. Academic Press; 2016 [cited 2022 Dec 9]. p. 47–72. (Chemokines; vol. 570). Available from: https://www.sciencedirect.com/science/article/pii/S0076687915005182

250. Lorenzen E, Ceraudo E, Berchiche YA, Rico CA, Fürstenberg A, Sakmar TP, et al. G protein subtype–specific signaling bias in a series of CCR5 chemokine analogs. Sci Signal [Internet]. 2018 Oct 16 [cited 2021 Jun 23];11(552). Available from: https://stke.sciencemag.org/content/11/552/eaao6152

251. Rico CA, Berchiche YA, Horioka M, Peeler JC, Lorenzen E, Tian H, et al. High-Affinity Binding of Chemokine Analogs that Display Ligand Bias at the HIV-1 Coreceptor CCR5. Biophysical Journal. 2019 Sep;117(5):903–19.

252. Martins E, Brodier H, Rossitto-Borlat I, Ilgaz I, Villard M, Hartley O. Arrestin Recruitment to C-C Chemokine Receptor 5: Potent C-C Chemokine Ligand 5 Analogs Reveal Differences in Dependence on Receptor Phosphorylation and Isoform-Specific Recruitment Bias. Mol Pharmacol. 2020 Nov 1;98(5):599–611.

253. Pollok-Kopp B, Schwarze K, Baradari VK, Oppermann M. Analysis of Ligand-stimulated CC Chemokine Receptor 5 (CCR5) Phosphorylation in Intact Cells Using Phosphosite-specific Antibodies*. Journal of Biological Chemistry. 2003 Jan 24;278(4):2190–8.

254. Gaertner H, Lebeau O, Borlat I, Cerini F, Dufour B, Kuenzi G, et al. Highly potent HIV inhibition: engineering a key anti-HIV structure from PSC-RANTES into MIP-1β/CCL4. Protein Engineering, Design and Selection. 2008 Feb 1;21(2):65–72.

255. Wiktor M, Hartley O, Grzesiek S. Characterization of Structure, Dynamics, and Detergent Interactions of the Anti-HIV Chemokine Variant 5P12-RANTES. Biophysical Journal. 2013 Dec;105(11):2586–97.

256. Cerini F, Landay A, Gichinga C, Lederman MM, Flyckt R, Starks D, et al. Chemokine Analogues Show Suitable Stability for Development as Microbicides. JAIDS Journal of Acquired Immune Deficiency Syndromes. 2008 Dec;49(5):472.

257. Jin H, Kagiampakis I, Li P, LiWang PJ. Structural and functional studies of the potent anti-HIV chemokine variant P2-RANTES. Proteins: Structure, Function, and Bioinformatics. 2010;78(2):295–308.

258. Isaikina P, Tsai CJ, Petrovic I, Rogowski M, Meng-Dürr A, Grzesiek S. Preparation of a stable CCL5•CCR5•Gi signaling complex for Cryo-EM analysis. In: Methods in Cell Biology. First edition. Elsevier; 2022.

259. Shiraishi Y, Kofuku Y, Ueda T, Pandey S, Dwivedi-Agnihotri H, Shukla AK, et al. Biphasic activation of β-arrestin 1 upon interaction with a GPCR revealed by methyl-TROSY NMR. Nat Commun. 2021 Dec 9;12(1):7158.

260. Nguyen AH, Thomsen ARB, Cahill TJ, Huang R, Huang LY, Marcink T, et al. Structure of an endosomal signaling GPCR–G protein–β-arrestin megacomplex. Nat Struct Mol Biol. 2019 Dec;26(12):1123–31.

261. Inglese J, Koch WJ, Caron MG, Lefkowitz RJ. Isoprenylation in regulation of signal transduction by G-protein-coupled receptor kinases. Nature. 1992 Sep;359(6391):147–50.

262. Noble AJ, Wei H, Dandey VP, Zhang Z, Tan YZ, Potter CS, et al. Reducing effects of particle adsorption to the air–water interface in cryo-EM. Nat Methods. 2018 Oct;15(10):793–5.

263. Noble AJ, Dandey VP, Wei H, Brasch J, Chase J, Acharya P, et al. Routine single particle CryoEM sample and grid characterization by tomography. Brunger AT, editor. eLife. 2018 May 29;7:e34257.

264. Carragher B, Cheng Y, Frost A, Glaeser R m., Lander G c., Nogales E, et al. Current outcomes when optimizing ‘standard’ sample preparation for single-particle cryo-EM. Journal of Microscopy. 2019;276(1):39–45.

265. Dubochet J, Adrian M, Chang JJ, Homo JC, Lepault J, McDowall AW, et al. Cryo-electron microscopy of vitrified specimens. Quarterly Reviews of Biophysics. 1988 May;21(2):129–228.

266. Taylor KA, Glaeser RM. Retrospective on the early development of cryoelectron microscopy of macromolecules and a prospective on opportunities for the future. Journal of Structural Biology. 2008 Sep 1;163(3):214–23.

267. Chen J, Noble AJ, Kang JY, Darst SA. Eliminating effects of particle adsorption to the air/water interface in single-particle cryo-electron microscopy: Bacterial

RNA polymerase and CHAPSO. Journal of Structural Biology: X. 2019 Jan 1;1:100005.

268. Noble AJ, Wei H, Dandey VP, Zhang Z, Tan YZ, Potter CS, et al. Reducing effects of particle adsorption to the air–water interface in cryo-EM. Nat Methods. 2018 Oct;15(10):793–5.

269. D'Imprima E, Floris D, Joppe M, Sánchez R, Grininger M, Kühlbrandt W. Protein denaturation at the air-water interface and how to prevent it. Brunger AT, Kuriyan J, Glaeser RM, Skiniotis G, editors. eLife. 2019 Apr 1;8:e42747.

270. Klebl DP, Gravett MSC, Kontziampasis D, Wright DJ, Bon RS, Monteiro DCF, et al. Need for Speed: Examining Protein Behavior during CryoEM Grid Preparation at Different Timescales. Structure. 2020 Nov 3;28(11):1238-1248.e4.

271. Efremov RG, Leitner A, Aebersold R, Raunser S. Architecture and conformational switch mechanism of the ryanodine receptor. Nature. 2015 Jan;517(7532):39–43.

272. Johnson ZL, Chen J. Structural Basis of Substrate Recognition by the Multidrug Resistance Protein MRP1. Cell. 2017 Mar 9;168(6):1075-1085.e9.

273. Razinkov I, Dandey VP, Wei H, Zhang Z, Melnekoff D, Rice WJ, et al. A new method for vitrifying samples for cryoEM. Journal of Structural Biology. 2016 Aug 1;195(2):190–8.

274. Dandey VP, Wei H, Zhang Z, Tan YZ, Acharya P, Eng ET, et al. Spotiton: New features and applications. Journal of Structural Biology. 2018 May 1;202(2):161–9.

275. Ravelli RBG, Nijpels FJT, Henderikx RJM, Weissenberger G, Thewessem S, Gijsbers A, et al. Cryo-EM structures from sub-nl volumes using pin-printing and jet vitrification. Nat Commun. 2020 May 22;11(1):2563.

276. Rima L, Zimmermann M, Fränkl A, Clairfeuille T, Lauer M, Engel A, et al. cryoWriter: a blotting free cryo-EM preparation system with a climate jet and cover-slip injector. Faraday Discuss. 2022 Nov 8;240(0):55–66.

277. Arnold SA, Albiez S, Bieri A, Syntychaki A, Adaixo R, McLeod RA, et al. Blotting-free and lossless cryo-electron microscopy grid preparation from nanoliter-sized protein samples and single-cell extracts. Journal of Structural Biology. 2017 Mar 1;197(3):220–6.

278. Schmidli C, Albiez S, Rima L, Righetto R, Mohammed I, Oliva P, et al. Microfluidic protein isolation and sample preparation for high-resolution cryo-EM. Proceedings of the National Academy of Sciences. 2019 Jul 23;116(30):15007–12.

279. Peisley A, Skiniotis G. 2D Projection Analysis of GPCR Complexes by Negative Stain Electron Microscopy. G Protein-Coupled Receptors in Drug Discovery. 2015;29–38.

280. Punjani A, Rubinstein JL, Fleet DJ, Brubaker MA. cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nat Methods. 2017 Mar;14(3):290–6.

281. Schmidli C, Rima L, Arnold SA, Stohler T, Syntychaki A, Bieri A, et al. Miniaturized Sample Preparation for Transmission Electron Microscopy. JoVE. 2018 Jul 27;(137):57310.

282. Shukla AK, Manglik A, Kruse AC, Xiao K, Reis RI, Tseng WC, et al. Structure of active β-arrestin-1 bound to a G-protein-coupled receptor phosphopeptide. Nature. 2013 May;497(7447):137–41.

283. Mayer D, Damberger FF, Samarasimhareddy M, Feldmueller M, Vuckovic Z, Flock T, et al. Distinct G protein-coupled receptor phosphorylation motifs modulate arrestin affinity and activation and global conformation. Nat Commun. 2019 Dec;10(1):1261.