

Classification: Structural Biology and Molecular Physics

Long-range coupled motions underlie ligand recognition by a chemokine receptor

Krishna Mohan Sepuru^{1,3}, Vinay Nair^{4,5}, Priyanka Prakash⁴, Alemayehu A. Gorfe^{4,5}, and Krishna Rajarathnam^{1,2,3*}

¹Department of Biochemistry and Molecular Biology, ²Department of Microbiology and Immunology, and ³Sealy Center for Structural Biology and Molecular Biophysics, The University of Texas Medical Branch, Galveston, Texas 77555

⁴Integrative Biology and Pharmacology, McGovern Medical School, University of Texas Health Science Center at Houston, Houston, TX.

⁵MD Anderson Cancer Center UTHealth Graduate School of Biomedical Sciences, Houston, TX.

*To whom correspondence should be addressed: Krishna Rajarathnam, Department of Biochemistry and Molecular Biology, The University of Texas Medical Branch, Galveston, TX 77555, USA, Tel: +1-409-772-2238; E-mail: krrojara@utmb.edu

Abstract

Chemokines are unusual class-A GPCR agonists because of their large size (~10 kDa) and binding at two distinct receptor sites: N-terminal domain (Site-I, unique to chemokines) and a groove defined by extracellular loop/transmembrane helices (Site-II, shared with all small molecule class-A ligands). Whereas binding at Site-II triggers receptor activation, the role of Site-I is not known. Structures and sequence analysis reveal that the receptor N-terminal domains (N-domains) are flexible and contain intrinsic disorder. Using a hybrid NMR-MD approach, we characterized the role of Site-I interactions for the CXCL8-CXCR1 pair. NMR data indicate that the CXCR1 N-domain becomes structured on binding and that the binding interface is extensive with 30% of CXCL8 residues participating in this initial interaction. MD simulations indicate that CXCL8 bound at Site-I undergoes extensive reorganization on engaging Site-II with several residues initially engaged at Site-I also engaging Site-II. We conclude that structural plasticity of Site-I interactions plays an active role in driving ligand recognition by a chemokine receptor.

Key words: chemokine, CXCL8, GPCR, NMR, solution structure, Molecular Dynamics, long-range coupling, CXCR1

Chemokines, the largest subfamily of cytokines, play diverse and fundamental roles, from trafficking immune cells and organogenesis to combating infection and injury. Not surprisingly, a dysregulation in chemokine function has been implicated in the pathophysiology of various autoimmune and inflammatory diseases, and in cancer (Griffith et al., 2014; Nagarsheth et al., 2017). Chemokines mediate their function by activating receptors that belong to the G protein-coupled receptor (GPCR) class. Humans express around 50 chemokines and 20 receptors, which show a rich tapestry of interactions resulting in large differences in affinity, specificity, and activity. Some receptors are activated by multiple chemokines, and at the other end of the spectrum, some are strictly activated by a single chemokine. Chemokines, in turn, show a range of receptor specificity and activity. Diversity of interactions is further compounded by chemokines and chemokine receptors existing as monomers and dimers, biased G protein vs. β -arrestin signaling, interacting with glycosaminoglycans, post-translational modification, and binding to atypical chemokine receptors (Bhusal et al., 2020; Rajagopal et al., 2013; Rajarathnam and Desai, 2020).

At a structural level, within this bewildering complexity is also an elegant simplicity. Chemokines, despite large differences in sequence identity, share the same structural fold. The chemokine structure consists of a short stretch of N-terminal residues preceding the conserved cysteines, an N-terminal loop (N-loop) followed by three β -strands and an α -helix, and disulfides that link the N-terminal and N-loop residues to the structural core. As class A GPCRs, chemokine receptor structure consists of seven transmembrane (TM) helices, three extracellular (EC) and three intracellular (IC) loops, and a short extracellular N-terminal and an intracellular C-terminal domain. Therefore, the spectrum of functional phenotypes must be encoded at the amino acid sequence level on a conserved ligand and receptor structural scaffolds.

Chemokines are unusual agonists for class A GPCRs, as conventional agonists tend to be small molecules with a rigid scaffold such as neurotransmitters, hormones, and those related to smell and taste. Chemokines are much bigger (MW ~10 kDa) but functionally no different as they bind their receptors with nanomolar affinity and activate the same signaling pathways. Therefore, chemokines must have evolved to exploit the GPCR structural features. Chemokines bind the receptor at two distinct sites: the N-terminal domain (Site-I, unique to chemokines) and a groove defined by the receptor EC loops/TM helices (Site-II, shared with all class A ligands) (Crump et al., 1997; Rajagopalan and Rajarathnam, 2006; Thiele and Rosenkilde, 2014). Whereas chemokine binding at Site-II triggers receptor activation, the role of Site-I interactions is not known.

In this study, we provide a structural model of how Site-I interactions orchestrate CXCL8 monomer binding to the CXCR1 receptor. Chemokines are classified into CXC, CC, CX₃C, and C subfamilies on the basis of conserved cysteines in the N-terminus. CXCL8, as a member of a subset of seven CXC chemokines characterized by the N-terminal 'ELR' motif, activates CXCR1 and CXCR2 receptors. Whereas all ELR-chemokines are potent CXCR2 agonists, CXCL8 monomer alone functions as a potent CXCR1 agonist (Nasser et al., 2009). CXCR1 and CXCR2 are expressed in neutrophils and in non-immune cells such as hepatocytes and neuronal

cells, and these receptors play fundamental roles in combating infection, regulating pain, and in the trafficking and proliferation of cancer cells (Ha et al., 2017; Rajarathnam et al., 2019). CXCR1 and CXCR2 sequences show large differences for the N-domain indicating a prominent role for Site-I interactions in determining specificity and affinity.

Knowledge of the structures are essential to describe the molecular mechanisms by which Site-I interactions determine binding at Site-II. The Site-I N-domain residues are not observed in the CXCR1 structure (Park et al., 2012). Sequences predict intrinsic disorder is a shared feature of N-domain of not only chemokine receptors but all class A GPCRs (Latorraca et al., 2017; Raucci et al., 2014; Venkatakrishnan et al., 2014). Characteristic features of intrinsic disorder include a preponderance of polar and charged residues, lack of a hydrophobic core, engaging in transient but specific interactions, and becoming structured on binding a protein partner (Latysheva et al., 2015; Uversky, 2019; van der Lee et al., 2014; Wright and Dyson, 2015). The N-domain residues are also not observed in the crystal structures of inhibitor-bound chemokine receptor complexes (Liu et al., 2020; Wu et al., 2010). The structure of CXCL8-CXCR1 complex is not known, but cryoelectron microscopy (cryoEM) and crystal structures of CXCL8 bound to CXCR2 and also of chemokines bound to CXCR4, CCR5, and CCR6 receptors are known (Liu et al., 2020; Qin et al., 2015; Wasilko et al., 2020; Zheng et al., 2017). A significant segment of the receptor N-domain is also missing in these structures, suggesting Site-I interactions are either weak, lost once the chemokine engages Site-II residues, or that the receptor Site-I residues continue to be dynamic in the chemokine-bound form.

We propose a hybrid strategy that combines nuclear magnetic resonance (NMR) spectroscopy for determining the solution structure of the CXCL8 monomer-CXCR1 N-domain complex and extended molecular dynamics (MD) simulations to describe how CXCL8 bound at Site-I engages Site-II residues. Our model for Site-I interactions is as follows. The first step involves CXCR1 N-domain capturing CXCL8 by the fly-casting mechanism (Shoemaker et al., 2000), as the unstructured N-domain has a large capture radius. CXCL8 bound at Site-I is a distinct but transient intermediate, and in the next step, Site-I functions as an anchor for CXCL8 N-terminal residues to search for Site-II residues. CXCL8 binding the CXCR1 N-domain captures the structural features of the Site-I intermediate. The NMR structure of the Site-I complex reveals an extensive binding interface stabilized by packing and ionic interactions. Interestingly, MD simulations reveal that Site-II interactions are coupled to extensive reorganization of Site-I interactions, including disengagement, rearrangement, and several residues initially engaged in Site-I interactions alone engaging in both Site-I and Site-II interactions. Our data make a compelling case for long-range coupled motions in driving ligand recognition by the CXCR1 receptor.

RESULTS

Structure of the CXCL8-CXCR1 N-domain complex. The solution structure of the CXCL8 monomer bound to the CXCR1 N-domain Site-I complex was determined using 1310 distance and 106 angular restraints. Ensemble structures, structural details, and statistics are shown in Fig. 1, Supplementary Fig. 1, and Supplementary Tables 1 and 2. For the CXCL8 monomer, we used

the CXCL8 variant (residues 1-66), which is missing the last six C-terminal residues known to stabilize the dimer interface (Berkamp et al., 2017; Joseph and Rajarathnam, 2015; Joseph et al., 2018). For CXCR1 N-domain, we used a construct containing residues 1 to 29 (Joseph and Rajarathnam, 2015; Joseph et al., 2018). To distinguish CXCR1 and CXCL8 residues, receptor residues are annotated as a single letter amino acid and ligand residues are annotated as a three-letter amino acid. The interface is extensive with as many as 30% of the CXCL8 residues (20 out of 66) participating in the Site-I interaction. The structure reveals that CXCL8 is more structured in the bound form. The backbone r.m.s.d. (root mean square deviation) of CXCL8 in the bound and free forms, for residues Lys11 to Leu66, is 0.34 and 0.65Å, respectively. The side chains of the interface residues highlight the overall higher conformational flexibility of CXCL8 in the free compared to the bound form (Figs. 1C, 1D). The CXCR1 N-domain peptide is unstructured in the free form. In the bound form, residues D11 to P29 are structured (r.m.s.d. 0.63Å), which is supported by 16 long-range intramolecular NOEs and 90 intermolecular NOEs to the CXCL8 N-loop, 3rd β -strand (β_3), and C-terminal helical residues (Figs. 1E, 1F). A schematic of CXCL8 and CXCR1 residues that show intermolecular NOEs is shown in Fig. 1G. The first ten distal residues of the CXCR1 N-domain are unstructured, which is evident from only a few sequential and no intermolecular NOEs.

The structure reveals that the CXCR1 N-domain binds in an extended fashion with no defined secondary structure except for a type IV β -turn for residues F17 to M20. Extensive packing and 20 backbone-side chain and side chain-side chain H-bonds and salt bridges stabilize the complex. CXCR1 F12, F17, and P29 residues bind a hydrophobic groove defined by CXCL8 Tyr13, Phe17, Phe21, Trp57 and Val61 residues (Fig. 2A). This hydrophobic pocket is flanked by two ionic clusters (Figs. 2B, 2C). The first ionic cluster consists of a network of interactions between D11/Gln59, D13/Arg60, D14/Lys64, and D14/His18, and the second ionic cluster between D24/Arg47, E25/Lys11, and D26/Arg47. Interestingly, Lys15, which is unique to CXCL8 among ELR-chemokines, is not involved in ionic interactions, but is observed in the proximity of CXCR1 non-polar M20, P21, and P22 residues (Fig. 2D). NMR backbone ¹⁵N-relaxation measurements of the CXCL8-CXCR1 N-domain complex indicate that the CXCL8 N-loop residues undergo significant conformational fluctuations, CXCL8 N-terminal ELR residues are unstructured both in the free and bound forms, distal CXCR1 N-terminal residues 1 to 10 are unstructured, and CXCR1 residues 11 to 27 are structured but show significant dynamics (Joseph et al., 2018). These structural and dynamics characteristics collectively indicate that conformational flexibility, especially of the interface residues, promote Site-I complex formation.

MD simulations of the CXCL8 monomer-CXCR1 complex. We set out to characterize how Site-I interactions determine binding at Site-II using MD simulation started from CXCL8 bound at Site-I. To this end, we first generated a structural model of CXCL8 monomer bound at Site-I of the CXCR1 receptor by merging the NMR structure of CXCL8-CXCR1 N-domain complex and a previously determined CXCR1 structure (Park et al., 2012). Using inter and intra molecular NOEs as distance constraints, we generated a structural model using HADDOCK

(Supplementary Fig. 2). We then carried out MD simulation of the CXCL8-CXCR1 complex for 800ns. During the initial phase of the simulation, we used H-bonds derived from the NMR structure as distance restraints. The restraints were removed once stabilizing interactions with Site-II were formed.

The time evolution of the backbone r.m.s.d. demonstrates that the simulation has equilibrated after about 100ns (Supplementary Fig. 3). The largest structural reorganization occurred in the initial 50ns of the simulations and primarily involves the CXCR1 loops, while the structured regions of CXCL8 and the helices of CXCR1 are stable during the entire simulation. This can also be seen from the plots of secondary structure contents (Supplementary Fig. 4) and backbone root mean square fluctuations (r.m.s.f.; Supplementary Fig. 5). In contrast to the minimal fluctuation of the secondary structure elements, the termini of CXCL8 and the termini as well as the intra- and extracellular loops of CXCR1 experience large conformational fluctuations. As many of these flexible regions are involved in intermolecular interactions, the MD data provided several novel insights into how CXCL8 bound at Site-I searches and binds Site-II residues. In addition to ELR residues, several CXCL8 N-loop residues also bound at Site-II, indicating binding at two spatially distinct receptor sites are not discrete events but involves extensive crosstalk between the sites (Figs. 3-5; Supplementary Fig. 6). In particular, reorganization of Site-I interactions allow CXCL8 to swivel about the second ionic cluster, which increases the search radius for binding Site-II residues. We group CXCL8 residues on the basis of their interactions into four classes.

Group-I consists of residues Glu4, Leu5, Arg6, Ile10, His33, and Glu38. These residues (except Ile10), which are not involved in Site-I interactions, are now exclusively engaged in binding at Site-II. In the Site-I structure, ELR residues are unstructured and His33 is observed in the proximity of the CXCL8 N-terminal residues, and its imidazole side chain is in an environment that is different from the bulk solvent as inferred from pKa measurements (Sepuru and Rajarathnam, 2018). During the first 50 ns of the MD simulation, Glu4 and Arg6 are engaged in specific ionic interactions with ECL2 K197 and D194, respectively (Figs. 3, 4). These ion pairs are stable and retained throughout the MD run. Both D194 and K197 are unique to CXCR1 with the corresponding residues in CXCR2 being asparagines, indicating that Site-II binding interactions are different between the receptors. During the initial 50 ns, Leu5 is observed in the proximity of ECL2 N193 and T195, and subsequently binds to a pocket defined by ECL2 V190 and L191, and ECL3 R269 (Figs. 3, 4). Interestingly, His33 starts interacting with ECL3 R269 about the same time as Leu5, suggesting concerted interactions that are stable until the end of the MD run. The first three N-terminal residues (Ser1-Ala2-Lys3), which are not critical for activity (Clark-Lewis et al., 1993), are pointed away from the binding groove. In both the free and Site-I structures, the carboxylate group of Glu38 is H-bonded to Gln8 backbone amide proton (Fig. 3, Supplementary Fig. 7). During the MD simulations, Glu38 disengages in the first 10ns, and is observed to interact with R279 and R280 in the final MD structures. Ile10, which is packed against CXCR1 Y28 in the Site-I structure, disengages in the first 50 ns, and is initially observed

in a pocket defined by ECL2 I279, ECL3 R279 and R280, and by the end of the run, is essentially packed against ECL3 R279 (Fig. 3).

Group-II consists of residues Lys11, Thr12, and Ser14. These residues, involved in ionic interactions in the NMR Site-I structure, undergo a reorganization and also start interacting with Site-II residues during the MD run. In the Site-I structure, Lys11 interacts with E25 and Y27, Thr12 interacts with E25 and D26, and Ser14 interacts with E25. Towards the end of the MD run, Lys11 interacts with E25 and D26 at Site-I and with S276 and R279 at Site-II; Thr12 retains its interaction with E25 at Site-I and now also interacts with S276 at Site-II (Supplementary Fig. 8); and Ser14 interacts with S28 and P29 at Site-I and with C277 at Site-II (Figs. 3, 5). Humans carrying a S276T mutation show impaired neutrophil function and are more susceptible to infection (Swamydas et al., 2016), highlighting the importance of these interactions for receptor function. These observations are also unprecedented, and make a compelling case for structural plasticity and dynamics of the binding residues in coupling Site-I and Site-II interactions.

Group-III consists of residues Tyr13, Lys15, Phe17, Phe21, and Arg47. These residues are involved in ionic/packing interactions in the NMR Site-I structure, and undergo a reorganization during the MD run and are now engaged with different Site-I residues. In the Site-I structure, Tyr13 is packed against A23, D24, and E25; Lys15 is packed against M20, P21, and P22; and Arg47 is involved in ionic interactions with D24 and D26 (Fig. 2). During the MD run, Lys15 is found to be dynamic and its side chain spanning most of the binding interface from F12 to P29, and towards the end, is involved in ionic interactions with D13 and D14 (Figs. 3, 5). Tyr13 is also dynamic and interacts with P21, P22, A23, D24, and E25. During the MD run, Arg47 disengages from D24 and D26, and towards the end, is packed against M20, P21, P22, and A23. It is notable that interactions of Lys15 and Arg47 are switched between the NMR Site-I and final MD structures (Figs. 3, 5). Both Phe17 and Phe21 interact with N16 and P21 in the Site-I structures. Phe17 was observed to interact with several non-polar residues with longer residence times around M20 and P22, and Phe21 was observed to be packed against M20, P21, and P22 throughout the MD run (Fig. 3).

Group-IV consists of residues His18, Arg60, and Lys64. These residues are involved in ionic interactions with D11, D13, and D14 in the NMR Site-I complex, but disengage around 250ns and formed no new intermolecular interactions (Fig. 3). Therefore, Group-IV residues exclusively play a role in stabilizing Site-I interactions during the docking stage and play no role in Site-II interactions. Group-IV interactions result in CXCR1 N-terminal residues binding in an extended fashion as shown in Figs. 1 and 3c (at the beginning of the MD run). However, losing these interactions results in a more compact N-terminal domain (MD structures at 300 and 800 ns). MD structural snapshots reveal that Lys15 (Group-III) and its neighboring residues take the place of Group-IV residues for interacting with the acidic residues of the first ionic cluster (Fig. 3), CXCR1 A23 is packed against CXCL8 Tyr13, Arg47, and Leu 49, and that the stretch of CXCR1 residues from L15 to P22 undergo segmental motions with A23 and the first ionic cluster functioning as pivot points (Supplementary Fig. 9). These data also indicate that

interactions of Lys15 are transient, suggesting the corresponding receptor residues are unlikely to be structured if a crystal structure of the CXCL8-CXCR1 were to become available.

Discussion

Chemokine N-terminal ELR and N-loop residues, and 30s-loop, β_3 -strand, and C-helical residues that are in the proximity of ELR and N-loop residues identified in the current MD study mediate receptor function (Clark-Lewis et al., 1993, 1994; Kuschert et al., 1998; Lowman et al., 1996; Schraufstatter et al., 1995; Suetomi et al., 1999). We have previously shown that the conserved disulfides and the CXC motif that are in the proximity of ELR and N-loop residues are also critical for function (Joseph et al., 2010; Joseph et al., 2013b; Rajagopalan et al., 2007; Rajarathnam et al., 1999). N-terminal, N-loop, 30s-loop, disulfides, and the CXC motif are also conformationally dynamic, suggesting that cross-talk between these residues play an important role for function. We had proposed a ‘conformational-selection’ model, whereby the CXCL8 N-terminal, N-loop and 30s-loop, and the disulfides and as well as chemokine receptors exist as a conformational ensemble, such that their association occurs via a series of coordinated conditional and discrete binding steps (Joseph et al., 2013b). The structural plasticity and reorganization of Site-I interactions observed in the current MD run are in agreement with such a model.

Residue-specific dynamic measurements of the CXCL8-CXCR1 N-domain complex indicate that N-loop residues - Ile10, Thr12, Tyr13, Ser14, and Lys15 - undergo chemical exchange at slower millisecond-microsecond (ms- μ s) time scales (Joseph et al., 2018). Chemical exchange for a set of contiguous N-loop residues provides a structural basis for how a coupled network that is dynamically primed for conformational selection could mediate Site-II interactions. The observation that Ile10 belongs to Group-I, Thr12 and Ser14 belong to Group-II, and Tyr13 and Lys15 belong to Group-III is striking, providing crucial insights into how structural plasticity and conformational dynamics of these residues are integral to the search process and binding Site-II residues. Ile10Ala mutant is substantially impaired for receptor activity but shows only two-fold reduced affinity for the CXCR1 N-domain (Clark-Lewis et al., 1994; Ravindran, 2010), suggesting that the steric bulk of Ile10 is less important for the initial docking but more important for binding at Site-II. The residue corresponding to Ile10 tends to be non-polar or aromatic in both CXC and CC chemokines (Supplementary Fig. 10), suggesting that this residue in other chemokines plays a similar role in coupling Site-I and Site-II interactions.

Our MD data indicate that binding Site-II is not a simple event but is coupled to a complex reorganization of Site-I interactions (Figs. 3 to 5). The observation that these changes occur relatively fast (mostly within 250ns) also corroborates a pivotal role for conformational plasticity and that crosstalk between Site-I and Site-II is a highly dynamic process. A previous MD study using CXCL8 bound to CXCR1 supplemented with a homology modeled CXCR1 N-domain based on the bovine rhodopsin structure could not describe interactions of the ELR residues

(Liou et al., 2014). This highlights the importance of a Site-I structure that captures native interactions for characterizing how binding at Site-I determines Site-II interactions.

CryoEM structure of CXCL8 bound to CXCR2 in the presence of G protein (active form) was recently reported (Liu et al., 2020). The structure reveals that Glu4 is engaged in ionic interactions with three arginines located in ECL2 and ECL3, Leu5 is involved in packing interactions with two valines in the ECL2, and Arg6 forms a H-bond with the carbonyl oxygen of a threonine in ECL3. These interactions are different from those observed in our MD study of the CXCL8-CXCR1 complex. Glu4 and Arg6 are engaged in specific ionic interactions with a ECL2 lysine and aspartate (Figs. 3, 4). A conservative mutation of Arg6 to a lysine abolishes receptor activity (Clark-Lewis et al., 1993). Therefore, Arg6 interacting with a backbone carbonyl in the CXCL8-CXCR2 structure is unexpected considering the absolute requirement of the arginine guanidinium side chain for affinity and activity. The observed differences could be due to differences in CXCR1 vs. CXCR2 Site-I interaction or that CXCL8 binds to a different subset of Site-II receptor residues in the active vs. inactive forms. Note that our MD study was of CXCL8 binding to CXCR1 in the inactive form whereas the cryoEM structure was of CXCL8 binding to CXCR2 in the active form.

Comparing the CXCL8-CXCR1 N-domain structure to CXCL12-CXCR4, CCL5-CCR2, and CCL11-CCR1 N-domain Site-I structures show some similarities but also notable differences (Abayev et al., 2018; Millard et al., 2014; Veldkamp et al., 2008; Ziarek et al., 2017) (Fig. 6). Structures reveal that chemokines bind the most proximal receptor N-terminal residues up to the preceding residue of the conserved cysteine (Fig. 6E). This is notable considering receptor N-domains show low sequence homology that is evident even for related receptors CXCR1 and CXCR2 (Supplementary Fig. 11). Receptor N-domain sequence lengths vary, are acidic due to many Asp/Glu, and in addition, receptor N-domain tyrosines are also known to be sulfated (Ludeman and Stone, 2014). In the CXCL8-CXCR1 N-domain Site-I structure (Fig. 1), proximal N-domain residues (D11 to P29) mediate binding, whereas the distal N-domain residues are unstructured. However, in the CXCL12-CXCR4 N-domain structure, besides the CXCR4 proximal N-terminal residues (N11-K25), distal N-terminal residues (S5-D10) are also involved in binding (Ziarek et al., 2017). This set of distal N-terminal residues bind to CXCL12 β_1 residues, and has been labeled as Site 0.5 on the basis of a similar observation for CXCL12 binding to ACKR3 (Gustavsson et al., 2017). Site 0.5 interactions are not observed in our CXCL8-CXCR1 N-domain structure or in the other two CC chemokine Site-I structures, indicating these interactions come into play only for some chemokine-receptor pairs. The structures of CCL5-CCR5 and CCL11-CCR1 N-domain Site-I complexes show that fewer receptor residues mediate binding and also that these interactions involve different regions of the chemokine (Figs. 6C and 6D). Solution NMR studies have shown that the N-domain of the intact CXCR1 is in equilibrium with the membrane, and that only the free N-domain can bind the ligand (Park et al., 2011). Using time-resolved fluorescence, surface pressure, and red edge excitation shift measurements, it has also been shown that the CXCR1 N-domain interacts with the membrane surface (Haldar et al., 2010). These observations provide a glimpse of the

diversity in Site-I interactions, and also suggesting that coupling between Site-I and Site-II interactions will accordingly vary.

A solution structure of a CXCR1 N-domain peptidomimetic bound to dimeric CXCL8 is available (Skelton et al., 1999). The peptidomimetic corresponds to residues 9 to 29 containing a 6-amino hexanoic acid linker in place of residues 15-19. Structures reveal significant differences, as evident by NOEs from 20 CXCL8 and 17 CXCR1 residues for the monomer and from only 6 CXCL8 and 7 CXCR1 residues for the dimer (Supplementary Fig. 12). CXCR1 residues P21 to P29 bind in an extended fashion in the monomer-bound structure, whereas these residues adopt a more compact structure in the dimer-bound form (Supplementary Fig. 12). Backbone relaxation and amide exchange measurements show that the monomer is more dynamic (Grasberger et al., 1993; Joseph et al., 2013a; Joseph et al., 2018; Rajarathnam et al., 1995; Rajarathnam et al., 1994), indicating conformational fluctuations are critical for the higher affinity and extensive binding interactions observed for the monomer. Fewer interactions for the dimer compared to the monomer have also been observed in CXCL12 binding to the CXCR4 N-domain (Veldkamp et al., 2008; Ziarek et al., 2017).

Our MD data for the CXCL8-CXCR1 complex indicate that several Site-I interactions are transient, residues that mediate Site-I interactions continue to be dynamic even after Site-II is engaged, and structural plasticity allows the same residue binding at both sites. CryoEM and crystal structures of CXCR2 bound to CXCL8, CCR6 bound to CCL20, CCR5 bound to CCL5 antagonist (5P7-CCL5), CXCR4 bound to vMIPII (a promiscuous viral chemokine antagonist), and a constitutively active viral chemokine receptor US28 bound to CX₃CL1 are known (Burg et al., 2015; Liu et al., 2020; Qin et al., 2015; Wasilko et al., 2020; Zheng et al., 2017) (Fig. 7). The coordinates of the CXCL8-CXCR2 complex are yet to be released and the Site-I structure of the CXCL8-CXCR2 N-domain is not known. Comparison of NMR CCL5-CCR5 N-domain complex and 5P7CCL5-CCR5 crystal structures reveals that several N-domain residues that are structured in the former are unstructured in the latter (Fig. 7F). Structures of CCL20-CCR6, vMIPII-CXCR4, and CX₃CL1-US28 complexes reveal electron density for only a few N-domain residues. Site-I structure of vMIPII-CXCR4 N-domain is not known but the structure of the CXCL12-CXCR4 N-domain shows extensive Site-I interactions. These observations collectively indicate that conformational dynamics of receptor N-domain residues preclude these residues from being ‘visible’ in the crystal and cryoEM structures, and also suggest that loss and reorganization of Site-I interactions on binding Site-II could be a basic mechanism underlying ligand recognition by all chemokine receptors.

Our MD data show that the first three CXCL8 N-terminal residues are not located in the groove, and residues Glu4-Leu5-Arg6 interact with EC loop residues. ELR-CXC chemokines have as few as one (CXCL7) and as many as nine (CXCL5) residues preceding the ‘ELR’ motif (Supplementary Fig. 11), suggesting these residues finetune Site-II interactions. The cryoEM structures of CXCL8-CXCR2 and CCL20-CCR6 also reveal that the CXCL8 and CCL20 N-terminal are located closer to the surface (Liu et al., 2020; Wasilko et al., 2020). However, crystal structures of 5P7CCL5-CCR5 and vMIPII-CXCR4 complexes reveal that the chemokine

N-terminal residues are located deep in the TM groove and are in the proximity of toggle switch residues that trigger the conformational change for receptor activation (Wescott et al., 2016). These observations suggest that there is no absolute requirement that chemokine residues must be located in the proximity of toggle-switch residues for receptor activation.

The past decade has seen an explosion of GPCR including chemokine receptor structures that can be attributed to advances in receptor expression and folding, reconstitution with detergents, and constructs that are stable and more amenable to crystallization (Kruse et al., 2013; Rosenbaum et al., 2007). The latter technology includes replacing a flexible IC loop with T4 lysozyme, using nanobodies to stabilize a specific conformer, and/or solving the structure bound to agonists, antagonists, and/or small molecule inhibitors that were critical for determining chemokine receptor structures and generating structural models (Bhusal et al., 2020; Wu et al., 2010; Zheng et al., 2016). These structures represent snapshots and are not sufficient for providing a spatiotemporal description of how the chemokine bound at Site-I searches and binds the residues at Site-II. In this study, we show that a hybrid NMR-MD approach can provide this information and that long-range coupled motions between Site-I and Site-II orchestrate CXCL8 monomer binding to the CXCR1 receptor. On the basis of our current studies and available literature, we propose tunable coupled motions between Site-I and Site-II underlie the range of affinity, potency, and specificity observed for various chemokine-receptor pairs. We conclude that NMR-MD approach in complement with ever increasing number of structures, can provide molecular details that are crucial for understanding the basic principles that govern chemokine-receptor interactions and their roles in human pathophysiology, thus paving the way for designing highly specific and potent drugs.

Materials and Methods

Protein Expression and Purification. CXCL8(1–66) monomer and CXCR1 N-domain (1-29) peptide were cloned, expressed, and purified as described previously (Fernando et al., 2007).

Chemical shift assignments and structure calculation. Chemical shift assignments of CXCL8 and CXCR1 N-domain in the bound form were obtained using CBCA(CO)NH, HNCACB, CC(CO)NH, ¹⁵N-edited TOCSY-HSQC, and HCCH-TOCSY experiments using 0.4 mM ¹³C/¹⁵N-CXCL8 bound to unlabeled CXCR1 N-domain and unlabeled CXCL8 bound to ¹³C/¹⁵N-CXCR1 N-domain samples (Gardner and Kay, 1998). The spectra were collected at 303K on a Bruker 800 MHz spectrometer with a cryogenically cooled probe head. The samples were prepared in 50 mM phosphate pH 6.0 containing 10% D₂O and 0.01% sodium azide. The backbone dihedral angles were predicted using the TALOS software (Shen et al., 2009). The chemical shifts were referenced using DSS as the internal standard. The NMR data were processed using Bruker Topspin and analyzed by Sparky. Intramolecular distance restraints were obtained from ¹⁵N-edited NOESY-HSQC and ¹³C-edited NOESY-HSQC experiments optimized either for the aliphatic or the aromatic region. The distance constraints and angular restraints for CXCL8 alone were used to calculate structures of CXCL8 and CXCR1 N-domain in the complex using ARIA/CNS suite of programs (Rieping et al., 2007). Intermolecular NOE distance restraints were obtained from ¹³C (ω_2)-edited/¹²C (ω_3)-filtered NOESY-HSQC spectra of ¹⁵N/¹³C-CXCL8 monomer complexed to 3-fold excess of unlabeled CXCR1 N-domain and *vice versa*. Structure of the CXCL8-CXCR1 N-domain complex was calculated iteratively using distance, dihedral angle, and H-bonding restraints as input using ARIA/CNS and the PARALLHDG 5.3 force field. The quality of the calculated structures was assessed using PROCHECK.

Model of CXCL8 bound at CXCR1 N-domain. A model of CXCL8 bound at N-domain of the full length CXCR1 receptor was generated using HADDOCK as follows. Initially, the missing N-terminal residues 1 to 28 were appended to the CXCR1 structure (PDB ID: 2LNL; residues 29-324) using PyMOL. For Site-I interactions, all experimental intramolecular and intermolecular NOEs were used as unambiguous restraints. For Site-II interactions, constraints between CXCL8 N-terminal Glu4 and Arg6 and all CXCR1 ECL2 and ECL3 acidic and basic residues were used as ambiguous restraints. A total of 4000 structures were generated using rigid body docking in HADDOCK (Dominguez et al., 2003). The top 100 structures that had the best intermolecular energies were then subjected sequentially to semiflexible simulated annealing and explicit solvent refinement during which the CXCL8 N-terminal ‘ELR’ and CXCR1 ECL2 and ECL3 acidic and basic residues were allowed to move freely. The pair-wise “ligand interface root-mean-square deviation matrix” over all structures was calculated, and final structures were clustered using a cutoff value of 6.0 Å. The clusters were sorted using r.m.s.d. and “HADDOCK score” (the weighted sum of a combination of energy terms), and the best scored structure was used as the initial model for MD studies.

Molecular dynamics simulations. We embedded the modeled CXCL8 monomer-CXCR1 complex structure into a fully hydrated bilayer consisting of 120 16:0/18:1 phosphatidylcholine

(POPC) lipids per leaflet, using CHARMM GUI (Jo et al., 2008). The size of the water box was $x = 93.6\text{\AA}$, $y = 93.6\text{\AA}$, $z = 113.0\text{\AA}$ such that the minimum distance of a protein atom from the edge of the box was 10\AA . The protein was solvated with TIP3P waters. The system was charge-neutralized and the ionic strength was set to 150mM by adding KCl. The final size of the system was 101,509 atoms. The protein-bilayer-water system was energy minimized for 2000 steps with lipid and protein heavy atoms fixed and equilibrated first for 200ps and then for four 100ps equilibration steps with a gradually decreasing restraint force constant. The equilibrated system was simulated with a 2fs time-step using SHAKE (Ryckaert, 1977) to restrain all bonds involving hydrogens. The simulation was conducted using the CHARMM36m force field (Huang et al., 2017) and the latest version of the NAMD program (Phillips et al., 2005).

We analyzed the simulations by calculating r.m.s.d. for the backbone atoms of the complex and the TM helices of CXCR1 after aligning using the C_{α} atoms of the TM helices, and those of CXCR1 and CXCL8 were calculated following alignment on the respective initial structures. To evaluate the conformational fluctuations of the complex, we calculated time-averaged r.m.s.f. of individual CXCR1 and CXCL8 residues after backbone alignment to the average structure (Zhou et al., 2017). We identified the interface residues by calculating the contacts between the heavy atoms of CXCR1 and CXCL8, with two residues were defined to be in contact if any of their atoms are within 4\AA of each other. Using the contact information, we plotted a contact-map of the CXCR1-CXCL8 interface and calculated the time-averaged contact frequencies for individual residues.

Author contributions.

K. R., A. G., K.M.S., and V.N. designed research.

K.M.S. performed NMR experiments and structure determination under supervision of K.R.

V. N. and P.P. performed the MD simulations and analyses under supervision of A. G.

K. R. drafted the manuscript with contributions from A.G. and K.M.S.

All authors reviewed the final manuscript.

Acknowledgments. We thank Dr. Wang for NMR technical support. The authors acknowledge the Sealy Center for Structural Biology and Molecular Biophysics at the University of Texas Medical Branch at Galveston for providing research resources, and Texas Advanced Computing Center (TACC) and the Extreme Science and Engineering Discovery Environment (XSEDE grant No MCB150054) for providing computational resources. This work was supported in part by a grant from the National Institute of Health R01GM124233 to A.A.G. and a pilot grant from the Institute for Human Infections and Immunity at UTMB to K.R. V.N. is supported by UTHealth Innovation for Cancer Prevention Research Training Program Pre-Doctoral Fellowship (Cancer Prevention and Research Institute of Texas grant RP160015).

Data availability. The solution NMR ensemble and related data sets of the CXCL8-CXCR1 N-terminal domain complex have been deposited in the Protein Data Bank (PDB) and Biological Magnetic Resonance Bank (BMRB) database under the codes PDB 6XMN and BMRB 50331, respectively.

Competing Interests: The authors declare no competing interests.

Figure legends

Figure 1. Solution structure of the CXCL8-CXCR1 N-domain complex. (a) Twenty ensemble backbone C_{α} traces of CXCL8-CXCR1 N-domain structure. For the CXCR1 N-domain, only the structured residues (D11 to P29) are shown. CXCL8 and CXCR1 N-domain are shown in blue and red, respectively. (b) An ensemble of the CXCL8-CXCR1 N-domain structures is shown in two different orientations for highlighting the binding interface. An ensemble of the CXCL8 monomer structure in the free (c) and CXCR1-bound (d) forms. The sidechains of a subset of basic residues that mediate Site-I interactions are shown and labeled. The nitrogens of the side chain are shown in blue. The ensemble indicates that the binding interface residues are structured in the bound form and conformationally more dynamic in the free form. (e) An ensemble of CXCR1 N-domain structure in the bound form with the side chains shown as a stick. Side chain nitrogens and oxygens are shown in blue and red, respectively. The first ten residues are unstructured and not involved in CXCL8 binding. (f) ^{13}C -filter NOE strips representing intermolecular NOEs from CXCL8 N-loop, β_3 -strand, and α -helical residues to CXCR1 N-domain. (g) A schematic showing intermolecular NOE contacts between CXCR1 and CXCL8 N-loop, 40s turn/ β_3 and C-helix. Basic and acidic residues are in blue and red, respectively.

Figure 2. CXCL8-CXCR1 Site-I interactions. (a) Surface representation highlighting hydrophobic interactions in the CXCL8-CXCR1 N-domain complex. Key CXCL8 non-polar residues are labeled and shown in orange and CXCR1 is shown in green. (b) A schematic highlighting interactions of the first ionic cluster. CXCL8 residues are labeled in black and CXCR1 residues in magenta. Electrostatic representation of the first ionic cluster. CXCL8 residues are labeled in yellow and CXCR1 residues in magenta. (c) A schematic highlighting interactions of the second ionic cluster. CXCL8 residues are labeled in black and CXCR1 residues in magenta. Electrostatic representation of the second ionic cluster. CXCL8 residues are labeled in yellow and CXCR1 residues in magenta. (d) A schematic highlighting CXCL8 Lys15 side chain methylene groups packed against CXCR1 non-polar residues (shown as sticks and labeled in magenta).

Figure 3. Molecular dynamics of the CXCL8-CXCR1 complex. Heat maps of the average contact frequency between all CXCL8 and CXCR1 residues during the first 50ns (a) and last 100ns (b). Only residues that come in contact during the simulations are shown. (c) Snapshots of the CXCL8-CXCR1 complex at different time points (0, 300, and 800ns) during the MD simulation. CXCR1 is shown in orange and CXCL8 is shown as both surface and ribbon representations in green. The N-terminus of CXCL8 and CXCR1 are labeled in red and black, respectively.

Figure 4. Site-II interactions. Distance plots between Glu4: C^{δ} -K197: N^{ζ} (a), Arg6: C^{ζ} -D194: C^{γ} (b), Leu5: C^{δ} -R269: C^{δ} (c), and His33: N^{δ} -R269: C^{ζ} (d). The plots shows that CXCL8 N-terminal 'ELR' and His33 residues recognize and bind their specific CXCR1 partners within 150ns.

Structural snapshots of the CXCL8–CXCR1 complex at the beginning (**c**) and end (**d**) of MD snapshots, showing CXCL8 N-terminal ‘ELR’ and His33 residues (in red) recognizing their specific CXCR1 partners (in black).

Figure 5. MD profiles of Site-II interactions. Distance plots of (**a**) Lys11:N^ζ-D26:C^γ, (**b**) Thr12:C^β-S276:C^β, (**c**) Tyr13:C^δ-E25:C^α, (**d**) Tyr13:C^δ-S28:C^β, (**e**) Lys15:C^α-F12:C^β, (**f**) Lys15:N^ζ-D14:C^γ, (**g**) Arg47:C^δ-A23:C^β, and (**h**) Arg47:N^ζ-D24:C^γ.

Figure 6. Comparison of chemokine-receptor N-domain site-I structures. (**a**) CXCL8-CXCR1 (PDB ID: 6XMN), (**b**) CXCL12-CXCR4 (PDB ID: 2N55), (**c**) CCL5-CCR5 (PDB ID: 6FGP), and (**d**) CCL11-CCR3 (PDB ID: 2MPM). The structures were aligned pairwise against CXCL8-CXCR1 structure. The top and bottom panels show ribbon and surface representations, respectively. The chemokine receptor N-domains are shown as a ribbon (blue). (**e**) Sequences of receptor N-domain. Residues involved in binding are in blue and the conserved cysteine is in black.

Figure 7. Comparison of chemokine-receptor complex structures. (**a**) CXCL8-CXCR1; (**b**) CCL20-CCR6 (PDB ID: 6WWZ); (**c**) CCL5P7-CCR5 (PDB ID: 5UIW); (**d**) CXCR4-vMIP2 structure (PDB ID: 4RWS); and (**e**) US28-CX3C11 (PDB ID: 5WB2). The receptors and chemokines are shown in orange and purple, respectively. The receptor N-termini are denoted in red. (**f**) Sequences of receptor N-terminal domain. Residues involved in binding are in blue, and residues involved in binding in the NMR Site-I structures but not visible in the crystal structures are in red. The ‘DFDD’ residues of CXCR1 are underlined as MD studies indicate that these residues are involved in transient interactions and so unlikely to be observed in a crystal structure.

References

- Abayev, M., Rodrigues, J., Srivastava, G., Arshava, B., Jaremko, L., Jaremko, M., Naider, F., Levitt, M., and Anglister, J. (2018). The solution structure of monomeric CCL5 in complex with a doubly sulfated N-terminal segment of CCR5. *FEBS J* 285, 1988-2003.
- Berkamp, S., Park, S.H., De Angelis, A.A., Marassi, F.M., and Opella, S.J. (2017). Structure of monomeric Interleukin-8 and its interactions with the N-terminal Binding Site-I of CXCR1 by solution NMR spectroscopy. *J Biomol NMR* 69, 111-121.
- Bhusal, R.P., Foster, S.R., and Stone, M.J. (2020). Structural basis of chemokine and receptor interactions: Key regulators of leukocyte recruitment in inflammatory responses. *Protein Sci* 29, 420-432.
- Burg, J.S., Ingram, J.R., Venkatakrishnan, A.J., Jude, K.M., Dukkipati, A., Feinberg, E.N., Angelini, A., Waghray, D., Dror, R.O., Ploegh, H.L., *et al.* (2015). Structural biology. Structural basis for chemokine recognition and activation of a viral G protein-coupled receptor. *Science* 347, 1113-1117.
- Clark-Lewis, I., Dewald, B., Geiser, T., Moser, B., and Baggiolini, M. Platelet factor 4 binds to interleukin 8 receptors and activates neutrophils when its N terminus is modified with Glu-Leu-Arg. *Proc Natl Acad Sci USA* 90: 3574–3577 (1993).
- Clark-Lewis, I., Dewald, B., Loetscher, M., Moser, B., and Baggiolini, M. (1994). Structural requirements for interleukin-8 function identified by design of analogs and CXC chemokine hybrids. *J Biol Chem* 269, 16075-16081.
- Crump, M.P., Gong, J.H., Loetscher, P., Rajarathnam, K., Amara, A., Arenzana-Seisdedos, F., Virelizier, J.L., Baggiolini, M., Sykes, B.D., and Clark-Lewis, I. (1997). Solution structure and basis for functional activity of stromal cell-derived factor-1; dissociation of CXCR4 activation from binding and inhibition of HIV-1. *EMBO J* 16, 6996-7007.
- Dominguez, C., Boelens, R., and Bonvin, A.M. (2003). HADDOCK: a protein-protein docking approach based on biochemical or biophysical information. *J Am Chem Soc* 125, 1731-1737.
- Fernando, H., Nagle, G.T., and Rajarathnam, K. (2007). Thermodynamic characterization of interleukin-8 monomer binding to CXCR1 receptor N-terminal domain. *FEBS J* 274, 241-251.
- Gardner, K.H., and Kay, L.E. (1998). The use of ²H, ¹³C, ¹⁵N multidimensional NMR to study the structure and dynamics of proteins. *Annu Rev Biophys Biomol Struct* 27, 357-406.
- Grasberger, B.L., Gronenborn, A.M., and Clore, G.M. (1993). Analysis of the backbone dynamics of interleukin-8 by ¹⁵N relaxation measurements. *J Mol Biol* 230, 364-372.
- Griffith, J.W., Sokol, C.L., and Luster, A.D. (2014). Chemokines and chemokine receptors: positioning cells for host defense and immunity. *Annu Rev Immunol* 32, 659-702.
- Gustavsson, M., Wang, L., van Gils, N., Stephens, B.S., Zhang, P., Schall, T.J., Yang, S., Abagyan, R., Chance, M.R., Kufareva, I., *et al.* (2017). Structural basis of ligand interaction with atypical chemokine receptor 3. *Nat Commun* 8, 14135.
- Ha, H., Debnath, B., and Neamati, N. (2017). Role of the CXCL8-CXCR1/2 Axis in Cancer and Inflammatory Diseases. *Theranostics* 7, 1543-1588.
- Haldar, S., Raghuraman, H., Namani, T., Rajarathnam, K., and Chattopadhyay, A. (2010). Membrane interaction of the N-terminal domain of chemokine receptor CXCR1. *Biochim Biophys Acta* 1798, 1056-1061.

- Huang, J., Rauscher, S., Nawrocki, G., Ran, T., Feig, M., de Groot, B.L., Grubmuller, H., and MacKerell, A.D., Jr. (2017). CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods* 14, 71-73.
- Jo, S., Kim, T., Iyer, V.G., and Im, W. (2008). CHARMM-GUI: a web-based graphical user interface for CHARMM. *J Comput Chem* 29, 1859-1865.
- Joseph, P.R., Poluri, K.M., Gangavarapu, P., Rajagopalan, L., Raghuwanshi, S., Richardson, R.M., Garofalo, R.P., and Rajarathnam, K. (2013a). Proline Substitution of Dimer Interface β -strand Residues as a Strategy for the Design of Functional Monomeric Proteins. *Biophys J* 105, 1491-1501.
- Joseph, P.R., and Rajarathnam, K. (2015). Solution NMR characterization of WT CXCL8 monomer and dimer binding to CXCR1 N-terminal domain. *Protein Sci* 24, 81-92.
- Joseph, P.R., Sarmiento, J.M., Mishra, A.K., Das, S.T., Garofalo, R.P., Navarro, J., and Rajarathnam, K. (2010). Probing the role of CXC motif in chemokine CXCL8 for high affinity binding and activation of CXCR1 and CXCR2 receptors. *J Biol Chem* 285, 29262-29269.
- Joseph, P.R., Sawant, K.V., Isley, A., Pedroza, M., Garofalo, R.P., Richardson, R.M., and Rajarathnam, K. (2013b). Dynamic conformational switching in the chemokine ligand is essential for G-protein-coupled receptor activation. *Biochem J* 456, 241-251.
- Joseph, P.R.B., Spyrapoulos, L., and Rajarathnam, K. (2018). Dynamics-Derived Insights into Complex Formation between the CXCL8 Monomer and CXCR1 N-Terminal Domain: An NMR Study. *Molecules* 23.
- Kruse, A.C., Ring, A.M., Manglik, A., Hu, J., Hu, K., Eitel, K., Hubner, H., Pardon, E., Valant, C., Sexton, P.M., *et al.* (2013). Activation and allosteric modulation of a muscarinic acetylcholine receptor. *Nature* 504, 101-106.
- Kuschert, G.S., Hoogewerf, A.J., Proudfoot, A.E., Chung, C.W., Cooke, R.M., Hubbard, R.E., Wells, T.N., and Sanderson, P.N. (1998). Identification of a glycosaminoglycan binding surface on human interleukin-8. *Biochemistry* 37, 11193-11201.
- Latorraca, N.R., Venkatakrishnan, A.J., and Dror, R.O. (2017). GPCR Dynamics: Structures in Motion. *Chem Rev* 117, 139-155.
- Latysheva, N.S., Flock, T., Weatheritt, R.J., Chavali, S., and Babu, M.M. (2015). How do disordered regions achieve comparable functions to structured domains? *Protein Sci* 24, 909-922.
- Liou, J.W., Chang, F.T., Chung, Y., Chen, W.Y., Fischer, W.B., and Hsu, H.J. (2014). In silico analysis reveals sequential interactions and protein conformational changes during the binding of chemokine CXCL-8 to its receptor CXCR1. *PLoS One* 9, e94178.
- Liu, K., Wu, L., Yuan, S., Wu, M., Xu, Y., Sun, Q., Li, S., Zhao, S., Hua, T., and Liu, Z.J. (2020). Structural basis of CXC chemokine receptor 2 activation and signalling. *Nature* (in press).
- Lowman, H.B., Slagle, P.H., DeForge, L.E., Wirth, C.M., Gillece-Castro, B.L., Bourell, J.H., and Fairbrother, W.J. (1996). Exchanging interleukin-8 and melanoma growth-stimulating activity receptor binding specificities. *J Biol Chem* 271, 14344-14352.
- Ludeman, J.P., and Stone, M.J. (2014). The structural role of receptor tyrosine sulfation in chemokine recognition. *Br J Pharmacol* 171, 1167-1179.
- Millard, C.J., Ludeman, J.P., Canals, M., Bridgford, J.L., Hinds, M.G., Clayton, D.J., Christopoulos, A., Payne, R.J., and Stone, M.J. (2014). Structural basis of receptor sulfotyrosine

recognition by a CC chemokine: the N-terminal region of CCR3 bound to CCL11/eotaxin-1. *Structure* 22, 1571-1581.

Nagarsheth, N., Wicha, M.S., and Zou, W. (2017). Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nat Rev Immunol* 17, 559-572.

Nasser, M.W., Raghuvanshi, S.K., Grant, D.J., Jala, V.R., Rajarathnam, K., and Richardson, R.M. (2009). Differential activation and regulation of CXCR1 and CXCR2 by CXCL8 monomer and dimer. *J Immunol* 183, 3425-3432.

Park, S.H., Casagrande, F., Das, B.B., Albrecht, L., Chu, M., and Opella, S.J. (2011). Local and global dynamics of the G protein-coupled receptor CXCR1. *Biochemistry* 50, 2371-2380.

Park, S.H., Das, B.B., Casagrande, F., Tian, Y., Nothnagel, H.J., Chu, M., Kiefer, H., Maier, K., De Angelis, A.A., Marassi, F.M., *et al.* (2012). Structure of the chemokine receptor CXCR1 in phospholipid bilayers. *Nature* 491, 779-783.

Phillips, J.C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., Chipot, C., Skeel, R.D., Kale, L., and Schulten, K. (2005). Scalable molecular dynamics with NAMD. *J Comput Chem* 26, 1781-1802.

Qin, L., Kufareva, I., Holden, L.G., Wang, C., Zheng, Y., Zhao, C., Fenalti, G., Wu, H., Han, G.W., Cherezov, V., *et al.* (2015). Structural biology. Crystal structure of the chemokine receptor CXCR4 in complex with a viral chemokine. *Science* 347, 1117-1122.

Rajagopal, S., Bassoni, D.L., Campbell, J.J., Gerard, N.P., Gerard, C., and Wehrman, T.S. (2013). Biased agonism as a mechanism for differential signaling by chemokine receptors. *J Biol Chem* 288, 35039-35048.

Rajagopalan, L., Chin, C.C., and Rajarathnam, K. (2007). Role of intramolecular disulfides in stability and structure of a noncovalent homodimer. *Biophys J* 93, 2129-2134.

Rajagopalan, L., and Rajarathnam, K. (2006). Structural basis of chemokine receptor function--a model for binding affinity and ligand selectivity. *Biosci Rep* 26, 325-339.

Rajarathnam, K., Clark-Lewis, I., and Sykes, B.D. (1995). 1H NMR solution structure of an active monomeric interleukin-8. *Biochemistry* 34, 12983-12990.

Rajarathnam, K., and Desai, U.R. (2020). Structural Insights Into How Proteoglycans Determine Chemokine-CXCR1/CXCR2 Interactions: Progress and Challenges. *Front Immunol* 11, 660.

Rajarathnam, K., Schnoor, M., Richardson, R.M., and Rajagopal, S. (2019). How do chemokines navigate neutrophils to the target site: Dissecting the structural mechanisms and signaling pathways. *Cell Signal* 54, 69-80.

Rajarathnam, K., Sykes, B.D., Dewald, B., Baggiolini, M., and Clark-Lewis, I. (1999). Disulfide bridges in interleukin-8 probed using non-natural disulfide analogues: dissociation of roles in structure from function. *Biochemistry* 38, 7653-7658.

Rajarathnam, K., Sykes, B.D., Kay, C.M., Dewald, B., Geiser, T., Baggiolini, M., and Clark-Lewis, I. (1994). Neutrophil activation by monomeric interleukin-8. *Science* 264, 90-92.

Rauci, R., Costantini, S., Castello, G., and Colonna, G. (2014). An overview of the sequence features of N- and C-terminal segments of the human chemokine receptors. *Cytokine* 70, 141-150.

Ravindran, A. (2010). The Structural basis of chemokine binding and affinity and receptor selectivity: Role of receptor N-terminal domain. Ph D thesis

- Rieping, W., Habeck, M., Bardiaux, B., Bernard, A., Malliavin, T.E., and Nilges, M. (2007). ARIA2: automated NOE assignment and data integration in NMR structure calculation. *Bioinformatics* 23, 381-382.
- Rosenbaum, D.M., Cherezov, V., Hanson, M.A., Rasmussen, S.G., Thian, F.S., Kobilka, T.S., Choi, H.J., Yao, X.J., Weis, W.I., Stevens, R.C., *et al.* (2007). GPCR engineering yields high-resolution structural insights into beta2-adrenergic receptor function. *Science* 318, 1266-1273.
- Ryckaert, J.P., Ciccotti, G., and Berendsen, H. J. (1977). Numerical integration of the Cartesian Equations of Motion of a System with Constraints: Molecular Dynamics of n-Alkanes. *J Comput Phys* 23, 327-341.
- Schraufstatter, I.U., Ma, M., Oades, Z.G., Barritt, D.S., and Cochrane, C.G. (1995). The role of Tyr13 and Lys15 of interleukin-8 in the high affinity interaction with the interleukin-8 receptor type A. *J Biol Chem* 270, 10428-10431.
- Sepuru, K.M., and Rajarathnam, K. (2018). Distinct Differences in Structural States of Conserved Histidines in Two Related Proteins: NMR Studies of the Chemokines CXCL1 and CXCL8 in the Free Form and Macromolecular Complexes. *Biochemistry* 57, 5969-5977.
- Shen, Y., Delaglio, F., Cornilescu, G., and Bax, A. (2009). TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* 44, 213-223.
- Shoemaker, B.A., Portman, J.J., and Wolynes, P.G. (2000). Speeding molecular recognition by using the folding funnel: the fly-casting mechanism. *Proc Natl Acad Sci U S A* 97, 8868-8873.
- Skelton, N.J., Quan, C., Reilly, D., and Lowman, H. (1999). Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. *Structure* 7, 157-168.
- Suetomi, K., Lu, Z., Heck, T., Wood, T.G., Prusak, D.J., Dunn, K.J., and Navarro, J. (1999). Differential mechanisms of recognition and activation of interleukin-8 receptor subtypes. *J Biol Chem* 274, 11768-11772.
- Swamydas, M., Gao, J.L., Break, T.J., Johnson, M.D., Jaeger, M., Rodriguez, C.A., Lim, J.K., Green, N.M., Collar, A.L., Fischer, B.G., *et al.* (2016). CXCR1-mediated neutrophil degranulation and fungal killing promote *Candida* clearance and host survival. *Sci Transl Med* 8, 322ra310.
- Thiele, S., and Rosenkilde, M.M. (2014). Interaction of chemokines with their receptors--from initial chemokine binding to receptor activating steps. *Curr Med Chem* 21, 3594-3614.
- Uversky, V.N. (2019). Intrinsically Disordered Proteins and Their “Mysterious” (Meta)Physics. *Front Phys* 7, 1-10.
- van der Lee, R., Buljan, M., Lang, B., Weatheritt, R.J., Daughdrill, G.W., Dunker, A.K., Fuxreiter, M., Gough, J., Gsponer, J., Jones, D.T., *et al.* (2014). Classification of intrinsically disordered regions and proteins. *Chem Rev* 114, 6589-6631 .
- Veldkamp, C.T., Seibert, C., Peterson, F.C., De la Cruz, N.B., Haugner, J.C., 3rd, Basnet, H., Sakmar, T.P., and Volkman, B.F. (2008). Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci Signal* 1, ra4.
- Venkatakrishnan, A.J., Flock, T., Prado, D.E., Oates, M.E., Gough, J., and Madan Babu, M. (2014). Structured and disordered facets of the GPCR fold. *Curr Opin Struct Biol* 27, 129-137.

- Wasilko, D.J., Johnson, Z.L., Ammirati, M., Che, Y., Griffor, M.C., Han, S., and Wu, H. (2020). Structural basis for chemokine receptor CCR6 activation by the endogenous protein ligand CCL20. *Nat Commun* *11*, 3031.
- Wescott, M.P., Kufareva, I., Paes, C., Goodman, J.R., Thaker, Y., Puffer, B.A., Berdoug, E., Rucker, J.B., Handel, T.M., and Doranz, B.J. (2016). Signal transmission through the CXC chemokine receptor 4 (CXCR4) transmembrane helices. *Proc Natl Acad Sci U S A* *113*, 9928-9933.
- Wright, P.E., and Dyson, H.J. (2015). Intrinsically disordered proteins in cellular signalling and regulation. *Nat Rev Mol Cell Biol* *16*, 18-29.
- Wu, B., Chien, E.Y., Mol, C.D., Fenalti, G., Liu, W., Katritch, V., Abagyan, R., Brooun, A., Wells, P., Bi, F.C., *et al.* (2010). Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science* *330*, 1066-1071.
- Zheng, Y., Han, G.W., Abagyan, R., Wu, B., Stevens, R.C., Cherezov, V., Kufareva, I., and Handel, T.M. (2017). 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 e1005.
- Zheng, Y., Qin, L., Zacarias, N.V., de Vries, H., Han, G.W., Gustavsson, M., Dabros, M., Zhao, C., Cherney, R.J., Carter, P., *et al.* (2016). Structure of CC chemokine receptor 2 with orthosteric and allosteric antagonists. *Nature* *540*, 458-461.
- Zhou, Y., Prakash, P., Liang, H., Cho, K.J., Gorfe, A.A., and Hancock, J.F. (2017). Lipid-Sorting Specificity Encoded in K-Ras Membrane Anchor Regulates Signal Output. *Cell* *168*, 239-251 e216.
- Ziarek, J.J., Kleist, A.B., London, N., Raveh, B., Montpas, N., Bonnetterre, J., St-Onge, G., DiCosmo-Ponticello, C.J., Koplinski, C.A., Roy, I., *et al.* (2017). Structural basis for chemokine recognition by a G protein-coupled receptor and implications for receptor activation. *Sci Signal* *10*.

Figure 1

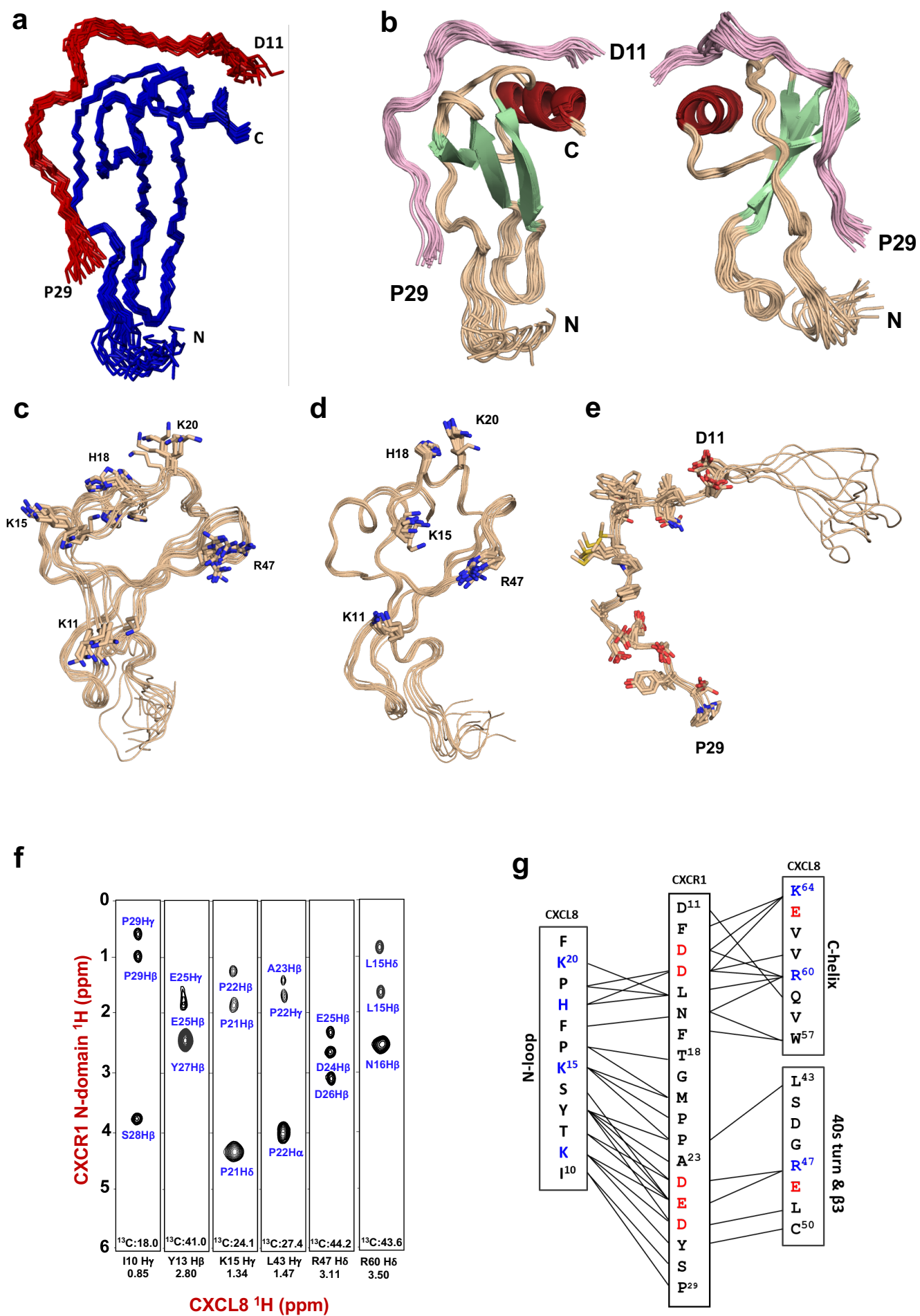


Figure 2

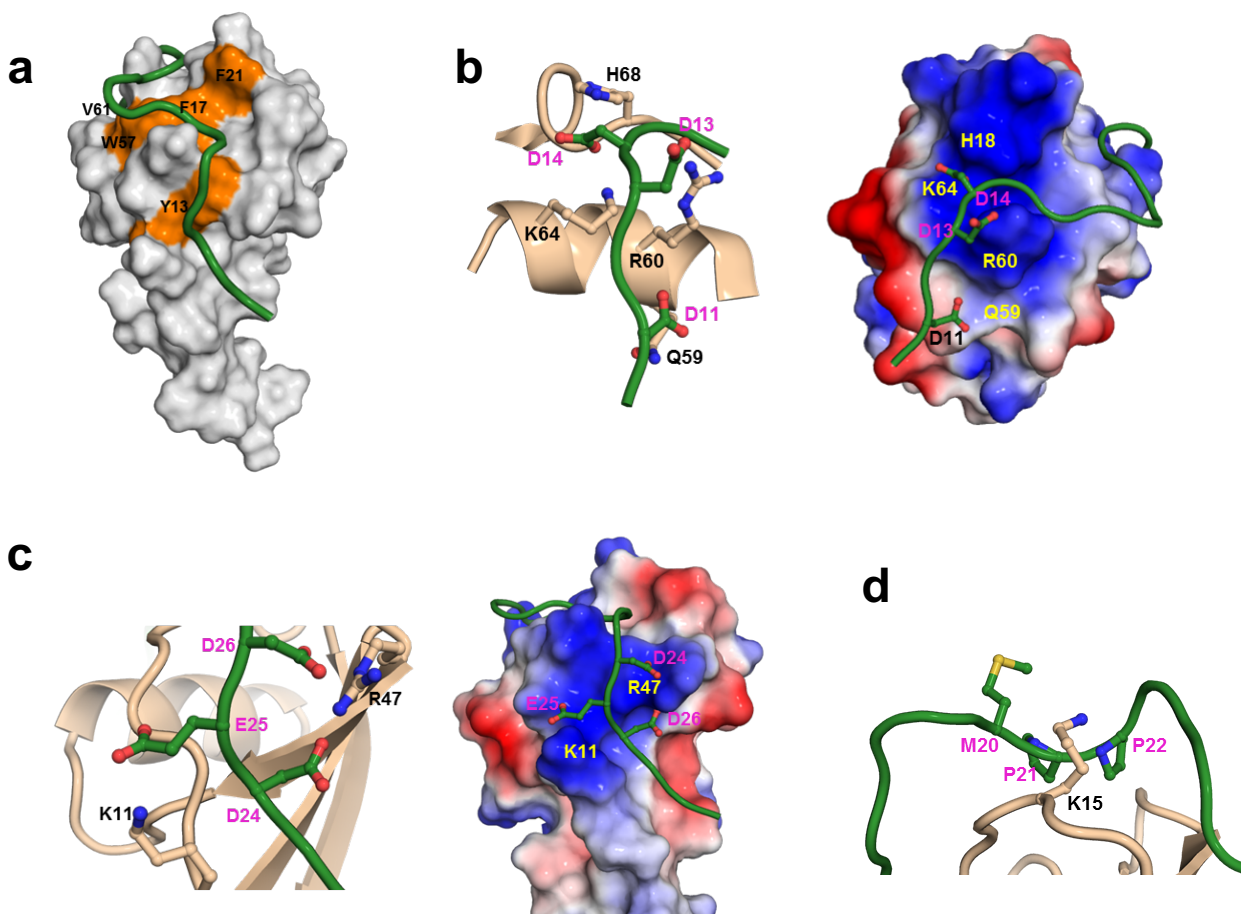


Figure 3

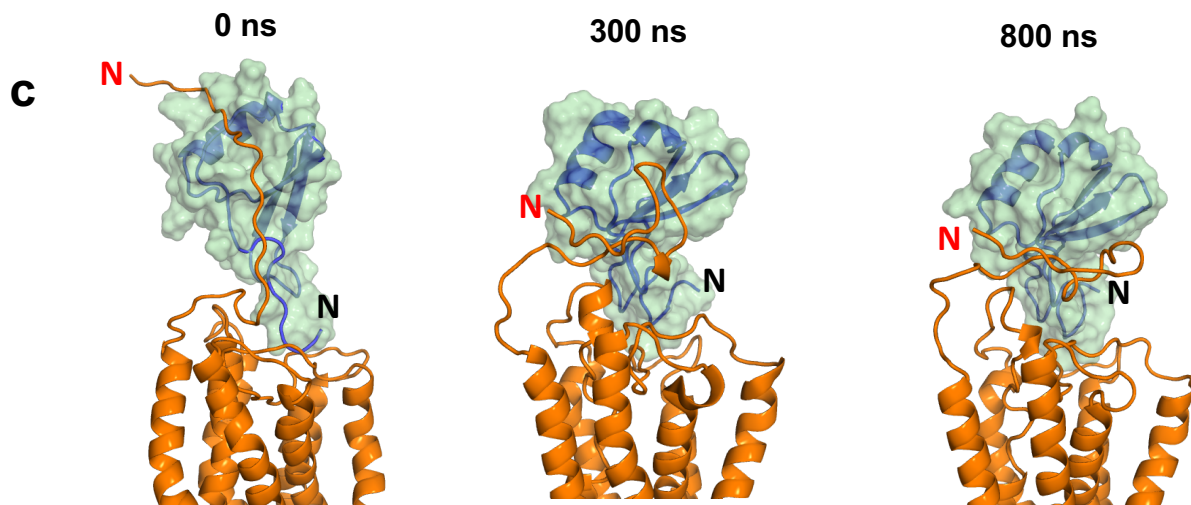
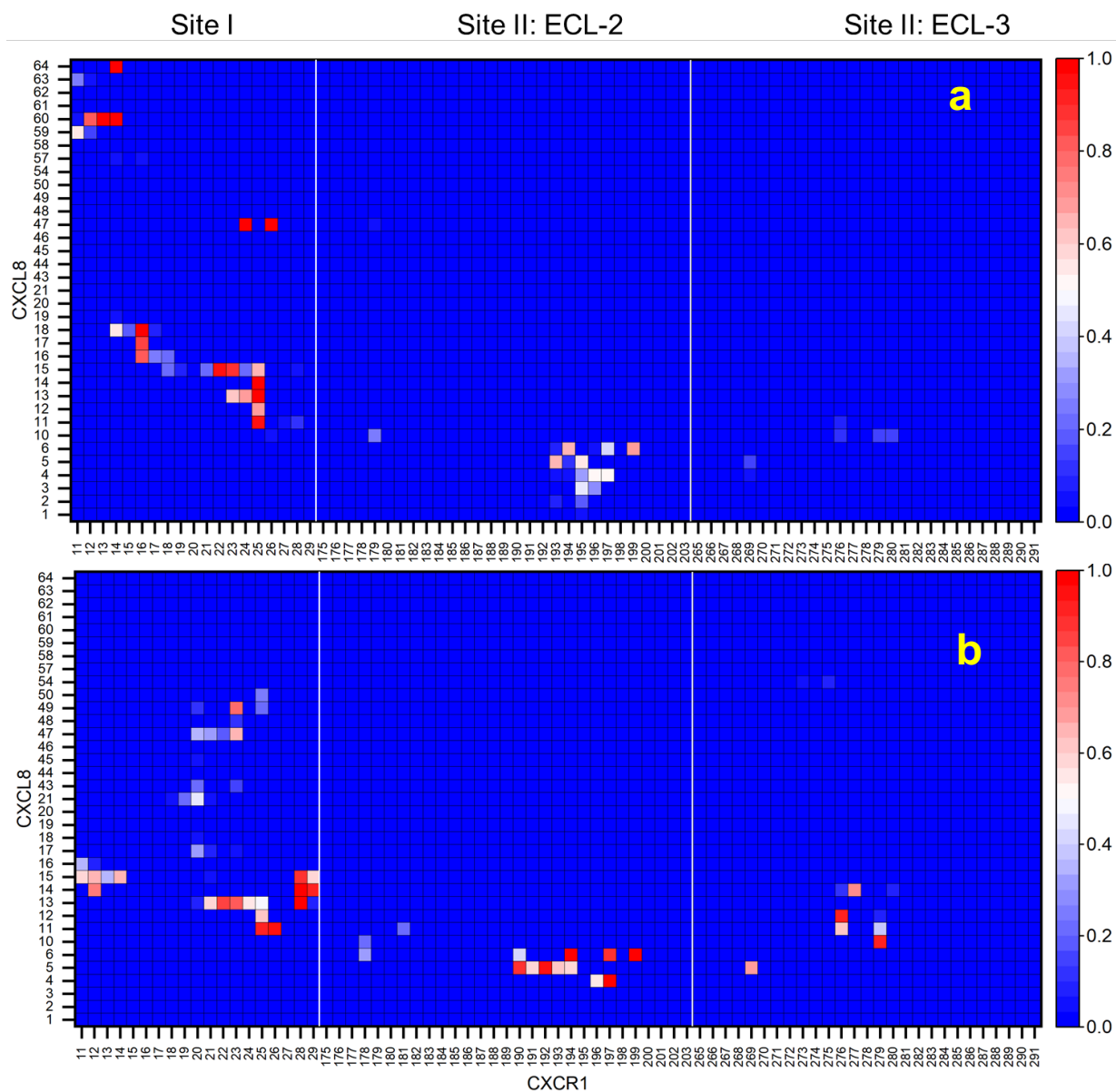


Figure 4

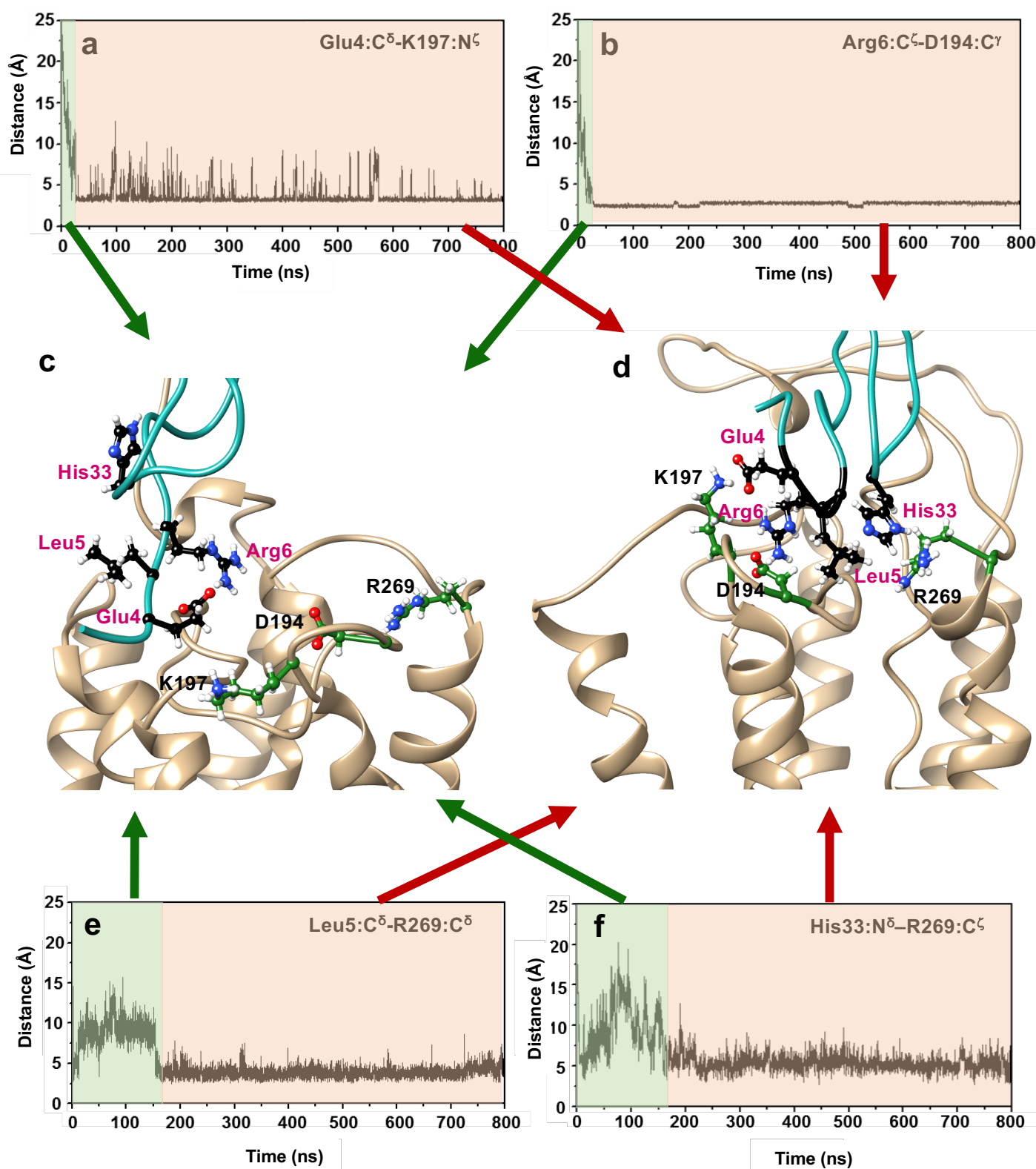


Figure 5

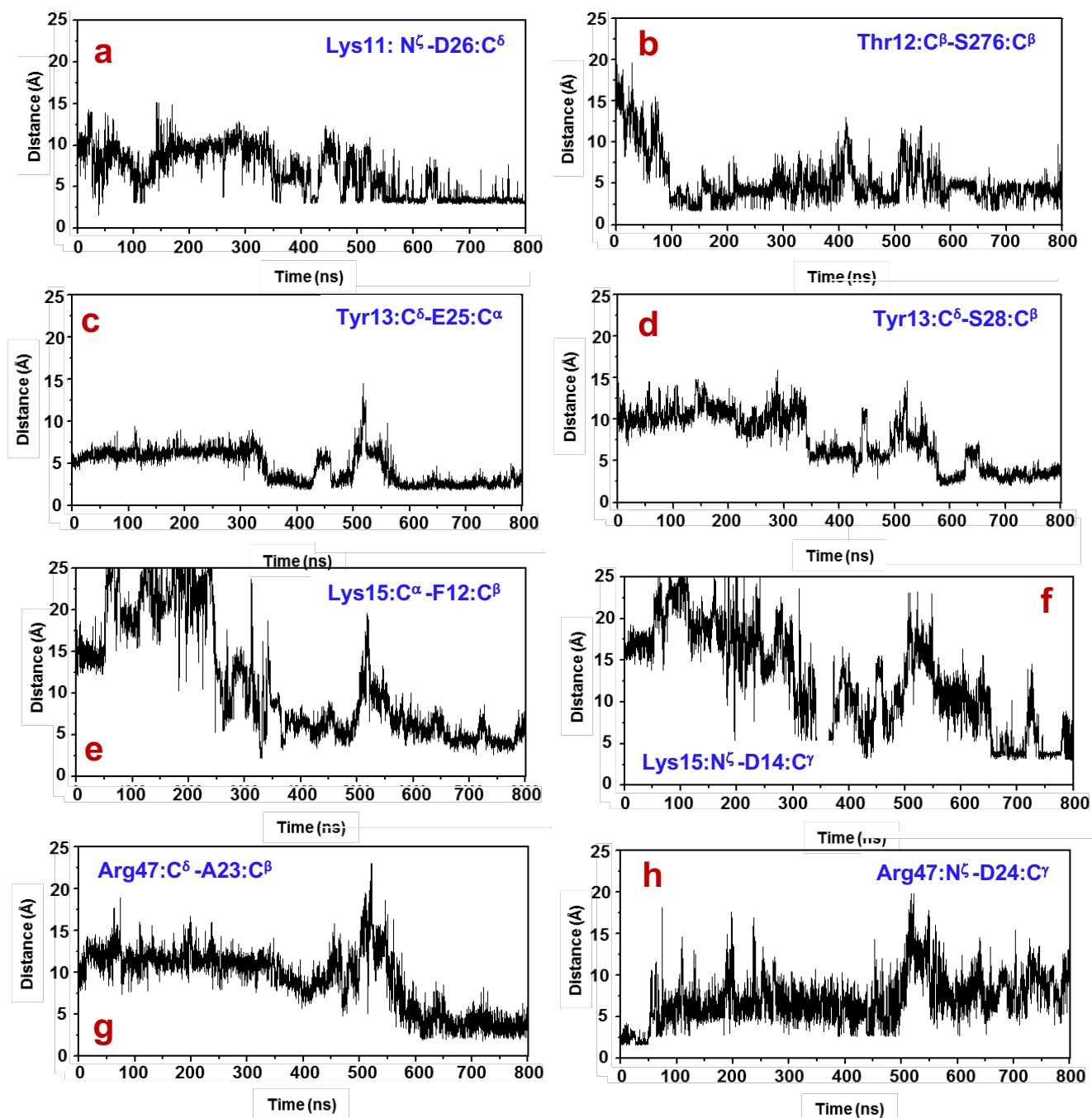


Figure 6

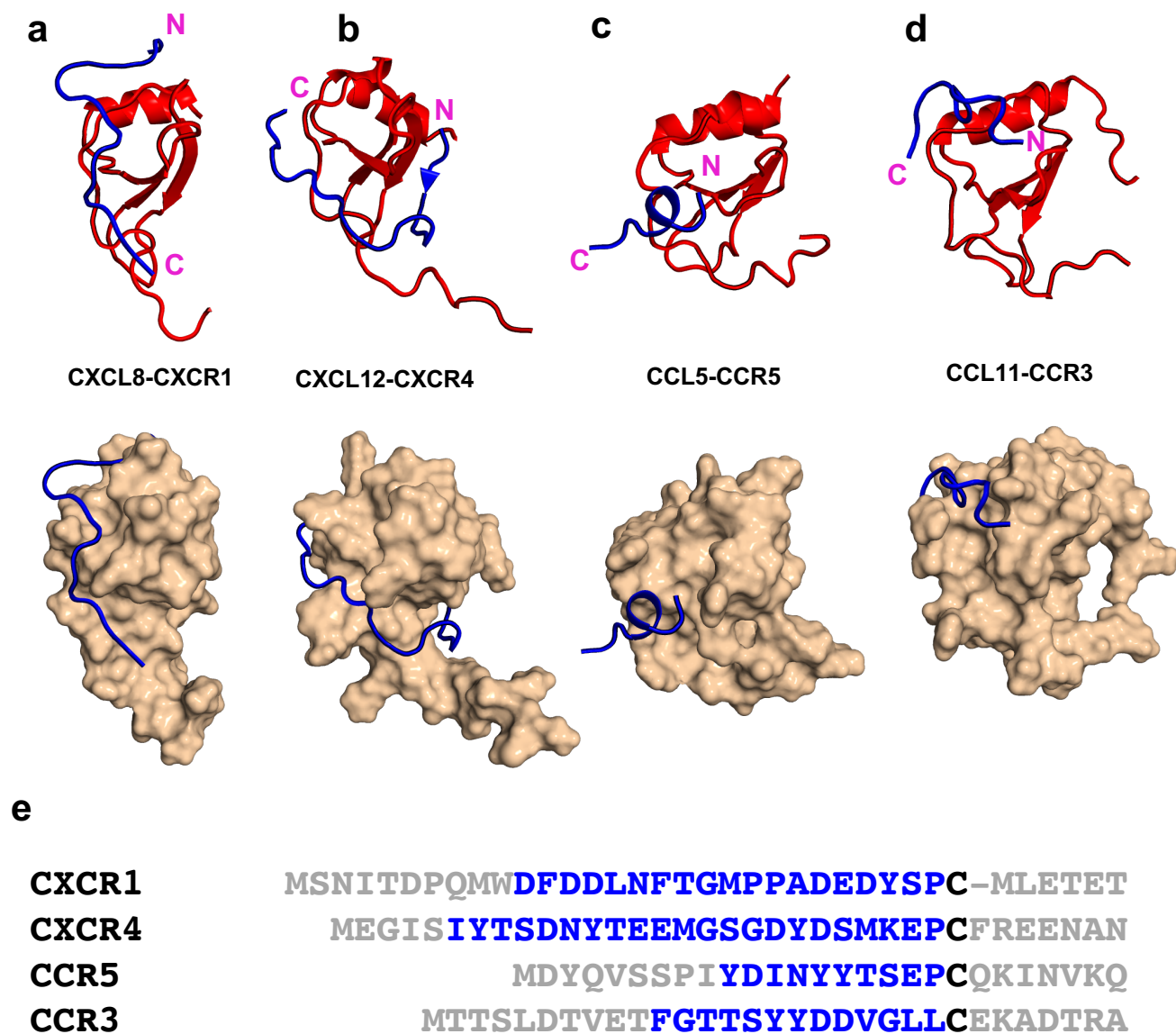
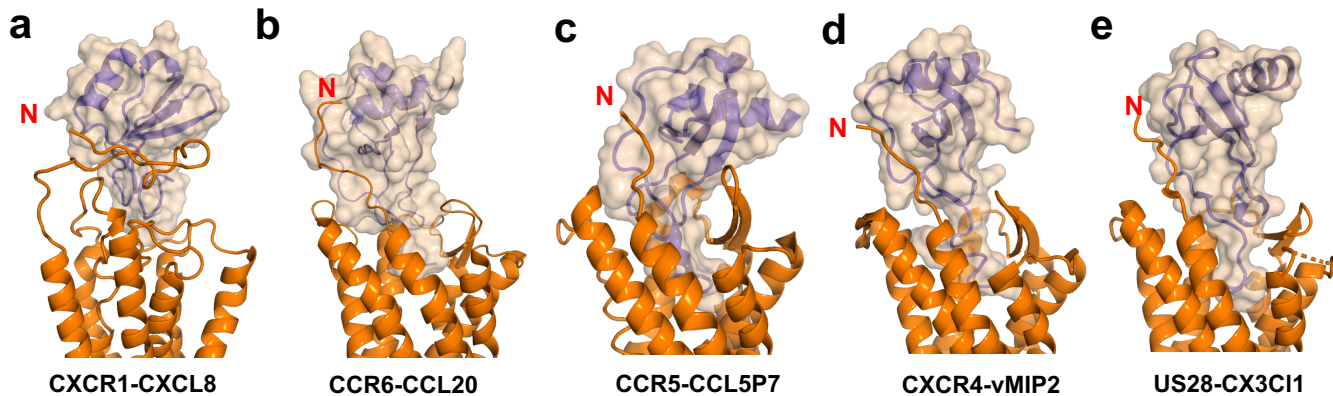


Figure 7



f

CXCR1	MSNITDPQMW <u>DFDDL</u> NFT GMPPA DEDYSPC -MLETET
CCR6	MSGESMNFSDVFDSS EDYFVSV NTSY YSD SEMLL CSLQ EV RQ
CCR5	MDYQVSSPI YDIN Y TSE PCQ KIN VKQ
CXCR4	MEG IS IY TS DNY TE EM G SG DY DS M KE PC F RE E N A N
US28	MTPTTTT AEL TTE F DY DE D A T PC V F T D V L N