Published in final edited form as:

Drug Discov Today Technol. ; 9(4): e293–e299. doi:10.1016/j.ddtec.2012.04.003.

NMR in the Analysis of Functional Chemokine Interactions and Drug Discovery

Joshua J. Ziarek and Brian F. Volkman*

Department of Biochemistry, Medical College of Wisconsin, 8701 West Watertown Plank Road, Milwaukee, Wisconsin 53226 USA

Abstract

The involvement of chemokines and chemokine receptors in a great variety of pathological indications underscores their utility as therapeutic targets. In general, chemokine-mediated migration and signaling requires three distinct interactions: self-association, glycosaminoglycan (GAG) binding, and activation of G protein-coupled receptors (GPCRs). Solution-state nuclear magnetic resonance (NMR) spectroscopy has long been used to determine the apo structure of chemokines and monitor complex formation; however, it has never contributed directly to drug discovery efforts that are traditionally focused on the previously inaccessible chemokine receptors. Our lab recently demonstrated that NMR structures can be successfully utilized to direct drug discovery against chemokines. The ease of collecting chemokine structural data coupled with the increased efficiency of structure-based drug discovery campaigns makes chemokine-directed therapies particularly attractive. In addition, recent advances in sample preparation, spectrometer hardware, and pulse program development are allowing researchers to examine interactions with previously inaccessible partners – including full-length chemokine receptors. These developments will facilitate exploration of novel ways to modulate chemokine activity using structure-guided drug discovery.

Introduction

The chemokine family of small, secreted molecules is irreplaceable in directing organogenesis, embryogenesis, and maintaining the immune response. Aberrant chemokine signaling contributes to a variety of pathologies, such as: atherosclerosis, Crohn's disease, neurological disease, and cancer [1–3]. Therefore, chemokines constitute high priority targets for therapeutic intervention. In this new age of chemical biology, the efforts of structural biologists are beginning to bear fruit by identifying and engineering protein specific treatments. NMR spectroscopy plays a crucial role in this effort. Since Clore and colleagues solved the NMR structure of interleukin-8 (IL-8/CXCL8) in 1990 [4], more than 30 other chemokines have been solved by NMR. The utility of solution NMR for analysis of protein structure, dynamics and interactions has increased steadily and chemokines are ideal targets, due to their small size, ease of recombinant production and isotope labeling, and stable structure. A major advantage of NMR is its ability to characterize intermolecular interactions in great detail without requiring full 3D structure determination. It is also the only technique that can probe changes in molecule dynamics, at atomic resolution, over

© 2012 Elsevier Ltd. All rights reserved.

*Corresponding author: bvolkman@mcw.edu, 414-955-8400.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

timescales ranging from picoseconds to seconds. This allows researchers to explore how protein structure and motion affect the three interactions necessary for chemokine function: 1) oligomerization, 2) glycosaminoglycan (GAG)-binding and 3) receptor association. Although NMR has been a major player in the identification of drug binding sites and compound optimization for more than a decade, it is now being applied for small molecule development in chemokine research. Our review aims to summarize recent advances in chemokine structure-function relationships and discuss the value of new knowledge for structure-based drug development.

Chemokine oligomerization and conformational dynamics

In contrast to X-ray crystallography, which gives a single ‘snap shot’ of protein structure, NMR can probe structural changes that occur on timescales covering more than 12 orders of magnitude. However, NMR demands unique sample requirements. Most importantly, a sample needs to contain NMR-active nuclei, such as ^1H , ^{13}C , and ^{15}N . Although the natural abundance of ^1H is near 99%, the large number of protons within macromolecules quickly makes unambiguous assignment near impossible. $^{13}\text{C}/^{15}\text{N}$ isotopically-labeled proteins provide additional correlations that can improve spectral resolution and atom assignment using multi-dimensional experiments. The position of each peak within a spectrum is known as its chemical shift. In order to monitor an interaction or determine a structure by NMR, the chemical shift of each observable nuclei must be identified or assigned. A practical protocol for semi-automated NMR assignment and structure determination was recently summarized by Ziarek and colleagues [5].

At micromolar-to-millimolar protein concentrations, most chemokines self-associate into dimers, though higher order oligomers and extended polymers have also been reported. However, chemokine activity is typically maximal at low nanomolar concentrations and the chemokine monomer appears to be responsible for full receptor activation. Conversely, it is now widely understood that chemokine oligomerization plays a vital functional role by enabling high-affinity GAG binding, thereby maintaining a chemotactic concentration gradient [6].

Concentration-dependent perturbations of chemokine NMR signals shifts can be used to identify residues in the dimer interface, in a process known as chemical shift mapping (Fig. 1). Chemical shift mapping plays an integral role in analysis of protein complexes by providing interface details without solving the structure. Changes in the rotational correlation time (determined from NMR relaxation times) or the translational self-diffusion coefficient (measured by pulsed-field gradient NMR diffusion methods) have also been used to detect self-association and assess the dimer K_d value. In addition to dimerization, some chemokines (such as CCL2 and CXCL10) form tetramers [7, 8] and a few others (CCL3, CCL4, and CCL5) have been shown to form helical polymers [9]. Interestingly, the highly conserved monomeric structure self-associates using two distinct dimer interfaces that generally correspond to either the CC or CXC subfamily (Fig. 2A). CC-type dimers form an interface via the N-loop and produce a two-stranded anti-parallel β -sheet. Dimerization of CXC-type chemokines occurs through the β 1-strands resulting in a six-stranded sheet and also involves contacts through the C-terminal helices.

Notwithstanding the distinct quaternary arrangements associated with the CC and CXC chemokine subfamilies, several interesting variations have recently been described. For example, under certain crystallization conditions CCL2 can form a CXC-type dimer and, conversely, CXCL12 can employ a CC-type interface [10]. Not surprisingly, the conserved tertiary fold also supports heterodimer formation within (CCL2/CCL8 [11] and CXCL4/CXCL8 [12]) and between (CCL5/CXCL4 [13]) chemokine subfamilies. Chemical shift

mapping and translational diffusion studies suggest that CC and CXC heterodimers are restricted to CC- and CXC-type interfaces, respectively [11, 14]; the mode of assembly for the first cross-family dimer remains to be seen. The C chemokine XCL1 is unique in that it lacks one of the two disulfide bonds found in all other chemokines. This sequence adaptation permits XCL1 to readily interconvert between two unrelated structural species in the native state [15]. Recent evidence suggests that the transition from the canonical chemokine fold to the dimeric β -sandwich structure involves complete unfolding [16]. In light of its unprecedented native state conformational equilibrium, XCL1 has been termed a 'metamorphic' protein [17].

While chemokine aggregation and GAG binding are clearly linked, the effect of oligomerization on receptor recognition and signaling is less certain. Dimerization of typical CC chemokines occludes the receptor binding surface, while CXC-type monomers and dimers both appear to be capable of GPCR binding. For at least one chemokine, there appears to be a structural and functional link between receptor binding and dimerization. Veldkamp and colleagues showed that the affinity of CXCL12 dimerization is increased more than 100-fold in the presence of a peptide corresponding to the CXCR4 N-terminus [18]. Interestingly, CXCL12 relaxation analysis has since identified a region undergoing microsecond to millisecond motions that encompasses both the receptor binding and dimerization interfaces [19]. By extrapolating the relaxation parameters to infinite dilution, Baryshnikova and Sykes demonstrated that CXCL12 undergoes conformational exchange independent of dimerization. Taken together, this suggests that binding of the receptor may activate a conformational switch to promote the self-association of CXCL12. Future studies exploring the ligand-induced main chain and side chain dynamics of chemokines would be useful in assessing chemokine allostery.

Analysis of complex formation

Identifying the oligosaccharide binding interface

Secreted chemokines oligomerize and form high-affinity complexes with cell surface proteoglycans or GAGs. This maintains the immobilized chemokines in a concentration gradient that orients cell migration toward the origin of secretion. Proudfoot and colleagues have exploited GAG-insensitive chemokines as anti-inflammatory agents (as reviewed in [20]); however, the molecular basis for GAG recognition is poorly understood. The carboxylate and sulfate groups of polyanionic GAGs are electrostatically complementary to the predominantly basic chemokines. However, other energetic contributions are required to explain the high-affinity binding of acidic chemokines CCL3 and CCL4 [21]. Unfortunately, the highly heterogeneous nature of GAGs has resulted in a poor structural characterization of the interaction; GAG species cannot be easily separated or synthesized in sufficient purity for NMR and crystallography studies. Additionally, long chain GAG molecules tend to form insoluble chemokine complexes. Chemical shift mapping with short oligosaccharides has implicated the role of a classical GAG binding motif found in the 40s loop of many chemokines (BBXB; where the B is basic and X indicates any amino acid), but additional epitopes undoubtedly participate in the interaction (Fig. 2B). Two chemokine/disaccharide structures have been solved thus far, but it is difficult to extrapolate the results to recognition of larger GAGs [22, 23]. Thus, we still lack detailed knowledge of how chemokines are recognized and immobilized by GAGs. However, new methods to engineer isotopically-labeled oligosaccharides of defined length and sulfation appear promising [24, 25]. NMR studies using labeled GAGs may reveal the specific saccharide moieties responsible for binding. Chemokine-specific GAG mimetics that block binding to cell-surface oligosaccharides could have therapeutic value as anti-inflammatory agents.

Chemokine structures with receptor N-terminal peptides

Chemokine receptor activation is a two-step process. The receptor N-terminus first binds to the chemokine (site 1). Flexible N-terminal residues of the chemokine then dock into the transmembrane pocket of the GPCR (site 2). To sidestep the difficulties associated with purifying and refolding full-length receptors, various studies have used chemical shift perturbations induced by an N-terminal receptor peptide (Fig. 2C). Site 1 in the CC, CXC, and CX3C chemokines analyzed by this approach is consistently located in the cleft formed by residues in the N-loop and β 3-strand. 3D NMR structures of two chemokine-peptide complexes define specific site 1 interactions that include this pocket and extend further along the β -sheet. The structure of CXCL8 was solved in complex with a CXCR1 peptide consisting of residues 9–29 (with residues 15–19 replaced by hexanoic acid) [26]. Three CXCL12/CXCR4 structures were solved using a constitutively-dimeric CXCL12 protein and a CXCR4 peptides comprising residues 1–38 [27]. Central to both complexes is the proximity of a receptor tyrosine side chain to the cleft formed by the N-loop and β 3-strand.

Most chemokine receptors are predicted to contain one or more sulfotyrosine residues in the N-terminal extracellular domain [28–30], and this posttranslational modification is thought to enhance ligand binding affinity. CXCR4 is O-sulfated at up to three tyrosine positions, and the NMR structure shows that each one occupies a specific binding site on the CXCL12 surface (Fig. 3A). CXCR4 sulfotyrosine (sY) 21 binds to the consensus site 1 cleft of CXCL12, while sY12 docks into a site on the β -sheet near the dimer interface. Each sulfotyrosine forms a combination of charged and hydrophobic interactions with the chemokine. Although it has not been confirmed experimentally, the CXCR1 tyrosine residue is also predicted to be sulfated, and may enhance the interaction with its chemokine ligands in a similar manner to the CXCL12/CXCR4 complex [28]. Additional structural studies will be needed to define the degree to which sulfotyrosine recognition is conserved across the family, since this post-translational modification may be a common element encoding receptor-chemokine affinity and specificity.

Monitoring interactions with full-length chemokine receptors

The recently determined CXCR4 crystal structure illustrates the size of the site 2 transmembrane pocket and the binding mode of several small molecule inhibitors [31]. Unfortunately, no structure of a GPCR-chemokine complex has been solved, so many molecular details of chemokine recognition and receptor activation remain hidden. Difficulties in receptor expression and purification make chemical shift mapping an unlikely route to examine full-length receptors because of the high protein concentrations required. A promising alternative is transferred cross relaxation (TCS), which employs selective NMR excitation of a substoichiometric receptor protein and magnetization transfer to nearby (< 6 Å) ^1H nuclei in the isotopically-labeled chemokine. Shimada and colleagues have used TCS to assess both the CXCL12/CXCR4 and CCL3/CCR5 interactions [32, 33]. By monitoring the CXCL12/CXCR4 interaction in the presence of the small molecule inhibitor AMD3100, they demonstrated the first structural evidence that chemokine activation is a two-step process. Addition of the inhibitor does not prevent CXCL12 association, but rather inhibits the chemokine N-terminus from associating with the transmembrane region and activating the receptor. Recent evidence suggests some chemokine receptors utilize agonist bias whereby disparate ligands induce unique functional outputs [34, 35]. In conjunction with the crystal structures, TCS measurements could be used to model the chemokine-GPCR complex and may be particularly attractive for characterizing promiscuous receptor-ligand interactions of the chemokine system without solving the structure of every complex. Advances in membrane protein expression have yielded several functional chemokine receptors, including: CXCR4, CCR5, and CCR1 [32, 33, 36]. Although still challenging, solution-state NMR analysis of integral membrane proteins is increasingly feasible due to

improved biochemical and spectroscopic methods. Future NMR studies of GPCR conformational dynamics will undoubtedly provide new insights into how the receptor structure responds to chemokine and inhibitor binding.

Structure-based drug design

Most chemokine drug discovery efforts employ high-throughput functional screening of large compound libraries to identify antagonists that bind the site 2 transmembrane pocket. Unfortunately, the binding surface of chemokine receptors presents an expansive landscape for inhibitors to antagonize; for example, the recently solved CXCR4 crystal structure possesses a surface area approximately 1600 Å². Like other protein:protein interface drug targets, the chemokine receptors have displayed poor hit rates and yielded relatively few active inhibitors. Experimentally determined protein structures can be screened in silico against large compound databases to prioritize molecules for functional tests and improve hit rates more than 200-fold [37]. Unfortunately, structural information of chemokine receptors is currently limited to CXCR4, and only five other unique GPCR structures have been solved. However, as noted previously, the GPCR transmembrane pocket is the second of two distinct chemokine recognition sites. Site 1, on the surface of the chemokine ligand, is readily accessible to structural analysis. Traditional drug discovery suggests that targeting the chemokine would be unsuccessful due to its small size and relatively flat topology. However, structure-guided approaches have identified high-affinity compounds for proteins of similar architecture [38, 39].

We recently tested the hypothesis that small molecule inhibitors targeting CXCL12 could be identified by in silico screening against the CXCL12: CXCR4 N-terminus NMR structure [27]. Rather than screening the entire chemokine surface, we focused on a sulfotyrosine binding pocket. Sulfation increases CXCL12 affinity 5-fold for the full-length receptor [40] and 22-fold for the CXCR4 peptide [41]. The sulfotyrosine binding pocket or chemokine cleft, structurally defined by sY21 and D20 (Fig. 3B), was subjected to high throughput docking of over 1.5 million lead-like small molecules from the ZINC database. Five of the best scoring compounds were titrated into ¹⁵N-CXCL12 and binding was monitored by 2D NMR; chemical shift mapping suggested that four of the five compounds bound specifically to the sY21 pocket. The highest affinity compound, ZINC 310454, inhibited the CXCR4-mediated calcium response in THP-1 monocytes. Subsequent structure-activity relationship (SAR) by NMR analysis of ZINC 310454 and related compounds demonstrated that the initial hit binds the sY21 recognition cleft and verified a crucial role for the carboxylate moiety. As shown in Figure 3C, substitution of the naphthyl group with a phenyl ring only resulted in the loss of the R12 chemical shift perturbation consistent with the increased distance from the smaller ring. This work demonstrates the potential of targeting the chemokine ligand with inhibitors and highlights the utility of NMR in identifying potential drug binding sites and monitoring interactions. The next major step in developing chemokine-directed inhibitors is a structure of a chemokine in complex with a small molecule antagonist. This will not only validate that structure-based drug discovery can yield site-specific inhibitors of chemokines, but will be crucial in optimizing compound hits to true lead-like molecules.

A combined bioinformatic and NMR analysis suggests that structure-based drug design against chemokines may be applicable to the entire family. The importance of receptor sulfotyrosine modifications in chemokine recognition has been demonstrated in CXCR4, CCR5, CCR2b, CXCR3, and CX3CR1 [40, 42–45]. Although no specific tyrosine sulfation sequence motif has been identified, 15 of the 19 chemokine receptors possess tyrosine residues in the extracellular N-terminus suggesting that this modification is a common feature. Primary and tertiary sequence analysis of previously determined chemokine

structures identified the presence of residues homologous to the CXCL12 sY21 binding pocket. NMR-monitored titration of sulfotyrosine ($\text{H}_2\text{N-Tyr}[\text{SO}_3]-\text{CO}_2$) into four chemokines (CXCL12, XCL1, CCL5, and CX3CL1) produced perturbations that clustered to the putative sulfotyrosine binding pocket – implying a conserved binding moiety [46]. To enable targeting of other chemokines, future efforts should focus on solving complexes with sulfated N-terminal peptides or developing alternative methods to determine the sulfotyrosine binding pocket of additional chemokines. Alternatively, the high structural conservation of chemokines suggests homology modeling may successfully guide in silico drug discovery.

Conclusions

Chemokines and their receptors have been the focus of drug discovery efforts for more than a decade, but limited structural knowledge has hampered the development of effective inhibitors. NMR has an important role that has evolved from simply determining chemokine structures to characterizing the complexities of oligomerization, GAG binding and receptor recognition. With the realization that small-molecule inhibitors can target chemokines directly and block receptor activation, synthetic ligands that modulate oligomeric state or alter GAG-binding can also be envisioned. Each new type of chemokine ligand has potential utility as a diagnostic or therapeutic tool. Improvements in NMR pulse program design and isotopic labeling now allow examination of large protein complexes. Advances in membrane protein expression and purification will soon place the direct investigations of chemokine receptors by NMR within reach. The strength of NMR to probe protein interactions at atomic resolution without requiring structure determination will be especially useful for understanding how receptors recognize multiple chemokines and why some of these interactions produce disparate functional outcomes.

Acknowledgments

JJZ is funded by a Postdoctoral Fellowship from the Medical College of Wisconsin Cancer Center and BFV is supported by National Institute of Health grants AI058072, GM097381 and AI063325.

References

1. Zlotnik A. Chemokines and cancer. *Int J Cancer*. 2006; 119(9):2026–2029. [PubMed: 16671092]
2. Savarin-Vuaillet C, Ransohoff RM. Chemokines and chemokine receptors in neurological disease: raise, retain, or reduce? *Neurotherapeutics*. 2007; 4(4):590–601. [PubMed: 17920540]
3. Zernecke A, et al. Chemokines in atherosclerosis: an update. *Arteriosclerosis, Thrombosis, and Vascular Biology*. 2008; 28(11):1897–1908.
4. Clore GM, et al. Three-dimensional structure of interleukin 8 in solution. *Biochemistry*. 1990; 29(7):1689–1696. [PubMed: 2184886]
5. Ziarek JJ, et al. Binding site identification and structure determination of protein-ligand complexes by NMR a semiautomated approach. *Methods in Enzymology*. 2011; 493:241–275. [PubMed: 21371594]
6. Proudfoot AE, et al. Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines. *Proc Natl Acad Sci U S A*. 2003; 100(4):1885–1890. [PubMed: 12571364]
7. Swaminathan GJ, et al. Crystal structures of oligomeric forms of the IP-10/CXCL10 chemokine. *Structure*. 2003; 11(5):521–532. [PubMed: 12737818]
8. Lubkowski J, et al. The structure of MCP-1 in two crystal forms provides a rare example of variable quaternary interactions. *Nature Structural Biology*. 1997; 4(1):64–69.
9. Czaplewski LG, et al. Identification of amino acid residues critical for aggregation of human CC chemokines macrophage inflammatory protein (MIP)-1alpha, MIP-1beta, and RANTES.

- Characterization of active disaggregated chemokine variants. *Journal of Biological Chemistry*. 1999; 274(23):16077–16084. [PubMed: 10347159]
10. Murphy JW, et al. Heterologous quaternary structure of CXCL12 and its relationship to the CC chemokine family. *Proteins*. 2010; 78(5):1331–1337. [PubMed: 20077567]
 11. Crown SE, et al. Heterodimerization of CCR2 chemokines and regulation by glycosaminoglycan binding. *Journal of Biological Chemistry*. 2006; 281(35):25438–25446. [PubMed: 16803905]
 12. Dudek AZ, et al. Platelet factor 4 promotes adhesion of hematopoietic progenitor cells and binds IL-8: novel mechanisms for modulation of hematopoiesis. *Blood*. 2003; 101(12):4687–4694. [PubMed: 12586630]
 13. von Hundelshausen P, et al. Heterophilic interactions of platelet factor 4 and RANTES promote monocyte arrest on endothelium. *Blood*. 2005; 105(3):924–930. [PubMed: 15459010]
 14. Nesmelova IV, et al. Platelet factor 4 and interleukin-8 CXC chemokine heterodimer formation modulates function at the quaternary structural level. *Journal of Biological Chemistry*. 2005; 280(6):4948–4958. [PubMed: 15531763]
 15. Tuinstra RL, et al. Interconversion between two unrelated protein folds in the lymphotactin native state. *Proceedings of the National Academy of Sciences of the United States of America*. 2008; 105(13):5057–5062. [PubMed: 18364395]
 16. Tyler RC, et al. Native-State Interconversion of a Metamorphic Protein Requires Global Unfolding. *Biochemistry*. 2011
 17. Murzin AG. Biochemistry. Metamorphic proteins. *Science*. 2008; 320(5884):1725–1726. [PubMed: 18583598]
 18. Veldkamp CT, et al. Recognition of a CXCR4 sulfotyrosine by the chemokine stromal cell-derived factor-1alpha (SDF-1alpha/CXCL12). *J Mol Biol*. 2006; 359(5):1400–1409. [PubMed: 16725153]
 19. Baryshnikova OK, Sykes BD. Backbone dynamics of SDF-1alpha determined by NMR: interpretation in the presence of monomer-dimer equilibrium. *Protein Sci*. 2006; 15(11):2568–2578. [PubMed: 17075134]
 20. Johnson Z, et al. Interaction of chemokines and glycosaminoglycans: a new twist in the regulation of chemokine function with opportunities for therapeutic intervention. *Cytokine Growth Factor Rev*. 2005; 16(6):625–636. [PubMed: 15990353]
 21. Kuschert GS, et al. Glycosaminoglycans interact selectively with chemokines and modulate receptor binding and cellular responses. *Biochemistry*. 1999; 38(39):12959–12968. [PubMed: 10504268]
 22. Shaw JP, et al. The X-ray structure of RANTES: heparin-derived disaccharides allows the rational design of chemokine inhibitors. *Structure*. 2004; 12(11):2081–2093. [PubMed: 15530372]
 23. Murphy JW, et al. Structural and functional basis of CXCL12 (stromal cell-derived factor-1 alpha) binding to heparin. *J Biol Chem*. 2007; 282(13):10018–10027. [PubMed: 17264079]
 24. Laguri C, et al. (13)C-Labeled Heparan Sulfate Analogue as a Tool To Study Protein/Heparan Sulfate Interactions by NMR Spectroscopy: Application to the CXCL12alpha Chemokine. *Journal of the American Chemical Society*. 2011; 133(25):9642–9645. [PubMed: 21634378]
 25. Langeslay DJ, et al. Detection of the 1H and 15N NMR resonances of sulfamate groups in aqueous solution: a new tool for heparin and heparan sulfate characterization. *Analytical Chemistry*. 2011; 83(20):8006–8010. [PubMed: 21913677]
 26. Skelton NJ, et al. Structure of a CXC chemokine-receptor fragment in complex with interleukin-8. *Structure*. 1999; 7(2):157–168. [PubMed: 10368283]
 27. Veldkamp CT, et al. Targeting SDF-1/CXCL12 with a ligand that prevents activation of CXCR4 through structure-based drug design. *J Am Chem Soc*. 2010; 132(21):7242–7243. [PubMed: 20459090]
 28. Liu J, et al. Tyrosine sulfation is prevalent in human chemokine receptors important in lung disease. *American Journal of Respiratory Cell and Molecular Biology*. 2008; 38(6):738–743. [PubMed: 18218997]
 29. Simpson LS, et al. Regulation of chemokine recognition by site-specific tyrosine sulfation of receptor peptides. *Chem Biol*. 2009; 16(2):153–161. [PubMed: 19246006]
 30. Stone MJ, et al. Tyrosine sulfation: an increasingly recognised post-translational modification of secreted proteins. *N Biotechnol*. 2009; 25(5):299–317. [PubMed: 19658209]

31. Wu B, et al. Structures of the CXCR4 chemokine GPCR with small-molecule and cyclic peptide antagonists. *Science*. 2010; 330(6007):1066–1071. [PubMed: 20929726]
32. Kofuku Y, et al. Structural basis of the interaction between chemokine stromal cell-derived factor-1/CXCL12 and its G-protein-coupled receptor CXCR4. *J Biol Chem*. 2009; 284(50):35240–35250. [PubMed: 19837984]
33. Yoshiura C, et al. NMR Analyses of the Interaction between CCR5 and Its Ligand Using Functional Reconstitution of CCR5 in Lipid Bilayers. *J Am Chem Soc*. 2010; 132(19):6768–6777. [PubMed: 20423099]
34. Zidar DA, et al. Selective engagement of G protein coupled receptor kinases (GRKs) encodes distinct functions of biased ligands. *Proceedings of the National Academy of Sciences of the United States of America*. 2009; 106(24):9649–9654. [PubMed: 19497875]
35. Drury LJ, et al. Monomeric and dimeric CXCL12 inhibit metastasis through distinct CXCR4 interactions and signaling pathways. *Proceedings of the National Academy of Sciences of the United States of America*. 2011
36. Allen SJ, et al. Expression, purification and in vitro functional reconstitution of the chemokine receptor CCR1. *Protein Expr Purif*. 2009; 66(1):73–81. [PubMed: 19275940]
37. Carlsson J, et al. Ligand discovery from a dopamine D(3) receptor homology model and crystal structure. *Nature Chemical Biology*. 2011
38. Shuker SB, et al. Discovering high-affinity ligands for proteins: SAR by NMR. *Science*. 1996; 274(5292):1531–1534. [PubMed: 8929414]
39. Arkin MR, et al. Binding of small molecules to an adaptive protein-protein interface. *Proc Natl Acad Sci U S A*. 2003; 100(4):1603–1608. [PubMed: 12582206]
40. Farzan M, et al. The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 alpha association and HIV-1 entry. *J Biol Chem*. 2002; 277(33):29484–29489. [PubMed: 12034737]
41. Seibert C, et al. Sequential tyrosine sulfation of CXCR4 by tyrosylprotein sulfotransferases. *Biochemistry*. 2008; 47(43):11251–11262. [PubMed: 18834145]
42. Farzan M, et al. Tyrosine sulfation of the amino terminus of CCR5 facilitates HIV-1 entry. *Cell*. 1999; 96(5):667–676. [PubMed: 10089882]
43. Fong AM, et al. CX3CR1 tyrosine sulfation enhances fractalkine-induced cell adhesion. *Journal of Biological Chemistry*. 2002; 277(22):19418–19423. [PubMed: 11909868]
44. Preobrazhensky AA, et al. Monocyte chemotactic protein-1 receptor CCR2B is a glycoprotein that has tyrosine sulfation in a conserved extracellular N-terminal region. *Journal of Immunology*. 2000; 165(9):5295–5303.
45. Colvin RA, et al. CXCR3 requires tyrosine sulfation for ligand binding and a second extracellular loop arginine residue for ligand-induced chemotaxis. *Molecular and Cellular Biology*. 2006; 26(15):5838–5849. [PubMed: 16847335]
46. Ziarek JJ, et al. Sulfotyrosine recognition as marker for druggable sites in the extracellular space. *International Journal of Molecular Sciences*. 2011; 12(6):3740–3756. [PubMed: 21747703]
47. Duma L, et al. Recognition of RANTES by extracellular parts of the CCR5 receptor. *J Mol Biol*. 2007; 365(4):1063–1075. [PubMed: 17101151]
48. Wang X, et al. Oligomeric structure of the chemokine CCL5/RANTES from NMR, MS, and SAXS data. *Structure*. 2011; 19(8):1138–1148. [PubMed: 21827949]
49. Veldkamp CT, et al. The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci*. 2005; 14(4):1071–1081. [PubMed: 15741341]
50. Gozansky EK, et al. Mapping the binding of the N-terminal extracellular tail of the CXCR4 receptor to stromal cell-derived factor-1alpha. *J Mol Biol*. 2005; 345(4):651–658. [PubMed: 15588815]

Figure 1. Chemical shift mapping can efficiently identify a binding interface without structure determination

A) The unique chemical shift positions of each CXCL12 nuclei reflect their distinct physicochemical environment. B) Titration of an unlabeled CXCR4(1–38) peptide into ^{15}N -CXCL12 results in chemical shift changes; nuclei that contribute most substantially to the binding interaction experience the greatest chemical shift perturbations. The change in chemical shift induced by a three-fold excess of CXCR4(1–38) are plotted as a function of CXCL12 primary sequence. C) The residues with the largest chemical shift perturbations are highlighted yellow on the CXCL12 (PDB 1SDF) surface to reveal the binding site for the CXCR4 N-terminal domain.

Figure 2. Chemokine interaction surfaces identified by chemical shift mapping

Chemokines must support three distinct binding interactions for biological activity: self-association, GAG-mediated immobilization in the extracellular matrix, and GPCR recognition. The upper (A–C) and lower (D–F) panels correspond to CCL5 and CXCL12, respectively. Amino acids involved in CC and CXC-type dimerization (A and D, respectively), heparin binding (B,E) and receptor N-terminus recognition (C,F) are highlighted on each structure in blue, magenta, and orange, respectively. Shift perturbations were reproduced from the following references: A, [47]; B, [48]; C, [47]; D, [49]; E, [23]; and F, [50].

Figure 3. The sulfotyrosine binding pocket of chemokines can be targeted by small molecule inhibitors

A) The majority of chemokine receptors possess tyrosine residues in the extracellular N-terminus that are known or predicted to be modified by O-sulfation. The NMR structure of dimeric CXCL12 in complex with a triply-sulfated CXCR4(1–38) peptide (PDB 2K05) illustrates how the three sulfotyrosine positions tie down the receptor N-terminus. B) We used D20 and sY21 from PDB 2K05 to define the spatial and physical parameters of the CXCL12 binding pocket for in silico screening. Over 1.5 million molecules were computationally screened and top scoring compounds were titrated into ^{15}N -CXCL12 and monitored by 2D NMR. C) Chemical shift mapping of our most potent hit, 3-

(Naphthalene-2-carbonylthiocarbamoylamino)benzoic acid (ZINC 310454; magenta), were well-correlated with the predicted docking pose. NMR analysis of a compound analog that replaced the naphthyl group with a phenyl ring reinforced the accuracy of the ZINC 310454 docking pose. Elimination of the outer ring of the naphthyl group (blue atoms) with a phenyl ring produced identical perturbations (magenta) except for the loss of R12 (blue) consistent with the increased distance from the bound ligand.

\$watermark-text

\$watermark-text

\$watermark-text