

Structure-Based Identification of Novel Ligands Targeting Multiple Sites within a Chemokine–G-Protein-Coupled-Receptor Interface

Emmanuel W. Smith,^{†,‡} Amanda M. Nevins,^{‡,‡} Zhen Qiao,[§] Yan Liu,[§] Anthony E. Getschman,[‡] Sai L. Vankayala,^{||} M. Trent Kemp,^{||} Francis C. Peterson,[‡] Rongshi Li,[§] Brian F. Volkman,^{*,‡} and Yu Chen^{*,†}

[†]Department of Molecular Medicine, University of South Florida, 12901 Bruce B. Downs Boulevard, Tampa, Florida 33612, United States

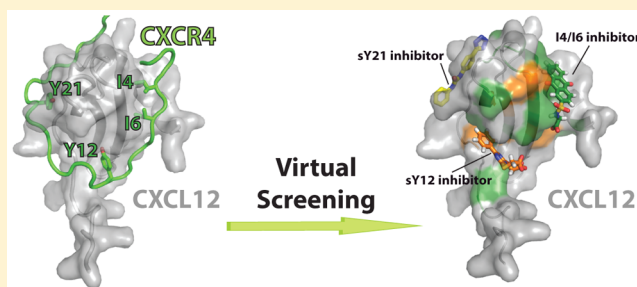
[‡]Department of Biochemistry, Medical College of Wisconsin, 8701 Watertown Plank Road, Milwaukee, Wisconsin 53226, United States

[§]Department of Pharmaceutical Sciences, Center for Drug Discovery, College of Pharmacy, and Cancer Genes and Molecular Regulation Program, Fred and Pamela Buffett Cancer Center, University of Nebraska Medical Center, 986805 Nebraska Medical Center, Omaha, Nebraska 68198, United States

^{||}Department of Chemistry, University of South Florida, 4202 East Fowler Avenue, Tampa, Florida 33620, United States

Supporting Information

ABSTRACT: CXCL12 is a human chemokine that recognizes the CXCR4 receptor and is involved in immune responses and metastatic cancer. Interactions between CXCL12 and CXCR4 are an important drug target but, like other elongated protein–protein interfaces, present challenges for small molecule ligand discovery due to the relatively shallow and featureless binding surfaces. Calculations using an NMR complex structure revealed a binding hot spot on CXCL12 that normally interacts with the I4/I6 residues from CXCR4. Virtual screening was performed against the NMR model, and subsequent testing has verified the specific binding of multiple docking hits to this site. Together with our previous results targeting two other binding pockets that recognize sulfotyrosine residues (sY12 and sY21) of CXCR4, including a new analog against the sY12 binding site reported herein, we demonstrate that protein–protein interfaces can often possess multiple sites for engineering specific small molecule ligands that provide lead compounds for subsequent optimization by fragment based approaches.



■ INTRODUCTION

Protein–protein interactions (PPIs) are appealing yet challenging targets in drug discovery efforts to modulate protein pathways in disease.¹ In contrast to enzymatic active sites that usually possess stable conformations and deep pockets, PPIs are often highly dynamic structures presenting a collection of shallow binding surfaces and are thereby frequently deemed “undruggable”.^{1,2} However, studies have shown that PPIs are not necessarily flat and that their binding is mediated by certain grooves or hot spots, where specific localized interactions provide most of the affinity.^{3,4} These binding hot spots are primarily influenced by hydrophobic interactions and, to a lesser extent, electrostatic interactions and hydrogen bonds.² Additionally, reports have emerged where small molecules have successfully targeted PPIs by binding to such hot spots.⁵ Therefore, drug discovery efforts against binding hot spots are considered a viable strategy in targeting PPIs, especially considering that a small molecule can occupy a surface of 300–1000 Å², while a hot-spot area is around 600 Å².² With the recent development of fragment-based approaches, there is

significant interest in investigating whether weak binders such as fragments can be developed for such hot spots and subsequently be used in lead optimization. In particular, with the relative promiscuity displayed by fragments and the conformational plasticity of many PPIs, there is uncertainty about whether weak-binding ligands can interact with the binding hot spots at PPIs with adequate specificity.⁶

The chemokine CXCL12 is involved in extensive protein–protein interactions with the CXCR4 receptor and, like other PPIs, presents an attractive yet difficult target for drug discovery.⁷ Chemokines are small chemotactic cytokines that bind and activate G-protein-coupled receptors (GPCRs), thus guiding receptor-expressing cells to tissues of constitutive chemokine expression.⁸ While this usually aids in the immune response by attracting leukocytes that express chemokine

Special Issue: Computational Methods for Medicinal Chemistry

Received: December 31, 2015

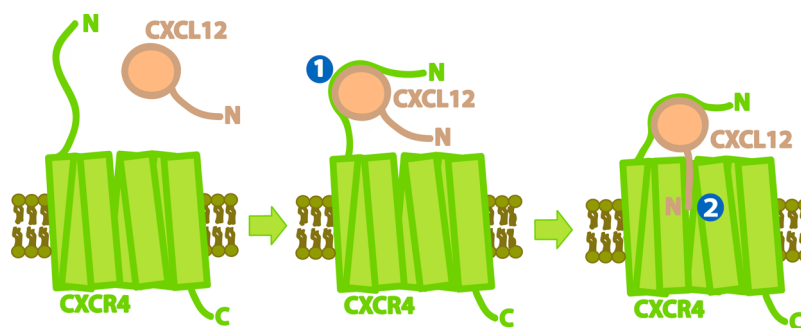


Figure 1. Two-site/two-step binding of CXCL12 to CXCR4. The first step (1) is recognition of the core domain of CXCL12 by the flexible, extracellular, N-terminal tail of CXCR4. The second step (2) is docking of the flexible, N-terminal tail of CXCL12 into CXCR4 leading to receptor activation.

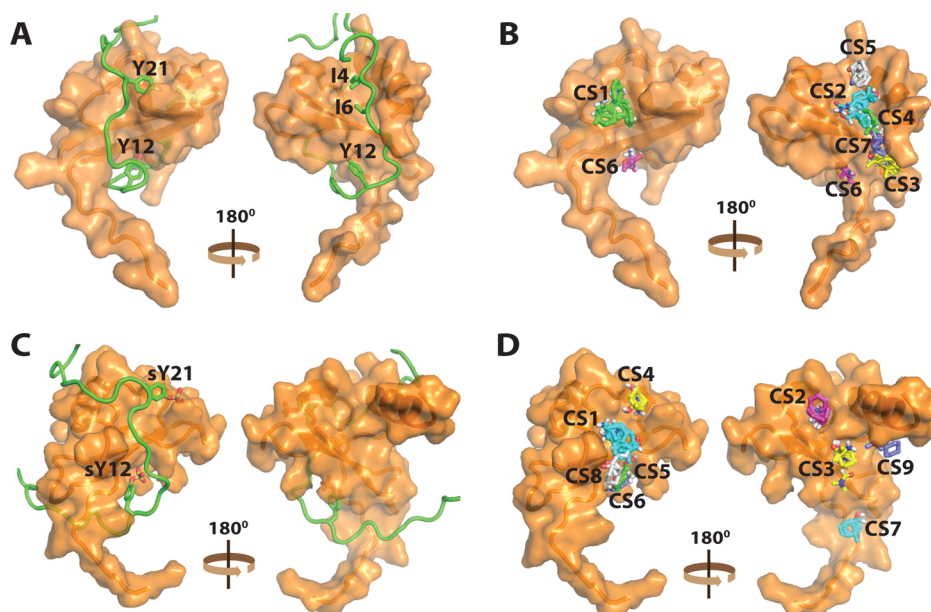


Figure 2. CXCL12–CXCR4 PPI and druggable hot spots. (A) Hot spots that mediate N-terminal CXCR4 peptide binding to monomeric CXCL12 (L55C/I58C) (PDB code 2N55). To the left, tyrosines Y12 and Y21 of CXCR4 N-terminus recognize and bind to potential hot spots on the surface of CXCL12. To the right, isoleucines I4 and I6 of the CXCR4 N-terminus bind to hot spots on the surface of CXCL12 (PDB code 2N55). (B) FTMap analysis of potentially druggable hot spots on monomeric CXCL12 (L55C/I58C) (PDB code 2N55). Probe clusters were sorted based on the number of probes in each cluster (cluster strength, CS). To the left, a large probe cluster, CS1 (green, 27 probes), was found at a site near the Y12 binding site and a small probe cluster, CS6 (purple, 6 probes), overlaps with the Y12 binding site. To the right, five probe clusters, CS2 (cyan, 15 probes), CS3 (yellow, 14 probes), CS4 (green, 12 probes), CS5 (gray, 11 probes), and CS7 (blue, 5 probes), were found at sites overlapping with the I4 and I6 binding site and surrounding area. (C) Hot spots that mediate sulfated N-terminal CXCR4 peptide binding to CXCL12 (PDB code 2K05, only monomer A depicted). (D) FTMap analysis of potentially druggable hot spots on CXCL12 (PDB code 2K05, only monomer A used). Probe clusters were sorted based on the number of probes in each cluster (cluster strength, CS). Five probe clusters, CS1 (cyan, 25 probes), CS4 (yellow, 13 probes), CS5 (green, 9 probes), CS6 (blue, 5 probes), and CS8 (gray, 3 probes), were found at sites overlapping with the sY12 and sY21 binding sites and surrounding areas. Probe clusters CS2 (purple, 15 probes) and CS3 (yellow, 15 probes) were found at sites that would normally interact with I4 and I6 in the monomeric CXCL12 complex.

receptors, cancer cells are known to hijack the same mechanism by expressing chemokine receptors in order to metastasize to tissues where chemokine expression is high.⁹ CXCL12 binds to CXCR4 in a two-step/two-site process (Figure 1). First, the extracellular flexible N-terminal domain of the receptor recognizes and binds to the surface of CXCL12, and then CXCL12 “docks” its own flexible N-terminal domain into the receptor resulting in activation and downstream signaling (Figure 1).¹⁰ Most drug discovery efforts for the CXCL12/CXCR4 interaction have been focused on targeting the second activation step by developing antagonists that would bind in the deep hydrophobic pocket of the receptor and

compete with the N-terminal domain of CXCL12.¹¹ The first step, however, represents a viable alternative target and a prototypic example of PPIs, mediated by a binding surface area covering a large portion of the chemokine. The flexible N-terminus wraps more than halfway around the chemokine, creating an interface between the CXCR4 N-terminus and CXCL12 of 2093 Å² and utilizing almost 25% of the chemokine surface area.¹² Because the binding energy is distributed across a large interaction area, targeting this interface with small molecule inhibitors is likely to be a major challenge.

Previous analyses have demonstrated potential binding hot spots mediating the binding between the receptor N-terminus

and CXCL12, which are facilitated by O-sulfation of tyrosine residues (Y21 and Y12) on the N-terminal domain of CXCR4 which then recognize specific sites on the surface of CXCL12 and increase the affinity of the interaction.^{12–14} Chemical modifications and mutagenic analysis of the CXCR4 N-terminus identified sulfated tyrosines sY21 and sY12 as the residues providing the CXCL12 contacts with the greatest binding energy.^{13,15} We have recently utilized virtual screening to identify novel ligands targeting the sY21,^{10,16,17} and sY12 (A. E. Getschman et al., unpublished results) binding hot spots separately, with binding affinities in the low micromolar range. These studies have demonstrated that the sulfotyrosine-binding sites on CXCL12 can indeed be utilized in targeting the initial PPI of chemokine signaling.

The previous successes led us to question whether there were other binding hot spots on CXCL12 and how they could be identified and targeted. In this study we performed FTMap¹⁸ analysis on the NMR structure of constitutively monomeric CXCL12 (L55C/I58C) in complex with an unsulfated N-terminal receptor peptide (PDB code 2N55) revealing a third potential binding hot spot on CXCL12 that usually binds to the I4/I6 residues of the receptor peptide. Then, through docking to single and multiple chemokine conformations, novel small molecule ligands have been uncovered that interact specifically with the I4/I6 binding site. In addition, we present a compound developed based on a previously discovered virtual screening hit for the sY12 binding site, further confirming that specific ligands can be engineered to target the different binding hot spots of the PPI between CXCL12 and CXCR4. These results demonstrate that CXCL12 can be targeted at multiple sites with small molecule ligands and offer insights into computational approaches in targeting chemokine–GPCR interactions as well as other PPIs.

■ RESULTS

Computational Analysis Identifies Additional Binding Hot Spots. Potentially druggable hot spots on protein surfaces can be identified via computational solvent mapping programs,¹⁹ such as FTMap.²⁰ These programs dock, in silico, small organic probes onto the surface of a protein while sampling a large number of conformations. Areas where multiple probes cluster may be indicative of druggability. The FTMap¹⁸ solvent mapping server was used to analyze the druggability of putative hot spots on the surface of CXCL12. As CXCL12 exists as both monomer and dimer in vivo,²¹ two structures were used for this analysis, corresponding to a monomeric and dimeric state respectively (Figure 2). We first focused on the NMR complex structure of constitutively monomeric CXCL12 (L55C/I58C) bound to a nonsulfated, 40-residue long (p40), N-terminal CXCR4 peptide (PDB code 2N55) (Figure 2A). Two mutations, L55C and I58C, were used to engineer a disulfide bond locking the protein conformation in the monomeric state and preventing dimerization.²² Probe clusters were ranked based on the number of probes in each cluster (cluster strength (CS), CS1–7) and were found at several sites that represent potential binding hot spots, including the binding site for Y12 (CS6) and a cleft near the binding site for Y21 from the receptor (CS1) (Figure 2A,B). The highest density of probe clusters (five clusters: CS2, CS3, CS4, CS5, and CS7) was found at and around a site that recognizes two isoleucine residues, I4 and I6, of the CXCR4 receptor, suggesting that this site might be more druggable than the Y12 and Y21 binding sites in the monomer

(Figure 2A,B). FTMap analysis was also performed using monomer A from the NMR complex structure of dimeric CXCL12 bound to a triply sulfated (sY7, sY12, sY21), 38-residue long (p38), N-terminal CXCR4 peptide (PDB code 2K05) (Figure 2C).¹² Probe clusters were again ranked based on the number of probes in each cluster (cluster strength (CS), CS1–9), and five probe clusters were found at sites overlapping with and in between the sY12 and sY21 binding sites, also corresponding to the Y12 and Y21 binding sites in the monomeric NMR structure (Figure 2C,D). Most probe clusters occupy hot spots at the cleft between the sY12 and sY21 binding sites (CS1, CS5, CS8). One probe cluster (CS4) occupies the sY21 binding site, and one (CS6) occupies the sY12 binding site. These findings further support that the (s)Y12/(s)Y21 recognition sites are potential binding hot spots. In the dimer structure, however, the I4/I6 binding site is masked by conformational changes in the protein as well as intermolecular interactions across the dimer interface. Still, a few probe clusters (CS2, CS3) were found in the area that would usually interact with I4/I6 of CXCR4 (Figure 2D) but not as many as were found in the constitutively monomeric CXCL12 (L55C/I58C) NMR structure (Figure 2B), suggesting that the constitutively monomeric CXCL12 (L55C/I58C) NMR structure may be a better template for virtual screening experiments against that site.

To gain additional insight into the properties of the I4/I6 binding site compared to other hotspots and analyze its recognition epitopes on the CXCR4 peptide, we also used Schrodinger Prime MM-GBSA^{23,24} to calculate the energetic contribution of each residue side chain on the receptor N-terminus to the PPI using the same constitutively monomeric NMR CXCL12 (L55C/I58C) complex structure (PDB code 2N55) and also compare it to monomer A from dimeric CXCL12 in complex with the sulfated CXCR4 peptide (PDB code 2K05) (Figure S1 in Supporting Information). Whereas the calculations demonstrated that I4 is the residue contributing most to the binding of the nonsulfated N-terminus of CXCR4 to CXCL12 (PDB code 2N55), the results concerning other residues and binding hot spots correlated less well with FTMap analysis and experimental data, as the calculations deemphasized the contributions of some important residues, such as Y21 (Supporting Information).^{12,25} We therefore focused mainly on FTMap for our binding hot spot analysis.

Novel Small Molecule Ligand Discovery against the I4/I6 Binding Site Using Rigid Docking. On the basis of FTMap analysis of constitutively monomeric CXCL12 (L55C/I58C) (Figure 2B) and the energetic contribution calculations for each residue of the N-terminal CXCR4 peptide (Figure S1), the I4/I6 binding site of monomeric CXCL12 appeared to be a promising binding hot spot for small molecule ligands. In order to experimentally probe the druggability of this newly identified candidate hot spot, we performed virtual screening experiments against the NMR monomer structure using DOCK 3.5.54 and the ZINC small molecule database.^{26–28} Conformation 1 of the NMR ensemble was used as a rigid template. From the top scoring compounds, 12 were screened experimentally using 2D ¹H–¹⁵N SOFAST-HMQC spectroscopy to test the binding of potential ligands to WT-CXCL12, which existed as monomer under the experimental conditions. Of the 12 compounds experimentally analyzed in this screen, 3 showed binding to WT-CXCL12 (1 (ZINC C04181455),²⁶ 2 (ZINC C40310216),²⁶ and 3 (ZINC C16480049)²⁶) (Figure S2),

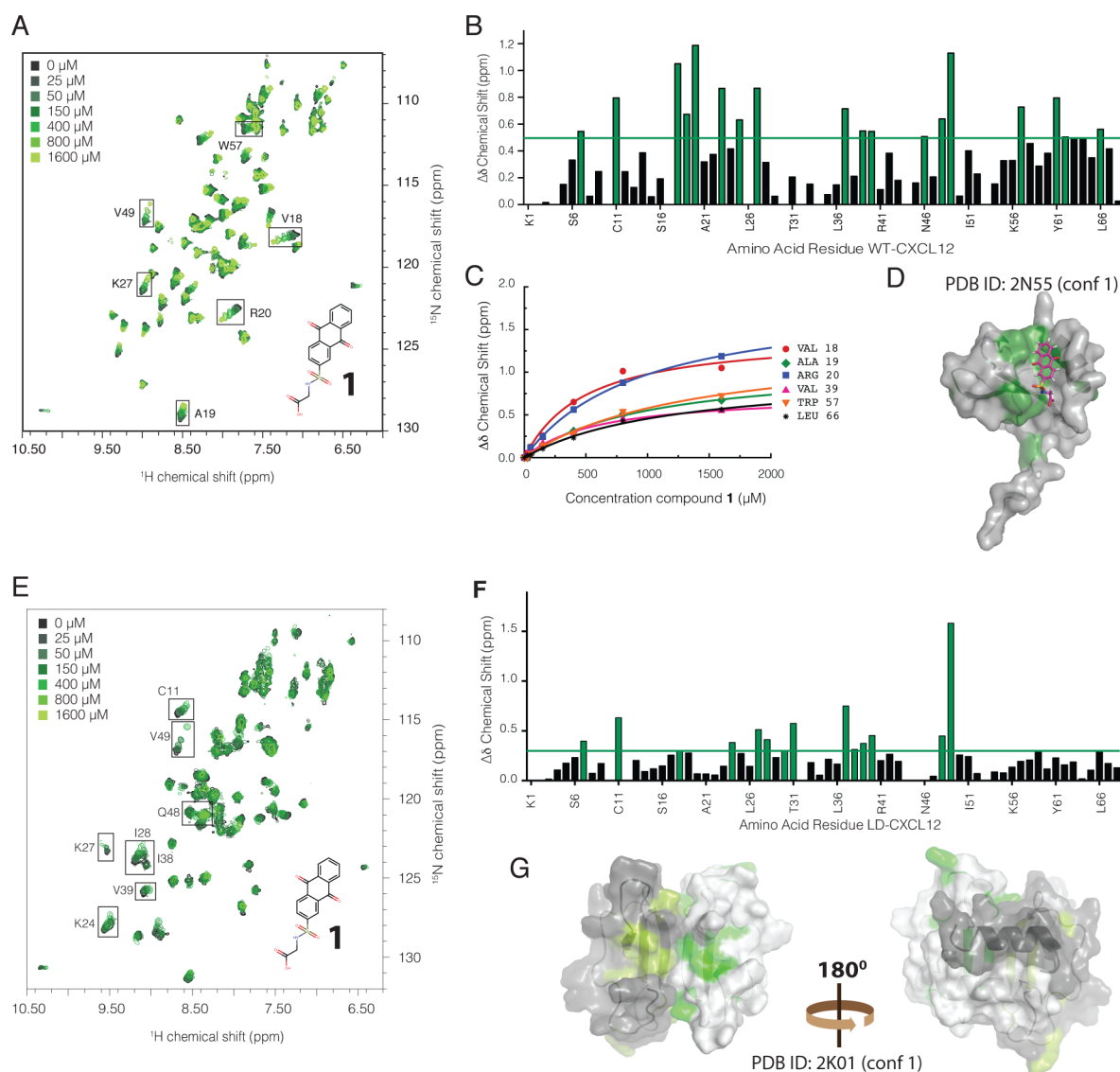


Figure 3. Novel small molecule ligand discovery against the I4/I6 binding site using rigid docking. (A) ^1H – ^{15}N HMQC overlay of titration of WT-CXCL12 with **1** (inset) used for chemical shift mapping showing significant chemical shift perturbations for a subset of residues. (B) Quantification of the change in total chemical shift ($^1\text{H}/^{15}\text{N}$ $\Delta\delta$ chemical shift) for individual amino acid residues of CXCL12 with those experiencing the largest shifts highlighted in green. (C) Fitted curves for selected residues used to calculate the affinity (K_d) of **1** for WT-CXCL12. (D) Mapping of significantly perturbed residues onto the NMR structure (conformation 1) of constitutively monomeric CXCL12 (L55C/I58C) (PDB code 2N55) used to identify **1** in rigid virtual screening, corresponds to the predicted binding site for **1**. (E) ^1H – ^{15}N HMQC overlay of titration of locked dimeric CXCL12 (L36C/A65C) with **1** (inset) used for chemical shift mapping showing significant chemical shift perturbations for a subset of residues. (F) Quantification of the change in total chemical shift ($^1\text{H}/^{15}\text{N}$ $\Delta\delta$ chemical shift) for individual amino acid residues of LD-CXCL12 (locked dimer) with those experiencing the largest shifts highlighted in green. (G) Mapping of significantly perturbed residues onto the NMR structure (conformation 1) of dimeric CXCL12 (L36C/I65C) (PDB code 2K01). The two monomers are colored in gray and white with corresponding perturbed residues colored in light green and dark green, respectively.

two of which (**1** and **2**) induced chemical shift perturbations to residues associated with the targeted I4/I6 binding site, for a final hit rate of 16.7% (Figure S2). In particular, **1** displayed concentration-dependent binding with the K_d estimated to be 1.2 ± 0.3 mM. The largest chemical shift perturbations were observed for residues surrounding the I4/I6 binding site of CXCL12 (Y7, C11, V18, R20, V23, H25, K27, V39, Q48, W57, Y61, and L66) (Figure 3A–D, S4A). For **2**, the concentration-dependent binding was weaker with an estimated K_d of 1.7 ± 0.7 mM.

To further probe the specificity of **1**, 2D ^1H – ^{15}N SOFAST-HMQC spectroscopy was carried out using the engineered locked dimer of CXCL12 (L36C/A65C). For the locked dimer,

engineered residues at the dimer interface (L36C/A65C) link both monomers together into the dimeric configuration through a pair of symmetric intermolecular disulfide bonds.¹² The I4/I6 binding site is only present in the monomeric state of CXCL12 and not accessible in the dimeric state of the protein. In particular, it has been found that in the CXCL12 dimer, the α -helix constituting a major part of the I4/I6 binding site adopts a different orientation relative to the rest of the protein including but not limited to the β -sheet that is also part of the I4/I6 binding site (Figure 2A,C).^{29,30} These conformational changes in the dimer flatten the binding surface of the I4/I6 binding site, making it significantly less druggable. A specific ligand for the I4/I6 site in WT-CXCL12, as suggested

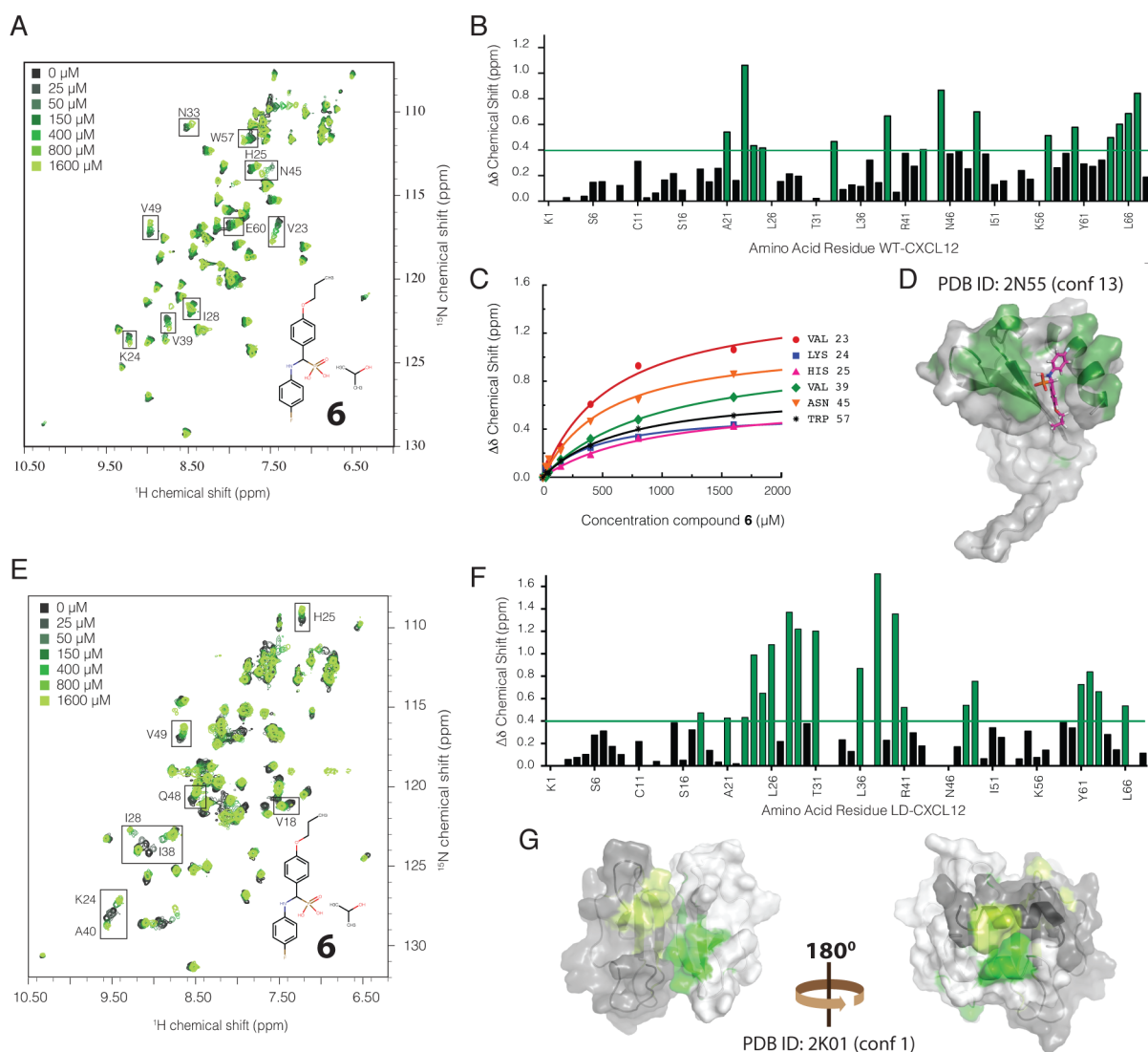


Figure 4. Novel small molecule ligand discovery against the I4/I6 binding site using ensemble docking. (A) ^1H - ^{15}N HMQC overlay of titration of WT-CXCL12 with **6** (inset) used for chemical shift mapping showing significant chemical shift perturbations for a subset of residues. (B) Quantification of the change in chemical shift ($\Delta\delta$ chemical shift) for individual amino acid residues of CXCL12 with those experiencing the largest shifts highlighted in green. (C) Fitted curves for selected residues used to calculate the affinity (K_d) of **6** for WT-CXCL12. (D) Mapping of significantly perturbed residues onto the NMR structure (conformation 13) of constitutively monomeric CXCL12 (L55C/I58C) (PDB code 2N55) used to identify **6** in ensemble virtual screening, corresponds to the predicted binding site for **6**. (E) ^1H - ^{15}N HMQC overlay of titration of locked dimeric CXCL12 (L36C/A65C) with **6** (inset) used for chemical shift mapping showing significant chemical shift perturbations for a subset of residues. (F) Quantification of the change in total chemical shift ($^1\text{H}/^{15}\text{N}$ $\Delta\delta$ chemical shift) for individual amino acid residues of LD-CXCL12 (locked dimer) with those experiencing the largest shifts highlighted in green. (G) Mapping of significantly perturbed residues onto the NMR structure (conformation 1) of dimeric CXCL12 (L36C/I65C) (PDB code 2K01). The two monomers are colored in gray and white with corresponding perturbed residues colored in light green and dark green, respectively.

by the NMR studies above, should only bind to the I4/I6 site in the monomeric state but not in the CXCL12 dimer. During the experiment with the locked dimer, **1** induced modest shift perturbations (<0.5 ppm) for multiple residues (Y7, K24, K27, I28, N30, I38, V39, A40, Q48) and larger (>0.5 ppm) shifts for a small group (C11, T31, Q37, V49) (Figure 3E,F). The linear (i.e., nonsaturable) concentration dependence of the largest shifts (e.g., V49) was consistent with nonspecific binding. The small perturbations of K24, K27, and I28 in the $\beta 1$ strand and I38, V39, and A40 in the $\beta 2$ strand surround a cluster of positively charged residues from both monomers (Figure 3G) that bind other negatively charged molecules³¹ and may attract **1** through nonspecific electrostatic interactions. Importantly, several residues in the I4/I6 binding site such as W57 and Y61

in the C-terminal helix were not perturbed significantly by the binding of **1** to the dimer, further suggesting that interactions between **1** and this site occur uniquely with the monomeric state of WT CXCL12 (Figure 3B,F).

Novel Small Molecule Ligand Discovery against the I4/I6 Binding Site Using Ensemble Docking. In order to more fully capture the flexibility of CXCL12, all 20 conformations of the NMR structure were used in an ensemble docking experiment with each conformation serving as a rigid template for virtual screening. The docking results of all 20 virtual screens were then combined, and compounds were selected from the top-ranking list for experimental analysis. Twenty-four compounds were analyzed using 2D ^1H - ^{15}N SOFAST-HMQC spectroscopy to test binding of the

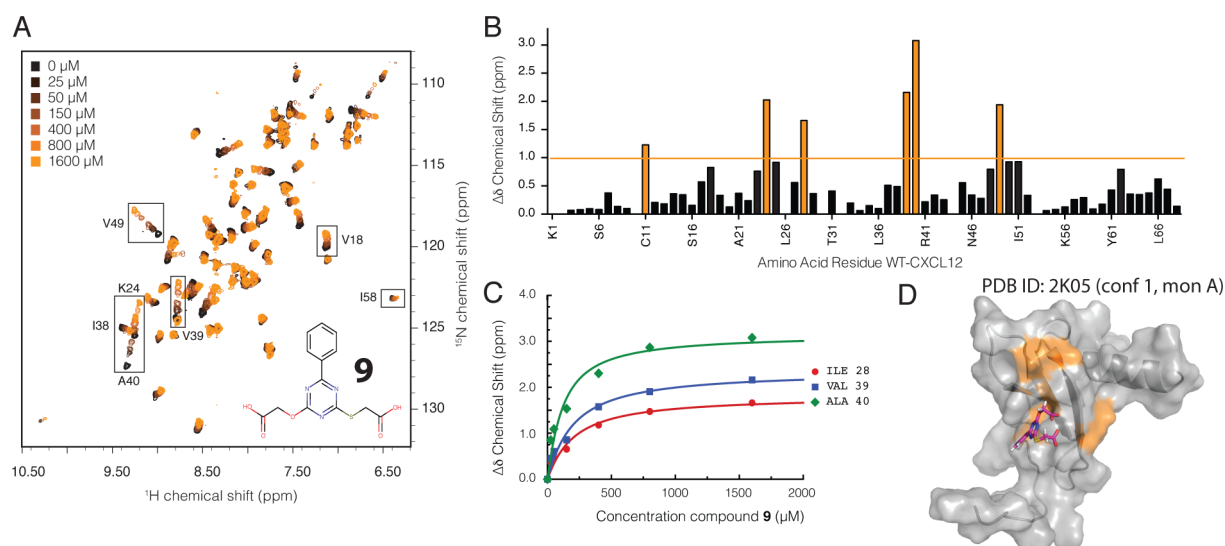


Figure 5. Successful targeting of the sY12 binding site of CXCL12. (A) ¹H-¹⁵N HMQC overlay of titration of WT-CXCL12 with **9** (inset) used for chemical shift mapping showing significant chemical shift perturbations for a subset of residues. (B) Quantification of the change in chemical shift ($\Delta\delta$ chemical shift) for individual amino acid residues of CXCL12 with those experiencing the largest shifts highlighted in orange. (C) Fitted curves for selected residues used to calculate the affinity (K_d) of **9** for WT-CXCL12. (D) Mapping of significantly perturbed residues onto the NMR structure (monomer A, conformation 1) of dimeric CXCL12 (PDB code 2K05) corresponds to the binding site for **9** as predicted by docking.

compounds to WT-CXCL12. From the 24 compounds screened, five showed binding to WT-CXCL12 (Figure S3). Among those five, **4** (ZINC C44978491)²⁶ and **5** (ZINC C69492022)²⁶ were determined to bind nonspecifically (Figure S3A,B). The rest, **6** (ZINC C12998741),²⁶ **7** (ZINC C15782120),²⁶ and **8** (ZINC C07362052),²⁶ induced chemical shift perturbations to residues associated with the I4/I6 binding site, resulting in a final hit rate of 12.5% (Figure S3C–E). In particular, **6** induced concentration-dependent chemical shift perturbations at a number of residues, the largest of which were associated with the I4/I6 binding site of CXCL12 (A21, V23, K24, H25, N33, V39, K43, N45, V49, W57, E60, K64, A65, L66, and N67), and had an estimated K_d of 1.1 ± 0.3 mM (Figure 4A–D, S4B). **7** and **8** displayed weaker affinities with estimated K_d values of 4.1 ± 2 mM and 4.1 ± 1 mM, respectively. The docking template that identified **6** corresponded to conformation 13 of the NMR ensemble, similar to **4** and **5**, and was unlike conformation 1 used in the original rigid docking that identified **1** (Figure S3). The docking templates that identified **7** and **8** corresponded to, respectively, conformations 5 and 16 of the NMR ensemble (Figure S3). This suggests that the compounds recognize slightly different protein conformations and that multiple conformations of the NMR structure can, and should, be used in virtual screening to identify novel ligands, even though some conformations may outperform others as docking template.

6 was also subjected to the binding analysis by 2D ¹H-¹⁵N SOFAST-HMQC spectroscopy using the locked dimer. Interestingly, whereas it exhibited some nonspecific interactions with several binding surfaces on CXCL12, **6** appeared to bind specifically to the β -sheet region of the positively charged dimer interface, with K_d of 74 ± 10 μM, suggesting that these binding surfaces represent new hot spots in the dimer (Figure 4E–G). Significantly, several residues in the I4/I6 binding site, such as W57, were much less perturbed in the engineered dimer than the WT, again indicating specific binding to the target site in the WT monomer.

Development of a Specific Ligand Targeting the sY12 Binding Site of CXCL12 Monomer.

It should be mentioned that the binding site for **6** in the engineered dimer overlaps with the sY12 binding pocket in the monomer, for which we have also been developing ligands through virtual screening (A. E. Getschman et al., unpublished results). Whereas those results will be described in detail in a separate publication, it will be interesting to compare binding to the sY12 site in the monomer versus dimer for the current study. For this purpose we turned to 2-((4-(carboxymethoxy)-6-phenyl-1,3,5-triazin-2-yl)thio)-acetic acid (**9**), designed and synthesized (Scheme S1) based on a novel inhibitor scaffold targeting sY12 binding site in the monomer. **9** was found to bind to the WT sY12 binding site by 2D ¹H-¹⁵N SOFAST-HMQC spectroscopy, with the largest chemical shift perturbations associated with residues surrounding the sY12 binding site (C11, K24, I28, V39, A40, and V49) and an estimated K_d of 200 ± 55 μM (Figure 5 and Figure S4C). These results demonstrate that the sY12 binding site is accessible in the monomer. **6** likely recognizes key structural features unique to the dimer, such as residues across the interface. The binding of **9** to the sY12 site also suggests that specific interactions between small molecules and this site can be recapitulated by designed compounds, thus paving the way for future fragment-based lead optimization efforts.

DISCUSSION

Multiple Binding Hot Spots for Targeting the CXCL12 PPI. Through a structure-based virtual screening strategy used in this study and together with the results from previous studies,^{10,16,17} we have identified novel small molecule ligands, which despite their weak affinities (micromolar to millimolar) are capable of binding to three distinct sites on the extensive PPI of the CXCL12 chemokine (sY21, sY12, and I4/I6 binding sites) with its receptor, CXCR4 (Figure 6). These results suggest that these PPI binding hot spots possess enough unique structural features, such as the shape of the pocket, the spatial arrangement of polar functional groups for hydrogen bonding interactions, significant hydrophobic surfaces, and a large

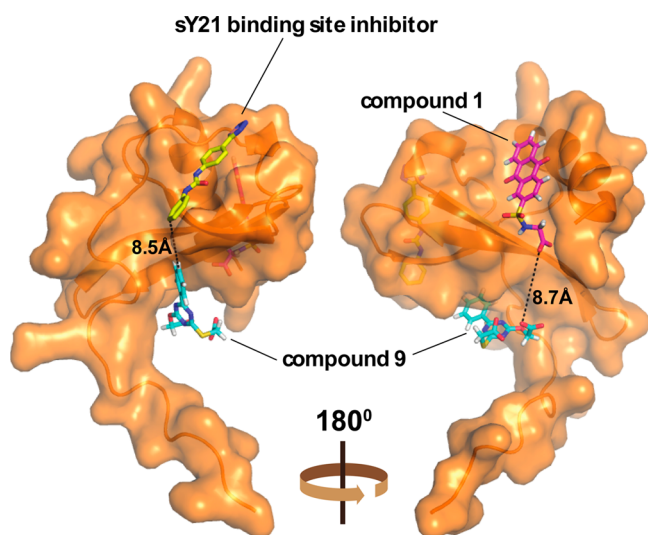


Figure 6. Novel small molecule ligands targeting sY12, sY21, and I4/I6 binding sites on CXCL12. sY21 binding site inhibitor (yellow) was retrieved from crystallographic complex structure (PDB code 4UAI) and superimposed to monomeric CXCL12 (PDB code 2N55). 1 (purple) and 9 (cyan) bind to the I4/I6 and Y12 binding sites, respectively, as predicted by docking. A distance of 8.5 Å separates the sY21 binding site inhibitor from 9, and a distance of 8.7 Å separates 1 from 9.

number of positively charged side chains, that can distinguish between weak binders including fragments and are thus suitable for fragment-based inhibitor discovery. Moreover, 9, located in the sY12 binding pocket, is 8.5 and 8.7 Å away from the sY21 binding site inhibitor and 1 in the I4/I6 pocket, respectively (Figure 6). The close proximity of these molecules to one another suggests that a fragment-based linking or growing strategy can be applied to develop ligands that can target multiple hot spots and achieve higher binding affinities. In particular, molecules such as 9 can provide a central scaffold that extends from the sY12 pocket into the binding sites for sY21 and I4/I6 from CXCR4.

Each of the three binding sites also plays a different role in the function of CXCL12 and correspondingly has varied ligand-binding characteristics that may contribute to the different affinities and activities of the individual ligands targeting each site. On the basis of previous mutagenesis and chemical modification studies, the sY21 binding site appears to be one of the main contributors to the binding affinity of the interactions with CXCR4.^{12,13} Correspondingly, the ligands identified against this site are the strongest binders among all three series of novel CXCL12 ligands, with affinities in the low micromolar range.^{10,16,17} The sY12 binding site is also involved in the CXCR4 interaction but is equally important functionally for its association with heparin in the extracellular matrix.³¹ Compounds targeting the sY12 binding site are less effective in disrupting CXCL12–CXCR4 interactions when compared with the ligands for the sY21 binding site but are instead able to compete with heparin for CXCL12 binding (A. E. Getschman et al., unpublished results). These molecules provide an alternative opportunity to interfere with *in vivo* chemotaxis, since heparin binding is used to protect CXCL12 from proteolysis and to establish the chemokine gradient necessary for adhesive migration.³¹ Meanwhile the focus of the current study, the I4/I6 binding site, is only accessible in the monomeric state and not in the dimeric state, as described

earlier. The I4/I6 binding site may therefore play a role in the different activities inherent to the CXCL12 monomeric and dimeric states. The conformational variability of this site may also partially explain the relatively low affinity of the ligands targeting this site, due to the high entropic penalty imposed on the protein upon ligand binding. Ligands targeting the I4/I6 site may provide valuable chemical probes to study CXCL12 by shifting the conformation equilibrium toward the monomeric state, whereas this unique binding pocket offers additional protein surfaces to augment chemokine binding for small molecules targeting multiple binding hot spots.

Our results also demonstrate the ability of computational approaches in identifying PPI binding hot spots and uncovering novel ligands using structural information from various experimental sources including NMR. In particular, they suggest that computational solvent mapping techniques such as FTMap can be used to uncover multiple druggable binding sites of PPIs. The successes of the virtual screening experiments demonstrate that the distinct binding features of PPIs can be recognized and utilized by computational docking programs to select specific binders. However, whereas both FTMap and binding energy analysis suggest the I4/I6 binding site to be the most druggable, its suitability as a hot spot for small molecule ligand binding may be affected by its overall conformational instability as discussed above. In addition, although the relative binding energy analysis offered insights into the importance of I4 in the interactions between CXCR4 and CXCL12, the calculations may have been unable to accurately capture the conformations of the small protein and its peptide ligand and therefore the subsequent contributions of some residues to binding.

NMR Structure in Ligand Discovery against a Flexible Target. In this study, we describe the use of both a rigid and an ensemble virtual screening strategy to identify novel ligands against a potentially transient binding pocket (i.e., the I4/I6 binding site) in an NMR structure. In our previous studies, NMR and X-ray crystal structures were used to uncover the ligands for the sY21,¹⁷ and sY12 binding sites, respectively (A. E. Getschman et al., unpublished results). Because of the structural ambiguity, in particular concerning side chain conformations, NMR structures have been considered less suitable for molecular docking experiments than crystal structures.³² However, similar to previous studies using an NMR ensemble to target protein active sites,³³ our results demonstrate that NMR structures, including different conformations of the ensemble, can be successfully utilized for identifying novel small molecule ligands targeting PPIs. Our particular studies have benefitted from the fact that the NMR structures being used are complexes of the target protein (CXCL12) with a biological ligand (CXCR4 peptide). In comparison, another virtual screening campaign using the apo NMR structure of a different chemokine, CCL21, was unsuccessful (unpublished data). Our analysis therefore indicates that complex NMR structures contain enough structural information for virtual screening. Additionally, it is possible that certain conformations in the NMR ensemble may be more suitable virtual screening templates than others because of higher experimental accuracy or more druggable features. However, systematic analysis is required to compare the performance of various NMR conformations in virtual screening and to develop methods to identify those that can lead to higher hit rates.

CONCLUSION

Our virtual screening approach has proven effective for identifying ligands against multiple sites on the surface of a small protein, CXCL12. Since each site also contributes to the functional interaction with the CXCR4 receptor, their proximity should enable fragment-based linking/growing/merging strategies to develop potent and specific inhibitors of CXCL12 by simultaneous targeting of multiple sites in future studies. Our work also suggests similar computational approaches can be applied to identify and target the PPI binding hot spots of other proteins using fragment-based methods.

EXPERIMENTAL SECTION

FTMap Analysis. FTMap analysis was performed using the FTMap computational map server.¹⁸ The constitutively monomeric CXCL12 structure complexed to a nonsulfated N-terminal CXCR4 peptide (L55C/I58C) (PDB code 2N55) was uploaded into the FTMap server and ran according to instructions provided (www.ftmap.bu.edu).¹⁸ The results were visually inspected using PyMol.³⁴ Monomer A from dimeric CXCL12 structure complexed to a sulfated N-terminal CXCR4 peptide (PDB code 2K05) was analyzed following the same procedure.

Virtual Screening. Conformation 1 of the NMR ensemble of the constitutively monomeric form of CXCL12 (L55C/I58C) (PDB code 2N55) complexed to the N-terminal CXCR4 peptide (p40) was used for the rigid docking virtual screening experiment. Residues I4 and I6 from the CXCR4 peptide were used to specify the binding pocket and generate the matching spheres. The spheres were chemically labeled for matching based on ionization states and hydrogen bonding properties of nearby protein residues. Both the fragment (575 530 compounds) and leadlike (4 552 896 compounds) subsets of the ZINC small-molecule database (<http://zinc.docking.org>)²⁶ were docked into the binding pocket using DOCK 3.5.54^{27,28} and sorted based on score. The top 1000 compounds were visually inspected for complementarity, and 12 compounds were chosen for testing based on complementarity and purchase availability. The compounds were 1, 2, 3, *N*-(1*H*-1,3-benzodiazol-2-yl)-2-[(oxolan-2-yl)methyl]amino-1,3-thiazole-4-carboxamide (ZINC C44900490),²⁶ (2*S*)-2-[3-(4-chloro-1*H*-indol-1-yl)propanamido]propanoic acid (ZINC C40310216),²⁶ *N*-{4-[3-(3-cyanophenyl)prop-2-enoyl]phenyl}ethane-1-sulfonamide (ZINC C65565279),²⁶ *N*-{4-[3-(3-hydroxyphenyl)prop-2-enoyl]phenyl}ethane-1-sulfonamide (ZINC C65566657),²⁶ 1-(4-methylbenzoyl)-*N*-(5-sulfanylidene-4,5-dihydro-1*H*-1,2,4-triazol-4-yl)piperidine-3-carboxamide (ZINC C28875196),²⁶ 2-[(2*E*)-3-(4-methoxyphenyl)prop-2-enamido]propanoic acid (ZINC C03887304),²⁶ 3-(2,5-dioxo-1-phenylimidazolidin-4-yl)propanoic acid (ZINC C06529806),²⁶ (2*S*)-2-[[3,4-dimethylphenyl]carbamoyl]amino]propanoic acid (ZINC C00534027),²⁶ and (2*E*)-3-(4-oxo-3,4-dihydroquinazolin-2-yl)prop-2-enoic acid (ZINC C17060315).²⁶ Purchased compounds were tested directly against the protein without further purity analysis. The purity of 1 was subsequently determined to be >95% by ¹H NMR.

For the ensemble virtual screening experiment, the same NMR structure (PDB code 2N55) was used, and each conformation from the ensemble was extracted and treated as a separate rigid docking experiment. Residues I4 and I6 from the CXCR4 peptide were used to specify the binding pocket in each conformation and generate the matching spheres. The spheres were chemically labeled for matching based on ionization states and hydrogen bonding properties of nearby protein residues. Both the fragment and leadlike subsets of the ZINC small-molecule database (<http://zinc.docking.org>) were docked into the binding pocket of each conformation. The results from all experiments were then combined and sorted by score. The overall docking score for each compound against a particular receptor conformation was used in the ranking of the combined results without additional scaling or other modifications (such as considering the internal energy of each receptor conformation). The top 2000 compounds were matched to their respective protein conformation

and then visually inspected for complementarity. While certain conformations in the NMR ensemble (i.e., conformation 18) were not matched with any compounds in the top ranking results, conformation 13 was the receptor template for the majority of the top 2000 compounds (50% for the fragment subset and 77% for the leadlike subset), possibly because of the particular arrangement of positively charged functional groups that increased the positive electrostatic potential in the binding pocket. Twenty-four compounds were finally chosen for testing based on complementarity and purchase availability. The compounds were 4, 5, 6, 7, 8, 2-(2-fluoro-4-sulfamoylphenoxy)acetic acid (ZINC C36948567),²⁶ 2-[(carbamothioylamino)carbamoyl]cyclohexane-1-carboxylic acid (ZINC C19860584),²⁶ *N*-(4-methylpyridin-2-yl)pyrrolidine-1-sulfonamide (ZINC C19909049),²⁶ 4-[(4,6-dioxohexahydropyrimidin-2-ylidene)amino]benzoic acid (ZINC C55134791),²⁶ 2-(2,3-dihydro-1-benzofuran-5-yl)-5-oxo-1-(pyridin-3-ylmethyl)pyrrolidine-3-carboxylic acid (ZINC C69779089),²⁶ 2-(3-[(cyclopropylcarbamoyl)methyl]-sulfanyl)-5-(4-methylphenyl)-4*H*-1,2,4-triazol-4-yl)acetic acid (ZINC C13035420),²⁶ 4-methanesulfonamido-*N*-[(pyridin-3-yl)methyl]-benzamide (ZINC C01063900),²⁶ 2-(2-carbamimidamido-4-oxo-4,5-dihydro-1,3-thiazol-5-yl)-*N*-(3-nitrophenyl)acetamide (ZINC C20028245),²⁶ 2-(2,3-dihydro-1-benzofuran-5-yl)-1-[(1-methylpyrazol-4-yl)methyl]-5-oxopyrrolidine-3-carboxylic acid (ZINC C69492035),²⁶ 2-(2-carbamimidamido-4-oxo-4,5-dihydro-1,3-thiazol-5-yl)-*N*-(4-nitrophenyl)acetamide (ZINC C20064260),²⁶ 3-(4-methoxyphenyl)-3-[(4-nitrophenyl)formamido]propanoic acid (ZINC C00099591),²⁶ 2-(2-carbamimidamido-4-oxo-4,5-dihydro-1,3-thiazol-5-yl)-*N*-(3-methoxyphenyl)acetamide (ZINC C13637710),²⁶ 3-(6-methyl-2,4-dioxo-1,2,3,4-tetrahydropyrimidine-5-sulfonamido)benzoic acid (ZINC C39361382),²⁶ *N*-(3-cyanobenzenesulfonyl)-2-(2,5-dimethylphenyl)acetamide (ZINC C69592086),²⁶ *N*-{[1-[2-(3,6-dioxo-1,2,3,6-tetrahydropyridazin-1-yl)acetyl]piperidin-3-yl]methyl}-methanesulfonamide (ZINC C48353628),²⁶ 4-({3-[(2-methylprop-2-en-1-yl)oxy]phenyl}sulfamoyl)-1*H*-pyrrole-2-carboxylic acid (ZINC C66587652),²⁶ 1-[4-(methanesulfonamido)-3-methylphenyl]-5-oxopyrrolidine-3-carboxylic acid (ZINC C65463160),²⁶ 4-chloro-*N*-[(furan-2-yl)methyl]-3-methanesulfonamidobenzamide (ZINC C61341511),²⁶ and 2-[8-[(4-fluorophenyl)amino]-1,3-dimethyl-2,6-dioxo-2,3,6,7-tetrahydro-1*H*-purin-7-yl]acetic acid (ZINC C12443262).²⁶ Purchased compounds were tested directly against the protein without further purity analysis. The purity of 6 was subsequently determined to be >90% by ¹H NMR. Due to the lack of material, we were unable to further purify this compound and retest it. Because of the overall relatively low affinity of all the ligands against the target binding site including 1 and the features of the binding site, we deem it unlikely that the perturbations were caused by impurities of 6.

Synthesis of Compound 9. Aqueous sodium hydroxide (50% w/v, 180 mg, 4.5 mmol) was added dropwise to a suspension of *N*-(4-chlorobenzoyl)-*N'*-carbamoyl thiourea (10, 500 mg, 2.24 mmol) in water (5 mL) as shown in Scheme S1. The reaction mixture was stirred for 90 min at room temperature. The product was precipitated by addition of glacial acetic acid (0.3 mL). The precipitate was filtered and washed with distilled H₂O (3 × 5 mL) and then suspended in refluxing ethanol (5 mL) and filtered. 400 mg (87%) of the 6-phenyl-4-thioxo-3,4-dihydro-1,3,5-triazin-2(1*H*)-one (triazine, 11) was obtained after the solid was dried under vacuum at 95 °C overnight. ¹H NMR (500 MHz, DMSO-*d*₆) δ 13.19 (s, 1H), 12.78 (s, 1H), 8.09 (d, *J* = 7.7 Hz, 2H), 7.68 (t, *J* = 6.8 Hz, 1H), 7.56 (t, *J* = 7.7 Hz, 2H).

Under an argon atmosphere and at room temperature, potassium carbonate (828 mg, 6 mmol, 3 equiv) was added to a stirred solution of 11 (410 mg, 2 mmol, 1 equiv) and ethyl bromoacetate (835 mg, 5 mmol, 2.5 equiv) in *N,N*-dimethylformamide (10 mL). The resulting mixture was stirred at room temperature for 2 h. Water (20 mL) and ethyl acetate (50 mL) were then added. The aqueous phase was extracted three times with ethyl acetate (3 × 25 mL). The organic phase was washed with brine three times (3 × 100 mL) and then dried with anhydrous sodium sulfate and filtered. The filtrate was concentrated and chromatographed on a silica gel column using hexane and ethyl acetate as the eluent (4:1 to 1:1). Ethyl 2-((4-(2-

ethoxy-2-oxoethoxy)-6-phenyl-1,3,5-triazin-2-yl)thio)acetate (**12**, 560 mg, 74% yield) was obtained as colorless oil. ^1H NMR (500 MHz, CDCl_3) δ 8.45 (d, J = 8.0 Hz, 2H), 7.58 (t, J = 7.3 Hz, 1H), 7.49 (t, J = 7.6 Hz, 2H), 4.99 (s, 2H), 4.26 (dq, J = 21.2, 7.1 Hz, 4H), 3.98 (s, 2H), 1.29 (td, J = 7.1, 2.3 Hz, 6H).

1.0 N aqueous LiOH solution (2 mL) was added to a solution of diester **12** (100 mg, 0.27 mmol) in THF (8 mL) at 0 °C. The resulting mixture was stirred for 1 h at the same temperature and evaporated to dryness. The residue was acidified to pH 2 with aqueous hydrochloric acid (1.0 N) to form a precipitate. The precipitate was purified by preparative HPLC to give 2-((4-(carboxymethoxy)-6-phenyl-1,3,5-triazin-2-yl)thio)acetic acid (**9**, 25 mg, 29.4%) (Scheme S1). HPLC purity: 96.1%. ^1H NMR (500 MHz, $\text{DMSO}-d_6$) δ 8.36 (d, J = 7.4 Hz, 2H), 7.66 (t, J = 7.4 Hz, 1H), 7.57 (t, J = 7.7 Hz, 2H), 4.85 (s, 2H), 3.97 (s, 2H). ^{13}C NMR (100 MHz, $\text{DMSO}-d_6$) δ 183.31 (s), 171.47 (s), 170.23 (s), 169.79 (s), 169.58 (s), 134.65 (s), 133.78 (s), 129.33 (s), 129.14 (s), 64.95 (s), 33.86 (s).

Docking Pose Prediction of Compound 9. Monomer A of conformation 1 in the NMR ensemble of the constitutively dimeric form of CXCL12 (PDB code 2K05) complexed to the triply sulfated N-terminal CXCR4 peptide (p38) was used to predict the binding pose of **9** through docking. The sY12 residue from the CXCR4 peptide was used to specify the binding pocket and generate the matching spheres. The spheres were chemically labeled for matching based on ionization states and hydrogen bonding properties of nearby protein residues. **9** was then docked into the matching spheres, and the best pose was manually chosen based on complementarity.

Purification and Sample Preparation of Recombinant [^{15}N]-CXCL12WT. Human [^{15}N]-CXCL12WT was purified as an N-terminal His6SUMO fusion protein in *E. coli* as described previously.^{17,35} Cells were grown in Terrific Broth and induced with 1 mM isopropyl β -D-1-thiogalactopyranoside (IPTG) prior to being harvested. Cell pellets were lysed, and lysates were clarified by centrifugation (12 000g for 20 min). The resolubilized inclusion body pellets were then loaded onto Ni-NTA resin, and after 1 h proteins were eluted with 6 M guanidinium chloride, 50 mM Na_3PO_4 (pH 7.4), 300 mM NaCl, 500 mM imidazole, 0.2% sodium azide, and 0.1% β -mercaptoethanol. The eluate was pooled and refolded for 4 h before cleavage of the His₆-SUMO fusion tag by Ulp1 protease overnight. The His₆-SUMO fusion tag and chemokine were separated using cation-exchange chromatography (SP Sepharose Fast Flow resin GE Healthcare UK Ltd.), and the eluate was subjected to reverse-phase high-performance liquid chromatography as a final purification. Proteins were frozen, lyophilized, and stored at -20 °C.

Automated Preparation of NMR Samples. Samples for NMR analysis were prepared in groups of 8 (or fewer) compounds using a Pal liquid handling robot configured for loading of 3 mm NMR sample cells (Leap Technologies). For every 8 compounds, 33 mL of 50 μM [^{15}N]-CXCL12WT in a solution of 25 mM deuterated MES (pH = 6.8), 10% (v/v) D_2O , and 0.02% (w/v) NaN_3 was prepared in a glass vial. Upon arrival and initial inventory each compound was brought up to a concentration of 100 mM in deuterated DMSO and stored in amber glass vials. For each compound an amount of 20 μL each of 80 mM and 7.5 mM stocks was prepared in plastic vials from the initial 100 mM stock. Compounds were capped and arranged in a plastic rack (6 $\times$ 9) starting with the 7.5 mM stock, then the 80 mM stock for each compound, and 100 μL of deuterated DMSO. A second rack was filled with 6 glass vials per compound to be used for mixing samples of 25 μM , 50 μM , 150 μM , 400 μM , 800 μM , and 1600 μM compound with 50 μM [^{15}N]-CXCL12WT prior to final transfer of 215 μL to 3 mm NMR tubes. Racks (compounds and NMR tubes) and protein were placed in the appropriate positions on the LEAP Pal, and syringes were washed (acetonitrile and H_2O) and monitored for alignment throughout sample preparation. Upon completion of sample preparation, NMR tubes were capped and samples that exhibited signs of protein precipitation or compound insolubility were noted prior to further analysis.

NMR Spectroscopy. All NMR spectra were acquired on a Bruker DRX 600 MHz spectrometer equipped with a ^1H , ^{15}N , ^{13}C , TXI cryoprobe and SampleJet autosampler at 298 K (25 °C). Experiments

were performed using 50 μM [^{15}N]-CXCL12WT with compound concentrations ranging from 0 μM to 1600 μM prepared as described above and monitored using ^1H - ^{15}N SOFAST-heteronuclear multiple quantum coherence (HMQC) experiments, and chemical shift assignments were acquired from previously published sources.³⁶ Spectra were processed using in-house scripts, and chemical shift tracking was performed using a combination of TitrView and CARA software.²² The combined $^1\text{H}/^{15}\text{N}$ chemical shift perturbations were calculated as $((5\Delta\delta_{\text{H}})^2 + (\Delta\delta_{\text{NH}})^2)^{0.5}$, where $\Delta\delta_{\text{H}}$ and $\Delta\delta_{\text{NH}}$ represent the respective amide proton and nitrogen chemical shifts. Equilibrium dissociation constants (K_d) were determined using nonlinear fitting of the calculated $^1\text{H}/^{15}\text{N}$ chemical shift perturbations as a function of compound concentration to a single-site quadratic equation (protein concentration was held constant at 50 μM).³⁷ For each compound, the residues with the largest chemical shift perturbations were fitted individually and the resulting K_d values and their respective errors were averaged to produce the reported affinity and standard deviation.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jmedchem.5b02042.

Binding energy analysis methods and results, ligand discovery, and docking poses (PDF)

■ AUTHOR INFORMATION

Corresponding Authors

*B.F.V.: phone, (414) 955-8400; fax, (414) 955-6510; e-mail, bvolkman@mcw.edu.

*Y.C.: phone, (813) 974-7809; fax, (813) 974-7357; e-mail, ychen1@health.usf.edu

Author Contributions

[†]E.W.S. and A.M.N. contributed equally.

Author Contributions

The manuscript was written through contributions of all authors. All authors have given approval to the final version of the manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work is funded by Grant R01 GM097381 to B.F.V. and by Grant CA173056 to R.L.

■ ABBREVIATIONS USED

PPI, protein–protein interaction; CXCL12, CXC chemokine ligand 12; CXCR4, CXC chemokine receptor 4

■ REFERENCES

- (1) Nero, T. L.; Morton, C. J.; Holien, J. K.; Wielens, J.; Parker, M. W. Oncogenic protein interfaces: small molecules, big challenges. *Nat. Rev. Cancer* **2014**, *14*, 248–262.
- (2) Zinzalla, G.; Thurston, D. E. Targeting protein-protein interactions for therapeutic intervention: a challenge for the future. *Future Med. Chem.* **2009**, *1*, 65–93.
- (3) Keskin, O.; Gursay, A.; Ma, B.; Nussinov, R. Principles of protein-protein interactions: what are the preferred ways for proteins to interact? *Chem. Rev.* **2008**, *108*, 1225–1244.
- (4) Dundas, J.; Ouyang, Z.; Tseng, J.; Binkowski, A.; Turpaz, Y.; Liang, J. CASTp: computed atlas of surface topography of proteins with structural and topographical mapping of functionally annotated residues. *Nucleic Acids Res.* **2006**, *34*, W116–118.

- (5) Wells, J. A.; McClendon, C. L. Reaching for high-hanging fruit in drug discovery at protein-protein interfaces. *Nature* **2007**, *450*, 1001–1009.
- (6) Chen, Y.; Shoichet, B. K. Molecular docking and ligand specificity in fragment-based inhibitor discovery. *Nat. Chem. Biol.* **2009**, *5*, 358–364.
- (7) Sun, X.; Cheng, G.; Hao, M.; Zheng, J.; Zhou, X.; Zhang, J.; Taichman, R. S.; Pienta, K. J.; Wang, J. CXCL12/CXCR4/CXCR7 chemokine axis and cancer progression. *Cancer Metastasis Rev.* **2010**, *29*, 709–722.
- (8) Zlotnik, A.; Yoshie, O. The chemokine superfamily revisited. *Immunity* **2012**, *36*, 705–716.
- (9) Chatterjee, S.; Behnam Azad, B.; Nimmagadda, S. The intricate role of CXCR4 in cancer. *Adv. Cancer Res.* **2014**, *124*, 31–82.
- (10) Smith, E. W.; Getschman, A. E.; Peterson, F. C.; Ziarek, J. J.; Li, R.; Volkman, B. F.; Chen, Y. Structural analysis of a novel small molecule ligand bound to the CXCL12 chemokine. *J. Med. Chem.* **2014**, *57*, 9693–9699.
- (11) Xu, C.; Zhao, H.; Chen, H.; Yao, Q. CXCR4 in breast cancer: oncogenic role and therapeutic targeting. *Drug Des., Dev. Ther.* **2015**, *9*, 4953–4964.
- (12) Veldkamp, C. T.; Seibert, C.; Peterson, F. C.; De la Cruz, N. B.; Haugner, J. C., 3rd; Basnet, H.; Sakmar, T. P.; Volkman, B. F. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. *Sci. Signaling* **2008**, *1*, ra4.
- (13) Veldkamp, C. T.; Seibert, C.; Peterson, F. C.; Sakmar, T. P.; Volkman, B. F. Recognition of a CXCR4 sulfotyrosine by the chemokine stromal cell-derived factor-1 α (SDF-1 α /CXCL12). *J. Mol. Biol.* **2006**, *359*, 1400–1409.
- (14) Simpson, L. S.; Zhu, J. Z.; Widlanski, T. S.; Stone, M. J. Regulation of chemokine recognition by site-specific tyrosine sulfation of receptor peptides. *Chem. Biol.* **2009**, *16*, 153–161.
- (15) Farzan, M.; Babcock, G. J.; Vasilieva, N.; Wright, P. L.; Kiprilov, E.; Mirzabekov, T.; Choe, H. The role of post-translational modifications of the CXCR4 amino terminus in stromal-derived factor 1 α association and HIV-1 entry. *J. Biol. Chem.* **2002**, *277*, 29484–29489.
- (16) Ziarek, J. J.; Liu, Y.; Smith, E.; Zhang, G.; Peterson, F. C.; Chen, J.; Yu, Y.; Chen, Y.; Volkman, B. F.; Li, R. Fragment-based optimization of small molecule CXCL12 inhibitors for antagonizing the CXCL12/CXCR4 interaction. *Curr. Top. Med. Chem.* **2012**, *12*, 2727–2740.
- (17) Veldkamp, C. T.; Ziarek, J. J.; Peterson, F. C.; Chen, Y.; Volkman, B. F. Targeting SDF-1/CXCL12 with a ligand that prevents activation of CXCR4 through structure-based drug design. *J. Am. Chem. Soc.* **2010**, *132*, 7242–7243.
- (18) Kozakov, D.; Grove, L. E.; Hall, D. R.; Bohnuud, T.; Mottarella, S. E.; Luo, L.; Xia, B.; Beglov, D.; Vajda, S. The FTMap family of web servers for determining and characterizing ligand-binding hot spots of proteins. *Nat. Protoc.* **2015**, *10*, 733–755.
- (19) Ung, P. M.; Ghanakota, P.; Graham, S. E.; Lexa, K. W.; Carlson, H. A. Identifying binding hot spots on protein surfaces by mixed-solvent molecular dynamics: HIV-1 protease as a test case. *Biopolymers* **2016**, *105*, 21–34.
- (20) Brenke, R.; Kozakov, D.; Chuang, G. Y.; Beglov, D.; Hall, D.; Landon, M. R.; Mattos, C.; Vajda, S. Fragment-based identification of druggable “hot spots” of proteins using Fourier domain correlation techniques. *Bioinformatics* **2009**, *25*, 621–627.
- (21) Veldkamp, C. T.; Peterson, F. C.; Pelzek, A. J.; Volkman, B. F. The monomer-dimer equilibrium of stromal cell-derived factor-1 (CXCL 12) is altered by pH, phosphate, sulfate, and heparin. *Protein Sci.* **2005**, *14*, 1071–1081.
- (22) Ziarek, J. J.; Getschman, A. E.; Butler, S. J.; Taleski, D.; Stephens, B.; Kufareva, I.; Handel, T. M.; Payne, R. J.; Volkman, B. F. Sulfopeptide probes of the CXCR4/CXCL12 interface reveal oligomer-specific contacts and chemokine allostery. *ACS Chem. Biol.* **2013**, *8*, 1955–1963.
- (23) Greenidge, P. A.; Kramer, C.; Mozziconacci, J. C.; Wolf, R. M. MM/GBSA binding energy prediction on the PDBbind data set: successes, failures, and directions for further improvement. *J. Chem. Inf. Model.* **2013**, *53*, 201–209.
- (24) Prime, version 4.2; Schrödinger, LLC: New York, 2015.
- (25) Brelot, A.; Heveker, N.; Montes, M.; Alizon, M. Identification of residues of CXCR4 critical for human immunodeficiency virus coreceptor and chemokine receptor activities. *J. Biol. Chem.* **2000**, *275*, 23736–23744.
- (26) Irwin, J. J.; Shoichet, B. K. ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **2005**, *45*, 177–182.
- (27) Irwin, J. J.; Shoichet, B. K.; Mysinger, M. M.; Huang, N.; Colizzi, F.; Wassam, P.; Cao, Y. Automated docking screens: a feasibility study. *J. Med. Chem.* **2009**, *52*, 5712–5720.
- (28) Lorber, D. M.; Shoichet, B. K. Hierarchical docking of databases of multiple ligand conformations. *Curr. Top. Med. Chem.* **2005**, *5*, 739–749.
- (29) Veldkamp, C. T.; Ziarek, J. J.; Su, J.; Basnet, H.; Lennertz, R.; Weiner, J. J.; Peterson, F. C.; Baker, J. E.; Volkman, B. F. Monomeric structure of the cardioprotective chemokine SDF-1/CXCL12. *Protein Sci.* **2009**, *18*, 1359–1369.
- (30) Baryshnikova, O. K.; Sykes, B. D. Backbone dynamics of SDF-1 α determined by NMR: interpretation in the presence of monomer-dimer equilibrium. *Protein Sci.* **2006**, *15*, 2568–2578.
- (31) Murphy, J. W.; Cho, Y.; Sachpatzidis, A.; Fan, C.; Hodsdon, M. E.; Lolis, E. Structural and functional basis of CXCL12 (stromal cell-derived factor-1 α) binding to heparin. *J. Biol. Chem.* **2007**, *282*, 10018–10027.
- (32) Kroemer, R. T. Structure-based drug design: docking and scoring. *Curr. Protein Pept. Sci.* **2007**, *8*, 312–328.
- (33) Damm, K. L.; Carlson, H. A. Exploring experimental sources of multiple protein conformations in structure-based drug design. *J. Am. Chem. Soc.* **2007**, *129*, 8225–8235.
- (34) The PyMOL Molecular Graphics System, version 1.3; Schrödinger, LLC: New York.
- (35) Takekoshi, T.; Ziarek, J. J.; Volkman, B. F.; Hwang, S. T. A locked, dimeric CXCL12 variant effectively inhibits pulmonary metastasis of CXCR4-expressing melanoma cells due to enhanced serum stability. *Mol. Cancer Ther.* **2012**, *11*, 2516–2525.
- (36) Schanda, P.; Kupce, E.; Brutscher, B. SOFAST-HMQC experiments for recording two-dimensional heteronuclear correlation spectra of proteins within a few seconds. *J. Biomol. NMR* **2005**, *33*, 199–211.
- (37) Ziarek, J. J.; Peterson, F. C.; Lytle, B. L.; Volkman, B. F. Binding site identification and structure determination of protein-ligand complexes by NMR a semiautomated approach. *Methods Enzymol.* **2011**, *493*, 241–275.