Structural and Functional Basis of CXCL12 (Stromal Cell-derived Factor-1α) Binding to Heparin*

James W. Murphy‡1, Yoonsang Cho‡, Aristidis Sachpatzidis‡, Chengpeng Fan‡, Michael E. Hodsdon‡, and Elias Lolis†2

From the ‡Department of Pharmacology, Yale University School of Medicine, New Haven, Connecticut 06520-8066 and the †Department of Laboratory Medicine, Yale University School of Medicine, New Haven, Connecticut 06520-8035

Received for publication, September 12, 2006, and in revised form, January 19, 2007 Published, JBC Papers in Press, January 29, 2007, DOI 10.1074/jbc.M608796200

CXCL12 (SDF-1α) and CXCR4 are critical for embryonic development and cellular migration in adults. These proteins are involved in HIV-1 infection, cancer metastasis, and WHIM disease. Sequestration and presentation of CXCL12 to CXCR4 by glycosaminoglycans (GAGs) is proposed to be important for receptor activation. Mutagenesis has identified CXCL12 residues that bind to heparin. However, the molecular details of this interaction have not yet been determined. Here we demonstrate that soluble heparin and heparan sulfate negatively affect CXCL12-mediated in vitro chemotaxis. We also show that a cluster of basic residues in the dimer interface is required for chemotaxis and is a target for inhibition by heparin. We present structural evidence for binding of an unsaturated heparin disaccharide to CXCL12 attained through solution NMR spectroscopy and x-ray crystallography. Increasing concentrations of the disaccharide altered the two-dimensional 1H,15N-HSQC spectra of CXCL12, which identified two clusters of residues. One cluster corresponds to β-strands in the dimer interface. The second includes the amino-terminal loop and the α-helix. In the x-ray structure two unsaturated disaccharides are present. One is in the dimer interface with direct contacts between residues His25, Lys27, and Arg31 of CXCL12 and the heparin disaccharide. The second disaccharide contacts Ala20, Arg21, Asn30, and Lys64. This is the first x-ray structure of a CXCL class chemokine in complex with glycosaminoglycans. Based on the observation of two heparin binding sites, we propose a mechanism in which GAGs bind around CXCL12 dimers as they sequester and present CXCL12 to CXCR4.

Chemokines are a superfamily of 8–11-kDa secreted chemoattractant cytokines, which are modulated by glycosaminoglycans (GAGs),3 activate chemokine receptors, and direct cellular migration (1). The 43 known human chemokines are divided into two major and two minor families. All have a three-dimensional structure composed of a three-stranded β-sheet followed by an α-helix, which is stabilized by conserved cysteine residues forming typically two disulfide bonds (Fig. 1A) (2). The chemokine CXCL12 (CXCL12) is the sole physiological agonist for CXCR4 (3). CXCL12 has been observed as a dimer in several crystal structures (4, 5). However, it has been reported as both a monomer and as a dimer in solution (6, 7). The oligomeric state is modulated by pH, concentration, and the presence of divalent cations and heparin (8). The CXCL12/CXCR4 signaling axis is critically involved in a wide variety of physiological functions. These include margination of neutrophils into the site of infection (9, 10), embryonic development (11), mobilization and directed migration of stem cells (12, 13), and neurological function (14). CXCR4 is involved in HIV-1 infection and HIV-associated dementia (15, 16), is implicated in cancer metastasis (17–19), and genetic mutants are responsible for WHIM disease (20). There is mounting evidence that glycosaminoglycans modulate the activity of chemokines via direct interactions (21). Although the complete functionality of these interactions is not yet completely understood, it is believed that GAGs and proteoglycans sequester chemokines (22). This sequestering increases dimerization and local concentrations, in turn leading to the formation of a chemokine gradient (21–23). In addition, heparan sulfate presents CXCL12 to CXCR4-expressing leukocytes (22). Another important functional aspect of heparin binding to CXCL12 is protection from proteolysis by dipeptidyl peptidase IV (CD26), which removes the NH2-terminal two residues thus inactivating CXCL12 (24).

CXCL12 was originally shown to possess affinity for heparin and heparan sulfate by binding to heparin affinity resin (25). It was proposed that positively charged CXCL12 and other chemokines interact with negatively charged GAGs through nonspecific electrostatic interactions. However, this interaction was subsequently shown to be specific (26). CXCL12 bound independently of its receptor to CXCR4-negative cells through interaction with cell surface glycosaminoglycans. Amara et al. (26) proceeded to use site-directed mutagenesis of CXCL12 to identify a BBXB motif on the first β-strand (Lys24, His25, and Lys27) that was critical for this in vitro interaction. This triple mutant retained its in vitro biological activity of stimulating

‡ This work was supported in part by National Institutes of Health Grant RO1 AI065029 (to E. L.), the Pilot Project Funding from the Yale Cancer Center (to M. E. H.), the Campbell Foundation (to E. L.), and the Patterson Foundation (to M. E. H.). The costs of publication of this article were defrayed in part by advertisement (E. L.). The authors declare no conflict of interest.

1 Supported by a Neuropsychopharmacology Training Program by National Institutes of Health Grant T32 NS007136.

2 To whom correspondence should be addressed: 333 Cedar St., New Haven, CT 06510-8066. Tel.: 203-785-6233; Fax: 203-737-2027; E-mail: elias.lolis@yale.edu.

3 The abbreviations used are: GAG, glycosaminoglycan; HEK-293, human embryonic kidney 293; HSQC, heteronuclear single quantum coherence.
intracellular calcium mobilization. However, the mutated CXCL12 had significantly decreased affinity for heparin (26). Interestingly, mutants of the CC chemokines CCL2 (MCP-1), CCL4 (MIP-1β), and CCL5 (RANTES) devoid of heparin binding retain in vitro chemotactic activity but lose in vivo activity (27).

A model of the CXCL12-heparin complex was based on the mutagenesis studies (28). Sadir et al. (28) calculated an electrostatic potential map of CXCL12 and docked a heparin tetrasaccharide into the positively charged groove. Energy minimization produced a structure in which heparin was modeled to bind quite well in the putative binding site between the subunits of the CXCL12 dimer. In the proposed model, heparin interacted primarily with the basic residues Lys³, Lys₂⁴, His⁵⁵, Lys²⁷, Arg⁴¹, and Lys⁴³ and secondarily with Asn⁴⁶ and Gln⁴⁸. We determined that heparan sulfate as well as high and low molecular weight heparins have a concentration-dependent negative effect on CXCL12 driven chemotaxis of CCR4+ CEM-CRCF cells. Our mutational data revealed a cluster of negatively charged residues involved in both chemotaxis and heparin binding. To elucidate the interactions between the proposed residues and heparin, we used NMR spectroscopy to monitor backbone chemical shift perturbations of CXCL12 during a titration with a heparin disaccharide. By using the fully sulfated, unsaturated disaccharide (heparin disaccharide I-S) derived from a fully sulfated heparin (Fig. 1B), we were able to use the minimum size that contains the basic structural elements required for CXCL12 binding. One chemical difference to be noted between this disaccharide and a typical heparin oligomer is the presence of a double bond between C4 and C5 of the uronic acid, making the disaccharide non-physiological. However, the relative positions of the sulfate moieties should remain unchanged. Therefore we used this disaccharide as a model for CXCL12-heparin interactions. The addition of the disaccharide affected two clusters of amino acids in a dose-dependent manner in NMR spectroscopy experiments. We then used x-ray crystallography to define the exact binding position of the disaccharides. The results from NMR spectroscopy combined with the crystal structure of CXCL12 in complex with the disaccharides define specific molecular interactions between CXCL12 and heparin. Our results not only support previous findings that heparin binds to basic residues in the dimer interface (26, 28) but also identify an additional binding site.

**EXPERIMENTAL PROCEDURES**

**Reagents**—Luria-Bertani medium, ampicillin, isopropyl 1-thio-β-d-galactopyranoside, Triton X-100, and dithiothreitol were purchased from American Bioanalytical (Natick, MA). ¹⁵N-ammonium chloride (98% ¹⁵N), heparin disaccharide I-S (containing an unsaturated hexauronate), oxidized and reduced glutathione, protease inhibitor mixture, ammonium sulfate, Trizma base (Tris base), and deuterium oxide were purchased from Sigma.

**CXCL12 Expression, Folding, and Purification**—The gene encoding human CXCL12 was cloned into the Ndel and XhoI restriction sites of the pET-22b expression vector (Novagen®). The resulting plasmid was transformed into Escherichia coli BL21(DE3). 1.5 liters of Luria-Bertani media containing 100 μg/ml ampicillin were inoculated and grown to A₆₀₀ of 0.6 and then induced with 1.0 mM isopropyl 1-thio-β-d-galactopyranoside. Induced cultures were grown for an additional 4 h at 37 °C and harvested by centrifugation for 10 min at 5000 × g. Cells were resuspended in 1X phosphate-buffered saline (pH 7.4) with 1% Triton X-100, lysed using a French Press and centrifuged for 20 min at 30,000 × g. CXCL12 was found exclusively in inclusion bodies. The inclusion bodies were washed three times in wash buffer A (100 mM Tris-HCl, pH 8.0, 5 mM EDTA, 5 mM dithiothreitol, 2 mM urea, 2% Triton X-100) and once with wash buffer B (100 mM Tris-HCl, pH 8.0, 5 mM EDTA, 5 mM dithiothreitol). Washed inclusion bodies were solubilized in 6 M guanidine HCl and diluted 1:100 into a refolding buffer (100 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.2 mM oxidized glutathione, 1 mM reduced glutathione) and stirred at 4 °C overnight. Precipitated material was removed by filtration. Refolded protein was bound to a SP-Sepharose column and eluted with a NaCl gradient. Fractions containing CXCL12 were pooled and further purified by reverse phase-HPLC. The peak containing CXCL12 was concentrated and lyophilized. CXCL12 was resuspended in sterile ddH₂O (pH 7.0) with 0.1 mM NaN₃ and protease inhibitor mixture prior to use. Protein concentration was determined by direct amino acid analysis at the W. M. Keck facility (Yale University).

¹⁵N-Labeled CXCL12—Isotopically labeled ¹⁵N-CXCL12 was produced similar to native protein with the exception of the substitution of LB medium with M9 minimal media containing 1.0 g/liters of ¹⁵NH₄Cl (98% ¹⁵N).

**Expression and Purification of Mutant CXCL12**—For the chemotaxis assays, wild-type and mutant CXCL12 were expressed in HEK-293 cells to ensure proper folding using a mammalian expression vector containing a secretion signal derived from a fully sulfated heparin (Fig. 1B), we were able to use the minimum size that contains the basic structural elements required for CXCL12 binding. One chemical difference to be noted between this disaccharide and a typical heparin oligomer is the presence of a double bond between C4 and C5 of the uronic acid, making the disaccharide non-physiological. However, the relative positions of the sulfate moieties should remain unchanged. Therefore we used this disaccharide as a model for CXCL12-heparin interactions. The addition of the disaccharide affected two clusters of amino acids in a dose-dependent manner in NMR spectroscopy experiments. We then used x-ray crystallography to define the exact binding position of the disaccharides. The results from NMR spectroscopy combined with the crystal structure of CXCL12 in complex with the disaccharides define specific molecular interactions between CXCL12 and heparin. Our results not only support previous findings that heparin binds to basic residues in the dimer interface (26, 28) but also identify an additional binding site.
with variable concentrations of heparin disaccharide I-S (0–12 mM). Experiments were carried out at 35 °C in a Varian INOVA 600 MHz spectrometer with a 5-mm triple resonance probe equipped with triple-axis (XYZ) pulsed magnetic field gradients. All pulse sequences were taken from the Varian BioPack user library. Spectra were processed and analyzed using the programs nmrPipe (31) and Sparky. Assignment of resonance peaks was done using three-dimensional 15N-TOCSY HSQC and 15N-NOESY HSQC NMR spectra of CXCL12 alone and previously published resonance assignments (33).

Crystallography—CXCL12 was concentrated to 12 mg/ml in water. Protein crystals were grown using the hanging drop method in the previously published condition (2M ammonium sulfate, 0.1 M Tris HCl, pH 8.5) at 18 °C (4). The crystals were soaked into 4-μl drops of heparin disaccharide I-S (16 mM) in 20% PEG-8000, 0.1 M Tris-HCl (pH 8.5) and incubated overnight at 18 °C. The crystals were protected in 20% PEG-8000, 0.1 M Tris-HCl (pH 8.5), and 25% glycerol for 1 min and frozen at −180 °C during data collection.

Data Collection and Processing—X-ray diffraction data were collected on an R-axis IV detector (Rigaku) at the macromolecular crystallography facility at the Yale University School of Medicine. Data were processed using MOSFLM (34) and SCALA (35). Phaser (36) was used for molecular replacement. The CCP4i software suite (37) and the Crystallography and NMR (CNS) (38) program suite were used for refinement and structure validation. Initial coordinates and structural libraries of the heparin disaccharide were created using the Dundee PRODRG server (39).

RESULTS

Mutagenesis of Wild-type CXCL12-Fc Affects Chemotaxis and Inhibition of Chemotaxis by GAGs—Recombinant, refolded CXCL12 produced in E. coli and Fc-tagged CXCL12 produced in HEK-293 cells had equivalent chemotactic activity (data not shown). A series of mutants of CXCL12 were designed to probe the effect of various surface areas in CXCR4 binding and activation (Fig. 1A). The EC50 values and efficacies of these mutants as compared with the wild-type protein are in Table 1. The chemotactic index is a measure of the chemotactic activity of each protein,
defined as the ratio of the number of cells migrated for each protein concentration to the number of cells migrated in response to assay media without chemokine. Efficacy (%) as used in Table 1 is a measure of the maximum activity of each mutant relative to the maximum chemotaxis of the wild type CXCL12.

The RFFESH motif (residues 12–17) was shown by others to have a role in receptor binding and activation through the use of chimeric chemokines with CXCL12 sequences, receptor binding assays, and intracellular calcium release upon agonist activation (6). We further explored this region with the single mutant R12Q and two double mutants, F13A/F14A and E15Q/H17N. Our results show increases in EC50 values ranging from 1.6 to 4.4-fold for the three mutants relative to wild-type CXCL12. Although our results cannot be directly compared with those of Crump et al. (6) due to differences in assays (chemotaxis versus Ca2+ mobilization), we agree that this region is important for CXCL12 function, with no particular residue making a large contribution relative to others. We were surprised, however, by the effect of Arg8 on the EC50. The maximum activity of three mutants R8Q, R12Q, and R8Q/R12Q is similar to wild type, but these mutants have correspondingly 60, 4, and 120 times higher EC50 values than the wild type. The positive charge of Arg8 is critical for activating CXCR4-mediated chemotaxis. A series of other mutants were designed to address the role of positively charged surface areas identified originally in the CXCL12 crystal structure (4). One of these, a quintuple mutant H25N/K27Q/R41Q/R47Q/Q48N that removed the charges showed a 35% reduction in efficacy with respect to wild type but a similar EC50 suggesting that the loss of a large surface area of positive potential has a moderate effect on receptor activation. Alternatively, because this mutant showed a clear resistance to heparin inhibition (see below), its reduced activity may be due to a lack of interaction with native GAGs essential for a maximum chemotactic effect. A triple mutant comprising residues T31V/A33G/A35G showed a small reduction of EC50 and efficacy suggesting a possible role of these residues in CXCL12-CXCR4 interaction.

Effect of GAGs on CXCL12-mediated Chemotaxis—As shown in Fig. 2, A and B, both heparan sulfate and high and low molecular weight heparin caused a clear reduction of the chemotactic activity of CXCL12 in all CXCL12 and GAG concentrations tested. The effect was significant at 10 or 100 μg/ml of GAG for 10 nM (p ≤ 0.0009) and 100 nM CXCL12 (p ≤ 0.0009). In B, post hoc t tests give p ≤ 0.0003 for the CXCL12 activity at 10 and 100 nM versus background. Both of these activities are significantly inhibited by all concentrations of heparin tested (10 nM CXCL12 is inhibited by 10 and 100 μg/ml of low molecular weight heparin with p ≤ 0.0005 and by 10 and 100 μg/ml of high molecular weight heparin with p = 0.0002 in both cases). 100 nM CXCL12 is inhibited by 10 and 100 μg/ml of low molecular weight heparin with p ≤ 0.0003 and by 10 and 100 μg/ml of high molecular weight heparin with p ≤ 0.0002.

### Table 1: Effect of CXCL12 mutations on chemotaxis

| Mutant                       | EC50a (nM) | Efficacyb (%) |
|------------------------------|------------|---------------|
| Wild-type                    | 7.9        | 100           |
| R8Q/R12Q                     | 960        | 110           |
| R8Q                          | 490        | 100           |
| R12Q                         | 35         | 92            |
| F13A/F14A                    | 21         | 87            |
| E15Q/H17N                    | 13         | 110           |
| T31G/A33G/A35G               | 18         | 81            |
| H25N/K27Q/R41Q/R47Q/Q48N     | 4.1        | 66            |

*M* EC50 was defined as the concentration of CXCL12 that gave 50% of the maximum response as determined by a concentration-response curve for each mutant.

*b* Efficacy was calculated as the ratio of the maximum response of each mutant to the maximum response of wild-type CXCL12 multiplied by 100.

**FIGURE 2. Effects of GAGs on CXCL12-mediated chemotaxis.** Chemotactic index is calculated as the number of cells migrating in the presence of CXCL12 with or without GAGs at a number of cells migrating in the absence of CXCL12. A, chemotaxis of heparan sulfate (0.1 to 100 μg/ml) preincubated with CXCL12 (10 and 100 nM). B, chemotaxis of low and high molecular weight heparin (10 and 100 μg/ml) were preincubated with CXCL12 (10 and 100 nM). Analysis of variance for data gives significance at *p* = 0.0001. In A, post hoc tests give *p* ≤ 0.0008 for the CXCL12 activity at 10 and 100 nM versus background. These activities are inhibited by all concentrations of heparan sulfate tested and the effect is significant for 10 nM (*p* ≤ 0.009) and 100 nM CXCL12 (*p* ≤ 0.009). In B, post hoc tests give *p* ≤ 0.0003 for the CXCL12 activity at 10 and 100 nM versus background. Both of these activities are significantly inhibited by all concentrations of heparin tested (10 nM CXCL12 is inhibited by 10 and 100 μg/ml of low molecular weight heparin with *p* ≤ 0.0005 and by 10 and 100 μg/ml of high molecular weight heparin with *p* = 0.0002 in both cases). 100 nM CXCL12 is inhibited by 10 and 100 μg/ml of low molecular weight heparin with *p* ≤ 0.0003 and by 10 and 100 μg/ml of high molecular weight heparin with *p* ≤ 0.0002.
both 10 and 100 nM CXCL12. These data are in agreement with FIGURE 3.

Structure of the CXCL12-Heparin Complex

Effects of GAGs on mutant CXCL12-mediated chemotaxis. Six mutants of CXCL12 were compared with wild-type. Samples with heparin were treated with 100 μg/ml of low molecular weight heparin. All proteins (wild-type and mutants) were tested at an optimal concentration for chemotaxis of CCRF-CEM cells (100 nM) with the exception of R8Q (1.0 μM). The activities of all proteins are significantly higher versus background (p < 0.001). The activities of all proteins (with the exception of the quintuple mutant; p = 0.02 for both with and without heparin) are significantly inhibited by treatment with 100 μg/ml of low molecular weight heparin (p = 0.005).

mutant showed a clear resistance to inhibition by heparin indicating that the GAG-CXCL12 interaction has been significantly disrupted. Four positive charges are neutralized in this mutant and the data support the suggestion that some or all of these residues, His25, Lys27, Arg41, Arg47 (and Gln48), are somehow involved in GAG binding. This is in agreement with previous mutagenesis studies (26). The most unexpected observation from this series of experiments came from the effect of heparin on the R8Q and R12Q single mutants. In both of these cases, the mutants showed a remarkable hypersensitivity to heparin with more than 80% of the activity lost upon binding a GAG versus only about 50% for the wild type. These data suggest that the absence of these arginines allows the formation of a GAG-CXCL12 complex with a conformation that increases inhibition.

Effect of Titration of Heparin Disaccharide I-S into a Solution of CXCL12 Monitored by NMR Spectroscopy—To investigate the interactions between CXCL12 and heparin, we used NMR spectroscopy to monitor chemical shifts affected by the addition of an unsaturated heparin disaccharide. 1H-15N-HSQC spectra were collected from a solution of 2.0 mM CXCL12 and increasing concentrations of disaccharide (0–12.0 mM). Fig. 4A shows the 1H-15N-HSQC spectrum of 2.0 mM apo-15N-CXCL12 with five overlaid spectra of the CXCL12:heparin disaccharide mixtures. The ratios of CXCL12:heparin disaccharide were 1:0.5, 1:1, 1:2, 1:4, and 1:6. For each resonance peak observed in the apo-CXCL12 spectra, a corresponding peak occurs in each spectrum from the titration, indicating that the overall structure of CXCL12 is maintained. Peaks due to the five arginine side chain δ1H-15N nuclei were aliased into the spectra at ~121 ppm. Only the side chain of Arg41 could be definitively assigned. The spectra from the titration overlay well with the apo spectrum. However, there were a variety of peaks of interest showing a concentration-dependent change in resonance positions. Fig. 4B shows a comparison of the absolute change in NMR chemical shift for each residue at a ratio of 2:1 disaccharide:CXCL12 compared with the apo spectrum. Chemical shift perturbations occur in the vicinity of residues identified by mutagenesis as required for heparin binding. This includes a region around His25 and Lys27 in the first β-strand and a somewhat smaller region around Arg41 in the second β-strand. In the three-dimensional structure of CXCL12, these two regions form a localized cluster along the dimer interface.

Increasing the concentration of disaccharide beyond a 1:2 ratio revealed chemical shift perturbations in two additional regions. One includes the series of His17, Val18, and Ala19 and the other includes residues Glu63 through Asn67 in the COOH-terminal α-helix. These two groups of residues cluster in the three-dimensional structure in a region comparable with the binding site of heparin in CXCL8 as determined by NMR spectroscopy (43). Hence, the NMR spectroscopic study indicates that there are two sites of disaccharide interaction, one corresponding to residues identified by mutation in the dimer interface and a second that corresponds to the site of interaction of heparin with CXCL8 (interleukin-8) (43).

Crystal Structure of the CXCL12-Heparin Disaccharide Complex—Attempts to grow crystals of CXCL12 in complex with various lengths of heparin oligosaccharides were unsuccessful. We therefore used the soaking method. Crystals of native CXCL12...
A Structure of the CXCL12-Heparin Complex

![Figure 4](image)

**FIGURE 4.** NMR spectroscopic studies of CXCL12 and heparin disaccharide interaction. A, NMR chemical shift changes induced by titration of heparin disaccharide I-S. 1H-15N-HSQC spectra of 2.0 mM 15N-CXCL12 with 1.0, 2.0, 4.0, 8.0, and 12.0 mM heparin disaccharide I-S (orange, yellow, light green, green, and blue, respectively) are overlaid on the spectrum of 2.0 mM apo 15N-CXCL12 (red). Resonance peaks with the largest NMR chemical shift changes are labeled. Peak labels indicate peaks for the apo spectrum. B, absolute NMR chemical shift change of each residue for the disaccharide titration. Absolute NMR chemical shift change for each ratio are calculated as $\frac{1}{2}(\frac{\text{Hppmapo} - \text{Hppmbound}}{\text{Hppmapo} - \text{Hppmbound} \times 10})$. 15N-CXCL12:heparin disaccharide I-S ratios as compared with 1:0 are colored red (1:0.5), orange (1:1), yellow (1:2), green (1:4), and blue (1:6). Changes in NMR chemical shifts for proline residues are reported as zero as they lack an amide proton. Regions of secondary structure are indicated below with block arrows representing β regions and zigzags representing helical regions. Two sets of residues (A/A’ and B/B’) that each form a cluster in the three-dimensional structure that are affected by titration of the disaccharide are indicated by brackets.

The disaccharide ligand is visible in the dimer interface nestled within a cluster of basic residues whose side chains are oriented into the pocket (Fig. 6B). This ligand is positioned where it had been predicted from mutagenesis studies (26) and is in an orientation that would allow the extension of a polysaccharide chain. CXCL12 directly contacts the sulfate and hydroxyl moieties of the heparin disaccharide. The disaccharide forms hydrogen bonds with His25 of chain B, Arg21 of both protein chains, and Lys27 of chain A (Table 3) (Fig. 6B).

A second disaccharide is bound to the exterior of the dimer and hydrogen bonded to Arg50, Ala21, and Lys26 of chain A and Asp30 of chain B. The location of this second ligand is stabilized by interactions with Val3 and Ser6 of chain A from another asymmetric unit, and is consistent with the second site identified by NMR spectroscopy. Table 3 lists the hydrogen bonds observed between the two heparin disaccharides and CXCL12.

**DISCUSSION**

Heparin and Heparan Sulfate Negatively Affect in Vitro Chemotaxis—The inhibitory effect of soluble GAGs on CXCL12-induced in vitro chemotaxis is similar to that previously observed in vivo.

Analysis of the structure revealed that electron density was not observed for the amino-terminal residues before Ser4 of chain B. Asn44 falls into the disallowed region of the Ramachan-
Structure of the CXCL12-Heparin Complex

for other chemokines. This can be explained by a sequestration mechanism that reduces chemokine availability and prevents their binding to their natural interaction partners (membrane-bound receptors and GAGs). The effect of the higher molecular weight GAGs is similar to inhibition of chemokine-mediated inflammation in vivo where soluble heparin is injected into the bloodstream (45). We have tested only heparan sulfate and high and low molecular weight heparin, and we assume that the other types of soluble GAGs will induce similar responses. The heparin disaccharide I-S did not exhibit any inhibitory activity (data not shown), presumably due to a low affinity interaction.

NMR Spectroscopy Identified Two Potential Heparin Binding Sites—Positively charged residues clustered in the dimer interface were reported as being required for CXCL12-heparin interaction (26). Our results indicate that residues surrounding His25, Lys27, and Arg41 compose one site of interaction in solution as their NMR chemical shifts are perturbed by the disaccharide. A second heparin binding site was observed in solution. This previously unidentified CXCL12 site corresponds to the heparin binding site of CXCL8 (43) and partially corresponds with the heparin binding sites of CXCL4 (platelet factor-4) (46), CXCL1 (growth related protein-α) (46), and CXCL10 (IP-10) (47). Furthermore, there appears to be a difference in the affinities of the two sites for the heparin disaccharide. Chemical shift changes are observed at lower CXCL12:heparin disaccharide I-S ratios for the site in the dimer interface than the region in the amino-terminal loop (residues 17–19).

Two Unsaturated Heparin Disaccharide Binding Sites Are Observed in the Crystal Structure—In the crystal structure one disaccharide is bound at the dimer interface and a second is bound to the amino-terminal loop and the α-helix. Much longer glycosaminoglycan chains are present in vivo. This disaccharide contains all the structural elements that are repeated in a longer oligomer with the exception of having an unsaturated hexuronate. With structural data now available we can make further conclusions about the role of amino acids previously thought to interact with heparin. Of the three basic residues in the BBXB sequence located in the first β-strand, His25 and Lys27 were found both to be hydrogen bonded to the disaccharide in the crystal structure. Computer modeling predicted Arg41 to bind heparin (28) and we found that Arg41 from both chains does indeed interact with the disaccharide. Lys43, on the second β-strand, was predicted by modeling to be important, yet mutagenesis of this single residue had no effect on binding (28). Although Lys43 does not interact with the disaccharide in our crystal structure, this does not rule out such interactions with longer glycosaminoglycans in vivo. A comparison of the bound unsaturated disaccharide to the previously published molecular model shows that

### TABLE 2
Data collection and refinement statistics

| Parameter                              | Value       |
|----------------------------------------|-------------|
| Space group                            | P2₁,2₁,2₁   |
| α, β, γ (°)                            | 90, 90, 90  |
| Wavelength (Å)                         | 1.5418      |
| Resolution range (Å)                   | 2.5-2.07    |
| lmax                                   | 7.2 (2.0)   |
| Completeness                           | 99.9% (100%)|
| Rfactor, free                          | 0.057 (0.367)|
| Redundancy                             | 6.3 (6.5)   |

- Height resolution shell is shown in parentheses.
- $R_{factor} = 100 \times \sum_{h} \left| I_{h} - \langle I_{h} \rangle \right| / \sum_{h} \left| I_{h} \right|$, where $I_{h}$ is the observed intensity, and $\langle I_{h} \rangle$ is the average intensity of multiple observations of symmetry-related reflections.
- $R_{free} = \sum_{h} \left| F_{h}^{obs} - F_{h}^{calc} \right| / \sum_{h} \left| F_{h}^{obs} \right|$, where $F_{h}^{obs}$ and $F_{h}^{calc}$ were calculated using the working and test reflection sets, respectively. 5% of the entire reflection was randomly taken as a test set.

FIGURE 5. Crystal structure of the CXCL12:heparin disaccharide I-S complex. A, schematic representation of CXCL12 with chain A in cyan and chain B in red. Two heparin disaccharide I-S ligand molecules are shown as sticks, one within the dimer interface and one on the outer portion of monomer A. Image on the right is a 90° rotation around the y axis. B, the disaccharide bound structure of CXCL12 (blue) is overlaid with the unbound form (PDB code 1A15) (4) (cyan). The amino termini are removed for clarity. Both are shown as schematics representing the secondary structure elements and the two heparin disaccharide I-S molecules are shown as sticks. The image on the right is a 90° rotation around the y axis.
The disaccharide is oriented slightly differently. This difference in orientation is most likely due to the fact that the disaccharide is small compared with the long glycosaminoglycan chains that were previously modeled (28). Additionally, as the disaccharide is unsaturated, it could bind in a slightly different conformation than a natural ligand.

The second disaccharide binds to the amino-terminal loop and \( \alpha \)-helix and forms hydrogen bonds to residues Ala21, Arg20, Asn30, and Lys64. This disaccharide is oriented such that continuation of the polysaccharide is possible without steric conflict. These results indicate that there are multiple sites of interaction on chemokines leading to an avidity effect, resulting in high affinity binding with long chain heparins.

It is interesting to note that there is a difference in the positions of chain A residues 1–6 and 29–36 as compared with the native structure. These two regions are linked by a disulfide bond. Whether heparin binding induces a relay of CXCL12 conformational changes on CXCR4, as suggested by the comparative analysis, to induce signaling by the NH\(_2\) terminus remains to be determined.

Fig. 7 summarizes the structural data from these experiments. The crystal structure of the complex is displayed and regions that are most affected by disaccharide titration monitored by NMR spectroscopy are highlighted. This three-dimensional view gives information about the relative locations of the two disaccharide binding sites in CXCL12.

### Comparison of the Heparin Binding Sites of CXCL12 to Those of Other Chemokines

The GAG binding sites of chemokines have been identified by mutagenesis (47), NMR (43, 46), and one crystal structure (48). For CXC chemokines, the heparin binding site of CXCL8 was identified as a cluster of residues with changes in chemical shift upon addition of heparin disaccharide I-S (43). This site is similar to the second site of CXCL12 and involves both CXCL8 residues 18–23 in the loop.

---

**TABLE 3**

| Donor, Protein Residue (Chain) | Atom | Ligand | Acceptor Atom | Distance Å |
|-------------------------------|------|--------|---------------|------------|
| His\(^{25}\) (B)             | ND1  | HIS-1  | O61(1)        | 2.42       |
| Lys\(^{27}\) (A)             | NZ   | HIS-1  | O-S2(1)       | 2.86       |
| Lys\(^{27}\) (A)             | NZ   | HIS-1  | O5(1)         | 3.04       |
| Arg\(^{41}\) (B)             | NH1  | HIS-1  | O-S2(2)       | 2.52       |
| Arg\(^{41}\) (A)             | NH1  | HIS-1  | O-S2(2)       | 2.52       |
| Lys\(^{51}\) (A)             | N    | HIS-2  | O61(1)        | 2.86       |
| Ala\(^{51}\) (A)             | N    | HIS-2  | O-S2(1)       | 2.89       |
| Asn\(^{51}\) (B)             | ND2  | HIS-2  | O-S2(2)       | 2.72       |
| Lys\(^{64}\) (A)             | NZ   | HIS-2  | O-S6(2)       | 3.18       |

* HIS-1 and HIS-2 are heparin disaccharides I-S molecules that are bound to the dimer interface and the interleukin-8-like site, respectively.
preceding the first β-strand and residues from the COOH-terminal α-helix. The interactions between a heparin dodecasaccharide and CXCL4 and CXCL1 were also determined using NMR spectroscopy (46). Like other chemokines, basic residues are involved and, as in CXCL8, chemical shift changes for residues at the loop immediately before the first β-strand implicate this region in contacts with the heparin oligomer. The heparin binding site of CXCL10 was determined using mutagenesis, binding experiments, and in vitro cellular assays (47). Our conclusion from this analysis is that CXCL12 has a unique heparin binding site in the dimer interface involving strands one and two in addition to a more common heparin binding site involving the loop region preceding the first β-strand and residues from the α-helix. Our interpretation is that there may be sites of varying affinities for disaccharides, each of which provides an avidity effect leading to high affinity interactions with a long chain GAG that binds CXCL12, inducing dimerization, and sequestering CXCL12.

The heparin binding sites of numerous CC chemokines have also been identified (27, 48–50). CC chemokines, with few exceptions, dimerize with a different topology and it would be interesting to compare the GAG binding sites. CCL5 (RANTES (regulated on activation normal T cell expressed and secreted)) was co-crystallized with heparin disaccharides (I-S and III-S) (48). In the CCL5-disaccharide complexes, a single disaccharide is bound at different locations to three crystallographically symmetric dimers of CCL5. Longer chain heparin molecules may have interactions with residues at all these locations. None of these three sites of interaction are similar to the CXCL12 sites of interaction we have identified, with the exception of His23 in CCL5 (corresponding to His25 in CXCL12). The CC and CXC chemokines appear to be different not only in their dimeric structures, but also in the location of the primary, tertiary, and quaternary sites involved with interactions with GAGs. The vast varieties of substituted glycosaminoglycans present in vivo may be responsible for selectivity of receptor binding for chemokines in different families and also within families.

One Binding Site of Heparin Partially Overlaps the Binding Site of the NH2-terminal Peptide of CXCR4—

The two-stage model of chemokine-chemokine receptor activation dictates an initial binding event followed by a conformational change leading to activation (51–53). Clore and coworkers (33) completed a study in which a peptide equivalent to residues 1–27 of CXCR4 was added to CXCL12 in solution. It was found that the pocket formed by the β-strands in the dimer interface, which includes residues Lys24, His25, Ala40, Arg41, Gln48, and Tyr61 of CXCL12, interacts with the CXCR4 peptide (33). This region is part of the heparin binding site. This supports a proposed mechanism for chemokine presentation by proteoglycans whereby the chemokine GPCR displaces proteoglycans bound to the chemokine agonist (22). Initially, secreted CXCL12 would be bound to GAGs attached to proteoglycans via multiple sites on CXCL12. This binding would increase dimerization of CXCL12 (8), protect CXCL12 from inactivation by CD26 (24), and increase the localized concentration of CXCL12. GAGs would then be displaced from CXCL12 by the amino-terminal region of CXCR4. CXCR4 could then be activated by the amino terminus of CXCL12 (6). Interestingly, the binding of GAGs to CXCL10 (47) and CCL2 (49) are also partially overlapping with their respective receptor binding sites. This suggests the chemokines from different families use similar mechanisms with different heparin binding sites.

Therapeutic intervention of diseases involving chemokines and their GPCRs traditionally targets the receptors with antagonists or otherwise modulates the signaling at the level of the
receptor. Due to complications arising from multiple chemokines activating a single receptor and vice versa, targeting the initial glycosaminoglycan-chemokine interaction with small GAG-derived molecules may prove to be an alternative avenue for disease therapy. For example, therapeutic pentasaccharide heparin is often used clinically, functioning as an anticoagulant as well as an anti-inflammatory agent (54). Likewise, small heparin-based molecules may function as a potential new class of anti-inflammatory compounds that do not stimulate anticoagulant activity. Our structure of the disaccharide-bound dimer can serve as the starting point for the design of such small molecule inhibitors of CXCL12-GAG interaction.

Acknowledgments—We thank Gregg Crichton and Paul Pepin for expert assistance with the crystallography, Sharon Mella for assistance with tissue culture, Demetrios Braddock for assistance with protein purification, and Camille Keeler for assistance with the NMR spectroscopy. We thank Dr. Lortat-Jacob for providing atomic coordinates of the modeled CXCL12-heparin tetradecasaccharide complex.

REFERENCES

1. Fernandez, E. J., and Lolis, E. (2002) Annu. Rev. Pharmacol. Toxicol. 42, 469–499
2. Lapidot, T., and Petit, I. (2002) Exp. Hematol. 30, 973–981
3. Bleul, C. C., Farzan, M., Choe, H., Parolin, C., Clark-Lewis, I., Sodroski, J., and Springer, T. A. (1996) Nature 382, 829–833
4. Dealwis, C., Fernandez, E. J., Thompson, D. A., Simon, R. J., Siani, M. A., and Lolis, E. (1998) Proc. Natl. Acad. Sci. U. S. A. 95, 6941–6946
5. Ohnishi, Y., Senda, T., Nandhagopalan, N., Sugimoto, K., Shioda, T., Nagal, Y., and Mitsu, Y. (2000) J. Interferon Cytokine Res. 20, 691–700
6. Veldkamp, C. T., Peterson, F. C., Pelzek, A. J., and Volkman, B. F. (2005) Protein Sci. 14, 1071–1081
7. White, F. A., Bhanagoo, S. K., and Miller, R. J. (2005) Nat. Rev. Drug Discov. 4, 834–844
8. Nagasawa, T., Tachibana, K., and Kishimoto, T. (1998) Semin. Immunol. 10, 179–185
9. Kucia, M., Reca, R., Miekus, K., Wanzeck, J., Wojakowski, W., Janowska-Wieczorek, A., Ratajczak, J., and Ratajczak, M. Z. (2004) Stem Cells (Durham) 23, 879–894
10. Kucia, M., Rajczak, J., and Ratajczak, M. Z. (2005) Biol. Cell 97, 133–146
11. Bajetto, A., Bonavia, R., Barbero, S., Florio, T., and Schettini, G. (2001) Front. Neuroendocrinol. 22, 147–184
12. Zhou, N., Luo, Z., Luo, J., Liu, D., Hall, J. W., Pomerantz, R. J., and Huang, Z. (2001) J. Biol. Chem. 276, 48286–48233
13. Kucia, M., Ratajczak, J., and Ratajczak, M. Z. (2005) Cell 127, 5776–5785
14. Mukhailov, D., Young, H. C., Linhardt, R. J., and Mayo, K. H. (1999) J. Biol. Chem. 274, 25317–25329
15. Campaillana, G. S., Lee, E. M., Sun, J., and Luster, A. D. (2003) J. Biol. Chem. 278, 17066–17074
16. Shaw, J. P., Johnson, Z., Borlat, F., Zwahlen, C., Kungl, A., Roulin, K., Harrenga, A., Wells, T. N., and Proudfoot, A. E. (2004) Nature 432, 896–907
17. Sun, J., Johnson, Z., Borlat, F., Zwahlen, C., Kungl, A., Roulin, K., Harrenga, A., Wells, T. N., and Proudfoot, A. E. (2004) Structure 12, 2081–2093
18. Kucia, M., Ratajczak, J., and Ratajczak, M. Z. (2005) J. Biol. Chem. 279, 22294–22305
19. Xanthou, G., Williams, T. J., and Pease, J. E. (2003) Eur. J. Immunol. 33, 2927–2936
20. Cai, S. H., Tan, Y., Ren, X. D., Li, X. H., Cai, S. X., and Du, J. (2004) Acta Pharmacol. Sin. 25, 152–160
21. Wells, T. N., Power, C. A., Lusti-Narasimhan, M., Hoogewerf, A. J., Cooke, R. M., Chung, C. W., Peitsch, M. C., and Proudfoot, A. E. (1996) J. Leukoc. Biol. 259, 53–60
22. Tyrrell, D. J., Kilfeather, S., and Page, C. P. (1995) Trends Pharmacol. Sci. 16, 198–204