Published in final edited form as:

Biochemistry. 2016 October 11; 55(40): 5746–5753. doi:10.1021/acs.biochem.6b00304.

Crystallographic structure of truncated CCL21 and putative sulfotyrosine-binding site

Emmanuel W. Smith[†,‡], **Eric M. Lewandowski**[†,‡], **Natasha A. Moussouras**[§], **Kyle G. Kroeck**[†], **Brian F. Volkman**[§], **Christopher T. Veldkamp**[‖], and **Yu Chen**[†,*]

[†]Department of Molecular Medicine, University of South Florida, Tampa, Florida 33612, United States

[§]Department of Biochemistry, Medical College of Wisconsin, Milwaukee, Wisconsin 53226, United States

[‖]Department of Chemistry, University of Wisconsin-Whitewater, Whitewater, Wisconsin 53190, United States

Abstract

CCL21 chemokine binds the G protein-coupled receptor CCR7, aiding not only in immune response but also in cancer metastasis. Compared with other chemokines, CCL21 has a unique extended unstructured C-terminus that is truncated in some naturally occurring variants. We have solved the X-ray crystallographic structure of a truncated CCL21 (residues 1–79) lacking the extended C-terminus, and identified, via 2D NMR, a putative sulfotyrosine-binding site that may recognize such post-translationally modified tyrosine residues on the receptor. Compared to the previously solved NMR structure of full-length CCL21, the crystal structure presents new druggable binding hot spots resulting from an alternative N-loop conformation. In addition, whereas the previous NMR structure did not provide any structural information after residue 70, the C-terminus of the truncated CCL21, ordered up to Ala77 in our crystal structure, is placed near the N-loop and sulfotyrosine-binding site, indicating that the extended C-terminus of the full length CCL21 can interact with this important region for receptor binding. These observations suggest a potential origin for the auto-inhibition of CCL21 activity that was recently described. The new crystal structure and binding hot spot analysis have important implications for the function of the CCL21 C-terminus and drug discovery.

Graphical Abstract

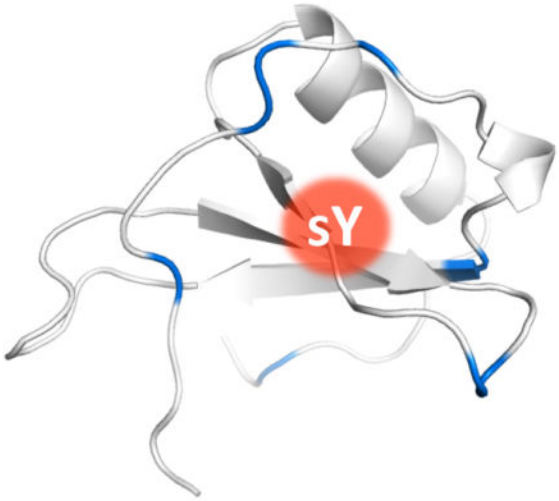
[*]Corresponding Author: Yu Chen. (813)-974-7809; Fax (813)-974-7357. ychen1@health.usf.edu.

[‡]**Author Contributions:**

These authors contributed equally.

NOTES

The coordinates have been deposited in the Protein Data Bank with access code 5EKI.



Chemokines are small chemotactic cytokines that bind and activate cognate G protein-coupled receptors (GPCRs), mediating the immune response through recruitment of receptor-expressing leukocytes [(1)]. They are classified into four subfamilies based on the composition of cysteine residues in the N-terminus [(1)]. CCL21 belongs to the CC subfamily and is predominately expressed in the lymph nodes, appendix, and spleen [(2)]. It binds and activates the CCR7 receptor, thus recruiting CCR7-expressing cells to the lymph nodes and other places throughout the body [(2)]. The CCL21-CCR7 signaling axis is also implicated in metastasis of cancer cells to the lymph nodes, and is associated with a poor clinical prognosis [(3)].

CCL19 also binds the CCR7 receptor, and although both CCL21 and CCL19 bind to CCR7 with equal potency, their binding mediates different responses [(4)]. This is due to an extended 40 amino acid C-terminus on CCL21 that is important in binding heparan sulfate, a glycosaminoglycan (GAG) that aids in adhesive migration (Fig. 1) [(5)]. This extended C-terminus contains an additional third disulfide bond, between C80 and C99, as opposed to the two conserved disulfide bonds typically found in other chemokines [(6)]. Truncation of the CCL21 extended C-terminus occurs in nature and alters the migratory response of dendritic cells by establishing soluble gradients, similarly to CCL19 that lacks the extended anchoring C-terminus (Fig. 1) [(5)]. A recent study has also revealed that the CCL21 C-terminus may induce the chemokine to adopt an inactive conformation, thus causing auto-inhibition [(7)]. The NMR structure of full-length CCL21 showed that while CCL21 adopts the typical chemokine motif [(6)], the extended C-terminus containing the third disulfide bond is unstructured after residue 70, offering no information on how the C-terminus may interact with the rest of the protein.

Chemokines bind and activate their receptors in a two-step/two-site process; the first step is recognition of the chemokine core domain by the flexible extracellular N-terminus of the receptor, and the second step is activation of the receptor through docking of the chemokine's flexible N-terminus to the receptor [(8)]. Tyrosine O-sulfation of the N-terminal tail of chemokine receptors has been shown to increase affinity of the receptor N-terminus

for the chemokine, thus mediating the first step of the chemokine-receptor association [9]. In the case of CXCL12-CXCR4, for example, the first recognition step is facilitated by sulfation of Tyr21 on the receptor N-terminal domain, which recognizes and binds a sulfotyrosine-binding site on the chemokine surface [10]. In other studies, sulfotyrosine was used to probe putative sulfotyrosine-binding sites on the surface of CXCL12, CCL5, CX3CL1, and XCL1 by 2D NMR [9]. However, no such studies have been performed with CCL21 and it is therefore not clear whether similar sulfotyrosine binding sites also exist in this unique chemokine.

Here, we describe the first crystallographic structure of a truncated version of CCL21 (residues 1–79) lacking the extended C-terminus, including evidence of a putative sulfotyrosine binding site revealed by NMR, and computational analysis of small molecule binding hotspots around that site. These results provide novel insights into the function of CCL21 and valuable information to guide future inhibitor discovery efforts targeting the CCL21 signaling pathway.

MATERIALS AND METHODS

Purification

The pQE30 (Qiagen) vector containing the His_6-SMT3 (SUMO) tagged truncated CCL21 sequence (residues 1–79) was transformed into BL21 (pREP4) cells. The cells were incubated overnight in LB media containing 50μg/ml ampicillin and 25μg/ml kanamycin. The overnight culture was diluted in 1L LB media plus 50μg/ml ampicillin and 25μg/ml kanamycin, and incubated at 37°C until the OD_{600} reached 0.6–0.8. Expression was induced with 0.5mM IPTG and incubation continued at 37°C for an additional 4 hours, after which the cells were pelleted by centrifugation (5,000 × g) for 10 min. The cell pellet was resuspended in 20ml of Buffer A (50mM sodium phosphate pH 7.4, 300mM sodium chloride, and 10mM imidazole) with 0.1% β-mercaptoethanol. The cells were then lysed via sonication and the inclusion body pellet, containing the truncated CCL21, was collected by centrifugation (15,000 × g) for 30 minutes. The inclusion body was resuspended in 20ml of Buffer B (6M guanidine HCl, 50mM sodium phosphate pH 8.0, 300mM sodium chloride, and 10mM imidazole) with 0.1% β-mercaptoethanol, and incubated at 37°C for 1 hour with rocking until the inclusion body fragments changed from a brown to clear color. Then the sample was centrifuged (40,000 × g) for 40 minutes and the supernatant was collected. The protein was bound to a HisTrap affinity column and eluted through a gradient of Buffer C (6M guanidine HCl, 100mM sodium acetate pH 4.5, 300mM sodium chloride, and 10mM imidazole) and 0.1% β-mercaptoethanol. The protein was then refolded by diluting the sample through dropwise addition to 12X the volume of refolding buffer (100mM Tris pH 8.0, 10mM reduced cysteine, and 0.5mM oxidized cysteine) at ~45–50 drops/min. After allowing refolding to take place overnight, the solution containing the refolded protein was concentrated to 25ml, and was further diluted 2-fold by adding 25mL of 100mM Tris buffer pH 8.0 in order to lower the concentration of guanidine HCl from ~500mM to ~ 250mM. 500μg of ULP1 protease was added and the sample was allowed to incubate at 30°C overnight. The precipitate was filtered out and then the sample was concentrated and dialyzed in Buffer D (100mM Tris pH 8.0, 300mM sodium chloride, and 10mM imidazole)

and passed through the HisTrap affinity column to remove the ULP1, the His_6-Smt3 tag, and any uncleaved protein. The cleaved protein was collected in the flow through, and was further concentrated and polished through gel filtration purification using the HiLoad 16/60 Superdex 75 column and eluted in water. The protein purity was tested by SDS-PAGE and found to be > 95% pure. The final protein concentration was 10mg/ml. CCL21WT was purified as previously described [6].

Crystallization

Hampton Crystal Screens 1 and 2, and the Qiagen JCSG I-IV and AmSO4 suites were screened for crystallization conditions at 20°C using the Phoenix Liquid Handling robot, which nano-dispensed 0.2μl of protein solution into 0.2μl of crystallization solution. Diffraction quality crystals appeared in a solution containing 3.0M ammonium sulfate and 1% MPD. Crystals were further optimized by adjusting the crystallization condition to 2.7M ammonium sulfate and 0.5% MPD, and the drop composition to 1μl of crystallization solution and 2μl of protein solution. The crystal used for data collection was soaked in cryo-protectant containing crystallization solution plus 25% glycerol.

X-ray crystal structure determination

Diffraction data were collected at the SER-CAT 22-BM beamline, at the Advanced Photon Source (APS) at Argonne National Laboratory. Processing was done with HKL2000 and programs from the CCP4 suite [11] were used for structure solution and refinement. Phaser Cell Analysis [12] was used to determine the asymmetric unit contained six copies of the protein. Balbes [13] was used for initial molecular replacement, with the sequence of the previously solved CCL21 NMR structure (PBD ID: 2L4N) serving as the search template. Balbes produced a solution with four copies of the protein in the asymmetric unit, and R_{Work} and R_{Free} values of 45% and 48%, respectively. This solution was then refined using Refmac5 [14] and the R-values fell to ~40%. Parrot [15] was used for density modification on the Refmac solution, and then Buccaneer [16] was run using the data from Parrot and the refined model from Balbes.

A trimmed version of the protein structure was made by removing all flexible side chains on the protein surface. The trimmed structure was then imported into Molrep [17], along with the output from Buccaneer, in an attempt to search for one of the two remaining copies of the protein in the asymmetric unit. Molrep successfully returned a solution with five copies, and the structure was subsequently refined with Refmac5, resulting in R_{Work} and R_{Free} values of 38% and 41%, respectively. Molrep was run again to search for the sixth copy of the protein, with the refined five-copy structure serving as the input. Molrep successfully returned a solution with all six copies in the asymmetric unit, and upon refinement with Refmac5, the R_{Work} and R_{Free} values improved to 35% and 38%, respectively.

After a six-copy solution was obtained, Coot [18] was used to further refine and build the completed structure. Several residues were missing from both the N and C-terminal ends of all six copies of the protein, and the flexible region from residues 31–34 needed to be rebuilt in all copies. After adding the terminal residues, rebuilding the flexible region, and further

refining the structure with Phenix [19] and Refmac5, the final R_{Work} and R_{Free} values were calculated at 20.5% and 24.8%, respectively.

2D ^{1}H-^{15}N HSQC Spectroscopy with sulfotyrosine titration

Uniformly ^{15}N labeled CCL21 (residues 1–79) used for the NMR sulfotyrosine titrations was purchased from Protein Foundry, LLC. The lyophilized ^{15}N CCL21 1–79 was dissolved to a concentration of 250μM in 25mM deuterated MES at pH 5.94 with 10% D_2O and 0.02% NaN_3. Sulfotyrosine (sY) used in these titrations was at a concentration of 500mM in 50mM deuterated acetate at pH 5.0 with 10% D_2O and 0.02% NaN_3. NMR spectroscopic data were collected at the NMR facility at the Medical College of Wisconsin on a Bruker Avance 600MHz spectrometer equipped with ^{1}H/^{13}C/^{15}N cryoprobe. ^{1}H-^{15}N heteronuclear single quantum coherence spectra were used to monitor a CCL21 1–79 sample titrated with incremental additions of sulfotyrosine at 0mM, 1mM, 5mM, 10mM, 15mM, 20mM, 30mM, 40mM, 60mM, 80mM, and 100mM concentrations. Using chemical shift assignments from the solved NMR structure[6],peaks were tracked using CARA. Amide ^{1}H-^{15}N chemical shift perturbations were computed as $[(5\Delta\δ_{NH})^2 + (\Delta\delta_N)^2]^{1/2}$, where $\Delta\delta_{NH}$ and $\Delta\delta_N$ are the total changes in backbone amide ^{1}H and ^{15}N chemical shifts in ppm, respectively, from 0mM to 100mM sY. Dose dependent chemical shift perturbations for CCL21 1–79 residues C9, Q14, and K16 upon titration with sY were fit to the following equation that accounts for ligand depletion:

$$\Delta\delta=\Delta\delta_{\max}\times\frac{(K_D+[\mathrm{CCL21}]+x)-\sqrt{(K_D+[\mathrm{CCL21}]+x)^2-4[\mathrm{CCL21}]x}}{2[\mathrm{CCL21}]}$$

where $\Delta\delta$ is the chemical shift perturbation, $\Delta\delta$ max is the maximum chemical shift perturbation at 100% bound CCL21, K_d is the CCL21 sY peptide dissociation constant, and x is the sY concentration. Using pro Fit 6.2 and the above equation, the K_d values and their respective errors were calculated and averaged to produce the reported affinity and standard deviation. Amino acids with the highest chemical shift perturbations were mapped onto the structure of CCL21 using PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.).

FTMap analysis

FTMap docks, *in silico*, 16 small organic molecules (ethane, ethanol, isopropanol, *tert*-butanol, acetonitrile, methylamine, *N*,*N*-dimethylformamide, dimethyl ether, benzaldehyde, benzene, cyclohexane, phenol, acetamide, acetone, acetaldehyde, and urea) as probes onto the surface of a protein. FTMap analysis was performed using the FTMap computational map server according to instructions provided (www.ftmap.bu.edu) [21]. The results were visually inspected with PyMOL (The PyMOL Molecular Graphics System, Version 1.8 Schrödinger, LLC.).

RESULTS

Crystal structure of a truncated CCL21

The truncated CCL21 was crystallized in the $P2_1$ space group with unit cell lengths of 65.75 Å, 58.24 Å, 66.05 Å, and γ = 119.9° (Table 1). The crystal structure was determined to 1.9 Å resolution, and provides new structural details important for our understanding of CCL21 intra- and inter-molecular interactions. There are six monomers contained within the asymmetric unit with a pseudo six-fold symmetry (Fig. 2A). Main chain residues of each monomer were clearly resolved in the electron density with the exception of residues in the N-terminus and a few in the C-terminus. Residues 6–77 were resolved for chain A, 6–76 for chain B and D, 7–76 for chain C and F, and 5–77 for chain E. Superimposition of each monomer in the asymmetric unit shows that overall variation of the backbone is negligible, with the most noticeable variation localized to a flexible 30's loop (residues Glu29-Pro37) and to a lesser extent in a 40's loop (residues Arg44-Glu50), the N-terminus, and the C-terminus (Fig. 2B). Variation among the monomers is also observed in the conformations of various flexible side chains, particularly those of lysine and arginine. When all monomers are superimposed to monomer A, they have an average RMSD of 0.188Å, aligning an average of 408 atoms. Focusing on monomer A demonstrates how truncated CCL21 adopts the typical chemokine fold (Fig. 2D). The flexible N-terminus and N-loop (Cys9-Val22), which are connected to the 30's loop and β3-strand by two adjacent disulfide bonds (C8–C34 and C9–C52), lead into to a three-stranded, antiparallel β-sheet (β1-β3) that is separated by loops. The 30's loop separates the β1-strand and β2-strand, and the 40's loop separates the β2-strand and β3-strand. An α-helix follows the β3-strand and ends in a C-terminal tail (Fig. 2D).

Oligomerization of certain chemokines can influence binding and downstream signaling effects [22]. Most CC chemokines that dimerize do so through an anti-parallel β-sheet formed between the N-termini of both monomers, while other CC chemokines are obligate monomers. The CC-like dimerization observed in crystallographic structures of other chemokines (Fig. 2C) was not established between any of the monomers in the asymmetric unit of the truncated CCL21 crystal structure (Fig. 2A), nor was it established between monomers from adjacent asymmetric units in the crystal-packing interface. Additionally, the proline in the N-terminus that has been suggested to promote CC chemokine dimerization is absent in CCL21 [23]. Convincingly, Love *et al.* showed through sedimentation equilibrium centrifugation that full-length CCL21 is monomeric [6]. This, along with the lack of any canonical CC chemokine dimerization motif in our structure, suggests CCL21 1–79 is a monomer, and that the hexameric arrangement shown in Fig. 2A is simply the result of crystal packing in the asymmetric unit.

NMR analysis of sulfotyrosine binding by CCL21

Similar to previous studies on CXCL12 and other chemokines, we performed a 2D ^{1}H-^{15}N HSQC assay on truncated CCL21 by titrating sulfotyrosine, and the resultant ^{1}H-^{15}N chemical shift perturbations were calculated and mapped onto the truncated CCL21 X-ray structure. Whereas some of the smaller perturbations display hallmarks of nonspecific binding (smaller perturbations and linear dose-response) or may be the effect of weak

specific binding at more than one site, all of the largest shift perturbations display evidence of specific binding (i.e. saturable binding curves). These large sulfotyrosine-induced shifts surround the pocket that was previously identified in other chemokines as a conserved sulfotyrosine-binding site (Fig. 3A–E) [9, 10], and the binding affinity of sulfotyrosine was estimated to be 38.9 mM (K_d). Furthermore, this site overlaps with a putative CCR7-binding site previously identified via ^{1}H-^{15}N HSQC studies on full-length CCL21 with a nonsulfated CCR7 N-terminal peptide (Fig. 4A–B) [6]. During the binding of the CCR7 receptor peptide into full-length CCL21, the largest shifts were localized to the N-loop, 40's loop, and β3 strand. As expected for weaker binding of a much smaller ligand, the sulfotyrosine-induced shifts localize to a similar but smaller region of the N-loop and 40's loop. The receptor peptide binding study also speculates that basic residues, such as K16 in the N-loop, might be involved in sulfotyrosine binding, as there is no basic residue in the β3 strand that would function similarly to R47 of CXCL12. The large shift perturbations exhibited by K16 and the neighboring Q14 lend support to that hypothesis. Overall, our results support previous findings on the conservation of a particular sulfotyrosine-binding site, and suggest that sulfation of the CCR7 N-terminus *in vivo* at one, or both, of its two tyrosine residues (Tyr8 and Tyr17) may contribute to site one recognition of CCL21, i.e., the binding of the receptor N-terminus to the chemokine core domain.

Computational analysis of CCL21 binding hot spots

Our recent study on CXCL12 has demonstrated that computational methods, such as FTMap [20, 21], can be utilized to identify binding hot spots for engineering small molecule ligands [24]. These programs dock a diverse set of small organic probes, such as ethanol and benzene, onto the protein surface. Areas where multiple probes cluster may indicate druggability. In addition to the sulfotyrosine recognition site detected by NMR, we performed a computational search for additional binding surfaces of CCL21 suitable for novel inhibitor discovery. FTMap [21] analysis was utilized to predict hot spots on the CCL21 surface and examine their intersection with the putative sulfotyrosine-binding site. This experiment was performed on both the full-length CCL21 NMR structure (conformation 1 from NMR ensemble in PDB ID: 2L4N) [6] and the truncated CCL21 X-ray structure. Comparison of the full-length CCL21 NMR structure [6] with the truncated crystal structure revealed a noticeable conformational difference in the N-loop (Cys9-Lys16) (Fig. 4A), a region implicated in sulfotyrosine-binding and binding to the CCR7 N-terminus (Fig. 4B) [6]. The N-loop in every NMR conformation is tilted internally towards the β3-strand, while in truncated CCL21 the N-loop is tilted externally, widening the space between the N-loop and the β3-strand, thus exposing several residues such as Cys9, Cys52, and Leu40 for potential receptor binding (Fig. 4A). Accompanying this change, differences are also observed in the 30's loop, which is linked to the N-loop through a disulfide bond. The structural similarities and variations between the NMR and crystal structures are reflected in the results of the computational analysis. Comparison of probe clusters docked onto the surface of both structures identified differential hot spots in the sulfotyrosine-binding site area between the two structures (Fig. 4C–D). Probe clusters were also ranked based on the number of probes in each cluster, represented by Cluster Strength (CS, CS01-5). The results suggest that the putative sulfotyrosine-binding site contains the majority of the druggable surface in both CCL21 structures, while presenting more binding hot spots in truncated

CCL21 than in full length CCL21. In the truncated CCL21 structure, there is a noticeable cluster (CS01 = 33 probes) present near the N-loop that is absent in the full-length structure, due to the widening of that site in the truncated structure (Fig. 4C–D). Hence, the structure of truncated CCL21 may represent a configuration more suitable for ligand and receptor binding, and may also provide a better template for future virtual screening experiments.

DISCUSSION

Influence of the C-terminus on CCL21 structure and function

Previous studies have shown that the unique C-terminus of CCL21 can affect its function through multiple mechanisms. The C-terminus interacts with the extracellular matrix associated GAGs and truncation of the C-terminus eliminates or reduces affinity for GAGs [25, 26]. This truncation is physiologically relevant as a serine protease does catalyze a similar truncation *in vivo*, leading to a soluble versus a stationary gradient [5, 27]. Dendritic cell responses to CCL21 are also regulated through another posttranslational modification, polysialylation of the CCR7 receptor [7]. In the absence of polysialic acid (PSA) posttranslational modifications in dendritic cells, CCL21 does not function as an effective chemoattractant; PSA's binding to CCL21's C-terminus greatly enhances CCR7 activation and chemotaxis [7, 28, 29]. One might think the enhancement occurs because PSA found on migrating cells displaces CCL21 bound GAGs that could be inhibitory [30], but this enhancement is observed during *in vitro* assays where extracellular matrix GAGs are not present, indicating full length CCL21 is in an auto-inhibited state relieved by PSA [7]. In addition, PSA is not functioning as a coreceptor that enhances CCL21's affinity for its receptor, as truncated CCL21 can activate CCR7 either lacking or containing the PSA modification [7]. These recently published observations indicate that full length CCL21 is auto-inhibited by its C-terminal domain.

The truncated CCL21 crystal structure presented here sheds new light on this auto-inhibition mechanism. The structural variations observed between the full-length and truncated CCL21 structures in the putative sulfotyrosine-binding region, particularly surrounding the N-loop, may reflect different conformational changes arising from interactions with the C-terminus (Fig. 4B). Specifically, the presence of the C-terminus appears to reduce the binding surfaces of a region previously shown to be interacting with CCR7 (Fig. 4B), as reflected by the decrease in druggability in the FTMap analysis (Fig. 4C–D). In addition, in comparison to the NMR structure where the C-terminus is completely disordered after residue 70, the new X-ray crystal model provides structural information for residues 71–77. This short segment, with three prolines, orients the C-terminus towards the N-loop, with a distance of ~18 Å between the Cα atoms of Ala77 and Cys9, further suggesting a possibility that the extended positively charged C-terminus in full-length CCL21 could potentially influence the conformation of the N-loop and adjacent residues, and affect function by interacting with sites in the receptor-binding area. However, it is certainly possible that the C-terminus can also affect CCL21 function by interacting directly with the receptor, and more details of its auto-inhibitory role will await future studies.

Whereas we cannot exclude the possibility that crystal packing contributes to the different N-loop conformation, we deem this unlikely. Comparing the truncated CCL21 crystal

structure with a previously determined crystal structure of CCL18, a related chemokine, also reveals that the two proteins adopt very similar conformations for the majority of the protein, including the N-loop region (Fig. 5A). Similar to CCL21, CCL18 was shown to function as a monomer, unlike many other CC-type chemokines. In addition, CCL18 lacks the extended C-terminus, thus resembling the truncated CCL21 more than the full-length CCL21 [23]. CCL18 was crystallized in the $P2_1$ space group with two monomers per asymmetric unit, and establishes different interactions at the crystal packing interface compared with truncated CCL21. Therefore, it is highly improbable that the similar configurations of these two related proteins would have both resulted from crystal packing. The N-loop conformation of the truncated CCL21 also resembles what is observed in the NMR structure of CCL19, which, like CCL18, does not have the extended C-terminus (Fig. 1, 5B) [31]. Taken together, the structural comparison suggests that the CCL21 conformation in the new crystal structure is most likely functionally relevant. This is further supported by the observation that several residues exposed by the new protein conformation, such as Cys9 and Lys75, are implicated in sulfotyrosine binding in our 2D NMR experiment (Fig. 3B). Previous NMR analysis also demonstrated that multiple residues, including those lining the putative sulfotyrosine-binding site, such as Tyr12, Gln14, Lys16, and Lys45, are perturbed by the presence of the C-terminus [7], suggesting that the C-terminus may indeed influence the conformation of this region.

Small molecule inhibitor discovery targeting CCL21

Due to the implication of the chemokine-receptor axis in various diseases, including cancer, efforts have been employed in studying chemokines and their receptors for the development of therapies [32]. The focus has mostly been placed on targeting the receptors, since chemokines were previously deemed “undruggable” due to their small size and lack of deep hydrophobic pockets. However, our group has successfully used a structure-based approach to identify inhibitors that bind to a sulfotyrosine-binding site on the surface of CXCL12, and interrupt its association with CXCR4, proving that such sites are indeed druggable [33, 34]. We have also presented the first complex crystallographic structure of CXCL12 bound by one of those inhibitors at the site normally occupied by sulfated tyrosine 21 (sY21) of the CXCR4 receptor [8]. Most recently, we have used FTMap analysis to identify an additional binding hot spot in CXCL12, and successfully identified novel small molecule ligands for this new site [24]. The crystal structure of CCL21 provides important information for similar structure-based inhibitor discovery efforts targeting this chemokine in the future. First, compared with NMR structures, crystal structures normally provide more accurate structural details, such as side chain conformations, making them better templates for computational modeling in most cases. Second, when compared with the previous full length CCL21 NMR structure, the truncated CCL21 crystal structure captures a different configuration that appears to present more druggable binding surfaces in a region important for its function. These features make the new CCL21 crystal structure a valuable tool for virtual screening, and such inhibitor discovery efforts are ongoing in our laboratory.

In conclusion, the X-ray structure of truncated CCL21 identified an alternative N-loop conformation not observed in the previously solved full-length CCL21 NMR structure [6]. The N-loop is part of a putative sulfotyrosine-binding site detected by NMR. The alternative

conformation makes the site appear more suitable for ligand binding and potentially more druggable, which will allow us to apply structure-based approaches to identify novel inhibitors against CCL21. New structural information for residues 71–76 further suggests potential interactions between the extended CCL21 C-terminus and the binding surfaces adjacent to the N-loop. These results offer important insights into the function of the unique C-terminus of CCL21.

Acknowledgments

The authors would like to thank Dr. Ruslan Sanishvili and all the presenters at the 2015 CCP4/APS School at Argonne National Laboratory for their assistance and insightful discussions during structure solution and refinement.

FUNDING SOURCES

Brian F. Volkman is a managing member of Protein Foundry, LLC. This work was supported by NIH grants R15 CA159202 (CTV), R01 GM097381 (BFV), and F30 CA210587 (NAM).

ABBREVIATIONS

CCL21	CC chemokine ligand 21
CCR7	CC chemokine receptor 7
CCL19	CC chemokine ligand 19
GPCRs	G protein-coupled receptors
GAG	glycosaminoglycan
sY	sulfotyrosine
PSA	polysialic acid
HSQC	heteronuclear single quantum coherence
NMR	nuclear magnetic resonance

References

1. Zlotnik A, Yoshie O. The chemokine superfamily revisited. Immunity. 2012; 36:705–716. [PubMed: 22633458]
2. Nagira M, Imai T, Hieshima K, Kusuda J, Ridanpaa M, Takagi S, Nishimura M, Kakizaki M, Nomiyama H, Yoshie O. Molecular cloning of a novel human CC chemokine secondary lymphoid-tissue chemokine that is a potent chemoattractant for lymphocytes and mapped to chromosome 9p13. The Journal of biological chemistry. 1997; 272:19518–19524. [PubMed: 9235955]
3. Ben-Baruch A. Organ selectivity in metastasis: regulation by chemokines and their receptors. Clin Exp Metastasis. 2008; 25:345–356. [PubMed: 17891505]
4. Zidar DA, Violin JD, Whalen EJ, Lefkowitz RJ. Selective engagement of G protein coupled receptor kinases (GRKs) encodes distinct functions of biased ligands. Proceedings of the National Academy of Sciences of the United States of America. 2009; 106:9649–9654. [PubMed: 19497875]
5. Schumann K, Lammermann T, Bruckner M, Legler DF, Polleux J, Spatz JP, Schuler G, Forster R, Lutz MB, Sorokin L, Sixt M. Immobilized chemokine fields and soluble chemokine gradients cooperatively shape migration patterns of dendritic cells. Immunity. 2010; 32:703–713. [PubMed: 20471289]

6. Love M, Sandberg JL, Ziarek JJ, Gerarden KP, Rode RR, Jensen DR, McCaslin DR, Peterson FC, Veldkamp CT. Solution structure of CCL21 and identification of a putative CCR7 binding site. Biochemistry. 2012; 51:733–735. [PubMed: 22221265]
7. Kiermaier E, Moussion C, Veldkamp CT, Gerardy-Schahn R, de Vries I, Williams LG, Chaffee GR, Phillips AJ, Freiberger F, Imre R, Taleski D, Payne RJ, Braun A, Forster R, Mechtler K, Muhlenhoff M, Volkman BF, Sixt M. Polysialylation controls dendritic cell trafficking by regulating chemokine recognition. Science (New York, NY). 2016; 351:186–190.
8. Smith EW, Liu Y, Getschman AE, Peterson FC, Ziarek JJ, Li R, Volkman BF, Chen Y. Structural analysis of a novel small molecule ligand bound to the CXCL12 chemokine. Journal of medicinal chemistry. 2014; 57:9693–9699. [PubMed: 25356720]
9. Ziarek JJ, Heroux MS, Veldkamp CT, Peterson FC, Volkman BF. Sulfotyrosine recognition as marker for druggable sites in the extracellular space. Int J Mol Sci. 2011; 12:3740–3756. [PubMed: 21747703]
10. Veldkamp CT, Seibert C, Peterson FC, De la Cruz NB, Haugner JC 3rd, Basnet H, Sakmar TP, Volkman BF. Structural basis of CXCR4 sulfotyrosine recognition by the chemokine SDF-1/CXCL12. Sci Signal. 2008; 1:ra4. [PubMed: 18799424]
11. Collaborative Computational Project N. The CCP4 suite: programs for protein crystallography. Acta crystallographica. Section D, Biological crystallography. 1994; 50:760–763. [PubMed: 15299374]
12. McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ. Phaser crystallographic software. J Appl Crystallogr. 2007; 40:658–674. [PubMed: 19461840]
13. Long F, Vagin AA, Young P, Murshudov GN. BALBES: a molecular-replacement pipeline. Acta crystallographica. Section D, Biological crystallography. 2008; 64:125–132. [PubMed: 18094476]
14. Murshudov GN, Vagin AA, Dodson EJ. Refinement of macromolecular structures by the maximum-likelihood method. Acta crystallographica. Section D, Biological crystallography. 1997; 53:240–255. [PubMed: 15299926]
15. Zhang KY, Cowtan K, Main P. Combining constraints for electron-density modification. Methods in enzymology. 1997; 277:53–64. [PubMed: 18488305]
16. Cowtan K. The Buccaneer software for automated model building. 1. Tracing protein chains. Acta crystallographica. Section D, Biological crystallography. 2006; 62:1002–1011. [PubMed: 16929101]
17. Vagin A, Teplyakov A. Molecular replacement with MOLREP. Acta crystallographica. Section D, Biological crystallography. 2010; 66:22–25. [PubMed: 20057045]
18. Emsley P, Cowtan K. Coot: model-building tools for molecular graphics. Acta crystallographica. Section D, Biological crystallography. 2004; 60:2126–2132. [PubMed: 15572765]
19. Adams PD, Afonine PV, Bunkoczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung LW, Kapral GJ, Grosse-Kunstleve RW, McCoy AJ, Moriarty NW, Oeffner R, Read RJ, Richardson DC, Richardson JS, Terwilliger TC, Zwart PH. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta crystallographica. Section D, Biological crystallography. 2010; 66:213–221. [PubMed: 20124702]
20. Brenke R, Kozakov D, Chuang GY, Beglov D, Hall D, Landon MR, Mattos C, Vajda S. Fragment-based identification of druggable 'hot spots' of proteins using Fourier domain correlation techniques. Bioinformatics. 2009; 25:621–627. [PubMed: 19176554]
21. Kozakov D, Grove LE, Hall DR, Bohnuud T, Mottarella SE, 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. [PubMed: 25855957]
22. Fernandez EJ, Lolis E. Structure, function, and inhibition of chemokines. Annu Rev Pharmacol Toxicol. 2002; 42:469–499. [PubMed: 11807180]
23. Liang WG, Ren M, Zhao F, Tang WJ. Structures of human CCL18, CCL3, and CCL4 reveal molecular determinants for quaternary structures and sensitivity to insulin-degrading enzyme. Journal of molecular biology. 2015; 427:1345–1358. [PubMed: 25636406]
24. Smith EW, Nevins AM, Qiao Z, Liu Y, Getschman AE, Vankayala SL, Kemp MT, Peterson FC, Li R, Volkman BF, Chen Y. Structure-Based Identification of Novel Ligands Targeting Multiple Sites

within a Chemokine-G-Protein-Coupled-Receptor Interface. Journal of medicinal chemistry. 2016; 59:4342–4351. [PubMed: 27058821]

25. Dyer DP, Salanga CL, Johns SC, Valdambrini E, Fuster MM, Milner CM, Day AJ, Handel TM. The anti-inflammatory protein TSG-6 regulates chemokine function by inhibiting chemokine:glycosaminoglycan interactions. The Journal of biological chemistry. 2016
26. Hirose J, Kawashima H, Swope Willis M, Springer TA, Hasegawa H, Yoshie O, Miyasaka M. Chondroitin sulfate B exerts its inhibitory effect on secondary lymphoid tissue chemokine (SLC) by binding to the C-terminus of SLC. Biochimica et Biophysica Acta (BBA) - General Subjects. 2002; 1571:219–224. [PubMed: 12090936]
27. Weber M, Hauschild R, Schwarz J, Moussion C, de Vries I, Legler DF, Luther SA, Bollenbach T, Sixt M. Interstitial dendritic cell guidance by haptotactic chemokine gradients. Science (New York, NY). 2013; 339:328–332.
28. Rey-Gallardo A, Delgado-Martin C, Gerardy-Schahn R, Rodriguez-Fernandez JL, Vega MA. Polysialic acid is required for neuropilin-2a/b-mediated control of CCL21-driven chemotaxis of mature dendritic cells and for their migration in vivo. Glycobiology. 2011; 21:655–662. [PubMed: 21199821]
29. Rey-Gallardo A, Escribano C, Delgado-Martin C, Rodriguez-Fernandez JL, Gerardy-Schahn R, Rutishauser U, Corbi AL, Vega MA. Polysialylated neuropilin-2 enhances human dendritic cell migration through the basic C-terminal region of CCL21. Glycobiology. 2010; 20:1139–1146. [PubMed: 20488940]
30. Hirose J, Kawashima H, Swope Willis M, Springer TA, Hasegawa H, Yoshie O, Miyasaka M. Chondroitin sulfate B exerts its inhibitory effect on secondary lymphoid tissue chemokine (SLC) by binding to the C-terminus of SLC. Biochim Biophys Acta. 2002; 1571:219–224. [PubMed: 12090936]
31. Veldkamp CT, Kiermaier E, Gabel-Eissens SJ, Gillitzer ML, Lippner DR, DiSilvio FA, Mueller CJ, Wantuch PL, Chaffee GR, Famiglietti MW, Zgoba DM, Bailey AA, Bah Y, Engebretson SJ, Graupner DR, Lackner ER, LaRosa VD, Medeiros T, Olson ML, Phillips AJ, Pyles H, Richard AM, Schoeller SJ, Touzeau B, Williams LG, Sixt M, Peterson FC. Solution Structure of CCL19 and Identification of Overlapping CCR7 and PSGL-1 Binding Sites. Biochemistry. 2015; 54:4163–4166. [PubMed: 26115234]
32. Raman D, Sobolik-Delmaire T, Richmond A. Chemokines in health and disease. Exp Cell Res. 2011; 317:575–589. [PubMed: 21223965]
33. Veldkamp CT, Ziarek JJ, Peterson FC, Chen Y, Volkman BF. Targeting SDF-1/CXCL12 with a ligand that prevents activation of CXCR4 through structure-based drug design. Journal of the American Chemical Society. 2010; 132:7242–7243. [PubMed: 20459090]
34. Ziarek JJ, Liu Y, Smith E, Zhang G, Peterson FC, Chen J, Yu Y, Chen Y, Volkman BF, Li R. Fragment-based optimization of small molecule CXCL12 inhibitors for antagonizing the CXCL12/CXCR4 interaction. Curr Top Med Chem. 2012; 12:2727–2740. [PubMed: 23368099]
35. Barinka C, Prahl A, Lubkowski J. Structure of human monocyte chemoattractant protein 4 (MCP-4/CCL13). Acta crystallographica. Section D, Biological crystallography. 2008; 64:273–278. [PubMed: 18323622]

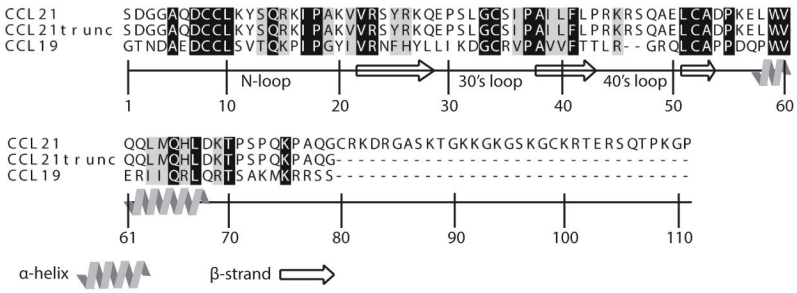


Figure 1. Sequence alignment of CCL21 and CCL19

Identical and similar residues are shaded in black and grey respectively. A truncated CCL21 (CCL21-trunc), lacking the extended C-terminus, was used in the current study.

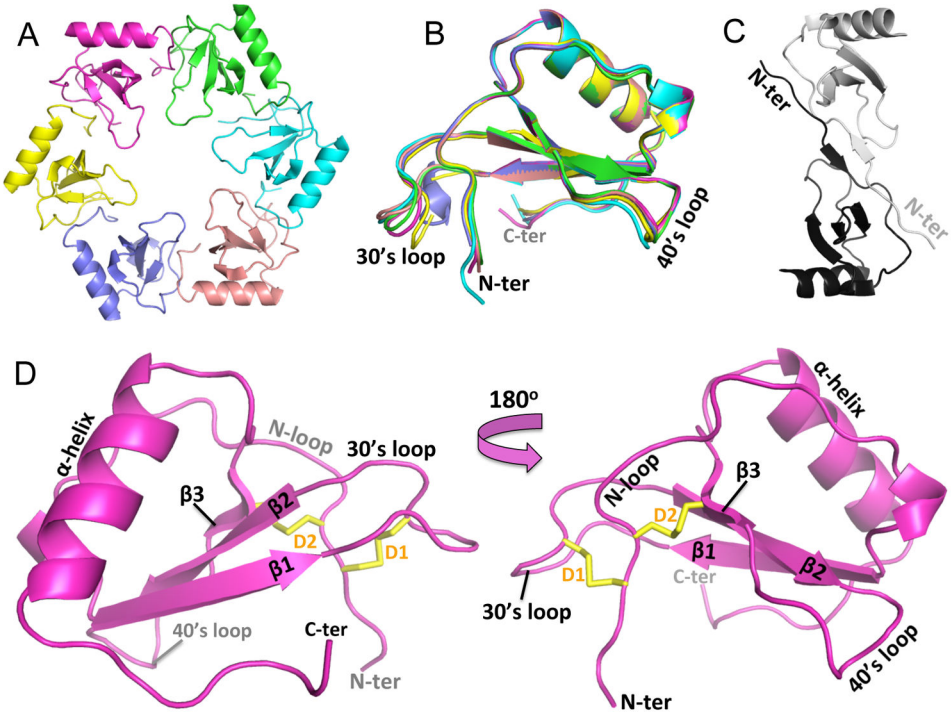


Figure 2. Crystal structure of truncated CCL21 (residues 1–79)

(A) Asymmetric unit containing six monomers. (B) Superimposition of the six monomers shows negligible backbone variation. (C) CC-like dimer formation in CCL13 crystal structure (PDB ID: 2RA4) [(35)]. (D) Monomer A from two side views. Disulfide bonds are depicted in yellow.

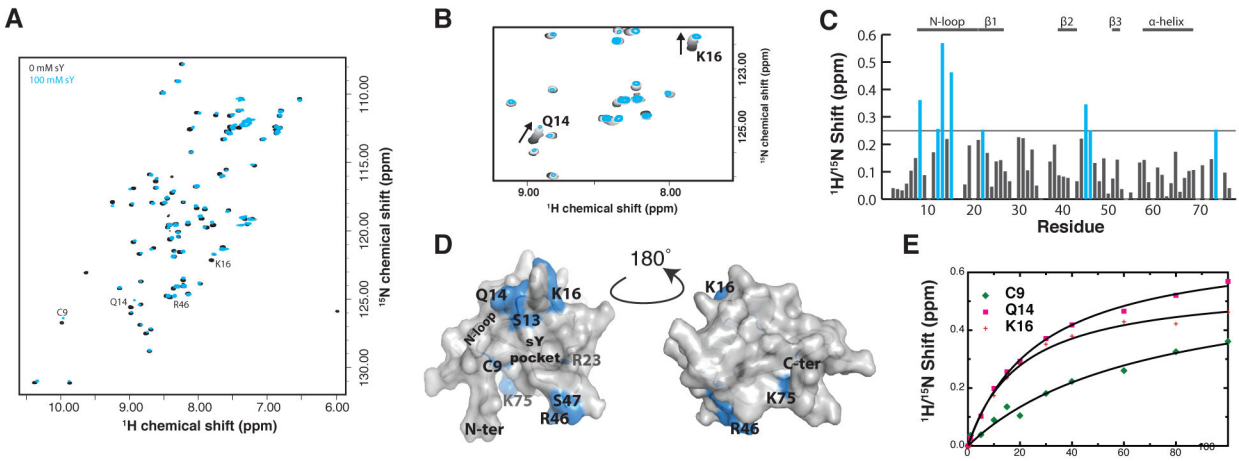


Figure 3. Truncated CCL21 titration with sulfotyrosine (sY)

(A) Overlay of truncated CCL21 (residues 1–79) HSQC spectra in the presence of 0mM sY (black) to 100mM sY (blue). Spectra were manually edited to remove sY streaks. (B) Enlarged section of (A) highlighting peak movement upon sY addition. (C) Total chemical shift perturbations from 0mM to 100mM sY plotted for each amino acid. Residues highlighted in blue had the largest chemical shifts, > 0.25 ppm. Chemical shifts from residues K11, V21, I36, E57, and Q61 were obscured under sY streaks and thus removed from the plot. (D) Residues highlighted in blue from (C) are mapped in blue on the structure of truncated CCL21. (E) K_d plots for residues with the greatest chemical shift perturbations.

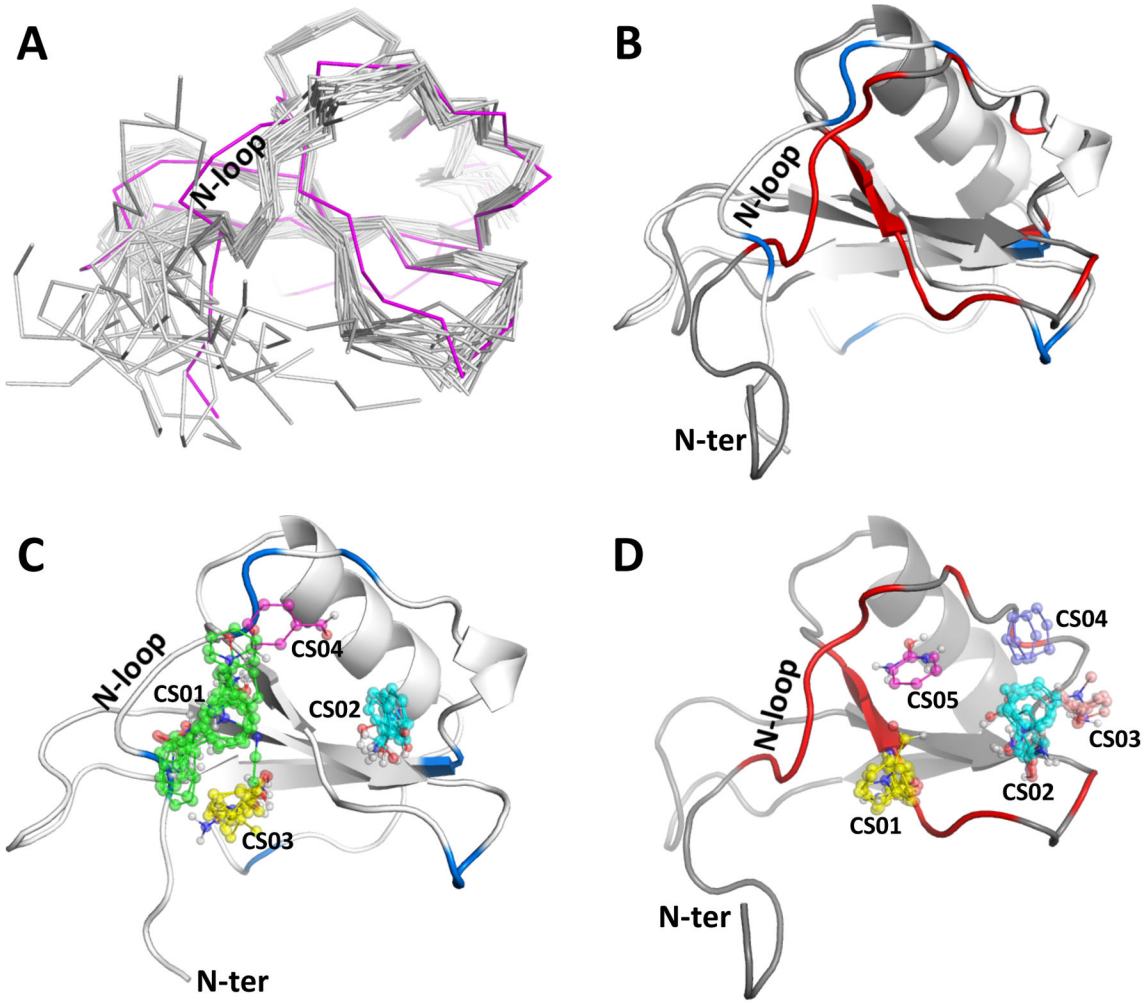


Figure 4. Comparison to full-length CCL21 NMR structure and analysis of druggability

(A) Superimposition of the X-ray structure of truncated CCL21 monomer A (purple) to the NMR ensemble of full-length CCL21 (grey, PDB ID: 2L4N) [(6)] shows N-loop difference. (B) X-ray structure of truncated CCL21 monomer A (light grey) with residues perturbed by sulfotyrosine titration (blue: peak shift > 0.25 ppm), superimposed to conformation 1 of the NMR full-length CCL21 (dark grey, PDB ID: 2L4N) [(6)] with residues perturbed by nonsulfated CCR7 peptide titration (red: peak shift > 1.5 ppm), as evidenced in a previous study [(6)], highlight the same general area suggesting it is the conserved sulfotyrosine-binding site. (C) FTMap analysis of the X-ray structure of truncated CCL21 shows four probe clusters (CS01-04) in potential druggable hot spots in the sulfotyrosine-binding site. CS01 contains 33 probes (green), CS02 contains 17 probes (cyan), CS03 contains 14 probes (yellow), and CS04 only contains one probe (magenta). (D) FTMap analysis of conformation 1 of NMR ensemble (PDB ID: 2L4N) shows five clusters (CS01-05) in potential druggable hot spots. CS01 has 16 probes (yellow), CS02 has 15 probes (cyan), CS03 has 9 probes (pink), CS04 has 2 probes (purple), and CS05 has 1 probe (magenta).

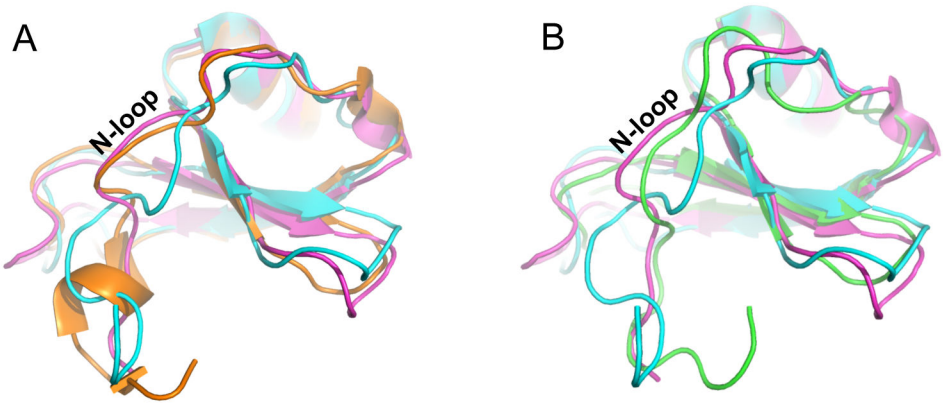


Figure 5. Comparison of the structures of CCL21, CCL18, and CCL19

The ribbon diagrams of the truncated CCL21 crystal structure and the full length CCL21 NMR structure are shown in magenta and cyan, respectively. (A) Superimposition of the CCL21 structures with the crystal structure of CCL18 (orange, PDB ID: 4MHE). (B) Superimposition of the CCL21 structures with the NMR structure of CCL19 (green, PDB ID: 2MP1).

Table 1

X-ray Data Collection and Refinement Statistics

Data Collection	
Space Group	$P2_1$
Cell Dimensions	
a, *b*, *c* (Å)	65.75, 58.24, 66.05
α,β,γ(°)	90.00, 119.94, 90.00
Resolution (Å)	50.00-1.90 (1.93–1.90)
No. Reflections	34110 (1705)
R_{merge} (%)	11.5 (59.9)
I/óI	16.5 (2.2)
Completeness (%)	99.9 (99.5)
Redundancy	7.0 (6.2)
Refinement	
Resolution (Å)	29.12-1.90
R_{work}/R_{free} (%)	20.5/24.8
No. Atoms	
Protein/Ligand/Water	3427/25/188
B-factors (Å^2)	
Protein/Ligand/Water	23.75/16.68/25.53
RMS Deviations	
Bond Lengths (Å)	0.004
Bond Angles (°)	0.81
Ramachandran Plot	
Most Favored Region (%)	94.2
Additionally Allowed (%)	5.8
Generously Allowed (%)	0.0

*Values in parentheses represent highest resolution shells