

Conservation of Unfavorable Sequence Motifs That Contribute to the Chemokine Quaternary State[†]

Ioannis Kagiampakis,[¶] Hongjun Jin,[§] Sinae Kim,^{||} Marina Vannucci,[⊥] Patricia J. LiWang,[‡] and Jerry Tsai^{*#}

Department of Biochemistry and Biophysics, Texas A&M University, College Station, Texas, 77843, Department of Cell Biology and Biochemistry, U.S. Department of Energy, Richland, Washington 99352, Department of Biostatistics, University of Michigan, Ann Arbor, Michigan 48109, Department of Statistics, Rice University, Houston, Texas 77251, School of Natural Sciences, University of California, Merced, California 95344, and Department of Chemistry, University of the Pacific, Stockton, California 95211

Received November 19, 2007; Revised Manuscript Received June 6, 2008

ABSTRACT: In the chemokine family, we characterize two examples of evolutionarily conserved unfavorable sequence motifs that affect quaternary structure. In contrast to the straightforward action of favorable sequences, these unfavorable motifs produce interactions disfavoring one outcome to indirectly promote another one but should not be confused with the broad sampling produced by negative selection and/or design. To identify such motifs, we developed a statistically validated computational method combining structure and phylogeny. This approach was applied in an analysis of the alternate forms of homodimerization exhibited in the chemokine family. While the chemokine family exhibits the same tertiary fold, members of certain subfamilies, including CXCL8, form a homodimer across the β 1 strand whereas members of other subfamilies, including CCL4 and CCL2, form a homodimer on the opposite side of the chemokine fold. These alternate dimerization states suggest that CCL4 and CCL2 contain specific sequences that disfavor CXCL8 dimerization. Using our computational approach, we identified two evolutionarily conserved sequence motifs in the CC subfamilies: a drastic two-residue deletion (Δ RV) and a simple point mutation (V27R). Cloned into the CXCL8 background, these two motifs were experimentally proven to confer a monomeric state. NMR analyses indicate that these variants are structured in solution and retain the chemokine fold. Structurally, the motifs retain a chemokine tertiary fold while introducing unfavorable quaternary interactions that inhibit CXCL8 dimerization. In demonstrating the success of our computational method, our results argue that these unfavorable motifs have been evolutionarily conserved to specifically disfavor one dimerization state and, as a result, indirectly contribute to favoring another.

Although protein interactions can act both favorably and unfavorably, they are usually thought to be favorable and act in specific ways to help bring about a certain outcome in protein structure. Interactions can also be unfavorable, but by their very nature, they are not often thought to actively contribute to protein structure. In addition, discrete motifs producing unfavorable interactions have not been widely studied, because they are predominantly nonspecific and produce negative results that cannot be interpreted. However, specific unfavorable interactions have been used in protein design to prevent certain outcomes from occurring (1, 2) and have been characterized in the termination of β -sheet structures (3). These unfavorable interactions should not be

confused with the results from negative selection, where positive interactions are directly removed (4), or general negative design (5), where certain interactions are disfavored (6). Although general negative design strategies produce unfavorable interactions, specific motifs are not used to achieve their goal. In this work, we want to identify evolutionarily conserved motifs that confer unfavorable interactions and characterize their role in determining protein quaternary structure. To accomplish this, we developed a predictive method combining structural with phylogenetic analysis that is validated using Bayesian variable selection (7). This approach was used in an investigation of how unfavorable interactions contribute to the chemokine dimerization state.

Chemokines are extracellular signaling proteins that play a general role in the innate and adaptive immune response, angiogenesis, cancer, and wound healing (8–11). In mammals, chemokines are separated into four main groups on the basis of the number of amino acids between the conserved cysteines (C) at the N-terminus of the protein (CXC, CC, CX3C, and C, where X is any amino acid). The largest group of chemokines consists of the CXC chemokines (21 human members) and the CC chemokines (29 human members) (12). With the exception of some transmembrane domains con-

[†] P.J.L., I.K., and H.J. were supported by NIH Grants RO1AI47832 and RO1AI070993 and Welch Foundation Grant A1472. J.T. and M.V. are supported by NIH/NIGMS Grant R01GM81631. M.V. is supported by NIH/NHGRI Grant R01HG003319 and NSF/DMS Grant DMS-0605001.

* To whom correspondence should be addressed. E-mail: JTsai@pacific.edu. Phone: (209) 946-2271. Fax: (209) 946-2607.

[‡] University of California.

[§] U.S. Department of Energy.

^{||} University of Michigan.

[⊥] Rice University.

[#] University of the Pacific.

[¶] Texas A&M University.

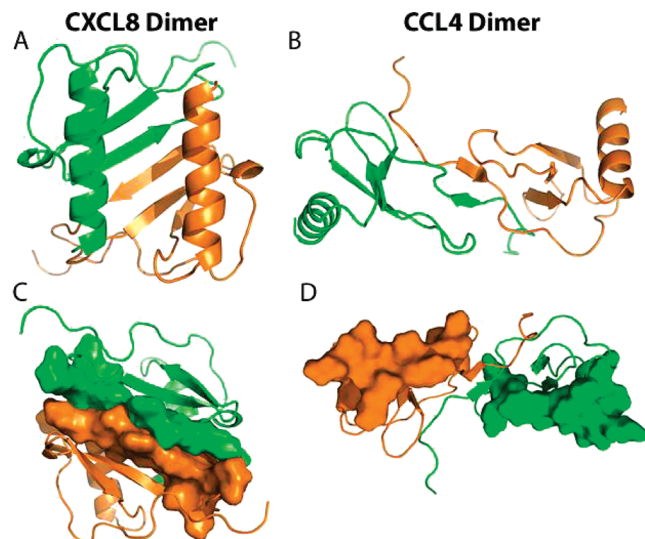


FIGURE 1: Comparing the quaternary structure of the CXCL8 and CCL4 homodimers. All images were created using PyMOL (66). (A) Showing PDB entry 3il8 (67), CXCL8 dimerizes using the β 1 strand and the α -helix. (B) Showing PDB entry 1hum (34), CCL4 dimerizes using the largely unstructured N-terminal region and the loops between strands β 2 and β 3. This interface forms eight interdomain hydrogen bonds: four from the small section of the antiparallel β -sheet formed by the β 0 strands and four more backbone–backbone hydrogen bonds with the β 3 strand. Apart from interactions with the β 0 and β 3 strands, CCL4's homodimer interacts primarily through hydrophobic interactions between residues that do not form canonical secondary structure elements. (C) CXCL8's β 1 strand residues involved in the antiparallel β -sheet quaternary interaction are shown in space filling models. Overall, the CXCL8 interface forms eight hydrogen bonds and the molecular area of interaction is 361 Å², where 303 Å² comes from side chain interactions and only 58 Å² comes from backbone interactions. Specific details are illustrated in Figure 2. (D) Space filling models of the corresponding residues shown on CCL4 that are involved in CXCL8's β 1 strand antiparallel β -sheet quaternary interaction.

taining chemokines, these proteins are generally small molecules (~8 kDa). While the overall level of amino acid sequence conservation of the family is low, they all share the same tertiary fold of a single α -helix across a three-stranded antiparallel β -sheet (Figure 1) suggesting that they most likely originated from a common ancestor. At the quaternary level, chemokines can form dimers, tetramers, or even heterodimers (13, 14). Possible roles of this oligomerization have been linked to chemokine function (15), such as suppression of specific surface-dependent neutrophil responses (16) and increasing *in vitro* antiproliferative effects (17). For homodimers, the quaternary structures of chemokines from different groups dimerize in different ways. As shown in Figure 1, CXCL8 chemokines interact across their β 1 strands to form a globular dimer, and in contrast, CCL4 dimerizes with its uncoiled N-terminus in a more extended conformation. While seemingly straightforward, attempts to disrupt the CXCL8 homodimer and retain the chemokine fold have required a set of mutations (13, 14, 16, 18, 19) from simultaneous changes in both the β 1 strand and α -helix to a chemical modification of L25 (Table 1). In this work, we used a directed, computational approach to discover evolutionarily conserved motifs that produce a correctly folded CXCL8 monomer.

Because all chemokines share a common tertiary structure yet the CXC and CC subfamilies generally interact with

different quaternary forms (Figure 1), we hypothesized that there must be motifs in the CC subfamily (dimerizing between their N-termini) that disfavor CXCL8 dimer formation across the β 1 strand but still conserve the chemokine fold. To discover these sequence elements, our approach to sequence analysis focused on increasing the magnitude of the signal of conserved differences and decreasing the noise from differences caused by random mutations. Therefore, we limited our sequence analysis structurally to only the CXCL8 dimerization region (Figure 2) and evolutionarily to two CC clades (the one including CCL4 and another including CCL2; see Figure S1). The results underwent rigorous statistical validation to separate true negative motifs from all the other noise produced in an alignment (7). On the basis of this work, we created three variants of CXCL8: two mutants that should induce a stable monomer and one control mutant that should not. Surprisingly, one of the predicted monomeric mutations involves a very drastic two-residue deletion within the dimerization interface. For the two evolutionarily conserved motifs from the CC family placed into a CXCL8 background, we experimentally prove they inhibit β 1 strand dimerization, while retaining the chemokine fold. The control exhibited only a moderate decrease in the level of dimerization. These results demonstrate that sequences conferring unfavorable interactions are evolutionarily conserved, and the implications of these findings related to chemokine quaternary structure are discussed.

EXPERIMENTAL PROCEDURES

CXCL8 Dimer and Monomer Surface Area Calculation. CXCL8's molecular surface area across its dimer interface (20) was calculated using a Voronoi polyhedron-based procedure (21). The residue interactions are provided in Table S1. We used all 30 NMR¹ structures of the CXCL8 dimer in PDB entry 2IL8 (22) for our calculations to allow for and sample the possible fluctuations across the dimer interface. Only contacts occurring in more than 15 of 30 structures were considered as an interaction and are shown in Figure 2. Values reported in Table S1 are averages.

Sequence Alignment. All the protein sequences of the vertebrate CC chemokines that were highly similar to CXCL8 were downloaded from the Cytokine Family cDNA Database (12). Sequences from signal peptides were not included. Using the phylogenetic tree shown in Figure S1, a ClustalW (23, 24) or DiAlign (25) (depending upon the circumstances) multiple-sequence alignment was created using all the sequences of CC chemokines for every single node and starting from the bottom of the tree. The alignment in the initial node was found to be significant in cases where the multiple-sequence alignments in the area of CXCL8's dimerization remained partially unchanged in the smaller subtrees. Using this procedure, two major subtrees with constant differences in the area of CXCL8's β 1 strand were constructed and named clade CCL4 (29 sequences in the blue subtree, Figure S1) and clade CCL2 (16 sequences in the red subtree, Figure S1). The sequences in these two clades

¹ Abbreviations: AUC, analytical ultracentrifugation; NMR, nuclear magnetic resonance; PDB, Protein Data Bank; SEC, size exclusion chromatography.

Table 1: CXCL8 and Its Variants^a

A) Sequence of CXCL8																								
Amino Acid	E	L	R	V	I	T	E	L	V	F	L	K	R	A	E	N	S							
Residue Number	24	25	26	27	28	37	38	43	62	65	66	67	68	69	70	71	72							
Secondary Structure	E	E	E	E	E	C	E	E	H	H	H	H	H	H	H	H	C							
B) BVS Prediction																								
Comparison with CCL4 clade	E	Y	Δ	Δ	I	T	A	I	Δ	F	L	K	L	A	E	N	S							
Comparison with CCL2 clade	E	Y	R	R	I	T	A	L	V	F	D	K	R	A	E	N	S							
C) CXCL8 Variants used in this study																		K_d (μM)		M/D				
Wild Type	E	L	R	V	I	T	E	L	V	F	L	K	R	A	E	N	S	~0.1		D				
ΔRV	E	L	Δ	Δ	I	T	E	L	V	F	L	K	R	A	E	N	S	--		M				
V27R	E	L	R	R	I	T	E	L	V	F	L	K	R	A	E	N	S	--		M				
R26A/V27A	E	L	A	A	I	T	E	L	V	F	L	K	R	A	E	N	S	18.2 ± 3.1		wD				
D) Previous Experiments																								
(a) Lusti-Narasimhan et al., 1996	E	L'	R	V	I	T	E	L	V	F	L	K	R	A	E	N	S	--		M				
(b) Lowman et al., 1997	R	L	R	V	R	T	E	L	V	F	L	K	R	A	E	N	S	7.6 ± 4.2		D				
	E	L	R	V	I	E	E	L	V	F	L	K	R	E	E	N	E	--		M				
	R	L	R	V	R	E	E	L	V	F	L	K	R	E	E	N	E	1800 ± 550		wD				
	R	L	R	V	R	E	E	L	V	H	E	K	R	E	E	N	E	1800 ± 500		wD				
(c) Jin et al., 2005	E	L	R	V	I	T	E	L	V	F	Δ	Δ	Δ	Δ	Δ	Δ	Δ	--		M				
	E	A	R	E	I	T	E	L	V	F	L	K	R	A	E	N	S	151.5		wD				
(d) Williams et al., 2005	E	Y	R	R	I	T	E	L	V	F	L	K	R	A	E	N	S	--		M				
	L	E	R	V	I	T	E	L	V	F	L	K	R	A	E	N	S	--		M				

^a In section A are listed the wild-type CXCL8 amino acid composition and the number and secondary structure of residues involved in the dimer interface shown in Figures 1A,C, and 2. Secondary structure is given by sheet (E), helix (H), and coil (C) as defined by DSSP (65). The periods between residue numbers represent breaks in the sequence. The statistical analysis is given in section B. BVS stands for Bayesian variable selection. Amino acid substitutions shown in black and underlined represent strongly identified sites, whereas those shown in gray are weakly identified sites (see Experimental Procedures). Sites not predicted are shown in gray with the wild-type amino acid. The deltas (Δ) are deleted residues in sections B–D. While analysis across all the residues was shown for completeness, this study focuses on only the analysis across the β1 strand (residues 24–28) and only those unique to the CC subfamily clade. For instance, we did not include the L25Y substitution, since it is seen in analysis with both clades. Section C lists CXCL8 variants used in this study. Mutations are shown in black, while the unchanged native sequence is shown in gray. Again, the complete wild-type sequence is shown along with the three variants used in this study: ΔRV, V27R, and R26A/V27A. In addition, we show the dimerization constant K_d (micromolar) and indicate monomer/dimer state, where M refers to monomer, D to dimer, and wD to weak dimer. Section D gives the same information for published mutations from previous experiments in the CXCL8 interface ordered by year the study was published. Following the same conventions, we show the amino acids to which residues are mutated in black and the native unchanged sequence in gray. L' represents an *N*-methylleucine.

were used to create the logo figures. To ensure the accuracy of the gap discovered at positions R26 and V27 of CXCL8 (Table 1), the alignment was repeated using DiAlign, which does not use gap penalties but the best diagonal without gaps in the alignment matrix (26). This alignment also predicted the two-residue deletion (data not shown). While the phylogenetic tree in Figure S2 suggests that these two mutations diverged from a common ancestor, it could equally be the case that a monomeric ancestor existed. This analysis considers only the differences between sequences and does not include bias about how they originated.

Validation by Bayesian Variable Selection. The method itself was discussed in more detail previously (7). We compared sequences only within the area of interaction defined above based on a simple hydropathy profile (27). To identify structural dissimilarities in protein alignments, we have applied the well-known variable selection method for classification with probit models proposed by Sha et al. (28). Bayesian variable selection is done via a binary vector with *p* entries that identifies the different sets of variables. The marginal posterior distribution of this binary vector is derived, and Markov chain Monte Carlo algorithms are used to sample from its posterior distribution. The method allows the identification of sets of discriminating variables and the

classification of future samples (via least-squares or Bayesian model averaging). A site is identified as significant when there is a high posterior probability that the site is dissimilar (has had a mutation). Conversely, those sites that are not identified are those that have a small probability of being dissimilar. Classification of strong and weak was determined on the basis of a comparison from two runs using different input parameters: strong sites were those identified by both runs and weak ones those identified by only one. The sensitivity of the method to the input parameters has been discussed previously (28). Because we are working from only one CXCL8 sequence, we needed to cross-validate our results. Once we identified the sites with a high posterior probability of being dissimilar, we tested how well we could discriminate between the groups (for example, CXCL8 from CCL2) on the basis of these sites. Taking one CCL2 sequence out, we try to “predict” whether this sequence is more similar to CXCL8 or to the other CCL2. If the answer should be the latter, it was considered a zero error.

Construction and Purification of Mutants. All the CXCL8 mutants were created in the pET-32a+ (Novagen) vector using Quikchange mutagenesis (Stratagene) and were confirmed via sequencing. Protein purification for all mutants

A) Quaternary Contact Plot.

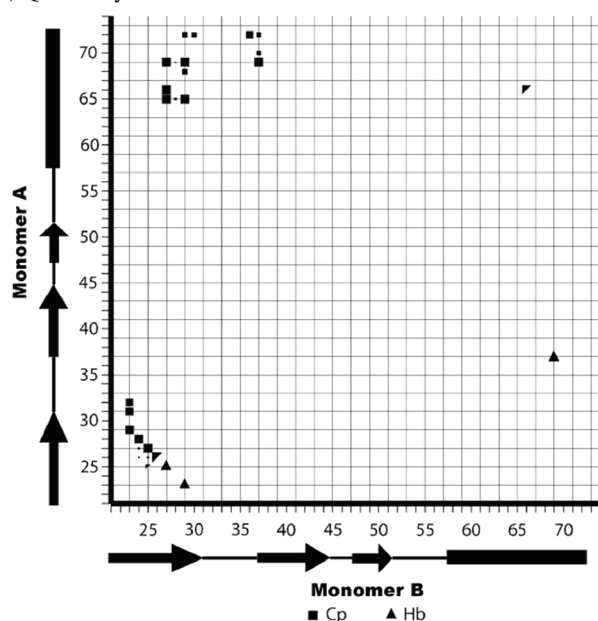
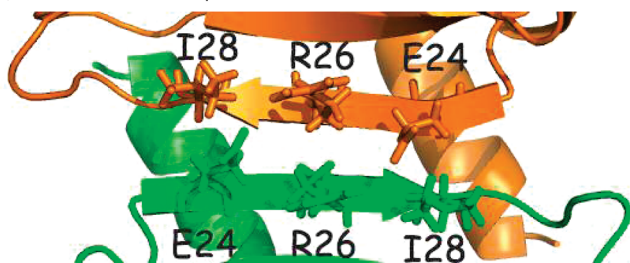
B) Residues across the β 1 strand interface.

FIGURE 2: CXCL8 homodimer interaction interface. (A) Quaternary contact plot. (B) Residues across the β 1 strand interface. For panel A, contacts were calculated using Qcontacts (20) and are listed in Table S1. Because the dimer interface is symmetric across the diagonal, we can plot the hydrophobic carbon-carbon contacts as filled squares (■) above the diagonal and the hydrogen bonds as filled triangles (▲) below the diagonal. The plot clearly shows the three areas of interaction across the dimer interface. The α -helices from each CXCL8 monomer form a loose hydrophobic network, and this interaction between helices is centered on the interaction between the two L66 residues. The next part of the CXCL8 interface involves the C-terminal, α -helical residues 65–72 interacting with residues 27–37 running from the β 1 strand into and through the turn between the β 1 and β 2 strands. In addition to the many hydrophobic contacts, two hydrogen bonds are formed (one from each monomer) between the side chain oxygen of T37 and the backbone carbonyl of A69. The last area across the two β 1 strands of the CXCL8 homodimer contributes >50% of the interaction surface area as well as hydrogen bonds to the interface. Each monomer's edge β 1 strand from residue 23 to 28 forms a regular antiparallel β -sheet interaction that unites a single six-stranded β -sheet with six backbone hydrogen bonds and a large, continuous area of interaction. For the whole interface, residue R26 sits on the axis of symmetry (interacts with the same residue on the opposing monomer) and coincidentally buries the most surface area. This is clearly shown in panel B, where residues E24, R26, and I28 are shown in the antiparallel sheet between the two β 1 strands.

and the wild type (WT) used a standard chemokine refolding protocol as described previously (29).

Size Exclusion Chromatography. The purified protein was first dissolved into 20 mM sodium phosphate buffer (pH 2.5) containing 150 mM NaCl, and then the pH of the solution was adjusted to 6.7 with 4 M NaOH. The solution was centrifuged at 12000 rpm for 20 min to remove any possible

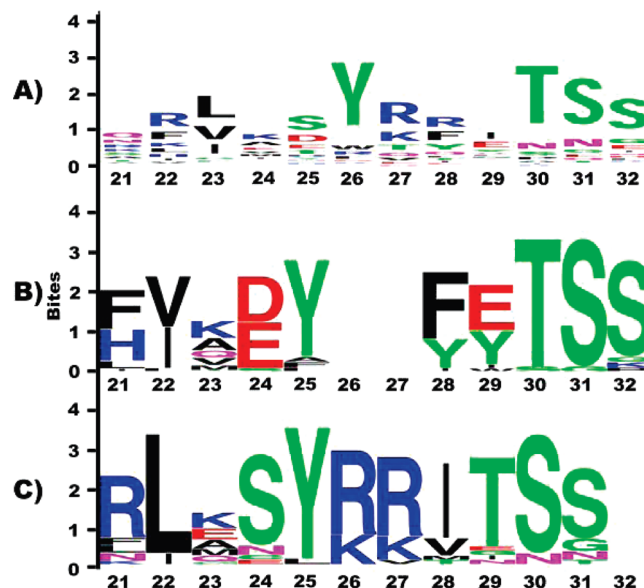


FIGURE 3: Logo plots of the β 1 strand region. Three multiple-sequence alignments are represented using Weblogo (68). The overall height of the stack indicates the conservation at that position in the multiple-sequence alignment, and the height of the individual single-letter codes within the stack indicates the relative frequency of the corresponding amino acid at that position. The x-axis values refer to the residue position of the mature human CXCL8. Color coding for the class of amino acids: green and purple for polar, blue for basic, red for acidic, and black for hydrophobic. (A) Problems of a global multiple-sequence alignment over the entire CC chemokine family with the CXCL8 sequence. One is the reduction of the magnitude of the signal (see the smaller letters of the Arg residues at positions 27 and 28), and another is the apparent +1 residue shift in the middle of the alignment as compared to the two following alignments (see the shift in the Tyr residue at position 26 and the Arg residues at positions 27 and 28). (B) Results of the multiple-sequence alignment comparing the CXCL8 sequence with the sequences from the clade surrounding CCL4. The deletion in the center at positions 26 and 27 is quite evident. (C) Results of the multiple-sequence alignment comparing the CXCL8 sequence with the sequences from the clade surrounding CCL2. Again, the conservation of the Arg at positions 26 and 27 is very apparent.

precipitation. The dissolved refolded protein (~0.5 mL) was loaded onto a Superdex G75 (HR 10/30) gel filtration column (GE Healthcare) equilibrated with 20 mM sodium phosphate buffer (pH 6.7) containing 150 mM NaCl. The protein was eluted with this buffer and run using the Pharmacia Akta system. The gel filtration column was calibrated with several standard marker proteins (Bio-Rad) in the same running buffer, and the molecular masses of chemokine variants were estimated using the calibrated standard curve. All sample elutions were performed at a rate of 0.4 mL/min; 1 mL fractions were collected, and the protein content was analyzed by 12% SDS-PAGE.

Analytical Ultracentrifugation (AUC). Sedimentation equilibrium data were collected on a Beckmann XL-A analytical ultracentrifuge using the An-60 Ti rotor at 25 °C. The following sets of speeds were used for each mutant: 26500 and 36500 rpm for R26A/V26A, 26500, 37500, and 48000 rpm for Δ RV, and 25000, 27500, 30000, 32000, 34000, 36000, and 38000 rpm for V27R. The absorbance for all three mutants was measured at 280 nm to determine the concentration of the protein during the sedimentation equilibrium. For V27R, the absorbance was measured at two additional wavelengths (235 and 245 nm). For each run,

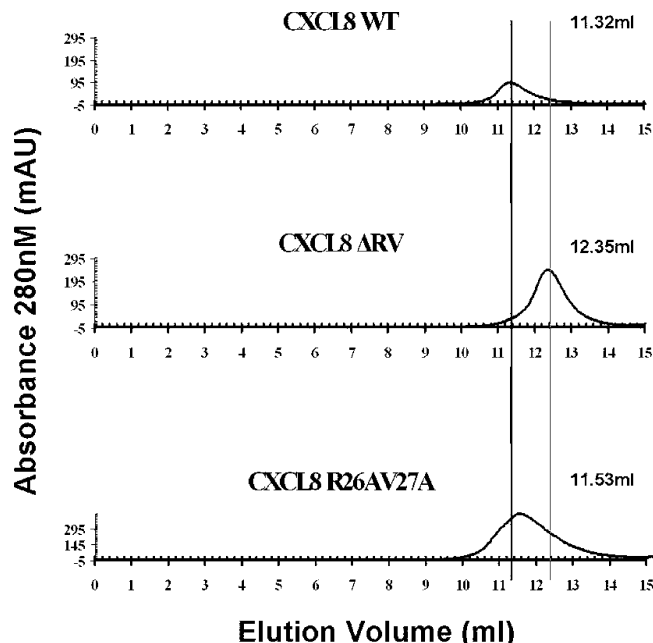


FIGURE 4: Size exclusion chromatography. SEC is shown for WT CXCL8, the Δ RV variant, and the R26A/V27A mutant. The black line shows the expected earlier position for the dimer, and the gray line represents the expected later position for the monomer. Elution volumes of the peak are shown on the x -axis.

protein was dissolved in sodium phosphate buffer with 150 mM NaCl at pH \sim 6.2. Initial protein concentrations were 110 μ M for R26A/V26A and 25, 50, and 111 μ M for Δ RV. For V27R, eight different concentrations between 50 and 500 μ M were used. The data were processed using the Ultra Scan Data analysis software (30). The calculated molecular masses and dissociation constants as well as their respective standard deviations was calculated on the basis of a Monte Carlo simulation of the initial global fitting over all the data collected.

NMR Spectroscopy. NMR samples of \sim 1 mM protein for all the variants were prepared in 150 mM NaCl, 20 mM sodium phosphate (pH 6.7), 0.01% NaN₃, 0.1 mM DSS, and a 95% H₂O/5% D₂O solvent mixture. The spectra were collected on a Varian Inova 600 MHz spectrometer at 25 $^{\circ}$ C. ¹H, ¹³C, and ¹⁵N chemical shifts were referenced to internal DSS (31). Sequence specific backbone assignments of ¹³C α , ¹³C β , and ¹⁵N were obtained from CBCA(CO)NH and CBCANH experiments (32). For structure comparison, ¹H–¹⁵N HSQC (33) spectra were acquired with 512* points in the ¹H (direct) dimension and 64* points in the ¹⁵N (indirect) dimension, where n^* represents n complex points. The spectral width for ¹H was 8000 Hz and for ¹⁵N was 1700.68 Hz. Total acquisition times were 64 ms for the ¹H and 37.6 ms for the ¹⁵N dimension. The relaxation experiments were performed according to the following. The spin-lattice (T_1) and spin-spin (T_2) values were determined for \sim 1mM ¹⁵N labeled CXCL8 Δ RV, from ¹⁵N–¹H-HSQC spectra collected using pulse sequences described above. Intensities of cross-peaks were obtained from peak-picking routines provided in the software PIPP package (70). Data from each amino acid residue was extracted in a straightforward fashion by measuring the intensities of cross-peaks in ¹⁵N–¹H-HSQC spectra as a function of a relaxation delay. The ¹⁵N T_1 data was collected with ¹⁵N delays of 0.02, 0.1, 0.12, 0.18, 0.25, 0.36, 0.52, 0.76, 1.00, and 1.25 s. T_2 data

was collected using ¹⁵N delays of 32, 48, 64, 80, 96, 128, 144, 160, 192, and 240 ms. The T_1 , T_2 data sets were processed using software nmrPipe (71). Data was excluded from the correlation time calculation if either T_1 or T_2 differed from the average value by >1 SD.

RESULTS

Structural Analysis of the Dimer Interfaces. We go into only as much detail comparing the CXCL8 and CCL4 interfaces as necessary to provide background to our analysis, since this comparison has been made previously (34) and in greater evolutionary and molecular detail (35). For the sake of consistency, we follow the same nomenclature. The monomers are 18% identical with each other in sequence, yet both exhibit the same fold (C α rmsd between CXCL8 and CCL4 of 1.6 \AA) as shown in Figure 1: a single α -helix over a β -sheet arranged in a Greek key motif (all the β -strands are antiparallel to each other and are numbered from the N-terminus β 0, β 1, β 2, and β 3). Even though they are so similar, each possesses certain unique features. CCL4 exhibits a small, extra section of β -sheet structure near the N-terminus, while CXCL8's α -helix is one turn longer at its C-terminal end. As Figure 1 clearly shows, CCL4 and CXCL8 dimerize quite differently. The CCL4 homodimer is elongated and interacts primarily through N-terminal β 0 strand residues 2–16 with two areas on the opposing monomer: the same residues as well as residues 46–51 of the β 3 strand (Figure 1B). In contrast, the CXCL8 dimer is more globular, and the dimerization interface appears more uniform (Figure 1A and highlighted in Figure 1C). From an analysis of quaternary interactions (20), the CXCL8 dimer interface combines interactions across three areas (Figure 2 and Table S1). Of the three, the interaction area that has drawn the most focus is the β 1 strand, since it contributes the most surface area and hydrogen bonds to the homodimer interface (35). In contrast, the β 1 strand does not contribute at all to the CCL4 homodimer, and as emphasized in Figure 1D, CCL4's β 1 strands do not contact each other in its dimer state. For these reasons, we have focused our sequence analysis on the β 1 strand to identify the inhibitory sequence elements in the CC family of chemokines that disfavor the CXC homodimerization state.

Identifying Conserved Sequences Conferring Negative Interactions. Structurally limiting our comparison to only a fragment of the CXCL8 sequence and evolutionarily to only CC chemokine subfamilies is unlike previous work that has tried to mutate the conserved residues found from an alignment over the entire chemokine sequence (13, 14, 16, 18, 19, 35). As shown in Figure 3A, the breadth of such an alignment diminishes the conservation signal of all residues and the global alignment misaligns this region by a single-residue shift toward the C-terminus. As a result, many attempts to create a stable monomer have been performed, and each has used multiple point mutations (Table 1). Instead, we performed a more focused and statistically rigorous analysis that considered only those residues involved in the β 1 strand of the CXCL8 dimer interface and also limited the comparison of the CXCL8 sequence to only an evolutionarily related subset or clade of the CC chemokine family. Although the approach is more thoroughly discussed in Experimental Procedures, we briefly summarize the steps and reasoning

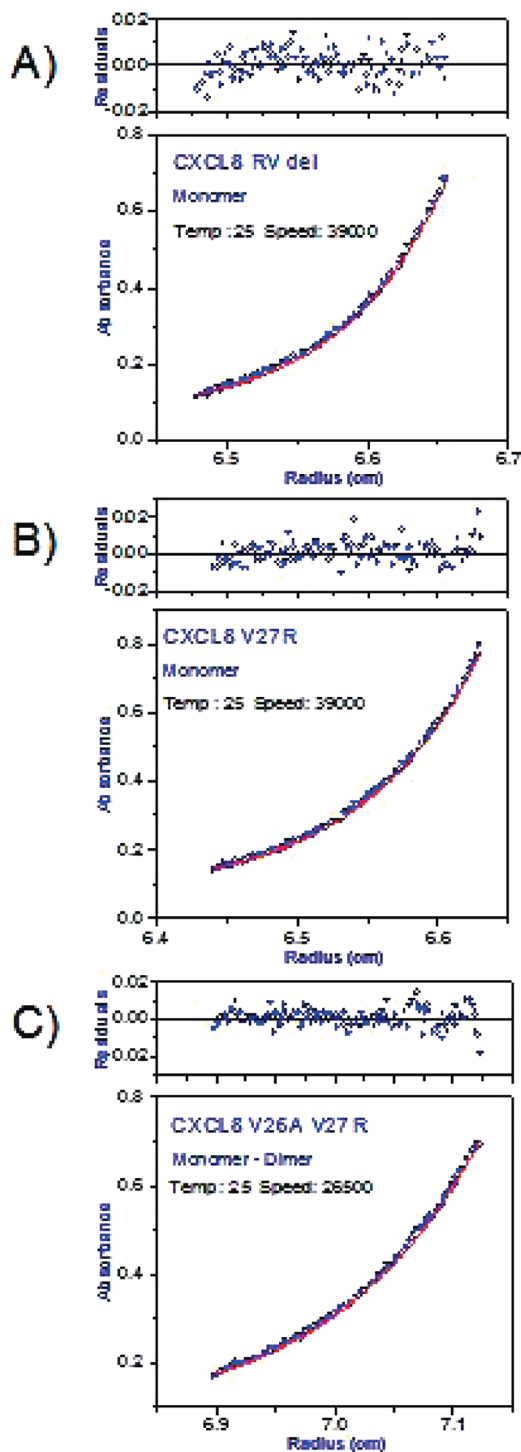


FIGURE 5: Individual analytical ultracentrifugation runs. Single AUC residuals (top) and fits (bottom) are shown for (A) ΔRV , (B) the V27R variant, and (C) the R26A/V27A mutant. Speed and temperature are shown on each plot. The wavelength is 280 nm for ΔRV , 281 nm for V27R, and 280 nm for R26A/V27A. As explained in Experimental Procedures, although a single run is shown here, the oligomerization data were fit for each variant over several runs. Both the ΔRV and V27R mutants fit a monomeric model with predicted molecular masses of 8.25 ± 0.52 and 8.74 ± 0.21 kDa, respectively. These values are close to the theoretical monomer molecular mass of these mutants. The R26A/V27A control fits best to a monomer–dimer equilibrium model with an estimated molecular mass of the monomer of 8.8 ± 0.6 kDa and a K_d of 18.2 ± 3.1 μM , which is a weaker dimer than the WT ($K_d \sim 0.1$ μM). The randomness of residuals shown above each plot indicates a good fit. The complete set of AUC runs at several different speeds and wavelengths is shown in Figure S2.

behind it. The clade analysis is based on the hypothesis that proteins with the same tertiary fold but different quaternary structure must have taken separate evolutionary paths from a common ancestral sequence and may have selected different mechanisms to achieve the same result. The choice of sequences to include in our clade analysis was defined by the branches in a phylogenetic tree of the CC chemokines (Figure S1). As the last step, the significance of our results was verified using a rigorous statistical method based on Bayesian variable selection (7) to filter out any insignificant results due to the low complexity of the sequence space.

On the basis of this procedure, we were able to identify a set of mutations that help us understand how the CC chemokine families negatively design against CXCL8 dimerization (summarized in Table 1). Because CCL4 possesses the same fold as CXCL8 yet dimerizes differently [as explained above (35)], our first comparison uses the clade surrounding CCL4. The sequence alignment shown in Figure 3B clearly indicates a two-residue deletion of R26 and V27 in the $\beta 1$ strand. While a number of other residues exhibit strong conservation, only the two-residue deletion and the L25Y mutation exhibited statistical significance (Table 1). To further test this approach, we also made the same limited sequence comparison of CXCL8 with the CCL2 sequence clade. Sharing the same chemokine fold, CCL2 dimerizes in a fashion similar to that of CCL4 (36) but also shares a low level of sequence identity with CXCL8. As corroborated by our statistical analysis, Figure 3C highlights the existence of the two Arg residues in the center of the $\beta 1$ strand created by a V27R mutation. Because we wanted to understand the sequence elements conferring negative interactions specific to each evolutionary branch, the L24Y mutation was disregarded since it was conserved across both the CCL4 and CCL2 clades. Therefore, our analysis produced two sets of negative interaction elements that break up a CXCL8 dimer yet retain the chemokine fold: a deletion of residues 26 and 27 (ΔRV) and a mutant creating a charge repulsion (V27R). As a control to test whether removal of the side chains or the rearrangement caused by the deletion is more important, the R26A/V27A double mutation was also introduced.

Monomeric State of the Mutants. The two variants of CXCL8 (ΔRV and V27R) along with R26A/V27A and the WT control were constructed and purified as described in Experimental Procedures. It is predicted that making these mutations in CXCL8 will abrogate the dimer. For the sake of clarity, we will further refer to these mutations using their abbreviated names indicated in the previous sentence and noted in Table 1. Size exclusion chromatography (SEC) and analytical ultracentrifugation (AUC) were used to investigate the quaternary state. In the SEC analysis, we compared the elution profiles of WT CXCL8 to those of the ΔRV and R26A/V27A mutants of CXCL8 (Figure 4). The major peak in the ΔRV profile eluted at a volume later than that of the WT that is quite similar to known CXCL8 monomeric mutants (29). In contrast, the R26A/V27A control variant of CXCL8 elutes only slightly later than wild-type CXCL8, but significantly before known monomeric variants. This similarity in elution profile to the WT suggests that the R26A/V27A variant forms a dimer in solution. To further verify the dimerization status of these variants, the more rigorous AUC experiment was carried out. A single run for each

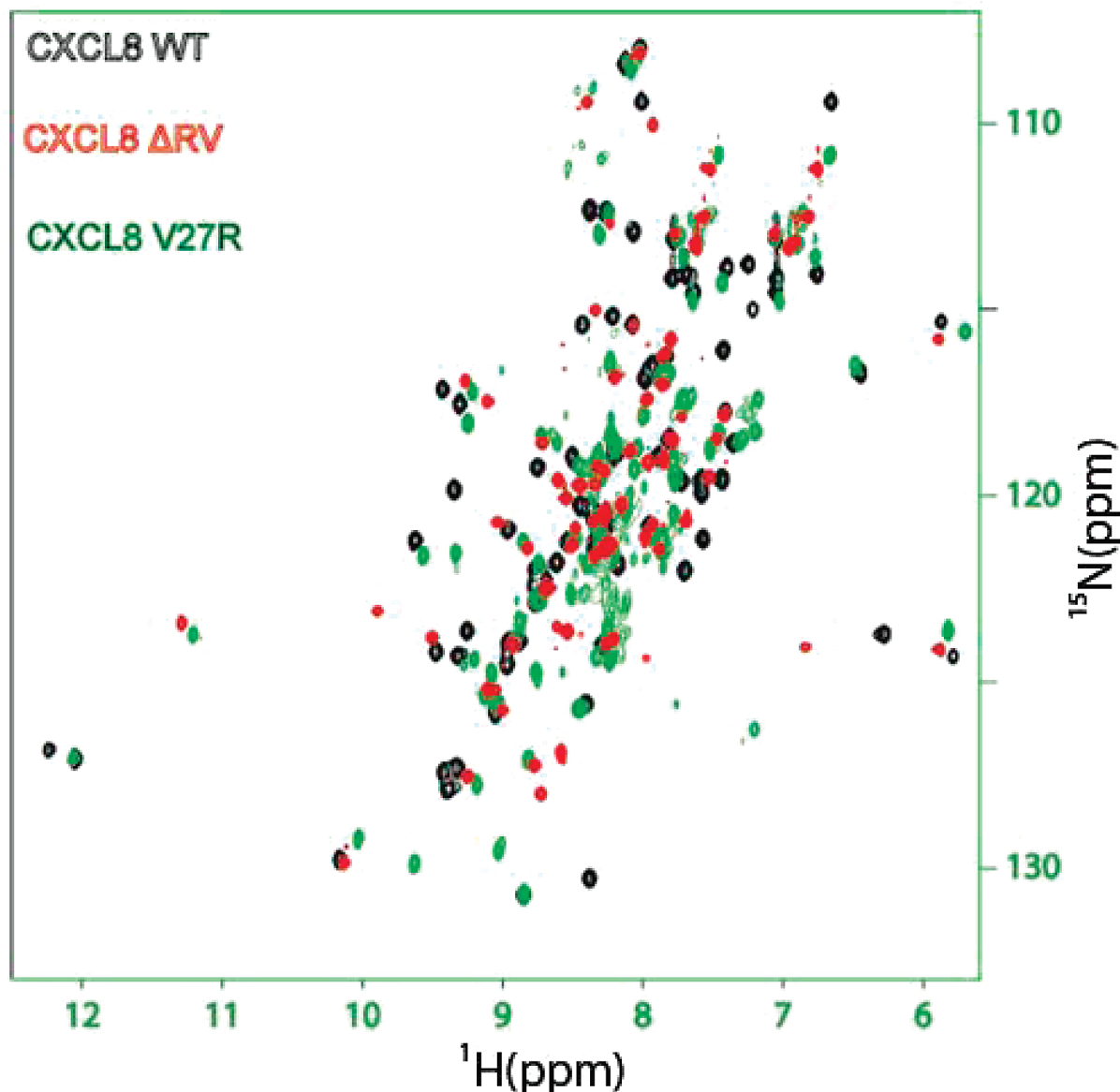


FIGURE 6: HSQC spectra. Comparison of ^{15}N – ^1H correlation spectra of wild-type CXCL8 to those of ΔRV and V27R. As labeled, data for wild-type CXCL8 are colored black. For contrast, peaks for ΔRV and V27R are colored red and green, respectively. Both of the mutant proteins display good dispersion of peaks, which implies that they are well-folded.

mutant is shown in Figure 5; however, for the sake of accuracy, the oligomerization state was modeled from an extensive set of AUC experiments (see Experimental Procedures) performed for each variant at different speeds, wavelengths, and concentrations (Figure S2). The result of the AUC analysis is that both the ΔRV and V27R mutants fit a monomeric model with predicted molecular masses of 8.25 ± 0.52 and 8.74 ± 0.21 kDa, respectively. These values are close to the theoretical monomer molecular masses of these mutants. These fits of ΔRV and V27R mutants were run at concentrations of up to 111 and 500 μM (see Experimental Procedures), respectively, which is far greater than the WT value of 0.194 μM (29) run under the same conditions. In comparison to other values of the WT, the K_d values of CXCL8 run under slightly different conditions ranging from 4 to 10 μM (18, 37, 38), these mutants are monomeric. Therefore, they show that the mutations strongly inhibit dimerization across the $\beta 1$ strand. The R26A/V27A control fits best to a monomer–dimer equilibrium model with an estimated molecular mass of the monomer of 8.8 ± 0.6

kDa and a K_d of 18.2 ± 3.1 μM , which is a weaker dimer than WT CXCL8 [$K_d = 0.1$ μM (29)]. Comparison to other values of WT CXCL8 K_d noted above substantiates that R26A/V27A still dimerizes in a manner comparable to that of WT CXCL8. These results show that the mutation disrupts but does not abolish dimerization. These SEC and AUC results clearly demonstrate that the two types of evolutionarily conserved sequence motifs confer negative interactions to inhibit CXCL8 dimerization across the $\beta 1$ strand.

NMR Evidence Indicating Chemokine Tertiary Structure. Because the quaternary state is sensitive to the tertiary state, we used NMR to determine whether the CXCL8 mutants formed a chemokine fold. First, HSQC ^{15}N – ^1H correlation spectra were compared. Figure 6 shows an overlay of the spectra of ΔRV , V27R, and WT CXCL8. The differences between these monomeric variants and WT CXCL8 are most likely due to the loss of intermolecular contacts across the dimer interface. We will first discuss the comparison of the HSQC spectra between the V27R mutant and WT, since it is only a single side chain mutation. We are confident that

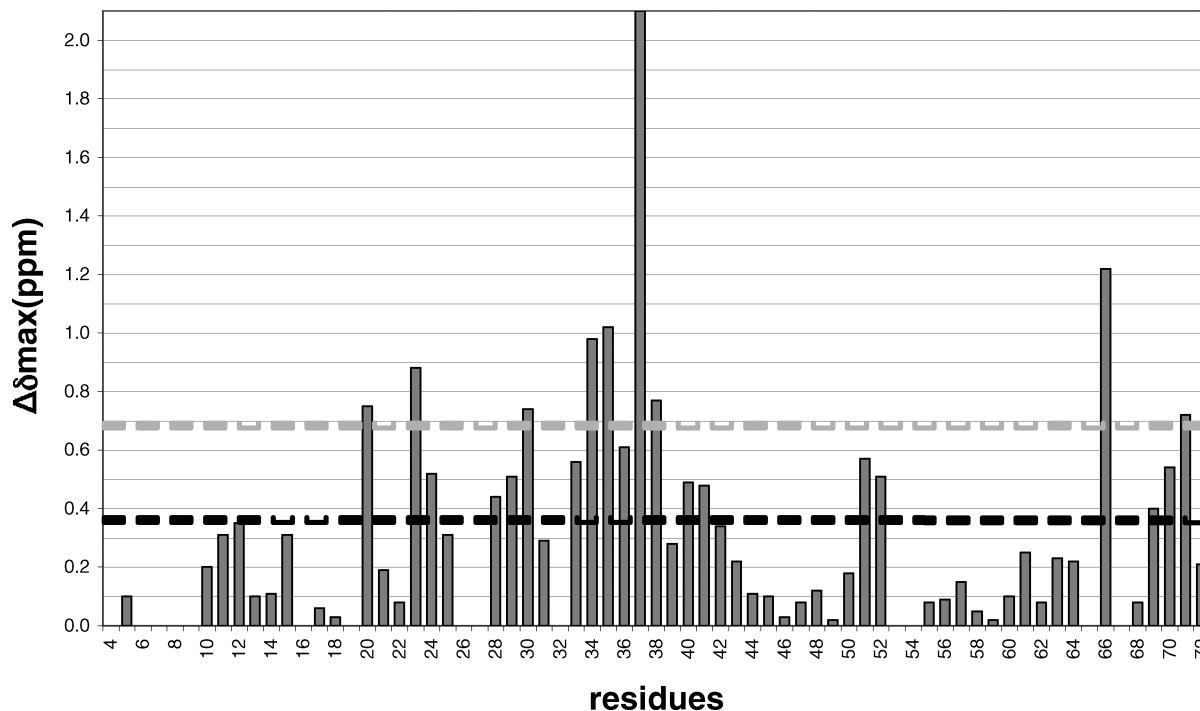


FIGURE 7: Chemical shift differences. The ^{15}N - and ^1H -weighted average chemical shift differences between WT CXCL8 and ΔRV are shown. The dashed black line represents the average chemical shift differences and the dashed gray line the standard deviation from the average. The weighted chemical shift differences (69) were calculated using the equation $\Delta\delta_{\text{obs}} = \{[(\Delta\delta_{\text{HN}})^2 + (\Delta\delta_{\text{N}}/5)^2]/2\}^{1/2}$. In this analysis of peak movement, the assignment of ^{15}N and ^1H resonances showed that many residues' resonances in the β_1 strand and the C-terminal portion of the α -helix change in the ΔRV mutant. All these peaks move in comparison to those of the WT, except I39. As discussed above, these expected changes in peak position indicate that the chemical environment around these amino acids has changed due to a lack of interface contacts and the two-residue deletion. In addition, the peaks corresponding to H33, C34, A35, and N36 have shifted in ΔRV , so the turn between the β_1 and β_2 strands has been affected, possibly because of the changes in the registry of the amino acid at the turn. The peaks corresponding to L51 and D52, in the turn between the β_1 strand and α -helix, have also moved. Lastly, the peaks from L66, A69, E70, and N71, which belong to the top part of the α -helix, all have shifted.

at NMR concentrations of ~ 1 mM (see Experimental Procedures), the V27R mutant is a monomer since the AUC data were recorded at 0.5 mM, close to the NMR concentration. In Figure 6, this monomeric V27R mutant exhibits good peak dispersion and displays a spectrum similar to those of other known CXCL8 monomers (29), so we conclude that the V27R mutant exhibits a chemokine fold.

Because AUC for the ΔRV mutant was maximally run at $1/10$ of the NMR concentration, we ran a ^{15}N relaxation experiment to verify that this mutant was a monomer. The ratio of ^{15}N relaxation parameters T_1 and T_2 can be used to estimate the size of a protein. For example, the wild-type CXCL8 dimer has been shown to have a T_1/T_2 ratio of 8.59 (39), and that of the CCL4 dimer is 8.3 (40). In contrast, monomeric variants of CCL4 have measured T_1/T_2 ratios of 3.9 and 4.0 (41). For this work, the CXCL8 mutant ΔRV was shown to have a T_1/T_2 ratio of 3.6, fully consistent with a monomeric chemokine. In Figure 6, the well-defined dispersion in the ^{15}N - ^1H correlation spectra is also representative of a folded protein. Unlike the simple point mutation in V27R, the deletion of two residues in the case of the ΔRV mutant causes an expected but significant change in the HSQC spectrum when compared to that of the WT. The deletion leads to loss of interactions across the protein-protein interface, and R26 and V27 also make several intramonomer contacts in the wild-type protein. When an analysis of tertiary contacts was performed (42), R26 and V27 make internal contacts with I28, T37, E38, I39, I40, V41, and K42 (data not shown). An assignment of ^{15}N , ^1H N,

$^{13}\text{C}\alpha$, and $^{13}\text{C}\beta$ resonances was carried out and confirmed that the residues contacting R26 and V27 in WT CXCL8 change environment in the ΔRV mutant (Figure 7). Although suggestive, the HSQC spectrum of the ΔRV variant does not provide conclusive evidence that this mutant retains the chemokine fold. Therefore, we turned to an NMR-based analysis of secondary structure (43) and oxidative state of cysteines (44) to prove that the ΔRV mutant folds into the chemokine structure shown in Figure 1. For secondary structure, we measured the $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ deviation from random coil values (Figure 8), which has been established as a strong predictor of secondary structure (45–47). While the β_3 strand is approximately the same length in ΔRV as in the WT, the β_1 and β_2 strands appear to be shorter. This potential fraying of the β_1 strand would contribute to inhibition of dimer formation. In addition, the deviations in Figure 8 show that the helix in ΔRV is not structured toward the C-terminus. Our T_1/T_2 ratios corroborate this result since they cannot be calculated past residue 65 in the ^{15}N relaxation experiment (data not shown). The unstructured C-terminal residues are consistent with the loss of intersubunit contacts made by the helix as previously suggested (13) and explicitly described in a CXCL8 monomer (48). However, despite these small variations, the two proteins overall appear to exhibit very similar secondary structures. Analysis of the chemical shift assignments with TALOS (49) predicts ϕ and ψ backbone torsion angles that corroborate our secondary structure analysis (data not shown). Furthermore, the ^{13}C NMR chemical shifts can distinguish between reduced, free

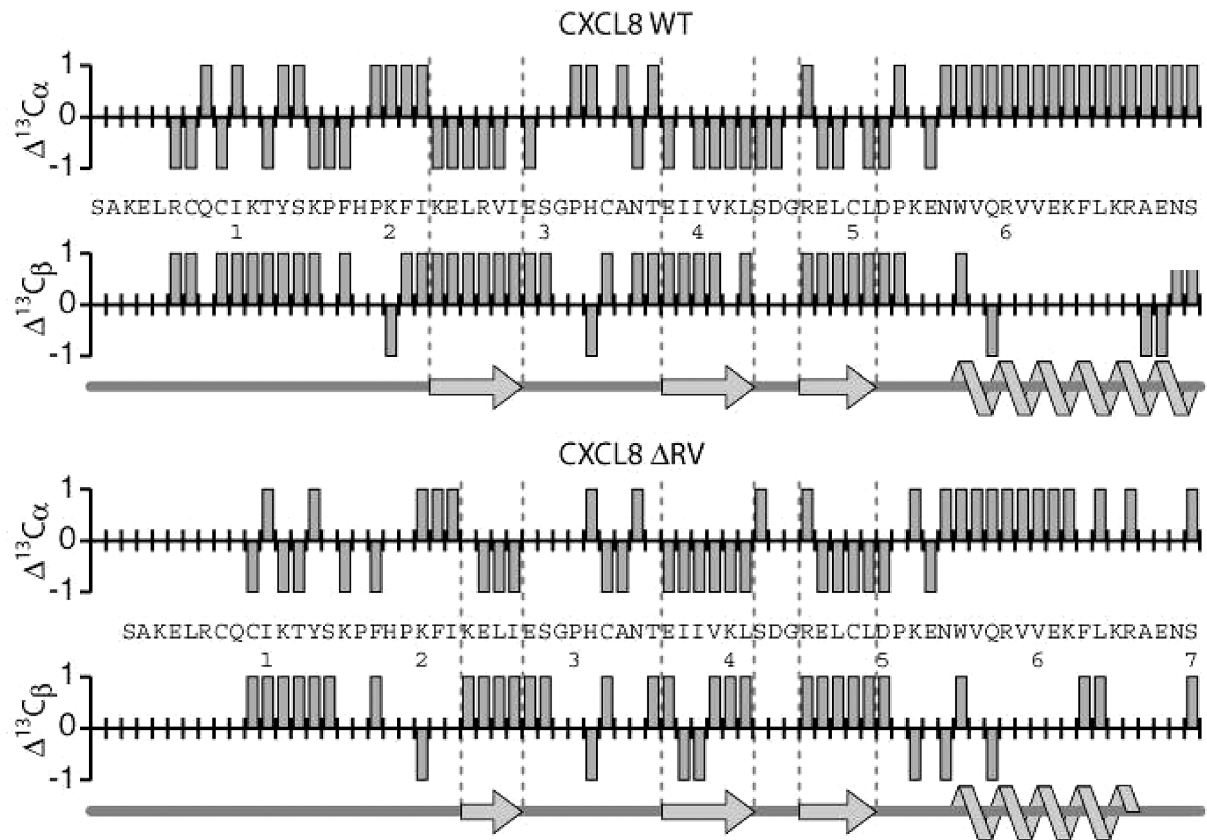


FIGURE 8: Chemical shift index values of $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ deviations from random coil chemical shift values relative to DSS (2,2-dimethyl-2-silapentane-5-sulfonate sodium salt) (31). Data for WT CXCL8 are shown on top and those for the ΔRV variant on the bottom. The vertical scale for the index is either -1 , 0 , or 1 . For each, the sequence is shown with the residue number underneath marked by tens. Also, the secondary structures of sheets are helices and are noted, and dashed lines are provided for reference. For better comparison, the ΔRV sequence is shifted over by two residues at the start. Note for the ΔRV variant, the $\beta 1$ strand is shorter and α -helix is four residues shorter than that of the wild type, as found for monomeric CXCL8 (13, 48).

Table 2: Cysteine ^{13}C Chemical Shifts					
^{13}C	CXCL8	C7	C9	C34 ^a	C50 ^a
C α	WT ^b	55.69	55.38	56.93	56.45
	² RV ^b	55.57	55.95	55.20	55.62
C β	WT ^b	39.49	44.23	42.15	48.29
	² RV ^b	39.27	42.49	41.77	46.06

^a These two residues are references to the WT CXCL8 sequence. Because of the two-residue deletion in ΔRV , these residues are actually residues 32 and 48 in the ΔRV sequence, respectively. ^b Chemical shift values are in parts per million.

cysteine and oxidized, disulfide-bonded cysteine (44). Since the chemical shift assignments for the backbone atoms of WT and ΔRV have been obtained using a series of three-dimensional experiments, we can easily compare the participation of cysteines in disulfide bonds between the two variants. As shown in Table 2, all of the $^{13}\text{C}\beta$ chemical shift values for the four cysteines in ΔRV are similar to those of WT CXCL8 in that they are greater than the cutoff value of 32 ppm for the reduced cysteine (44), which indicates that all cysteines in the ΔRV mutant are in the oxidized disulfide form, like the wild-type protein. In addition, the ΔRV 's cysteine $^{13}\text{C}\beta$ and $^{13}\text{C}\alpha$ chemical shift values deviate only slightly from the WT CXCL8 values. Not only do these results show that the ΔRV mutant's cysteines are disulfide-bonded, but they also indicate that the ΔRV disulfides inhabit an environment similar to that of the wild-type protein. Therefore, ΔRV very likely shares the same disulfide pattern as WT CXCL8. Previous studies on the formation of these

disulfide bonds have shown that they are integral to chemokine structure (50, 51). In much the same spirit that Anfinsen's experiments used disulfide bonding patterns to prove correct refolding (52), the fact that ΔRV makes the same two disulfide bonds as WT CXCL8 (from $\beta 0$ to $\beta 3$ strands and from $\beta 0$ to $\beta 2$ strands) indicates that the β -sheet structure is like the WT. This evidence of correct chemokine β -sheet structure in conjunction with good dispersion in the ^{15}N – ^1H correlation spectra and correct secondary structure formation strongly argues that, like the V27R mutant, the drastic ΔRV mutant retains a chemokine fold.

DISCUSSION

In this investigation of differential chemokine quaternary structure, we applied a hybrid sequence analysis to identify conserved motifs within one chemokine protein interface and have experimentally shown that these motifs inhibit dimerization across the $\beta 1$ strand. In an effort to maximize the signal over noise, the sequence analysis considered structure (localization of comparison to a particular portion involving the CXCL8 dimer interface) and phylogeny (limiting the comparison to clades consisting of proteins closely related in evolution). In comparing CXCL8's sequence with the CCL4 and CCL2 sequence clades, this analysis reveals that these two chemokine families conserved different motifs over their $\beta 1$ strands (Figure 3). For the CCL4 sequence clade, a surprising two-residue deletion is found in the middle of the $\beta 1$ strand, while for the CCL2 sequence clade, a simple

V27R point mutation is found. As shown in Figure 3, limiting the sequence analysis is important. Crystallographic evidence shows that certain CC chemokine families can dimerize like CXCL8 (53, 54). Including these sequences in our analysis would certainly cause ambiguity in the results, and as shown in Figure 3A, mutations suggested by a straightforward sequence alignment would not have found these sequence elements. In addition, because the double deletion was located in the middle of a strand, a structural alignment would have certainly assigned the missing residues to either side of the strand in loop regions and missed the importance of the deletions' position.

Placing the Δ RV and V27R negative interaction motifs in the CXCL8 background, we experimentally confirmed that these sequences affect CXCL8's ability to dimerize across the β 1 strand without changing the chemokine fold. AUC and SEC experiments indicate that the two CXCL8 variants form monomers and no longer dimerize (Figures 4 and 5, respectively). NMR HSQC analysis shows good peak dispersion for both of the CXCL8 variants (Figure 6), where the V27R variant displays a pattern similar to that of folded monomers found in previous work (29). Because the Δ RV mutation is more drastic and its HSQC spectrum is so different, it could be argued that the Δ RV mutation causes some rearrangement in chemokine structure in either the β -sheet, β 1 strand, or helix. The NMR evidence showing the correct disulfide pattern (Table 2) verifies that the Δ RV variant forms a chemokine β -sheet, at least among the β 0, β 2, and β 3 strands, while the Δ RV mutant's good dispersion in the ^{15}N – ^1H correlation spectra results from a folded structure (Figure 7). In Figure 8, the $^{13}\text{C}\alpha$ and $^{13}\text{C}\beta$ deviation from random coil analysis (43) indicates that the β 1 strand is formed and shows C-terminal fraying of the α -helix, which is indicative of a folded monomer in solution (48). In comparison with previous work on creating CXCL8 monomers (13, 14, 16, 18, 19, 35), these results demonstrate that our hybrid sequence analysis provides a straightforward and successful approach to identifying sequence motifs that inhibit dimer formation. Overall, our results show that the two conserved sequence motifs introduce unfavorable interactions into the CXCL8 interface and inhibit CXCL8 homodimer formation.

The Δ RV and V27R motifs both disrupt CXCL8-type dimerization by more than just affecting the β 1 strand interface. Ostensibly, Δ RV is a gross modification in the middle of a β -strand, yet it retains a chemokine fold. Because the R26A/V27A variant only weakens dimer formation by a $\Delta\Delta G$ of ~ 12.9 kJ/mol, the Δ RV's effects include the removal of the side chain interactions across the β 1 strand and the double deletion alters interactions of the downstream turn and α -helix (Figures 2 and 7). Both of these are involved in the dimer interface. As the double deletion derives from the CCL4 clade, this supposition is corroborated by the shortened β 1 strand and downstream loop shown by CCL4. Figure 8 provides a clearer, better explanation. The β 1 strand is shorter and the α -helix frayed toward the C-terminus in comparison to those of WT CXCL8, although the frayed α -helix is characteristic of a chemokine monomer (48). Such structural changes do not favor dimer formation. By contrast, as a modest, single-point mutation, the V27R substitution introduces a long positively capped side chain into the center of the CXCL8 interface that most likely creates an electro-

static repulsion with the adjoining R26 residue and a steric obstruction to the correct placement of the α -helices during dimer formation. A similar negative interaction has been found in a survey of protein β -sheet structures (3) and was used to inhibit dimerization in protein design (2). While this mutation has been shown in conjunction with the L25Y mutation to cause a chemokine monomer (16), our work is the first reported, single-natural amino acid substitution shown to disrupt the CXCL8 homodimer and retain the chemokine fold, where the other single-amino acid substitution was a non-natural L-methylleucine at residue 25 (14) (Table 1). These results for both the Δ RV and V27R negative interaction elements agree with previous work showing that mutations in both the β 1 strand and α -helix are necessary to disrupt the CXCL8 dimer (13).

The finding of conserved sequence motifs that disfavor a certain type of homodimerization has several implications. First, these results suggest that at least one role of these motifs in the CCL4 and CCL2 sequence clades is to inhibit the more regular protein–protein interaction involving the β 1 strand. As displayed in Figure 3, these motifs are strongly conserved. Δ RV is found in all 29 sequences in the CCL4 clade (blue subtree of Figure S1), while V27R is found in the majority of the 16 sequences in the CCL2 clade (red subtree of Figure S1). For the other instances in the CCL2 clade, the residue is another positively charged lysine residue. This conservation strongly suggests that there is some evolutionary pressure to retain these motifs. One such pressure could be to favor N-terminal homodimerization. Since crystallographic structures have shown that it is possible to see chemokine tetramers with both types of interfaces (55, 56), we can think of the various homodimerization states being in equilibrium. By disfavoring the β 1 strand interface, the Δ RV and V27R sequence motifs indirectly favor homodimerization at the N-terminal interface. Because the exact function of homodimerization with chemokines is still being debated (15, 18, 19, 41, 57–63), we cannot say for certain that the primary role of these motifs is to indirectly favor CC-type homodimerization. However, with new evidence of chemokine heterodimerization (64), the inhibitory motifs could take on new significance by hindering cross reactivity with other chemokines in a CXC-like dimer. Because of the striking evolutionary persistence and the clear experimental findings, our results do support the conclusion that the Δ RV and V27R sequence elements are examples of conserved motifs conferring unfavorable interactions against CXC-type dimerization.

ACKNOWLEDGMENT

We thank Rosemarie Swanson for her help with this work and J. Bradley Holmes for help with the surface area code. We would also like to thank Dr. Karl Koshlap and Xiangming Kong of the Biomolecular NMR lab at TAMU for help setting up the experiments.

SUPPORTING INFORMATION AVAILABLE

Further information about evolutionary relationships, complete AUC data, extra NMR analysis, and details of tertiary and quaternary contacts. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Hecht, M. H., Richardson, J. S., Richardson, D. C., and Ogden, R. C. (1990) De novo design, expression, and characterization of Felix: A four-helix bundle protein of native-like sequence. *Science* 249, 884–891.
- Wang, W., and Hecht, M. H. (2002) Rationally designed mutations convert de novo amyloid-like fibrils into monomeric β -sheet proteins. *Proc. Natl. Acad. Sci. U.S.A.* 99, 2760–2765.
- Richardson, J. S., and Richardson, D. C. (2002) Natural β -sheet proteins use negative design to avoid edge-to-edge aggregation. *Proc. Natl. Acad. Sci. U.S.A.* 99, 2754–2759.
- Kortemme, T., Joachimiak, L. A., Bullock, A. N., Schuler, A. D., Stoddard, B. L., and Baker, D. (2004) Computational redesign of protein-protein interaction specificity. *Nat. Struct. Mol. Biol.* 11, 371–379.
- Desjarlais, J. R., and Lazar, G. A. (2003) Negative design for improved therapeutic proteins. *Trends Biotechnol.* 21, 425–427.
- Havranek, J. J., and Harbury, P. B. (2003) Automated design of specificity in molecular recognition. *Nat. Struct. Biol.* 10, 45–52.
- Kim, S., Tsai, J., Kagiampakis, I., LiWang, P., and Vannucci, M. (2007) Detecting protein dissimilarities in multiple alignments using Bayesian variable selection. *Bioinformatics* 23, 245–246.
- Gerard, C., and Rollins, B. J. (2001) Chemokines and disease. *Nat. Immunol.* 2, 108–115.
- Gouwy, M., Struyf, S., Catusse, J., Proost, P., and Van Damme, J. (2004) Synergy between proinflammatory ligands of G protein-coupled receptors in neutrophil activation and migration. *J. Leukocyte Biol.* 76, 185–194.
- Rosenkilde, M. M., and Schwartz, T. W. (2004) The chemokine system: A major regulator of angiogenesis in health and disease. *APMIS* 112, 481–495.
- Schwarz, M. K., and Wells, T. N. (2002) New therapeutics that modulate chemokine networks. *Nat. Rev. Drug Discovery* 1, 347–358.
- Tanase, S., and Nomiyama, H. (1995) Cytokine Family Database. <http://cytokine.medic.kumamoto-u.ac.jp/> (accessed July 3, 2008).
- Lowman, H. B., Fairbrother, W. J., Slagle, P. H., Kabakoff, R., Liu, J., Shire, S., and Hebert, C. A. (1997) Monomeric variants of IL-8: Effects of side chain substitutions and solution conditions upon dimer formation. *Protein Sci.* 6, 598–608.
- Lusti-Narasimhan, M., Chollet, A., Power, C. A., Allet, B., Proudfoot, A. E., and Wells, T. N. (1996) A molecular switch of chemokine receptor selectivity. Chemical modification of the interleukin-8 Leu25 \rightarrow Cys mutant. *J. Biol. Chem.* 271, 3148–3153.
- Proudfoot, A. E., Handel, T. M., Johnson, Z., Lau, E. K., LiWang, P., Clark-Lewis, I., Borlat, F., Wells, T. N., and Kosco-Vilbois, M. H. (2003) Glycosaminoglycan binding and oligomerization are essential for the in vivo activity of certain chemokines. *Proc. Natl. Acad. Sci. U.S.A.* 100, 1885–1890.
- Williams, M. A., Cave, C. M., Quaid, G., Robinson, C., Daly, T. J., Witt, D., Lentsch, A. B., and Solomkin, J. S. (2005) Interleukin 8 dimerization as a mechanism for regulation of neutrophil adherence-dependent oxidant production. *Shock* 23, 371–376.
- Nesmelova, I. V., Sham, Y., Dudek, A. Z., van Eijk, L. I., Wu, G., Slungaard, A., Mortari, F., Griffioen, A. W., and Mayo, K. H. (2005) Platelet factor 4 and interleukin-8 CXC chemokine heterodimer formation modulates function at the quaternary structural level. *J. Biol. Chem.* 280, 4948–4958.
- Leong, S. R., Lowman, H. B., Liu, J., Shire, S., Deforge, L. E., Gillice-Castro, B. L., McDowell, R., and Hebert, C. A. (1997) IL-8 single-chain homodimers and heterodimers: Interactions with chemokine receptors CXCR1, CXCR2, and DARC. *Protein Sci.* 6, 609–617.
- Rajaratnam, K., Sykes, B. D., Kay, C. M., Dewald, B., Geiser, T., Baggiolini, M., and Clark-Lewis, I. (1994) Neutrophil activation by monomeric interleukin-8. *Science* 264, 90–92.
- Fischer, T. B., Holmes, J. B., Miller, I. R., Parsons, J. R., Tung, L., Hu, J. C., and Tsai, J. (2006) Assessing methods for identifying pair-wise atomic contacts across binding interfaces. *J. Struct. Biol.* 153, 103–112.
- Voronoi, G. F. (1908) Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *J. Reine Angew. Math.* 134, 198–287.
- Clore, G. M., Appella, E., Yamada, M., Matsushima, K., and Gronenborn, A. M. (1990) Three-dimensional structure of interleukin 8 in solution. *Biochemistry* 29, 1689–1696.
- Higgins, D. G., Thompson, J. D., and Gibson, T. J. (1996) Using CLUSTAL for multiple sequence alignments. *Methods Enzymol.* 266, 383–402.
- Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680.
- Morgenstern, B. (2004) DIALIGN: Multiple DNA and protein sequence alignment at BiBiServ. *Nucleic Acids Res.* 32, 33–36.
- Morgenstern, B., Frech, K., Dress, A., and Werner, T. (1998) DIALIGN: Finding local similarities by multiple sequence alignment. *Bioinformatics* 14, 290–294.
- Kyte, J., and Doolittle, R. F. (1982) A simple method for displaying the hydrophobic character of a protein. *J. Mol. Biol.* 157, 105–132.
- Sha, N., Vannucci, M., Tadesse, M. G., Brown, P. J., Dragoni, I., Davies, N., Roberts, T. C., Contestabile, A., Salmon, M., Buckley, C., and Falciani, F. (2004) Bayesian variable selection in multinomial probit models to identify molecular signatures of disease stage. *Biometrics* 60, 812–819.
- Jin, H., Hayes, G. L., Darbha, N. S., Meyer, E., and LiWang, P. J. (2005) Investigation of CC and CXC chemokine quaternary state mutants. *Biochem. Biophys. Res. Commun.* 338, 987–999.
- Demeler, B. (2005) UltraScan A Comprehensive Data Analysis Software Package for Analytical Ultracentrifugation Experiments. In *Modern Analytical Ultracentrifugation: Techniques and Methods* (Scott, D. J., Harding, S. E., and Rowe, A. J., Eds.) pp 210–229, Royal Society of Chemistry, Cambridge, U.K.
- Wishart, D. S., Bigam, C. G., Yao, J., Abildgaard, F., Dyson, H. J., Oldfield, E., Markley, J. L., and Sykes, B. D. (1995) ^1H , ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR. *J. Biomol. NMR* 6, 135–140.
- Grzesiek, S., Dobeli, H., Gentz, R., Garotta, G., Labhardt, A. M., and Bax, A. (1992) ^1H , ^{13}C , and ^{15}N NMR backbone assignments and secondary structure of human interferon- γ . *Biochemistry* 31, 8180–8190.
- Bax, A., Ikura, M., Kay, L. E., Torchia, D. A., and Tschudin, R. (1990) Comparison of different modes of two-dimensional reverse-correlation NMR for the study of proteins. *J. Magn. Reson.* 86, 304–318.
- Lodi, P. J., Garrett, D. S., Kuszewski, J., Tsang, M. L., Weatherbee, J. A., Leonard, W. J., Gronenborn, A. M., and Clore, G. M. (1994) High-resolution solution structure of the β chemokine hMIP-1 β by multidimensional NMR. *Science* 263, 1762–1767.
- Covell, D. G., Smythers, G. W., Gronenborn, A. M., and Clore, G. M. (1994) Analysis of hydrophobicity in the α and β chemokine families and its relevance to dimerization. *Protein Sci.* 3, 2064–2072.
- Handel, T. M., and Domaille, P. J. (1996) Heteronuclear (^1H , ^{13}C , ^{15}N) NMR assignments and solution structure of the monocyte chemoattractant protein-1 (MCP-1) dimer. *Biochemistry* 35, 6569–6584.
- Burrows, S. D., Doyle, M. L., Murphy, K. P., Franklin, S. G., White, J. R., Brooks, I., McNulty, D. E., Scott, M. O., Knutson, J. R., Porter, D., Young, P. R., and Hensley, P. (1994) Determination of the monomer-dimer equilibrium of interleukin-8 reveals it is a monomer at physiological concentrations. *Biochemistry* 33, 12741–12745.
- Paolini, J. F., Willard, D., Consler, T., Luther, M., and Krangel, M. S. (1994) The chemokines IL-8, monocyte chemoattractant protein-1, and I-309 are monomers at physiologically relevant concentrations. *J. Immunol.* 153, 2704–2717.
- Grasberger, B. L., Gronenborn, A. M., and Clore, G. M. (1993) Analysis of the backbone dynamics of interleukin-8 by ^{15}N relaxation measurements. *J. Mol. Biol.* 230, 364–372.
- Laurence, J. S., LiWang, A. C., and LiWang, P. J. (1998) Effect of N-terminal truncation and solution conditions on chemokine dimer stability: Nuclear magnetic resonance structural analysis of macrophage inflammatory protein 1 β mutants. *Biochemistry* 37, 9346–9354.
- Laurence, J. S., Blanpain, C., Burgner, J. W., Parmentier, M., and LiWang, P. J. (2000) CC chemokine MIP-1 β can function as a monomer and depends on Phe13 for receptor binding. *Biochemistry* 39, 3401–3409.
- Tsai, J., Gerstein, M., and Levitt, M. (1997) Simulating the minimum core for hydrophobic collapse in globular proteins. *Protein Sci.* 6, 2606–2616.

43. Wishart, D. S., Sykes, B. D., and Richards, F. M. (1992) The chemical shift index: A fast and simple method for the assignment of protein secondary structure through NMR spectroscopy. *Biochemistry* 31, 1647–1651.
44. Sharma, D., and Rajarathnam, K. (2000) ^{13}C NMR chemical shifts can predict disulfide bond formation. *J. Biomol. NMR* 18, 165–171.
45. Wishart, D. S., and Sykes, B. D. (1994) Chemical shifts as a tool for structure determination. *Methods Enzymol.* 239, 363–392.
46. Spera, S., and Bax, A. (1991) Empirical correlation between protein backbone conformation and $\text{C}\alpha$ and $\text{C}\beta$ ^{13}C nuclear magnetic resonance chemical shifts. *J. Am. Chem. Soc.* 113, 5490–5492.
47. Williamson, D., McLennan, I. J., Bax, A., Gamcsik, M. P., and Glickson, J. D. (1990) Two-dimensional NMR study of bleomycin and its zinc(II) complex: Reassignment of ^{13}C resonances. *J. Biomol. Struct. Dyn.* 8, 375–398.
48. Rajarathnam, K., Clark-Lewis, I., and Sykes, B. D. (1995) ^1H NMR solution structure of an active monomeric interleukin-8. *Biochemistry* 34, 12983–12990.
49. Cornilescu, G., Delaglio, F., and Bax, A. (1999) Protein backbone angle restraints from searching a database for chemical shift and sequence homology. *J. Biomol. NMR* 13, 289–302.
50. Rajagopalan, L., Chin, C. C., and Rajarathnam, K. (2007) Role of intramolecular disulfides in stability and structure of a noncovalent homodimer. *Biophys. J.* 93, 2129–2134.
51. Rajarathnam, K., Sykes, B. D., Dewald, B., Baggiolini, M., and Clark-Lewis, I. (1999) Disulfide bridges in interleukin-8 probed using non-natural disulfide analogues: Dissociation of roles in structure from function. *Biochemistry* 38, 7653–7658.
52. Anfinsen, C. B. (1973) Principles that govern the folding of protein chains. *Science* 181, 223–230.
53. Malik, Z. A., and Tack, B. F. (2006) Structure of human MIP-3 α chemokine. *Acta Crystallogr. F* 62, 631–634.
54. Meunier, S., Bernassau, J. M., Guillemot, J. C., Ferrara, P., and Darbon, H. (1997) Determination of the three-dimensional structure of CC chemokine monocyte chemoattractant protein 3 by ^1H two-dimensional NMR spectroscopy. *Biochemistry* 36, 4412–4422.
55. Lubkowski, J., Bujacz, G., Boque, L., Domaille, P. J., Handel, T. M., and Wlodawer, A. (1997) The structure of MCP-1 in two crystal forms provides a rare example of variable quaternary interactions. *Nat. Struct. Biol.* 4, 64–69.
56. Swaminathan, G. J., Holloway, D. E., Colvin, R. A., Campanella, G. K., Papageorgiou, A. C., Luster, A. D., and Acharya, K. R. (2003) Crystal structures of oligomeric forms of the IP-10/CXCL10 chemokine. *Structure* 11, 521–532.
57. Clark-Lewis, I., Kim, K. S., Rajarathnam, K., Gong, J. H., Dewald, B., Moser, B., Baggiolini, M., and Sykes, B. D. (1995) Structure-activity relationships of chemokines. *J. Leukocyte Biol.* 57, 703–711.
58. Zhang, Y., and Rollins, B. J. (1995) A dominant negative inhibitor indicates that monocyte chemoattractant protein 1 functions as a dimer. *Mol. Cell. Biol.* 15, 4851–4855.
59. Williams, G., Borkakoti, N., Bottomley, G. A., Cowan, I., Fallowfield, A. G., Jones, P. S., Kirtland, S. J., Price, G. J., and Price, L. (1996) Mutagenesis studies of interleukin-8. Identification of a second epitope involved in receptor binding. *J. Biol. Chem.* 271, 9579–9586.
60. Paavola, C. D., Hemmerich, S., Grunberger, D., Polsky, I., Bloom, A., Freedman, R., Mulkins, M., Bhakta, S., McCarley, D., Wiesent, L., Wong, B., Jarnagin, K., and Handel, T. M. (1998) Monomeric monocyte chemoattractant protein-1 (MCP-1) binds and activates the MCP-1 receptor CCR2B. *J. Biol. Chem.* 273, 33157–33165.
61. Fernando, H., Chin, C., Rosgen, J., and Rajarathnam, K. (2004) Dimer dissociation is essential for interleukin-8 (IL-8) binding to CXCR1 receptor. *J. Biol. Chem.* 279, 36175–36178.
62. Rajarathnam, K., Prado, G. N., Fernando, H., Clark-Lewis, I., and Navaro, J. (2006) Probing receptor binding activity of interleukin-8 dimer using a disulfide trap. *Biochemistry* 45, 7882–7888.
63. Jin, H., Shen, X., Baggett, B. R., Kong, X., and LiWang, P. J. (2007) The human CC chemokine MIP-1 β dimer is not competent to bind to the CCR5 receptor. *J. Biol. Chem.* 282, 27976–27983.
64. Crown, S. E., Yu, Y., Sweeney, M. D., Leary, J. A., and Handel, T. M. (2006) Heterodimerization of CCR2 chemokines and regulation by glycosaminoglycan binding. *J. Biol. Chem.* 281, 25438–25446.
65. Kabsch, W., and Sander, C. (1983) Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* 22, 2577–2637.
66. DeLano, W. L. (2002) *PyMOL*, Delano Scientific, San Carlos, CA.
67. Baldwin, E. T., Weber, I. T., St Charles, R., Xuan, J. C., Appella, E., Yamada, M., Matsushima, K., Edwards, B. F., Clore, G. M., and Gronenborn, A. M. (1991) Crystal structure of interleukin 8: Symbiosis of NMR and crystallography. *Proc. Natl. Acad. Sci. U.S.A.* 88, 502–506.
68. Crooks, G. E., Hon, G., Chandonia, J. M., and Brenner, S. E. (2004) WebLogo: A sequence logo generator. *Genome Res.* 14, 1188–1190.
69. McCornack, M. A., Cassidy, C. K., and LiWang, P. J. (2003) The binding surface and affinity of monomeric and dimeric chemokine macrophage inflammatory protein 1 β for various glycosaminoglycan disaccharides. *J. Biol. Chem.* 278, 1946–1956.
70. Garrett, D. S., Powers, R., Gronenborn, A. M., and Clore, G. M. (1991) A common sense approach to peak picking in two-, three-, and four-dimensional spectra using automatic computer analysis of contour diagrams. *J. Magn. Reson.* 95, 214–220.
71. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J. Biomol. NMR* 6 277–293.

BI702288A