Proc. Natl. Acad. Sci. USA
Vol. 88, pp. 502–506, January 1991
Biochemistry

Crystal structure of interleukin 8: Symbiosis of NMR and crystallography

(chemotaxis/molecular replacement/receptor binding/growth factors)

ERIC T. BALDWIN*†, IRENE T. WEBER*, ROBERT ST. CHARLES*, JIAN-CHENG XUAN*, ETTORE APPELLA‡, MASAKI YAMADA§, KOUJI MATSUSHIMA¶, BRIAN F. P. EDWARDS‖, G. MARIUS CLORE**, ANGELA M. GRONENBORN**, AND ALEXANDER WLODAWER*

*Crystallography Laboratory, National Cancer Institute–Frederick Cancer Research and Development Center, ABL-Basic Research Program, Frederick, MD 21702; ‡Laboratory of Cell Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892; §Research Laboratory, Dainippon Pharmaceutical Co., Ltd., Suita, Osaka 564, Japan; ¶Laboratory of Molecular Immunoregulation, Biological Response Modifiers Program, National Cancer Institute, Frederick, MD 21702; ‖Departments of Biochemistry and Physiology, Wayne State University School of Medicine, Detroit, MI 48201; and **Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892

Communicated by Brian W. Matthews, October 11, 1990 (received for review August 3, 1990)

ABSTRACT **The crystal structure of a host defense system chemotactic factor, interleukin 8, has been solved by molecular replacement using as a model the solution structure derived from nuclear magnetic resonance experiments. The structure was refined with 2 Å x-ray data to an *R* factor of 0.187 (0.217 at 1.6 Å). A comparison indicates some potential differences between the structure in solution and in the crystalline state. Our analysis also predicts that residues 4 through 9 on the amino terminus and the β-bend, which includes His-33, may be important for receptor binding.**

Nuclear magnetic resonance (NMR) has been established in the last 5 years as the only viable alternative to x-ray crystallography for determining detailed three-dimensional structures of macromolecules up to 20 kDa molecular mass (1, 2). Although test calculations using crambin (3), tendamistat (4), and ubiquitin (5) have established the feasibility of utilizing NMR-derived models as starting points for solving x-ray structures by molecular replacement, no unknown crystal structures have been previously solved by that route. We present here the results of such an investigation. The high-resolution x-ray structure of interleukin 8 (IL-8) was solved by a combination of NMR (6, 7) and crystallography, and the models obtained by these two techniques were compared.††

IL-8 is a small, chemotactic factor released by monocytes and a number of other cell types in response to inflammatory stimuli (8). The factor promotes the directed migration of neutrophils, basophils, and T cells (9) along a concentration gradient to the inflammatory focus. IL-8 has also been reported to initiate the oxidative burst response in neutrophils. The factor has been known as monocyte-derived neutrophil chemotactic factor (10) or neutrophil-activating protein (11, 12) among others, but it is now generally designated IL-8 (9). IL-8 is a member of a family of similar immune system proteins including the monocyte chemoattractant factor (13), human melanoma growth-stimulating activity (14), murine macrophage inflammatory protein 2 (15), and bovine platelet factor 4 (BPF4) (16). Each of these protein sequences shows a conserved region containing four cysteines that form two disulfide bridges. Under the experimental conditions of the structural investigations IL-8 is a dimer (6, 7) and BPF4 is a tetramer (17).

Materials and Methods

Recombinant human IL-8 that was used in the crystallographic studies was purified from *Escherichia coli* (18). All of the x-ray data used for this structure determination were collected on crystals that were grown using protein recycled from the NMR experiments of Clore and Gronenborn and coworkers (6, 7). Crystals belonging to the space group $P3_121$, $a = b = 40.3$ Å, $c = 90.1$ Å, were grown from 40% ammonium sulfate (pH = 8.5) as described (19). Identical crystals have also been grown using chemically synthesized protein. Recently, workers at Sandoz Pharmaceutical have reported that crystals apparently identical to the ones described here could be grown from polyethylene glycol at pH 6.5 (20). The crystals contain a monomer in the asymmetric unit. X-ray data extending to 1.6 Å resolution were collected with a Siemens area detector, using two crystals. The data set consisted of 11,178 reflections (5879 to 2.0 Å) and was 92% complete.

Two independent sets of atomic coordinates were used as starting points for molecular replacement calculations. The coordinates of BPF4 refined at 3 Å resolution were provided by B. Edwards before publication (17), whereas the restrained minimized mean NMR structure was contributed by G. M. Clore and A. Gronenborn (6, 7). The rotation function was solved for both models using the program MERLOT (21). When BPF4 was used as a search model, the signal was very weak with the correct peak never ranking higher than third in the peak list. The translation for this model was determined with COREL S (22) using a method described below but has not been refined with higher resolution data (23). The rotation function was, however, easily interpreted using the restrained minimized mean NMR structure. The correct orientation often ranked first in the peak list with rms peak heights as high as 6.2 σ. The rotation function solutions corresponded to the rotation necessary to orient the IL-8 NMR dimer two-fold axis approximately in the direction of the crystallographic two-fold axis in $P3_121$ ($P3_221$). Thus, it was apparent that the crystallographic two-fold axis related the two halves of the molecular dimer. This circumstance substantially restricts the translation freedom of the dimer and requires that its center of mass lie along the $x = y$ diagonal (24).

The translation function was solved using the rigid-body refinement option in the program COREL S (22) to refine the orientation of the model in all positions. The oriented model was translated in 2.5 Å increments along the two-fold axis. These calculations were performed with 15-9 Å data (63 reflections) in both enantiomorphic space groups and yielded an unambiguous answer in $P3_121$ ($R = 0.44$, correlation

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: IL-8, interleukin 8; BPF4, bovine platelet factor 4.

†To whom reprint requests should be addressed.

††The atomic coordinates have been deposited in the Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, Upton, NY 11973 (reference 3IL8).

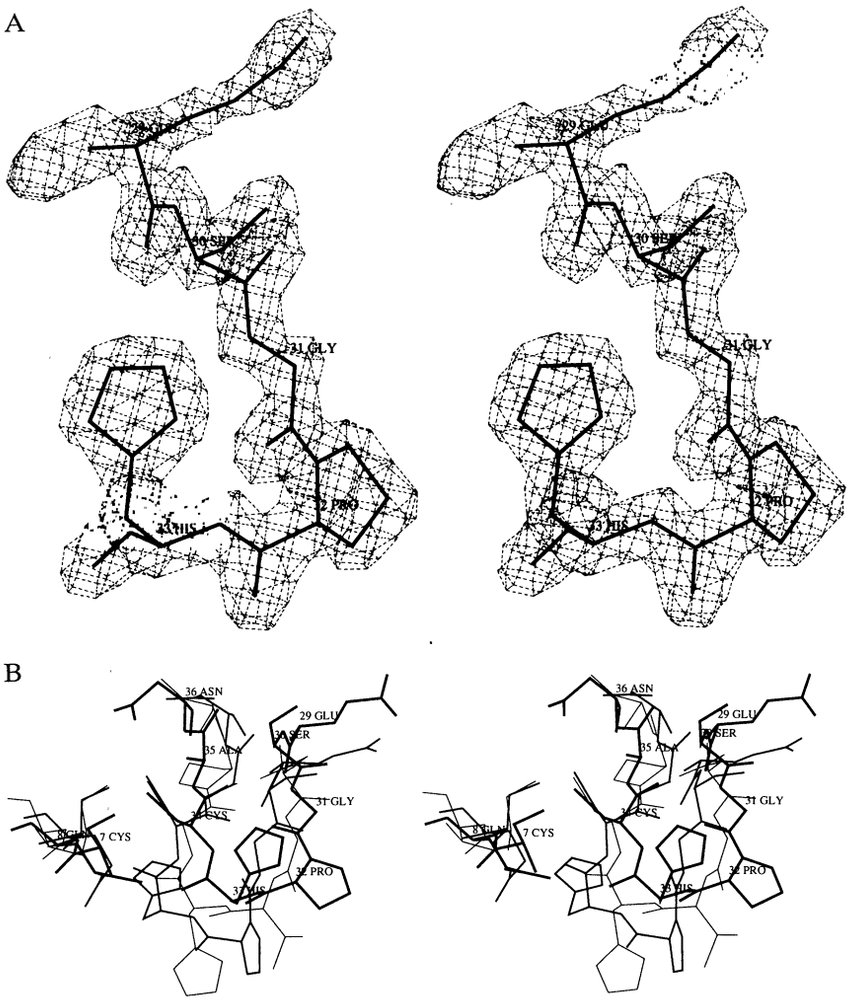

FIG. 1. (*A*) Electron density at 1.6 Å and the refined model for residues 29–33 in IL-8. The 2 F_o-F_c map has been contoured at 1.5 σ level. All atoms are well placed in the density. (*B*) Superposition of the loop 29–36 of the x-ray (thick lines) and NMR (medium lines) models of IL-8 as well as of the corresponding loop in BPF4 (thin lines). Even though the initial model was the IL-8 NMR structure, this loop refined away from this conformation.

coefficient = 0.64). In $P3_221$ the translation with the best *R* factor, 0.498, had a correlation coefficient of 0.428. The translation function calculations were checked with the molecular dynamics-least squares program X-PLOR (25). The X-PLOR translation function for the resolution range of 10-4 Å resolution gave a solution with a correlation coefficient of 0.41 (the next highest peak was 0.334). The *R* factor in $P3_221$ was 0.586 with a correlation coefficient of 0.334.

The refinement of this solution was continued with X-PLOR. Rigid body refinement lowered the *R* factor from 0.514 to 0.501 (10 to 4 Å), and the positional refinement with 10-3 Å data resulted in $R = 0.407$. Finally, a run of simulated annealing with heating to 2000 K reduced the *R* factor to 0.279. Further refinement was continued with the restrained least squares program PROLSQ (26). An implementation including fast Fourier transforms (27) and restraining intermolecular contacts (28) was used. Several cycles of refinement, interspersed with manual rebuilding using computer graphics and the addition of solvent molecules, led to the current model, characterized by an *R* factor of 0.187 at 2.0 Å (0.217 at 1.6 Å) and the deviation of bond lengths from the standard values of 0.019 (0.022). All main chain torsion angles are found in the allowed regions of the Ramachandran plot.

Results and Discussion

The current model of IL-8 based upon x-ray analysis should be still considered preliminary since data extending to resolution higher than 1.5 Å are being collected and solvent structure is not yet finalized, but the quality of the electron density maps is very good (Fig. 1*A*). Residues 4–72 are well placed in the density, with no breaks in the main chain when contoured at 1σ level. No electron density corresponding to the first three residues at the amino terminus can be seen and this part of the structure is considered to be disordered. Most side chains are also unambiguously located in their densities, whereas some surface residues show signs of possible disorder or multiple conformations.

The structure of crystalline IL-8 unambiguously corresponds to a dimeric molecule, with its two-fold axis of rotation bisecting the crystal *x* and *y* axes at $z = 0$ (Fig. 2). The molecular architecture of the monomer consists of three

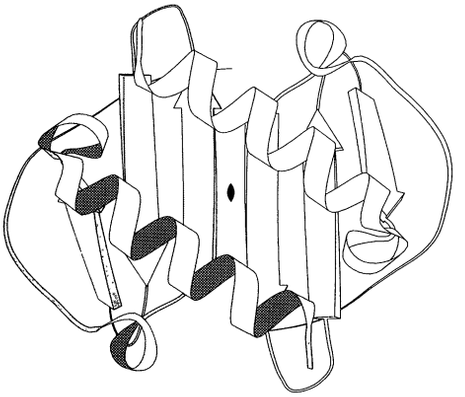

FIG. 2. Ribbon diagram of the structure of the human IL-8 dimer. One of the monomers is shaded, and the two-fold axis is indicated. The dimer is composed of a six-stranded β-sheet and two parallel α-helices.

antiparallel β-strands connected with loops and one long helix made of the carboxyl-terminal residues 57–72 (Fig. 3). The dimer is primarily stabilized by hydrogen bonds between the first β-strand (residues 23–29) in each molecule (Fig. 2) and by additional side chain interactions. Thus the dimeric molecule consists of a six-stranded sheet and of two antiparallel helices inclined at an angle of ≈60° to the β-strand direction. The center-to-center distance between the helices is about 12 Å. The two disulfide bridges, between Cys-7 and Cys-34 and between Cys-9 and Cys-50, have standard torsion angles and are located in clearly defined electron densities.

The crystal structure of IL-8 has been compared with the starting NMR model with the program ALIGN (30). Either the C_α coordinates of the two models or all main chain atoms were superimposed, and atoms that were deviating by more than 3 Å after best superposition were not used in the calculation of rms deviations. When the monomers were superimposed, the rms deviation for 63 pairs of C_α atoms was 0.88 Å, with residues 4–6, 10, 31–33, and 36 showing shifts larger than 2 Å (Table 1, Fig. 4). The residues at the amino terminus had standard deviations ≥5 Å about their positions in the 30 NMR coordinate sets and should be considered undefined. The main difference between the well-defined parts of the models is in the loop containing residues 29–36.

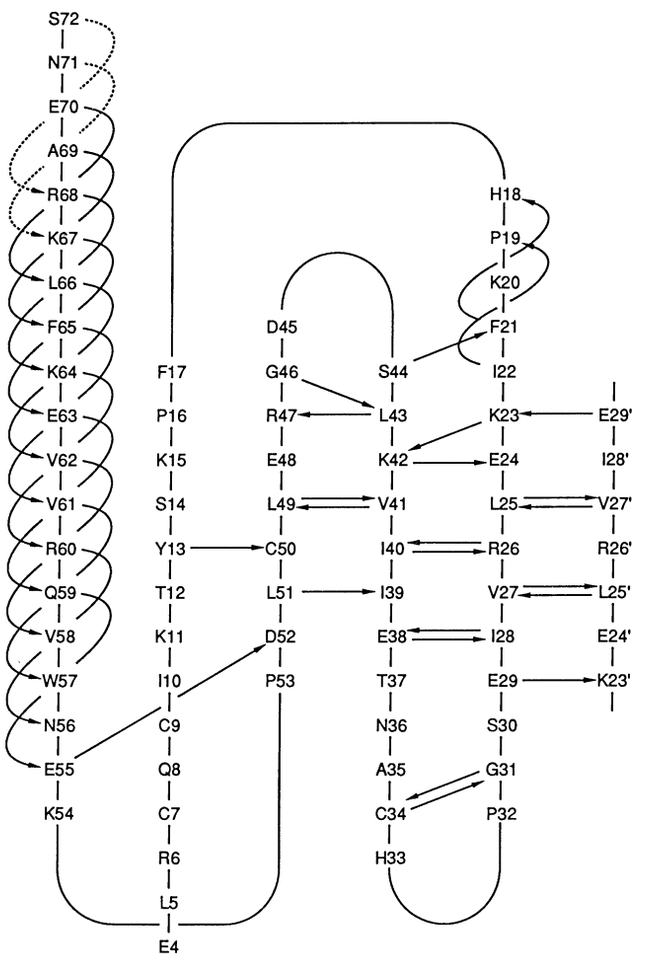

FIG. 3. Hydrogen bonding pattern of IL-8 in the crystal. The three strands of antiparallel β-sheet are related by the two-fold axis between the strands at residues 26. Thus, a six-stranded β-sheet is formed in the dimer. Beginning at Pro-19 there is a single turn of a 3_{10} helix (29). The hydrogen bond responsible for the low pK of His-18 between the ND1 nitrogen and the backbone amide of Lys-20 is slightly shorter in the x-ray structure. Some β-sheet stabilizing hydrogen bonds supported by the NMR structure are longer than 3.5 Å in the x-ray structure.

Table 1. Comparison of the overall rms deviation for α-carbons and main chain for BPF4, IL-8 NMR, and IL-8 x-ray

	IL-8 x-ray	IL-8 NMR	BPF4
IL-8 x-ray			
C_α		2.03 (97)	1.95 (100)
Main		1.95 (95)	2.02 (99)
IL-8 NMR			
C_α	0.88 (91)		1.65 (100)
Main	0.96 (92)		1.71 (100)
BPF4			
C_α	1.70 (100)	1.56 (100)	
Main	1.70 (99)	1.60 (99)	

When either monomers or dimers are compared, the NMR structure of IL-8 is more similar to BPF4 than is the x-ray structure of IL-8, and this is especially true when dimers are compared. The rms deviation for each set of structures is calculated with obvious outliers removed (>3σ). The percentage of atoms included in each calculation is indicated in parentheses. When the NMR and x-ray structures of IL-8 were compared, residues 4–6 and 31–33 were rejected by ALIGN.

(This loop is covalently attached to the amino-terminal strand through the disulfide bridge between residues 7 and 34.) The most striking difference between the two structures is in the position of His-33 (Fig. 1*B*). In the NMR structure, His-33 accepts a hydrogen bond through the NE2 atom from the main chain amide of Gln-8, whereas in the x-ray structure the NE2 atom donates a hydrogen bond to the carbonyl of Glu-29. The other histidine present in IL-8 (His-18) makes an identical hydrogen bond in both structures with the amide nitrogen of Lys-20. It is interesting to note that in solution both His-18 and His-33 have low pK values [3.7 and 4.9, respectively (6, 7)], which can be due to hydrogen bonding from a backbone amide to the imidazole ring. Although His-33 is unambiguously positioned in the electron density map (Fig. 1*A*), the crystal structure does not explain the low pK of this residue since its side chain is neither bound to an amide nitrogen nor adjacent to any charged residues. This result points to a genuine difference between the structures in the solution and crystalline states.

Superposition of the two dimers shows a much larger range of variations, with the rms deviation for the 134 C_α atoms of 2.03 Å (Table 1, Fig. 4). This is primarily due to the smaller twist of the β-sheet observed in the x-ray structure, which brings the two helices closer together, to 12 Å rather than the 14 Å seen in the NMR model. Although the total shift between the two structures is much larger for the dimers, the maximum shifts are lower, since the positions of the loop 29–36 are closer to each other in that case.

The currently available model of BPF4, refined with 3 Å data, was also compared with the x-ray model of IL-8 (Table 1). The overall rms deviation for the 62 pairs of C_α atoms is 1.70 Å for monomer superpositions. Although this deviation is much larger than between the NMR and x-ray models of IL-8, the maximum deviation is smaller, indicating that the differences are less localized. In particular, the main chain of the BPF4 at the equivalent IL-8 loop 29–36 is intermediate between the two IL-8 structures, with residues 31–32 being closer to the IL-8 x-ray structure and residues 33–35 closer to the IL-8 NMR structure. The side chain of His-50 (equivalent to His-33 in IL-8) is in a very different orientation (Fig. 1*B*).

Comparison of the NMR and x-ray structures of IL-8 shows that the deviations between these two models are comparable to those reported previously for other proteins. The deviations between two structures of tendamistat were 1.05 Å for the backbone atoms and 1.84 Å for all heavy atoms, with some of these effects clearly attributed to the influence of crystal contacts (31). Similarly, squash trypsin inhibitor, a

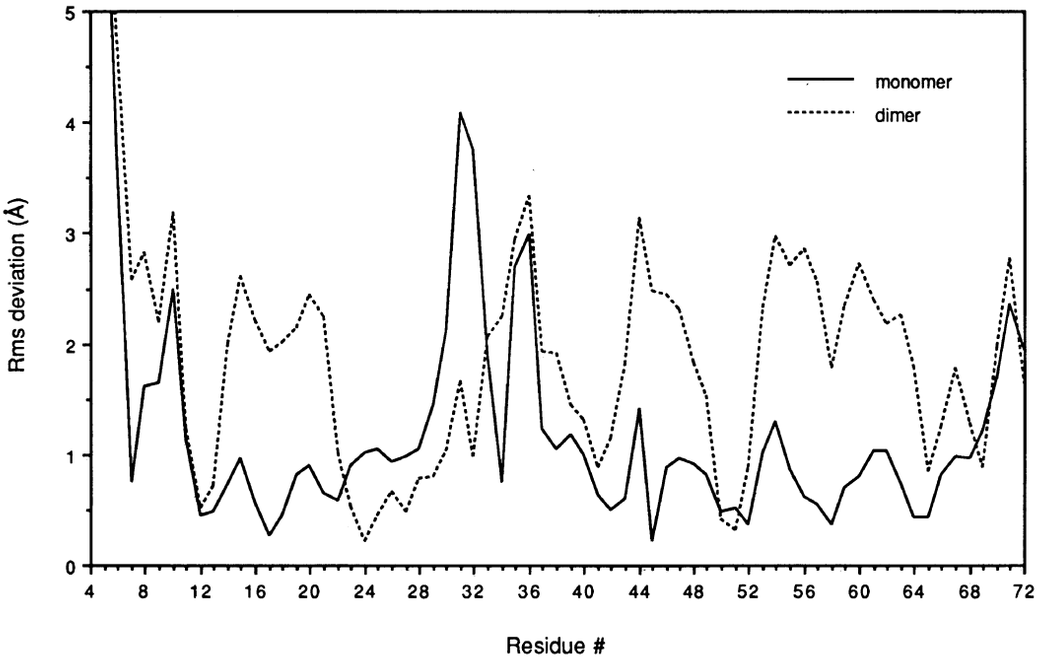

FIG. 4. Deviations between the C_α atomic positions of the NMR and x-ray models of the monomer and dimer of IL-8.

36-residue polypeptide, deviates by 0.96 Å for the backbone and 1.95 Å for all atoms, for the NMR structure in solution and the x-ray structure in a complex with trypsin (32). In this case, the differences may also be influenced by the complexed state of the protein.

Since the average NMR structure was derived from 30 dimer coordinates resulting from the NMR structure determination, an rms deviation, $\bar{u}^2$, about the mean coordinate for each atom was computed for the IL-8 model and these values were converted to pseudo-crystallographic temperature factors ($B = 8\pi^2\bar{u}^2$). The average temperature factors for the main chain and for the side chains alone are shown in Fig. 5 *A* and *B*, respectively. Where the rms deviation exceeded 1 Å in the NMR model, the *B* factor was set to 80 Å^2. The B factors follow similar trends for both models. The correlation coefficient for the x-ray and NMR main chain *B* factors is

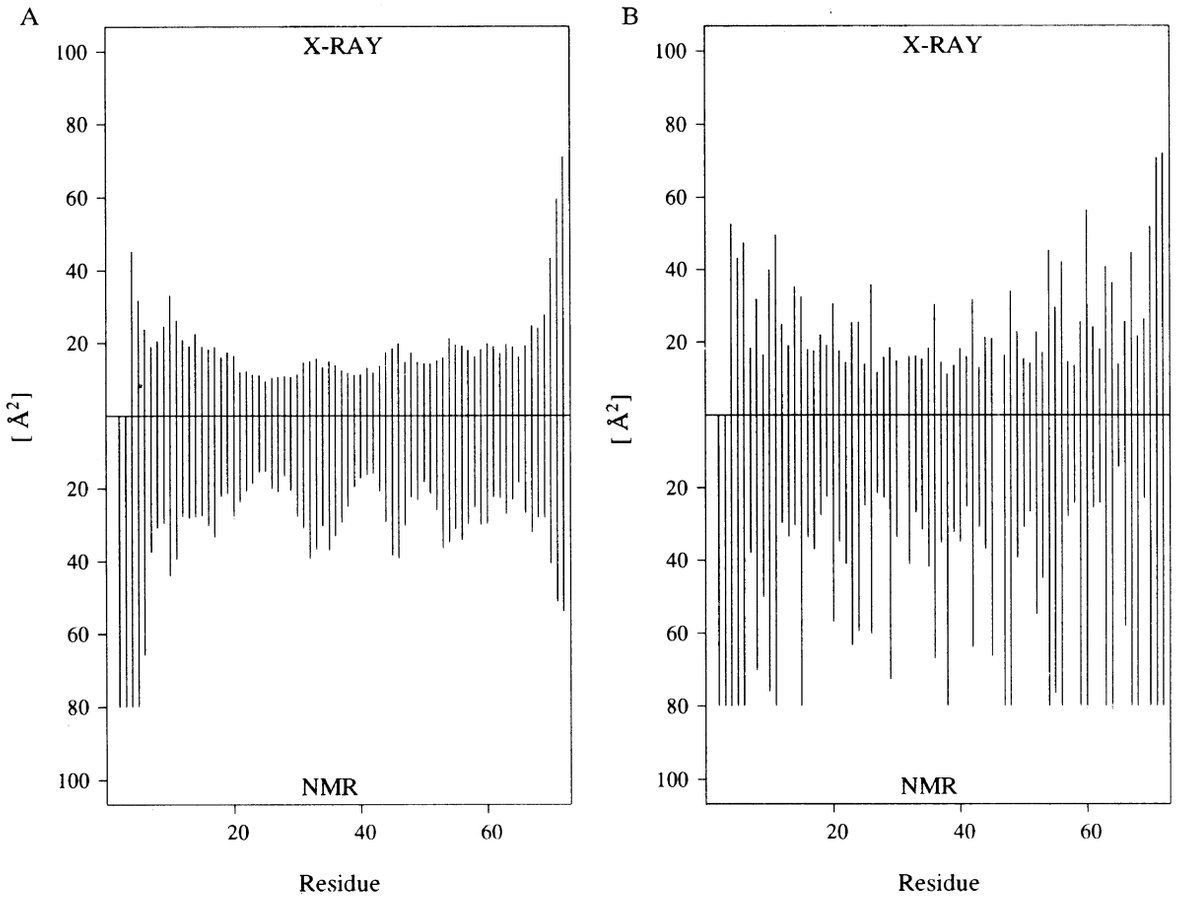

FIG. 5. Comparison of the main chain (*A*) and side chain (*B*) *B* factors for the IL-8 NMR and x-ray models. The mean atomic rms deviations of the 30 individual NMR studies about the mean coordinate positions have been converted to pseudo *B* factors (see text).

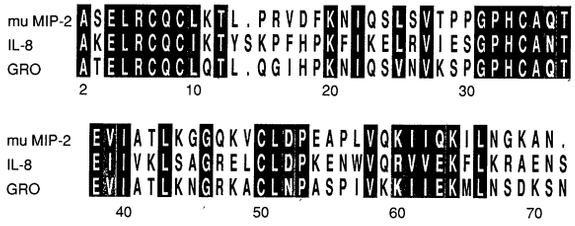

FIG. 6. Comparison of the amino acid sequences of human IL-8, murine macrophage inflammatory protein 2 (mu MIP-2), and human melanoma growth-stimulating activity (GRO). All of these proteins bind to the human IL-8 neutrophil receptor. The analysis indicates a number of identities at the amino terminus and at the region near His-33. Dark bars indicate amino acid identity among all the sequences, whereas shaded bars indicate a conservative substitution.

0.65, and it is 0.66 for the side chains. These data indicate that the flexibility of both structures is well correlated.

The major difference between these two structures lies in the quaternary arrangement of the two subunits of the dimer. The two helices are closer together in the x-ray structure than in the NMR model. Although we do not have an easy explanation for the difference, we note that major differences in domain associations due to the influence of crystal contacts have been reported in the past, in particular for immunoglobulins (33), but the types of interdomain interactions reported there were of a different nature than those for IL-8. Comparison with BPF4 shows the distance between the helices to be identical to the NMR model of IL-8 but different from the IL-8 crystal structure. Finally, attempts to use the crystal structure as a starting model for the NMR structure calculation showed that the smaller interhelix distance cannot be reconciled with the available NMR data. Therefore, the question of the source of these differences cannot be answered at this time.

In any case, the role of the dimer and the exact interhelix distance may be of secondary importance for explaining biological activity of IL-8. Though Clore *et al.* (6, 7) speculated that the IL-8 dimer may interact with its receptor through the two helices, by analogy with HLA (34), this need not necessarily be the case. Since the biological activity of IL-8 is maximal at 10 ng/ml, significant concentrations of monomeric IL-8 may be present under physiological conditions. Calculations of exposed hydrophobic surfaces (35) show identical proportions of 57% for both the monomer and the dimer, indicating that either may be stable in solution. Experiments to determine the monomer–dimer equilibrium constant are necessary. Human melanoma growth-stimulating activity (14), murine macrophage inflammatory protein 2 (15), and IL-8 all bind to the same receptor on neutrophils (36). Detailed comparison of the amino acid sequences (29) of these proteins revealed less homology along the helix (Fig. 6). However, several amino acids at the amino terminus (Glu-4 through Cys-9) and at the β-bend at His-33 (Gly-31 through Glu-38) are conserved. These loops are close together in three dimensions and project from the edge of the β-sheet into solution. We propose that these residues may play an important role for the way in which IL-8 binds to its receptor.

We thank Drs. R. Harrison and J. Sussman for help with structure solution, Dr. J. Moult for the calculation of hydrophobic surfaces, Dr. G. Andrews for synthetic IL-8, and Dr. A. Cerami for the murine macrophage inflammatory protein 2 sequence. The Advanced Scientific Computing Laboratory, Frederick Cancer Research and Development Center, provided a substantial allocation of time on their Cray X-MP supercomputer. This work was supported in part by National Cancer Institute, Department of Health and Human Services, Contract N01-CO-74101 with ABL and by a National Research Service Award postdoctoral award to E.T.B.

1. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids* (Wiley, New York).
2. Clore, G. M. & Gronenborn, A. M. (1989) *CRC Crit. Rev. Biochem. Mol. Biol.* **24,** 479–564.
3. Brünger, A. T., Campbell, R. L., Clore, G. M., Gronenborn, A. M., Karplus, M., Petsko, G. A. & Teeter, M. M. (1987) *Science* **235,** 1049–1053.
4. Braun, W., Epp, O., Wüthrich, K. & Huber, R. (1989) *J. Mol. Biol.* **206,** 669–676.
5. Weber, P. L., Ecker, D. J., Marsh, J., Cooke, S. T. & Mueller, L. (1988) *Trans. Am. Crystallogr. Assoc.* **24,** 91–105.
6. Clore, G. M., Appella, E., Yamada, M., Matsushima, K. & Gronenborn, A. M. (1989) *J. Biol. Chem.* **264,** 18907–18911.
7. Clore, G. M., Apella, E., Yamada, M., Matsushima, K. & Gronenborn, A. (1990) *Biochemistry* **29,** 1689–1696.
8. Matsushima, K. & Oppenheim, J. J. (1989) *Cytokine* **1,** 2–13.
9. Larsen, C. G., Anderson, A. O., Appella, E., Oppenheim, J. J. & Matsushima, K. (1989) *Science* **243,** 1464–1467.
10. Yoshimura, T., Matsushima, K., Tanaka, S., Robinson, E. A., Appella, E., Oppenheim, J. J. & Leonard, E. J. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 9233–9237.
11. Walz, A., Peveri, P., Aschaver, H. & Baggiolini, M. (1987) *Biochem. Biophys. Res. Commun.* **149,** 755–761.
12. Westwick, J., Li, S. W. & Camp, R. D. (1989) *Immunol. Today* **10,** 146–147.
13. Furutani, Y., Nomura, H., Notake, M., Oyamada, Y., Fuku, T., Yamada, M., Larson, C. G., Oppenheim, J. J. & Matsushima, K. (1989) *Biochem. Biophys. Res. Commun.* **159,** 249–255.
14. Richmond, A., Balentein, E., Thomas, H. G., Flaggs, G., Barton, D. E., Spiess, J., Bordoni, R., Fracke, U. & Derynck, R. (1988) *EMBO J.* **7,** 2025–2033.
15. Wolpe, S. D., Sherry, B., Juers, D., Davatelis, D., Yurt, R. W. & Cerami, A. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 612–616.
16. Poncz, M., Surrey, S., Larocco, R., Weiss, M. J., Rappaport, E. F., Conway, T. M. & Schwartz, E. (1987) *Blood* **69,** 219–223.
17. St. Charles, R., Walz, D. A. & Edwards, B. F. P. (1989) *J. Biol. Chem.* **264,** 2092–2099.
18. Furuta, R., Yamagishi, J., Kotani, H., Sakamoto, F., Fukui, T., Matsui, Y., Sohmura, Y., Yamada, M., Yoshimura, T., Larsen, C. G., Oppenheim, J. J. & Matsushima, K. (1989) *J. Biochem. (Tokyo)* **106,** 436–441.
19. Baldwin, E. T., Franklin, K. A., Appella, E., Yamada, M., Matsushima, K., Wlodawer, A. & Weber, I. T. (1990) *J. Biol. Chem.* **265,** 6851–6853.
20. Auer, M., Kallen, J., Schleischitz, S., Walkinshaw, M. D., Wasserbauer, E., Ehn, G. & Lindley, I. J. D. (1990) *FEBS Lett.* **265,** 30–32.
21. Fitzgerald, P. M. D. (1987) *J. Appl. Crystallogr.* **21,** 273–278.
22. Sussman, J., Holbrook, S. R., Church, G. M. & Kim, S.-H. (1977) *Acta Crystallogr. Sect. A* **33,** 800–804.
23. Baldwin, E. T., Sussman, J., St. Charles, R. & Wlodawer, A. (1990) in *Techniques in Protein Chemistry* II, ed. Villafranca, J. J. (Academic, New York), in press.
24. Fehlhammer, H., Schiffer, M., Epp, O., Colman, P. M., Lattman, E. E., Schwager, P., Steigemann, W. & Schramm, H. J. (1975) *Biophys. Struct. Mech.* **1,** 139–146.
25. Brünger, A., Kuriyan, J. & Karplus, M. (1987) *Science* **235,** 458–460.
26. Hendrickson, W. A. (1985) *Methods Enzymol.* **115,** 252–270.
27. Finzel, B. C. (1987) *J. Appl. Crystallogr.* **20,** 53–55.
28. Sheriff, S. (1987) *J. Appl. Crystallogr.* **20,** 55–57.
29. Richardson, J. S. & Richardson, D. C. (1988) *Science* **240,** 1648–1652.
30. Satow, Y., Cohen, G. H., Padlan, E. A. & Davies, D. R. (1986) *J. Mol. Biol.* **190,** 593–604.
31. Billeter, M., Kline, A. D., Braun, W., Huber, R. & Wüthrich, K. (1989) *J. Mol. Biol.* **206,** 677–687.
32. Holak, T. A., Bode, W., Huber, R., Otlewski, J. & Wilusz, T. (1989) *J. Mol. Biol.* **210,** 649–654.
33. Schiffer, M., Ainsworth, C., Xu, Z.-B., Carperos, W., Olsen, K., Solomon, A., Stevens, F. J. & Chang, C.-H. (1989) *Biochemistry* **28,** 4066–4072.
34. Bjorkman, P. J., Saper, M. A., Samraour, B., Bennett, W. S., Strominger, J. L. & Wiley, D. C. (1987) *Nature (London)* **329,** 506–512.
35. Janin J., Miller, S. & Chothia, C. (1988) *J. Mol. Biol.* **204,** 155–164.
36. Mukaida, N., Hisinuma, A., Zachariae, C. O. C., Oppenheim, J. J. & Matsushima, K. (1990) in *Proceedings of the 2nd International Symposium on Chemotactic Cytokines*, eds. Westwick, J., Kunkel, S. & Lindley, I. J. D. (Humana, Clifton, NJ), in press.