

## Modeling the three-dimensional structure of the monocyte chemo-attractant and activating protein MCAF/MCP-1 on the basis of the solution structure of interleukin-8

Angela M.Gronenborn and G.Marius Clore

Laboratory of Chemical Physics, Building 2, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, USA

**A model of the three-dimensional structure of the monocyte chemo-attractant and activating protein MCAF/MCP-1 is presented. The model is predicted based on the previously determined solution structure of interleukin-8 (IL-8/NAP-1) [Clore, G.M., Appella, E., Yamada, M., Matsushima, K. and Gronenborn, A.M. (1990) *Biochemistry* 29, 1689–1696]. Both proteins belong to a superfamily of cytokine proteins involved in cell-specific chemotaxis, host defense and the inflammatory response. The amino acid sequence identity between the two proteins is 24%. It is shown that the regular secondary structure elements of the parent structure can be retained in the modeled structure, such that the backbone hydrogen bonding pattern is very similar in the two structures. The polypeptide backbone is superimposable with an atomic r.m.s. difference of 0.9 Å and all side chains can be modeled by transferring the parent side chain conformation to the new structure. Thus, the deduced structure, like the parent one, is a dimer and consists of a six-stranded antiparallel  $\beta$ -sheet, formed by two three-stranded Greek keys, one from each monomer, upon which lie two symmetry-related antiparallel  $\alpha$ -helices,  $\sim 24$  Å long and separated by  $\sim 14$  Å. All amino acid sequence changes can be accommodated within the parent polypeptide framework without major rearrangements. This is borne out by the fact that the IL-8/NAP-1 and modeled MCAF/MCP-1 structures have similar non-bonding energies. These results strongly suggest that both proteins and all other members of the superfamily most likely have the same tertiary structure. Analysis of the distribution of the solvent-exposed residues can be interpreted in the context of the different receptors involved in mediating the specific responses to both proteins and suggests that the different activities of the two proteins, namely neutrophil (IL-8) versus monocyte (MCAF/MCP-1) activation and chemotaxis, reside in the specific arrangements of amino acid side chains pointing outwards from and lying in the cleft between the two exposed long  $\alpha$ -helices.**

*Key words:* interleukin-8/molecular modeling/monocyte chemo-attractant protein/NMR/solution structure

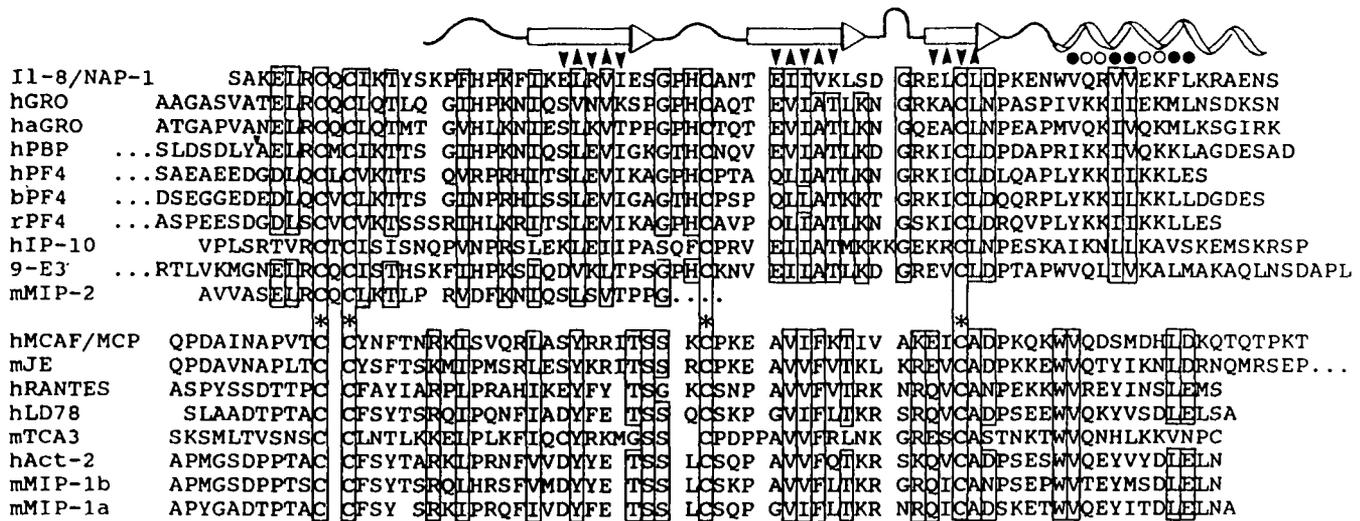
### Introduction

During recent years a large number of cytokines, involved in cell-specific chemotaxis, host defence and inflammatory responses, has been identified. Many of these cytokines appear to have overlapping spectra of bioactivities, making it difficult to attribute specific responses to individual molecules. Sequence comparisons between several purified chemotactic cytokines and c-DNA clones from activated T-cells or fibroblasts revealed the existence of a superfamily of related proteins, all of which represent polypeptide chains of  $\sim 8$ – $10$  kd and contain four

cysteine residues at near identical positions [see Wolpe and Cerami (1989) and Matsushima and Oppenheim (1990) for reviews]. A summary of the presently known sequences within this superfamily is presented in Figure 1. With the introduction of a minimal number of gaps, all sequences can be aligned with conservation of the four cysteines as well as a number of other residues. This superfamily of proteins has been further subdivided into two distinct families. The first is the  $\beta$ -thromboglobulin ( $\beta$ -TG) family which comprises, amongst others, platelet factor 4 (PF4) (Deuel *et al.*, 1977),  $\gamma$ -interferon-induced protein ( $\gamma$ -IP-10) (Luster *et al.*, 1985), melanoma growth stimulatory activity (MGSA/GRO) (Anisowicz *et al.*, 1987; Richmond *et al.*, 1988), macrophage inflammatory protein-2 (MIP-2) (Wolpe *et al.*, 1988) and neutrophil activation protein (NAP-1) also known as interleukin-8 (IL-8) or T-cell chemotactic factor (Schroeder *et al.*, 1987; Waltz *et al.*, 1987; Yoshimura *et al.*, 1987; Larsen *et al.*, 1989). In this family the first two cysteine residues are separated by a single residue in a Cys-X-Cys pattern, while in the second family they are adjacent to each other (i.e. Cys-Cys). Members of the second family include monocyte chemo-attractant protein-1, which is the product of the human JE gene (MCP-1/JE), also known as monocyte chemotactic and activating factor (MCAF) (Rollins *et al.*, 1988; Matsushima *et al.*, 1989), macrophage inflammatory protein-1 (MIP-1) (Wolpe *et al.*, 1988), RANTES (Schall *et al.*, 1988), and the product of an immune activation gene (Act-2) (Lipes *et al.*, 1988). Members within each family exhibit 25–55% sequence identity and the amino acid identity between members of the two families ranges from 21 to 31%, clearly indicating an evolutionary and functional relationship between all proteins belonging to both families.

Recently, we determined a high resolution three-dimensional structure of IL-8 (NAP-1) in solution by nuclear magnetic resonance (NMR) spectroscopy (Clore *et al.*, 1990), and showed that IL-8 forms a dimer whose general architecture is similar to the  $\alpha 1/\alpha 2$  domains of the human class I histocompatibility antigen HLA-A2 (Bjorkman *et al.*, 1987). Subsequently, the X-ray structure of IL-8 at 1.6 Å resolution was solved by molecular replacement using the solution NMR structure as a model (Baldwin *et al.*, 1991). There also exists a 3 Å resolution X-ray structure of a second member of this family, bovine platelet factor 4 (St Charles *et al.*, 1989) which aggregates to a tetramer, both in solution (Mayo and Chen, 1989) and in the crystal state (St Charles *et al.*, 1989). IL-8 in solution consists of a six-stranded antiparallel  $\beta$ -sheet formed by two three-stranded Greek keys, one from each monomer, with two symmetry-related antiparallel  $\alpha$ -helices,  $\sim 24$  Å long and separated by  $\sim 14$  Å, positioned prominently on top of this rather flat sheet. In light of the structural similarities between the  $\alpha 1/\alpha 2$  domains of HLA-A2 and IL-8 we speculated that these exposed helices might represent the primary site of interaction between IL-8, as well as other members of the superfamily, and their respective cellular receptors (Clore *et al.*, 1990).

The present paper addresses the proposed structural homology between members of the two cytokine families in more detail. In particular, we show that it is possible to model the structure



**Fig. 1.** Sequence alignment between presently known members of the superfamily of chemotactic cytokines. The known secondary structure elements are indicated above the aligned sequences. Residues within the  $\beta$ -sheet region that are pointing up ( $\blacktriangle$ ) or down ( $\blacktriangledown$ ) are marked, as well as hydrophobic ( $\bullet$ ) and hydrophilic ( $\circ$ ) ones in the  $\alpha$ -helix. The cleavage site for converting hPBP into NAP-2 is also indicated ( $\blacktriangledown$ ). IL-8/NAP-1, interleukin-8 or neutrophil attractant protein-1 (Yoshimura *et al.*, 1987; Schroeder *et al.*, 1987; Waltz *et al.*, 1987; Larsen *et al.*, 1989). hGRO, human growth related protein, haGRO, chinese hamster growth stimulatory protein (Anisowicz *et al.*, 1987; Richmond *et al.*, 1988); hPBP, platelet basic protein (Niewiarowski and Paul, 1981); hPF4, human platelet factor 4 (Poncz *et al.*, 1987). bPF4, bovine platelet factor 4 (Ciaglowski *et al.*, 1986); rPF4, rat platelet factor 4 (Doi *et al.*, 1987); hIP-10, human  $\gamma$ -interferon-induced protein (Luster *et al.*, 1985); 9-E3, rous sarcoma virus-induced protein (Sugano *et al.*, 1987); mMIP-2, murine macrophage inflammatory protein 2 (Wolpe *et al.*, 1988); hMCAF/MCP-1, human monocyte chemotactic and activating factor (Matsushima *et al.*, 1989); mJE (Rollins *et al.*, 1988); hRANTES (Schall *et al.*, 1988); hLD78 (Obaru *et al.*, 1986); mTCA3 (Burd *et al.*, 1987); hAct-2 (Lipes *et al.*, 1988). mMIP-1, murine macrophage inflammatory protein 1 (Wolpe *et al.*, 1988).

of a member of the second family, namely MCAF/MCP-1, based on the three-dimensional solution structure of IL-8, which belongs to the first family. The modeling can be carried out with very few changes in the polypeptide backbone, leaving the secondary structure elements essentially intact and resulting in a structure which is of equal quality, both in terms of geometry and energetics. In this regard, it should be noted that the solution and crystal structures of IL-8 exhibit some small but genuine differences relating to the loop comprising residues 31–36 and the relative orientation and spacing of the two  $\alpha$ -helices (Clore and Gronenborn, 1991). The implications of these differences with respect to the modeling are discussed. Further, we show that in several cases non-conservative changes between the two amino acid sequences are accompanied by compensatory changes at neighboring positions, leaving the tertiary structure essentially unchanged. In addition, the two different Cys configurations which initially led to the proposal of two distinct families can be accommodated easily within the same structural framework, exhibiting only marginal changes in local environment. We therefore believe that both families should be considered as one, with the two Cys arrangements representing solely two evolutionary-related variants of the same ancestral gene. Sequence comparison of the solvent-accessible amino acids suggests that the difference in the biological activities and specificity of the two proteins namely neutrophil (IL-8) versus monocyte (MCAF-MCP-1) activation and chemotaxis, resides in the residues pointing out from and located in the cleft between the two long C-terminal  $\alpha$ -helices.

**Modeling strategy**

The primary sequence alignment between the IL-8 and MCAF/MCP-1 sequences was taken from published sequence comparisons (Wolpe and Cerami, 1989; Leonard and Yoshimura, 1990). The MCAF/MCP-1 amino acid sequence (76 residues)

can be aligned starting at residue 5 with that of IL-8 (72 residues) incorporating two single amino acid deletions corresponding to positions 8 and 32 in the IL-8 sequence, resulting in sequence identity for 17 out of 72 residues (i.e. 24% sequence identity) and conservative changes for a further 12 residues. Main chain modeling was carried out starting at residue 10 in the MCAF/MCP-1 sequence which corresponds to position 6 in the IL-8 sequence, and proceeded to residue 73. Omission of the N-terminus is justified since residues 1–5 are ill-defined in the three-dimensional structure of IL-8 and are essentially in a random coil conformation in solution. The last three amino acid residues were also disregarded since the C-terminus of IL-8 ends two amino acids before that of MCAF/MCP-1 and the amino acid corresponding to the C-terminal serine of IL-8 is a proline in MCAF/MCP-1. Thus, the polypeptide backbone for residues 10–73 of MCAF/MCP-1 was taken directly from residues 6–71 of IL-8 with adjustments incorporated for the two regions (residues 10–11 and 34–35, numbering of MCAF/MCP-1) where the single amino acid deletions occur. The MCAF/MCP-1 side chains were placed into the main chain model using the following criteria. For identical residues in the parent molecule (IL-8) and MCAF/MCP-1 all side chain torsion angles were taken directly from the parent molecule. For a conservatively changed amino acid, the side chain was placed in an analogous position to that found in the parent molecule; thus, the  $\chi_1$  angle was transferred to the model. If more than one  $\chi$  angle could be transferred (e.g. in the pair Glu-Gln), the principle of maximal overlap for the heavy atom positions was employed. For pairs comprising Val, Ile or Thr both  $C_\gamma$  atoms were superimposed. For the one Phe-Ile pair the  $\chi_2$  angle was set to  $\sim 180^\circ$  since it has been found that aromatic to Ile substitutions frequently violate the maximal overlap principle at the  $C_\gamma$  position (Summers *et al.*, 1987). Finally, for all non-conservative changes, the MCAF/MCP-1 side chain was placed with the  $\chi_1$  angle of the parent molecule and with all further side chain torsion angles

adjusted to within  $\pm 20^\circ$  of the preferred rotamer position such that a maximal overlap of chemically equivalent atoms was achieved. Finally, the placement of side chains was visually inspected and it was ascertained that no serious steric clashes occurred. All interactive graphics modeling was carried out on a Silicon Graphics 4D/70 workstation using the program Quanta (Polygen Co.).

The modeling was carried out assuming a dimeric structure (as was found for IL-8), although there is no experimental evidence as to the aggregation state of MCAF/MCP-1. The rationale for this assumption lies in the hypothesis that any possible interaction between the receptor and the cytokine would involve the two antiparallel helices and the cleft between them. Thus the minimum aggregation state would have to be a dimer. Higher aggregation states, such as a tetramer as observed for bovine platelet factor (Mayo and Chen, 1989; St Charles *et al.*, 1989) were not explicitly considered, but since the tetrameric structure consists of a dimer of dimers, the topology of the putative interaction site would remain unchanged.

The resulting model was subjected to a restrained least squares regularization procedure using the program XPLOR (Brünger, 1988). The target function (Nilges *et al.*, 1988) that was minimized comprised covalent terms (i.e. bond lengths, bond angles and improper torsions for planarity and chirality, with force constants of 600 kcal/mol/Å<sup>2</sup> and 600 kcal/mol/rad<sup>2</sup>, for the bond and angular terms respectively), a quartic van der Waals repulsion term with the van der Waals radii set to 0.8 times their standard values, and weak harmonic restraints (with a force constant of 20 kcal/mol/Å<sup>2</sup>) to keep the C $\alpha$  atoms for residues 12–34 and 36–73 at the approximate positions of the parent structure. A total of 500 cycles of Powell minimization was carried out doubling the force constant for the van der Waals repulsion term every 100 cycles from a starting value of 0.5 kcal/mol/Å<sup>4</sup> to a maximum value of 4 kcal/mol/Å<sup>4</sup>. As described previously in applications involving three-dimensional structure determination by simulated annealing using NMR data, the bonding and non-bonding components of this target function have been calibrated empirically, such that the resulting deviations from idealized geometry are very small and the non-bonding contacts are good, as judged by a negative Lennard–Jones van der Waals energy (Nilges *et al.*, 1988). The regularized structure was then subjected to 2000 cycles of Powell energy minimization using the full all-hydrogen CHARMM (Brooks *et al.*, 1983; Polygen Co.) empirical energy function to generate the final model. The atomic r.m.s. deviations between the starting conformation and the final model of MCAF/MCP-1 were 0.84 Å for the backbone atoms and 1.51 Å for all atoms.

## Results and discussion

Homology modeling of unknown protein structures is based on the observation that the known three-dimensional structures of proteins determined to date can be categorized into a relatively limited number of families and that the tertiary structures of homologous proteins are evolutionarily more conserved than their primary amino acid sequences, which in turn are more conserved than the DNA sequence of the respective genes (Bajaj and Blundell, 1984; Chothia and Lesk, 1986; Blundell *et al.*, 1987). This manifests itself most clearly in the arrangement of regular secondary structure elements; i.e.  $\alpha$ -helices and  $\beta$ -strands are arranged in closely comparable topologies, with more pronounced changes restricted frequently to surface accessible loops (Chothia and Lesk, 1986). It was therefore assumed that the regular

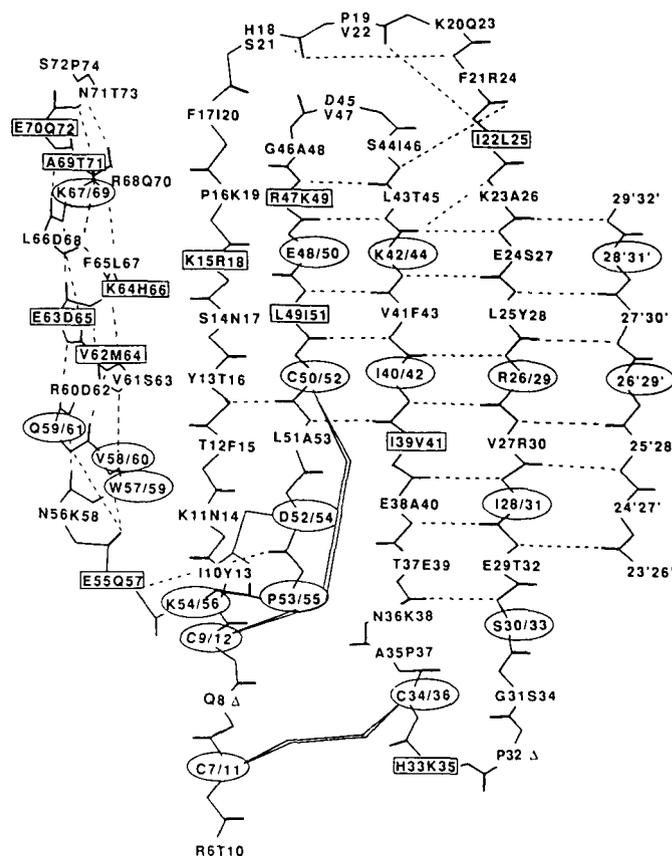
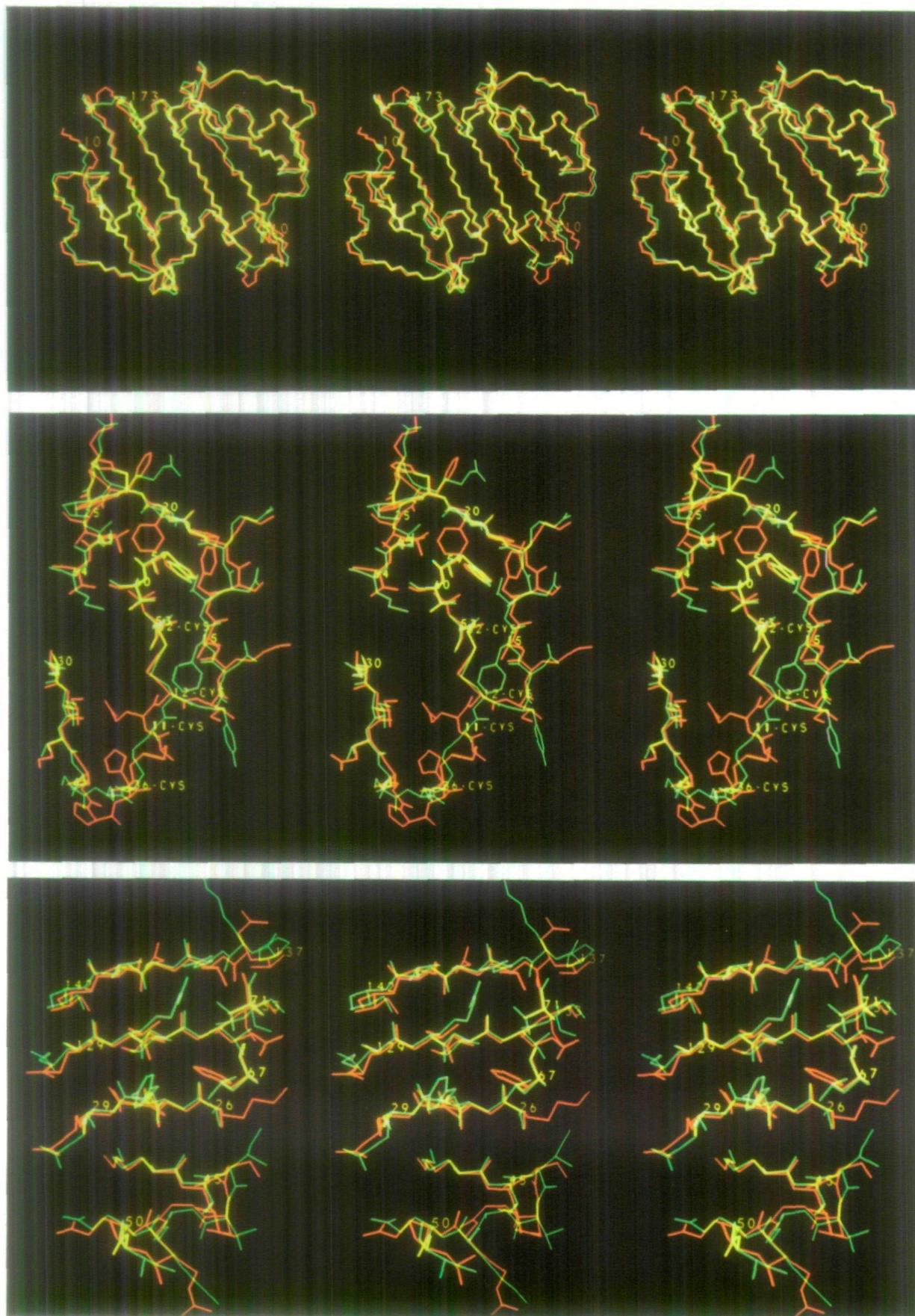


Fig. 2. Schematic illustration of the structural framework of IL-8 and MCAF/MCP-1. The triple-stranded antiparallel  $\beta$ -sheet is arranged as a Greek key with the dimer interface formed between strand I of one monomer and strand I' of the other. The amino acid sequence of both proteins is listed, including their respective positions. Identical amino acids are encircled, conservatively changed amino acids are enclosed by a box. The conserved backbone hydrogen bonds are indicated by dotted lines, those only present in IL-8 by thinner lines. The position of the disulfide bonds is also pointed out.

secondary structure elements observed in the solution structure of IL-8 would also be present in the homologous protein MCAF/MCP-1. The sequence alignment for both proteins, taking those secondary structure elements into account, is illustrated in Figure 2, which comprises a schematic representation of the IL-8 secondary structure, together with the amino acid sequences of the two proteins. It is apparent from this representation that the majority of identical residues are located within the Greek key triple-stranded  $\beta$ -sheet portion of the structure, the turn connecting the sheet with the C-terminal  $\alpha$ -helix, and the long  $\alpha$ -helix. In addition, conservative changes are also found most frequently within these three regions. The largest extent of amino acid variability is observed for the loop connecting the disulfide-bridged N-terminal cysteines with the  $\beta$ -sheet structure. Since we assume that the first nine amino acid residues of MCAF/MCP-1 are essentially random coil in solution, based on the NMR results for residues 1–5 in IL-8, and that Pro74 probably disrupts the C-terminal helix, the total number of residues for the modeled structure is 64. The amino acid sequence identity between IL-8 and MCAF/MCP-1 for this portion of the protein is 26%, and the similarity is 44% when conservative changes are taken into account. These conservative changes only include those pairs which belong to groups exhibiting the same charge and hydrophobic properties, as well as similar size. If a somewhat



**Fig. 3.** Superposition of the backbone atoms of the IL-8 NMR structure (red) and the modeled MCAF/MCP-1 structure (green) (A). Superposition of various side chains of the IL-8 NMR structure (red) and the modeled MCAF/MCP-1 structure (green) illustrating (B) the two disulfide bridges and the hydrophobic interactions involving residues of the C-terminal helix with residues of the  $\beta$ -sheet below and in the loop centered around Phe17 (IL-8)/Ile20 (MCAF/MCP-1), and (C) the dimer interface. (The residue numbering in the figure refers to MCAF/MCP-1.) All representations are stereoviews.

more generous interpretation of conservative changes is taken into account (French and Robson, 1983), including, for example, Glu-Ala, Glu-Thr, Glu-Ser, Leu-Phe, Ile-Phe and Val-Phe pairs, then the entire triple-stranded  $\beta$ -sheet would contain only conserved residues. It is also interesting to note that the amino acid substitution Glu29–Thr32 is accompanied by the reverse substitution Thr37–Glu39 and that these two amino acids are located opposite each other on two antiparallel  $\beta$ -strands, pointing upwards from the sheet in direction of the helices. Thus the space occupied by these two amino acid side chains is identical in both proteins. In the final modeled structure of MCAF/MCP-1 most backbone H-bonds in the regular secondary elements that were found for the experimentally determined IL-8 structure are retained, and those are marked in Figure 2 by the thicker dotted lines. The H-bonds within the two turn regions connecting strands 1 and 2 and strands 2 and 3 (residues 31–34 and 43–47 respectively in IL-8/NAP-1) are no longer present in the MCAF/MCP-1 structure. The loop around Cys 34–36 quite clearly has to have a somewhat different conformation given that the proline at position 32 in IL-8 is deleted in MCAF/MCP-1 and a sequence change Ala35–Pro37 occurs as well. The 3:5  $\beta$ -hairpin loop connecting strands 2 and 3 in IL-8 is retained in the modeled structure, although the typical turn residues Asn and Gly are replaced by Val and Ala. A superposition of the backbone atoms for the IL-8 NMR structure and the modeled MCAF/MCP-1 structure is shown in Figure 3(A).

Both proteins contain two disulfide bridges, one of which can be directly superimposed between the parent and modeled structure. The latter is a classical left-handed spiral conformation and connects Cys9–12 with Cys50–52. The region of both proteins around this disulfide bond is shown in Figure 3(B). The second disulfide bridge is a right-handed hook in the IL-8 structure (Figure 3B). This conformation is somewhat unusual and is additionally stabilized in the solution structure by a hydrogen bond between the backbone NH of Gln8 and the N<sup>ε2</sup> imidazole atom of His33, accounting for both the unusually low pK of 4.9 for His33 and the substantial downfield shift of the NH resonance of Gln8 (11.94 p.p.m.). Since the region around Cys7–11 contains a deletion of the amino acid between the two cysteines (Gln8 in IL-8/NAP-1) going from the IL-8 to the MCAF/MCP-1 sequence, this disulfide bridge conformation cannot be retained in the modeled structure. We therefore decided to look at alternative disulfide conformations that could be accommodated at this position and found that a left-handed spiral represents the best of several possibilities. Thus the disulfide bond between Cys11 and Cys36 was modeled in a left-handed conformation, with  $\chi$  angles of –65, –175, –97, –101 and 48° and a C<sup>α</sup>

to C<sup>α</sup> separation of 6.4 Å (Figure 3B). It is obvious that the additional His H-bond present in IL-8/NAP-1 cannot be present any more in MCAF/MCP-1 since the histidine is replaced by a lysine (Lys35) and a deletion occurs at the position of the glutamine. However, a topologically similar stabilizing interaction between the side chain of Lys35 and carbonyl group of Val9 and/or Ala7 is easily modeled. At this point it is interesting to mention that in the crystal structure of IL-8 the loop around His33 is in a different conformation to that found in solution (Baldwin *et al.*, 1990; Clore and Gronenborn, 1991) and, similarly, the equivalent loop in the X-ray structure of PF-4 (St Charles *et al.*, 1989) also exhibits a distinctly altered conformation from that in the two IL-8 structures. It therefore seems likely that the loop region comprising residues 31–34 in IL-8 or 34–36 will have different detailed structures for the different proteins.

IL-8 contains a helical turn from residues 18 to 22 which is further stabilized by a hydrogen bond between the N<sup>δ1</sup> imidazole atom of His18 and the backbone amide proton of Lys20. In MCAF/MCP-1 this type of interaction is maintained with a hydrogen bond between the O<sup>δ1</sup> atom of Ser21 (replacing the N<sup>δ1</sup> imidazole atom of the histidine) and the backbone amide of Gln23 (Figure 3B).

The positioning of the long  $\alpha$ -helix on top of the  $\beta$ -sheet in IL-8 is mainly accomplished by a set of specific hydrophobic interactions, with Trp57 being in close proximity to Tyr13, Phe17 and Leu51. This kind of van der Waals interaction is also present in the modeled structure of MCAF/MCP-1 and simply involves a different set of hydrophobic side chains. Thus Trp59 in MCAF/MCP-1 is surrounded by Phe15, Ile20 and Ala51 (Figure 3B). Likewise, Phe65, which forms the central anchor for the  $\alpha$ -helix on top of the sheet in IL-8, is involved in a hydrophobic interaction with Ile22 and Leu25. Again the modeled structure shows that this set of interactions is replaced by an equivalent constellation, with Leu67 now interacting with Leu25 and Tyr28 in MCAF/MCP-1 (Figure 3C). These two sets of hydrophobic clusters can be retained throughout the superfamily since each protein sequence contains conserved hydrophobic residues at the above-specified positions as evidenced from the sequence alignment presented in Figure 1.

All of the structural features described above for both IL-8/NAP-1 and MCAF/MCP-1 are the result of a high degree of sequence similarity as well as the preservation of the pattern of hydrophobic and hydrophilic residues. Consequently it is easily possible to accommodate all the interior amino acid changes with minimal disturbance of the general polypeptide fold. Thus the atomic r.m.s. difference for the backbone atoms between the IL-8 structure and the modeled MCAF/MCP-1 structure is only 0.9

**Table I.** Non-binding energies and deviations from idealized covalent geometry for the IL-8 NMR structure, the energy minimized IL-8 structure and the modeled MCAF/MCP-1 structure

Structure	Non-binding energies (kcal/mol)				Deviations from ideality		
	Total	van der Waals <sup>a</sup>	Electrostatic	H-bond	Bonds (Å)	Angles (°)	Improper <sup>b</sup> (°)
IL-8 NMR <sup>c</sup>	–1232	–474	–654	–104	0.011	2.458	0.485
IL-8 EM <sup>c</sup>	–2396	–507	–1713	–176	0.012	2.457	0.557
MCAF/MIP-1	–2357	–424	–1732	–201	0.008	1.958	0.258

<sup>a</sup>The van der Waals energy is calculated for the 6–12 Lennard–Jones potential of the CHARMM (Brooks *et al.*, 1983) empirical energy function.

<sup>b</sup>The improper torsion angles relate to planarity and chirality.

<sup>c</sup>IL-8 NMR is the restrained minimized average structure from Clore *et al.* (1990) which was obtained by averaging the coordinates of 40 individual simulated annealing structures (calculated on the basis of 1880 experimental distance restraints and 362 torsion angle restraints derived from NMR measurements) and subjecting the resulting mean structure to restrained minimization against a target function comprising terms for covalent geometry, terms for the experimental restraints, and a quartic van der Waals repulsion term for the non-bonded contacts. IL-8 EM is the structure obtained by 2000 cycles of energy minimization of IL-8 NMR. The atomic r.m.s. difference between IL-8-NMR and IL-8 EM is 0.51 Å for the backbone atoms and 0.59 Å for all atoms.

Å, and there are no significant differences in either the values of the non-bonding energies or the deviations from idealized covalent geometry between the IL-8 NMR structure and the modeled MCAF/MCP-1 structure (Table I). Remaining bad steric clashes would most certainly be reflected in a higher van der Waals energy and the fact that the energetics for the modeled structure are similar to the NMR structure can be taken as an indication that good packing within the protein interior has been achieved. It should be noted, however, that these energetic considerations do not prove unambiguously that the MCAF/MCP-1 model structure is correct. Rather, they indicate that the model and the assumptions upon which it is based are consistent with the constraints imposed by steric (van der Waals) and geometric (covalent) considerations.

It is also of interest to assess the changes in surface charge that arise from solvent-accessible amino acids on the exterior of the two proteins, especially in view of the probable importance

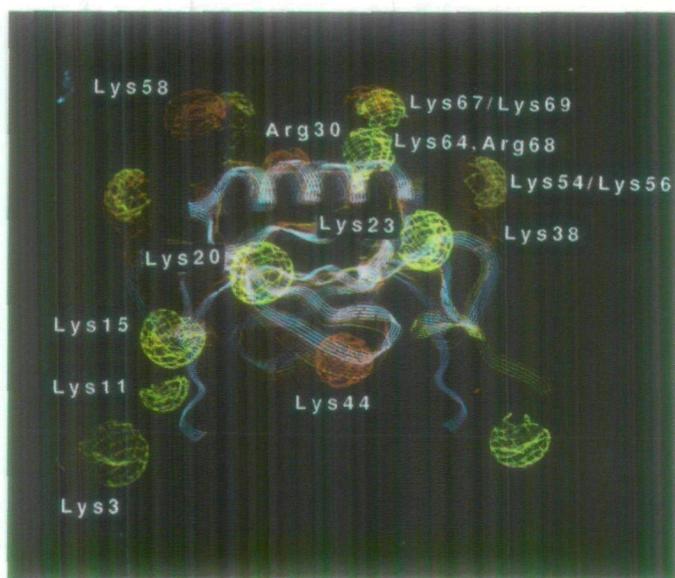


Fig. 4. Distribution of positively charged side chains in IL-8/NAP-1 (yellow) and MCAF/MCP-1 (red) overlaid onto a ribbon representation of both polypeptide chains.

of the exposed residues for receptor interaction. Comparing both negatively and positively charged amino acids, the most pronounced differences are observed for the positively charged amino acids. An illustration of the relative distribution of the positive charges is provided by Figure 4. In IL-8 there are two patches of positive charge on top of the two helices comprising Lys64 and Arg68, on the one hand, and Lys67 on the other. While one of these (Lys67–Lys69) is retained on MCAF/MCP-1, the other one is not. Instead a positively charged lysine is found at position 58 (MCAF/MCP-1) resulting in an alteration of the charge pattern on the surface of the two helices. In addition, a positive charge is found in the turn connecting the last  $\beta$ -strand of the Greek key to the  $\alpha$ -helix in both proteins (Lys54–Lys56). While in IL-8 this positive charge arises from a single amino acid side chain, it is extended in MCAF/MCP-1 into a larger patch formed by the two side chains of Lys56 and Lys38, which is located below Lys58 in the loop region connecting  $\beta$ -strands I and II. At the bottom of the  $\beta$ -sheet only one positive charge is found in MCAF/MCP-1, arising from Lys44. This has no equivalent in IL-8. Likewise IL-8 carries four positive charges in the N-terminal loop region (Lys3, Lys11, Lys15, Lys20 and Lys23) which have no equivalent in MCAF/MCP-1.

A comparison of the residues at the floor of the cleft between the two  $\alpha$ -helices for both proteins is presented in Figure 5. The most striking difference between the two proteins is the fact that in IL-8/NAP-1 all the residues that point upwards between the helices are hydrophobic, while in MCAF/MCP-1 charged and polar residues are found in prominent positions. Thus the central Leu25 and Val27 side chains of IL-8/NAP-1 are substituted by Tyr28 and Arg30 in MCAF/MCP-1, leaving only Val41 and Phe43 as hydrophobic residues at either end of the cleft. In addition, those amino acids on the helices that point across the center of the cleft are Leu66–Leu66' in IL-8/NAP-1 and Asp68–Asp68' in MCAF/MCP-1, again adding to the purely hydrophobic character of the cleft in IL-8/NAP-1 and the more polar characteristics found for MCAF/MCP-1.

In this regard it is interesting to note that the separation and angle between the two  $\alpha$ -helices in the crystal structure are 11.1 and 164° respectively (Baldwin *et al.*, 1991), compared with 14.8 and 172° in the solution structure (Clore *et al.*, 1990). The origin for this difference lies in the different degree of twisting of the

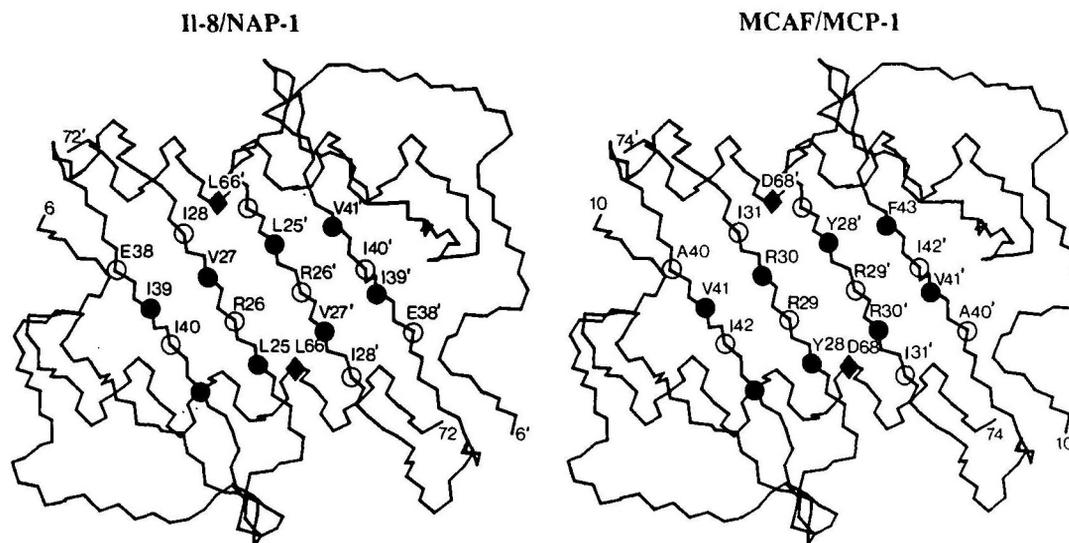


Fig. 5. Schematic illustration of the distribution of amino acids at the floor of the cleft between the two long  $\alpha$ -helices for IL-8/NAP-1 and MCAF/MCP-1. ●, Residues pointing upwards into the cleft; ○, residues pointing downwards from the sheet; ◆, residues pointing into the cleft from the helix.

underlying central strands of the  $\beta$ -sheet formed by strand I of one subunit and strand I' of the other: in the crystal structure the twist of these two strands is  $179^\circ$ , whereas in the solution structure the twist is  $168^\circ$ , characteristic of a typical  $\beta$ -sheet. As a result the cleft is somewhat larger in the solution structure than in the crystal one (Clare and Gronenborn, 1991). These differences are clearly genuine as the X-ray structure predicts at least 30 additional interproton distances  $<3.5 \text{ \AA}$  between residues of the two subunits, for which no corresponding nuclear Overhauser effects are observed in the NMR spectra (Clare and Gronenborn, 1991). If MCAF were modeled on the basis of the X-ray structure this would have a number of significant consequences. First, the carboxylate of Asp68 of one subunit would be separated by  $<3.5 \text{ \AA}$  from the carboxylate of Asp68' of the other, an interaction that would clearly be disfavored electrostatically, in contrast to the model based on the solution structure where these two negatively charged groups are separated by  $5\text{--}6 \text{ \AA}$ . Second, Arg30 would clash sterically with the overlying Asp68', while in the model based on the solution structure Arg30 can interact in an electrostatically favorable manner with Glu39 of its own subunit and Asp68' of the other. Third, the hydroxyl group of Tyr28 would clash with the side chain of Met64, while in the model based on the solution structure it is clearly solvent accessible.

In contrast to the distinct differences in the residues lining the cleft between the two helices in IL-8 and MCAF, those amino acids at the center of the cleft which point downwards from the sheet in the view presented in Figure 5 are conserved between IL-8/NAP-1 and MCAF/MCP-1. Therefore, if, as it seems likely, the two helices and the cleft between them are involved in the interaction with the receptor, the completely hydrophobic character of the cleft found in IL-8/NAP-1 would clearly provide a means for discriminating this from the more polar cleft found in MCAF/MCP-1. In this respect it is interesting to note that IL-8/NAP-1, hGRO and NAP-2, a cleavage product of hPBP, compete for the same receptor, while MCAF/MCP-1 does not. The most notable differences between the former three protein sequences and the latter (and other members in this class; see Figure 1) are found in the central residues between the helices: namely IL-8/NAP-1, hGRO and NAP-2 possess leucine and valine while MCAF/MCP-1 (and others) have tyrosine and arginine (or other charged residues) in their place. Likewise the exposed positively charged arginine at position 60 pointing upwards from the helix in IL-8/NAP-1 is replaced by a negatively charged aspartic acid in MCAF/MCP-1 and a variety of polymorphic amino acids in the other members of this family. Thus, these amino acid changes on the helices and the cleft between them may provide the basis for receptor discrimination.

### Acknowledgements

This work was supported by the Intramural AIDS Anti-Viral Program of the Office of the director of the National Institutes of Health.

### References

- Anisowicz, A., Bardwell, L. and Sager, R. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 7188–7192.
- Bajaj, M. and Blundell, T.L. (1984) *Annu. Rev. Biophys. Bioengng.* **13**, 453–492.
- Baldwin, E.T., Weber, I.T., St Charles, R., Zuna, J.C., Appella, E., Yamada, M., Matsushima, K., Edwards, B.F.P., Clare, G.M., Gronenborn, A.M. and Wlodawer, A. (1991) *Proc. Natl Acad. Sci. USA*, in press.
- Blundell, T.L., Sibanda, B.L., Sternberg, M.J.E. and Thornton, J.M. (1987) *Nature*, **326**, 347–352.
- Bjorkman, P.H., Saper, M.A., Samraoui, B., Bennett, W.S., Strominger, J.L. and Wiley, D.C. (1987) *Nature*, **329**, 506–512.
- Brooks, B.R., Brucoleri, R.E., Olafson, B.D., States, D.J., Swaminathan, S. and Karplus, M.J. (1983) *Comput. Chem.*, **4**, 187–217.
- Brünger, A.T. (1988) *XPLOR Manual*. Yale University, New Haven, CT.
- Burd, R.P., Freeman, G.J., Wilson, S.D., Berman, M., DeKruyff, R., Billings, P.R. and Dorf, M.E. (1987) *J. Immunol.*, **139**, 3126–3131.
- Chothia, C. and Lesk, A. (1986) *EMBO J.*, **5**, 823–826.
- Ciagłowski, R.E., Snow, J.W. and Waltz, D.A. (1986) *Arch. Biochem. Biophys.*, **250**, 249–256.
- Clare, G.M. and Gronenborn, A.M. (1991) *J. Mol. Biol.*, in press.
- Clare, B.M., Appella, E., Yamada, M., Matsushima, K. and Gronenborn, A.M. (1990) *Biochemistry*, **29**, 1689–1696.
- Deuel, T.F., Keim, P.S., Farmer, M. and Henrickson, R.L. (1977) *Proc. Natl Acad. Sci. USA*, **74**, 2256–2258.
- Doi, T., Greenberg, S.M. and Rosenberg, R.D. (1987) *Mol. Cell Biol.*, **7**, 898–904.
- French, S. and Robson, B. (1983) *J. Mol. Evol.*, **19**, 171–175.
- Larsen, C.G., Anderson, A.O., Appella, E., Oppenheim, J.J. and Matsushima, K. (1989) *Science*, **243**, 1464–1467.
- Leonard, E.J. and Yoshimura, T. (1990) *Immunol. Today*, in press.
- Luster, A.D., Ukeless, J.C. and Ravetch, J.V. (1985) *Nature*, **315**, 672–676.
- Lipes, M.A., Napolitano, M., Jeang, K.-T., Chang, N.T. and Leonard, W.J. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 9704–9708.
- Matsushima, K. and Oppenheim, J.J. (1989) *Cytokine*, **1**, 2–13.
- Matsushima, K., Larsen, C.G., DuBois, G.C. and Oppenheim, J.J. (1989) *J. Exp. Med.*, **169**, 1485–1490.
- Mayo, K.H. and Chen, M.-J. (1989) *Biochemistry*, **28**, 9469–9478.
- Niewiarowski, S. and Paul, D. (1981) In Gordon, E. (ed.), *Platelets in Biology and Pathology*. Elsevier North Holland, Vol. 2, pp 91–96.
- Nilges, M., Clare, G.M. and Gronenborn, A.M. (1988) *FEBS Lett.*, **229**, 317–324.
- Obaru, K., Fukuda, M., Maeda, S. and Shimada, K. (1986) *J. Biochem.*, **99**, 885–894.
- Poncz, M., Surrey, S., LaRocco, P., Weiss, M.J., Rappaport, E.F., Conway, T.M. and Schwartz, E. (1987) *Blood*, **69**, 219–223.
- Richmond, A., Balentien, E., Thomas, H.G., Flaggs, G., Barton, D.E., Spiess, J., Bordoni, R., Francke, U. and Derynck, R. (1988) *EMBO J.*, **7**, 2025–2033.
- Rollins, B.J., Morrison, E.D. and Stiles, C.D. (1988) *Proc. Natl Acad. Sci. USA*, **85**, 3738–3742.
- Schall, T.J., Jongstram, J., Dyer, B.J., Jorgensen, J., Clayberger, C., Davies, M.M. and Krensky, A.M. (1988) *J. Immunol.*, **141**, 1018–1025.
- Schroeder, J.M., Mrowietz, U., Morita, E. and Christopher, E. (1987) *J. Immunol.*, **139**, 3473–3483.
- St Charles, R., Walz, D.A. and Edwards, B.F.P. (1989) *J. Biol. Chem.*, **264**, 2092–2099.
- Sugano, S., Stoeckle, M.Y. and Hanafusa, H. (1987) *Cell*, **49**, 321–328.
- Summers, N.L., Carlson, W.D. and Karplus, M. (1987) *J. Mol. Biol.*, **196**, 175–198.
- Walz, A., Peveri, P., Arschaer, H. and Baggiolini, M. (1987) *Biochem. Biophys. Res. Commun.*, **149**, 755–761.
- Wolpe, S.D. and Cerami, A. (1989) *FASEB J.*, **3**, 2565–2573.
- Wolpe, S.D., Davatelis, G., Sherry, B., Beutler, B., Hesse, D., Nguyen, H.T., Moldawer, L.L., Nathan, C.F., Lowry, S.F. and Cerami, A. (1988) *J. Exp. Med.*, **167**, 570–581.
- Yoshimura, T., Matsushima, K., Tanaka, S., Robinson, E.A., Appella, E., Oppenheim, J.J. and Leonard, E. (1987) *Proc. Natl Acad. Sci. USA*, **84**, 9233–9237.

Received on August 30, 1990; revised and accepted on November 5, 1990