

Analysis of hydrophobicity in the α and β chemokine families and its relevance to dimerization

DAVID G. COVELL,¹ GARY W. SMYTHERS,¹ ANGELA M. GRONENBORN,²
AND G. MARIUS CLORE²

¹ Biomedical Supercomputing Laboratory, PRI/Dyncorp, Frederick Cancer Research and Development Center, National Cancer Institute, Frederick, Maryland 21702

² Laboratory of Chemical Physics, National Institutes of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520

(RECEIVED June 24, 1994; ACCEPTED September 1, 1994)

Abstract

The chemokine family of chemotactic cytokines plays a key role in orchestrating the immune response. The family has been divided into 2 subfamilies, α and β , based on the spacing of the first 2 cysteine residues, function, and chromosomal location. Members within each subfamily have 25–70% sequence identity, whereas the amino acid identity between members of the 2 subfamilies ranges from 20 to 40%. A quantitative analysis of the hydrophobic properties of 11 α and 9 β chemokine sequences, based on the coordinates of the prototypic α and β chemokines, interleukin-8 (IL-8), and human macrophage inflammatory protein-1 β (hMIP-1 β), respectively, is presented. The monomers of the α and β chemokines have their strongest core hydrophobic cluster at equivalent positions, consistent with their similar tertiary structures. In contrast, the pattern of monomer surface hydrophobicity between the α and β chemokines differs in a manner that is fully consistent with the observed differences in quaternary structure. The most hydrophobic surface clusters on the monomer subunits are located in very different regions of the α and β chemokines and comprise in each case the amino acids that are buried at the interface of their respective dimers. The theoretical analysis of hydrophobicity strongly supports the hypothesis that the distinct dimers observed for IL-8 and hMIP-1 β are preserved for all the α and β chemokines, respectively. This provides a rational explanation for the lack of receptor crossbinding and reactivity between the α and β chemokine subfamilies.

Keywords: α and β chemokines; chemotactic cytokines; dimeric structure; hMIP-1 β ; hydrophobicity; interleukin-8; tertiary structure

The chemokines comprise a large family of chemotactic cytokines and are made up of a polypeptide chain of ~8–10 kDa with 4 cysteine residues at near identical positions (Oppenheim et al., 1991; Schall, 1991; Baggiolini et al., 1994). The chemokines have been subdivided into 2 subfamilies, α and β , based on whether the first 2 cysteines are separated by 1 residue (α) or are adjacent (β), on function (the α chemokines are potent chemoattractants for neutrophils but not monocytes, whereas the reverse is true of the β chemokines), and on chromosomal location. There is also no receptor crossbinding and reactivity between the α and β subfamilies. Members within each subfamily exhibit 25–70% sequence identity, whereas the sequence identity between the 2

subfamilies ranges from 20 to 40%. The 3-dimensional structures of 2 members of the α subfamily, interleukin-8 (IL-8) (Clare et al., 1990; Baldwin et al., 1991; Clare & Gronenborn, 1991) and platelet factor-4 (PF4) (St. Charles et al., 1989), and 1 member of the β subfamily, human macrophage inflammatory protein-1 β (hMIP-1 β) (Lodi et al., 1994) have been solved by NMR and/or X-ray crystallography. All 3 proteins are multimeric. IL-8 and hMIP-1 β are homodimers, whereas PF4 is a homotetramer consisting of a dimer of dimers of the IL-8 type. Although the tertiary structure of the monomeric unit is very similar for all 3 proteins, as expected given their close sequence homology, the dimeric quaternary structures of the α and β chemokines are completely different (Lodi et al., 1994). The IL-8 dimer is globular in shape, whereas the hMIP-1 β dimer is elongated and cylindrical (Fig. 1). The IL-8 dimer comprises a 6-stranded antiparallel β -sheet, on top of which lie 2 antiparallel helices separated by ~14 Å (Clare et al., 1990), and the C₂ axis is located between residues 26 and 26' (equivalent to residue 29 of hMIP-1 β) at the center of strands β 1 and β 1'. In con-

Reprint requests to: David G. Covell, Biomedical Supercomputing Laboratory, PRI/Dyncorp, Frederick Cancer Research and Development Center, National Cancer Institute, Frederick, Maryland 21702; or G. Marius Clare or Angela M. Gronenborn, Laboratory of Chemical Physics, Building 5, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892-0520.

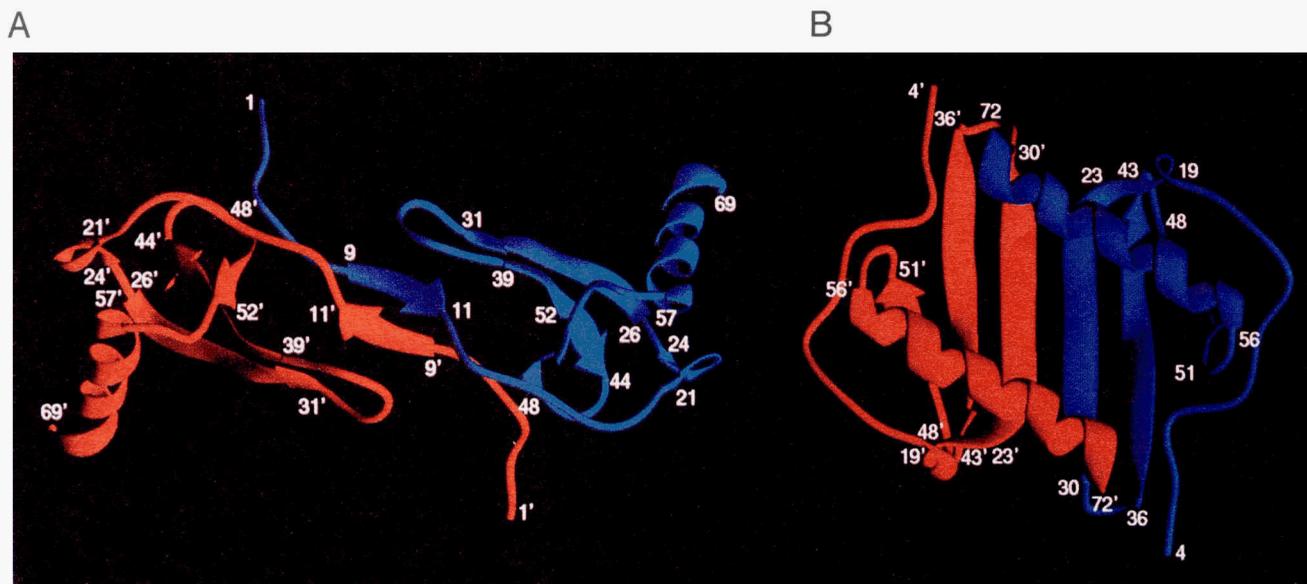


Fig. 1. Ribbon drawing of the hMIP-1 β and IL-8 dimers. The program RIBBONS (Carson, 1987) was used to generate the models and the coordinates are taken from Lodi et al. (1994) and Clore et al. (1990), respectively. In each case, 1 subunit is shown in red and the other in blue. The PDB accession numbers for the hMIP-1 β and IL-8 coordinates are 1HUM and 1IL8, respectively.

trast, in the hMIP-1 β dimer, the C_2 axis is located between residues 10 and 10', which are part of an additional mini-antiparallel β -sheet formed by strands β_0 and β_0' ; the 2 helices are 46 Å apart on opposite sides of the molecule, and strands β_1 and β_1' are ~30 Å apart and located on the exterior of the protein (Lodi et al., 1994). On the basis of calculations of solvation free energies of dimerization (Eisenberg & McLaglan, 1986), it was suggested that the stabilization of these different quaternary structures could be attributed to the burial of hydrophobic residues that form the dimer interface in hMIP-1 β and IL-8 (Lodi et al., 1994). This is further supported by the experimental finding that IL-8 dimerization is independent of pH and salt concentration (G.M. Clore & A.M. Gronenborn, unpubl. data), and that the hMIP-1 β dimer is stable at very low pH and is also independent of salt concentration (Lodi et al., 1994).

In this paper we present a quantitative cluster analysis of the hydrophobic properties of the α and β chemokines, based on concepts developed by Kauzmann (1959), with the aim of establishing the factors responsible for the stabilization of the tertiary structure and the determinants of the dramatically different dimeric structures. In addition, the method presented can also be used to predict possible sites of interactions of the dimeric forms of the α and β chemokines with their cell surface receptors. Although this analysis is based on the observation that protein-protein associations usually involve the more hydrophobic portions of each molecule, the precise interactions within any interface can, and most likely do, depend on properties other than hydrophobicity. In this regard, we note that in the case of hMIP-1 β , higher-order structures are principally stabilized by electrostatic interactions as evidenced by both their pH and salt dependence (Lodi et al., 1994). Similarly, the stabilization of the PF4 tetramer formed by a dimer of dimers of the IL8 type principally involves electrostatic interactions (Mayo & Chen, 1989; St. Charles et al., 1989).

Calculation strategy

Geometry

The quantitative analysis of hydrophobicity presented here is based on a simplified model consisting of only the C^α coordinates to represent the geometry of each monomer unit (Levitt, 1976). Reduced models represent a tradeoff between the accuracy of a complete molecular model and the requirements to rapidly and completely analyze the details of packing and solvent exposure for residues in the core and on the surface of the molecule. Simplified molecular models have been used widely to examine the details of protein stability, packing, and folding (Levitt, 1976; Covell & Jernigan, 1990; Dill, 1990; Skolnick & Kolinski, 1990; Covell, 1994; Young et al., 1994). Although models of greater atomic complexity offer an attractive alternative to the simple C^α model proposed here, the tools to complete an equivalent analysis are not yet developed. Further improvements would be needed in at least 2 areas to permit extension of this approach from a simple C^α model to an all-atom model. In particular atom-atom interaction potentials for the quantification of hydrophobicity are not well established and a simple scheme for identifying candidate regions on the target surface for hydrophobicity scoring has not yet been determined.

To assess surface hydrophobicity, positions exterior to each monomer unit are obtained by embedding the C^α positions into a face-centered cubic lattice (unit cell dimensions = 3.8 Å) and retaining only a shell of lattice points surrounding these coordinates (Jernigan et al., 1989). The thickness of the shell is established by removing lattice points that are either too close or too far from the protein. The outer boundary of the shell includes lattice points closer than 9.0 Å from any protein C^α coordinate, whereas the inner boundary includes only lattice points farther than 6.1 Å from any protein C^α position. The inner

boundary is based on observations on crystal complexes (Jernigan et al., 1989); the outer boundary is the limit for which the contact energy parameters are likely to be valid (Christenson & Claesson, 1989). These exterior positions can be thought of as defining the portion of a molecule's surface that would be accessible to an approaching monomer (represented by its C α coordinates), and thus parallels the concept of molecular surfaces based on their accessibility to a water molecule probe. The set of C α coordinates within a sphere of radius 9.0 Å from each shell point defines the nearest protein coordinates of each shell point. This set of coordinates is referred to as a cluster. A complete description of this procedure is provided by Young et al. (1994).

The C α coordinates comprising the core cluster of monomers are determined by simply counting the set of nearest neighbors ($d \leq 7.5$ Å) for each C α position (Covell & Jernigan, 1990). Coordinates with the greatest number of neighbors identify highly packed regions in the monomer core. The composition of amino acids in each surface and core cluster is used to score its total hydrophobicity as described below.

Cluster hydrophobicity

The hydrophobicity score for clusters of amino acids at the surface and the core of the monomer are determined using previously published residue-based pairwise contact potentials of Miyazawa and Jernigan (1985). Averages of these pairwise contact energies are used as the hydrophobicities for each residue type. The values used for this analysis are: F, -5.12; M, -4.91; I, -4.88; L, -4.65; W, -4.36; V, -4.17; C, -4.00; Y, -3.24; A, -2.82; H, -2.75; G, -2.34; T, -2.30; P, -2.22; R, -2.18; S, -2.07; Q, -1.98; E, -1.94; N, -1.90; D, -1.81; K, -1.50). These residue hydrophobicities can be understood in terms of the hydrophobic-hydrophilic designations of amino acids and the pairings that contribute to stability (Miyazawa & Jernigan, 1985; Covell & Jernigan, 1990; Covell, 1992). Although the hydrophobicity scale used here is slightly different from other published scales, it shows a strong correlation with the Tanford-Nozaki scale (Nozaki & Tanford, 1971) as shown by Cornette et al. (1987). Based on this strong correlation, the results of the analysis using the Miyazawa and Jernigan scale (1985) are not expected to be substantially different when using other hydrophobicity scales.

Sequence analysis

Members of the α and β subfamilies of chemokines were obtained using the GCG database software package (Genetics Computer Group, 1991), using the query sequences of IL-8 and hMIP-1 β to extract candidate sequences for later analysis. Table 1 lists the chemokine family members considered in this analysis. The list represents a selection of sequences identified by the BLAST program (Altschul et al., 1990) that have high scores against either IL-8 or hMIP-1 β . The selection of sequences for analysis was made from entries having between 100 and 29% sequence identity to either of the query sequences. Analysis was not done on the complete set of high scoring sequences, due to entries with identical or nearly identical sequences to those listed in Table 1.

Results and discussion

Stabilization of tertiary structure: Core clusters

The C α coordinates of the IL-8 (Clore et al., 1990) and hMIP-1 β (Lodi et al., 1994) monomers were used to define the backbone template for analysis of sequence members of the chemokine family. Sequences from each chemokine subfamily were then applied to the C α positions of the IL-8 and hMIP-1 β structures, and the aligned sequences for analysis on the hMIP-1 β coordinates are shown in Figure 2A, whereas alignments on the IL-8 coordinates are shown in Figure 2B. Each alignment is based on the best-fit structural superposition of the hMIP-1 β and IL-8 monomers presented by Lodi et al. (1994), which indicated that the C α coordinates of 59 residues could be superimposed with an RMS of 1.6 Å with sequence positions 1–30 and 33–68 of IL-8 being equivalent to positions 4–34 and 35–69 of hMIP-1 β . In the discussion to follow, the numbering scheme for each alignment is based on the initial entry in Figure 2A and B, respectively. When comparisons are made between the 2 sets of alignments, the equivalent positions between them will be identified.

In considering the results presented below, it is important to bear in mind that an analysis of surface-accessible residues for the hMIP-1 β and IL-8 monomers indicates that the distinction between core (as identified by lack of surface accessibility) and surface residues cannot be easily made on the basis of surface area alone. Thus, in the case of the hMIP-1 β monomer, there is only 1 completely buried residue (Phe 42) and 1 nearly completely buried residue (Ala 52, 3 Å²). Similarly, for the IL-8 monomer there are no completely buried residues, and the least-exposed residues are Cys 34 (7 Å²), Val 41 (12 Å²), and Val 58 (8 Å²). All the remaining residues for both monomers have over 26 Å² of accessible surface. Hence, any core cluster must necessarily involve residues with some amount of surface accessibility. In addition, the use of a C α model blurs the distinction between hydrophobic and hydrophilic residue types. However, the concept of using a cluster of nearest neighbors from which a hydrophobicity score is determined more correctly reflects the surface geometry of the folded protein and permits calculation of an average hydrophobicity score. The fact that these clusters also include surface-accessible residues that have both hydrophobic and hydrophilic character reflects the conditions observed in these proteins as well as that observed for most small globular proteins.

We first present the analysis of the C α coordinates of the hMIP-1 β monomer. Using the native hMIP-1 β sequence, the cluster of core amino acids having the greatest hydrophobicity (Fig. 2A) includes Val 25, Asp 27, Tyr 28, Tyr 29, Phe 42, Gln 49, Val 50, Cys 51, and Ala 52. The core cluster with the second greatest hydrophobicity score includes Cys 12, Ser 14, Tyr 15, Thr 16, Ala 39, Val 40, Val 41, Phe 42, and Cys 51. Overlap is observed between these clusters at positions 42 and 51, suggesting a single core of residues that provides greatest stabilization of the native hMIP-1 β structure. Figure 3A illustrates this point for the structure of the hMIP-1 β monomer by spectrally highlighting residues in these 2 top-scoring hydrophobic core clusters according to their hydrophobicity. Residues colored in red indicate core positions with the greatest hydrophobicity. A similar pattern is seen for sequence positions comprising the 2 top-scoring hydrophobic core clusters for the other sequences

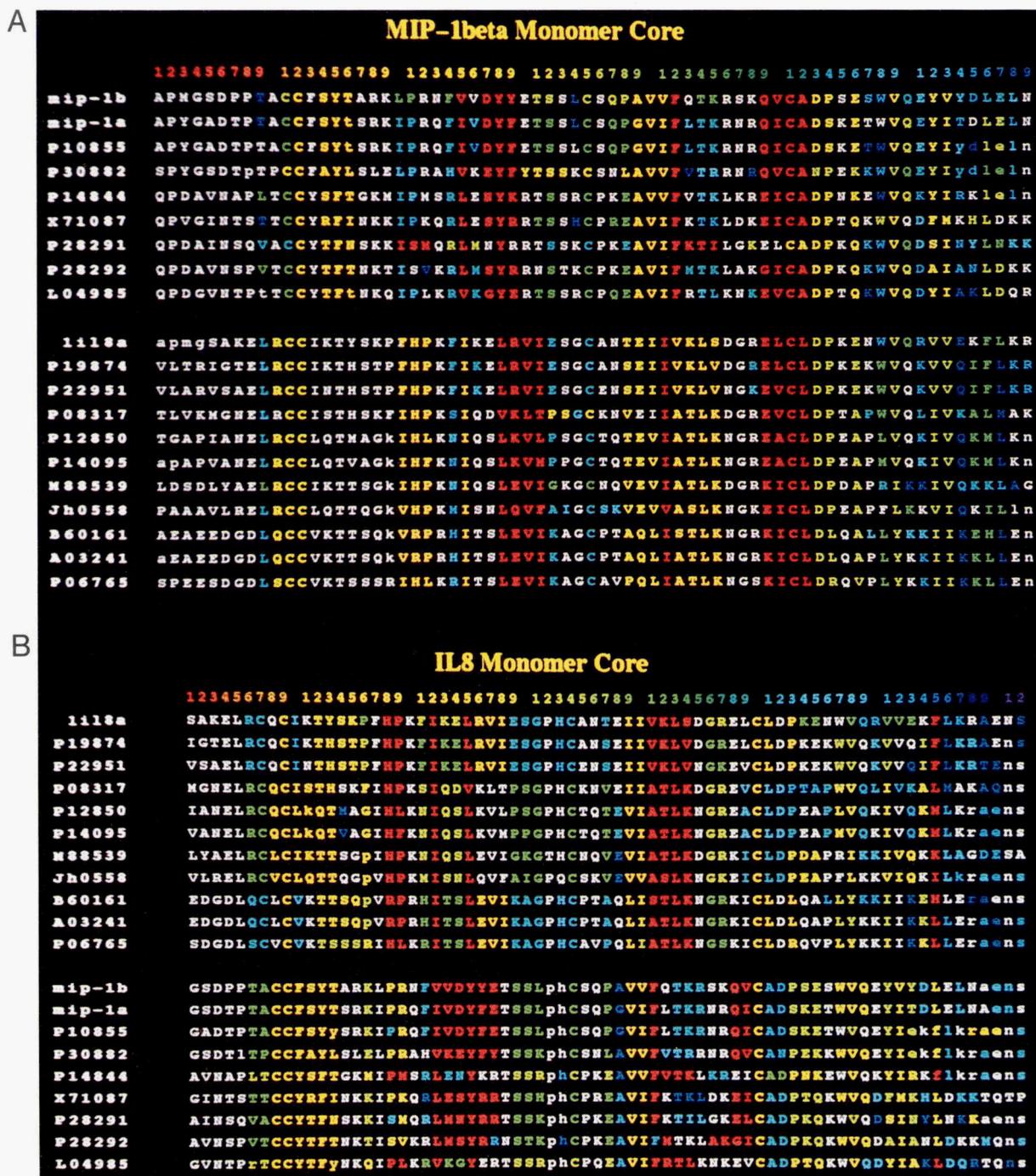


Fig. 2. Alignments of sequence members of each subfamily applied to the hMIP-1β coordinates (A) and the IL-8 coordinates (B). Sequence positions are highlighted according to the relative strength of hydrophobicity of each core cluster. Sequence positions in the core cluster of strongest hydrophobicity are indicated in red and positions in clusters of weaker hydrophobicity are indicated in colors approaching blue in the color spectrum. Only the top 5 scoring clusters are identified by color. Residues in lower scoring clusters are white. The row of sequence positions at the top of each panel is colored from left to right according to this hydrophobicity scale. Core clusters that contain positions found in lower scoring clusters are colored according to the higher-ranked cluster. Consequently, depending on the amino acid composition of a given cluster, 1 region may fall in or out of the top 5 clusters and can thus appear as colored or white, accordingly. In each panel, the native sequence and sequences of its subfamily are listed at the top of each table (e.g., panel A lists the hMIP-1β sequence and members of the β subfamily at the top with members of the α subfamily at the bottom. Panel B reverses this order, placing members of the α subfamily at the top and members of the β subfamily at the bottom.) The alignment is based on the structural superposition of hMIP-1β and IL-8 presented by Lodi et al. (1994). Lowercase letters indicate gaps in the alignment. In these cases, the amino acid of the reference structure is assumed to occupy this position. The alignment between subfamilies involves the insertion of Pro 32 and His 33 in the cases of aligning the β subfamily against the α subfamily of sequences, and deletion of this same pair in the alignment of the α subfamily against the β subfamily. Some of the sequences analyzed are shorter than the reference sequence. Lowercase letters are also used to identify the amino acids assigned to these positions at the chain termini.

Table 1. Members of the chemokine family selected for analysis

Identification	Name	Description	% Match to IL-8	% Match to hMIP-1 β
α Chemokine subfamily				
P10145	IL-8	Interleukin-8, neutrophil activation protein	100	26
P19874	RPF1	Rabbit neutrophil attractant/activation protein	88	26
P22951	AMCF-1	Pig alveolar macrophage chemotactic factor	79	19
P08317	PCEF-4	EMFI-chick embryo fibroblast protein 9E3 precursor	55	36
P12850	GRO-Mouse	Mouse growth-related protein	54	21
P14095	GRO-Rat	Rat growth-related protein	52	28
M88539		Small fusion protein	43	22
Jh0558	NAP-2	Neutrophil-activating peptide ENA-78	40	34
B60161	PF4	Platelet factor 4 (fragment)	34	23
A03241	PFHU4	Platelet factor 4 (human)	36	25
P06765	PLF4	Platelet factor 4 precursor	39	21
β Chemokine subfamily				
P14097	hMIP-1 β	Human macrophage inflammatory protein-1 β	26	100
P16619	MI10	Human tonsillar lymphocyte LD78 protein (hMIP-1 α)	38	63
P10855	MIP-1 α	Mouse macrophage inflammatory protein-1 α	32	55
P30882	SISD	Mouse T-cell-specific RANTES protein precursor	27	52
P14844	Rat MCP-1	Rat monocyte chemotactic protein	32	51
X71087	HSMCP3	Monocyte chemoattractant protein, MCP-3	31	48
P28291	Bovine MCP-1	Bovine monocyte chemotactic protein	40	44
P28292	Rabbit MCP-1	Rabbit monocyte chemotactic protein	27	41
L04985	GPIMCP1A	Monocyte chemoattractant protein	29	39

in the β chemokine family (top portion of Fig. 2A). Aside from conservation of cysteines, complete conservation of amino acid type for the core positions is observed for Tyr 28, Phe 42, and Ala 52 within the β subfamily. It is perhaps interesting to note that Trp 58, which packs orthogonally against Phe 42, is not included in the most hydrophobic core cluster of hMIP-1 β . This is due in large part to the fact that Trp 58 has a third of its surface area exposed to solvent.

The pattern of sequence positions involved with the core clusters for the α subfamily of sequences when applied to the hMIP-1 β coordinates (lower portion of Fig. 2A) is similar to that found for the β subfamily. Sequence positions comprising the core cluster of greatest hydrophobicity (indicated in red) overlap those found for the β subfamily at sequence positions 28, 29, and 40–51. Although the types of amino acids occupying these sequence positions in the α subfamily are not similar to those of the β subfamily, they share the strong hydrophobic character found at these positions in the β subfamily.

The aligned sequences for members of the α and β subfamilies onto the IL-8 coordinates are shown in Figure 2B. Residues for the native IL-8 sequence having the greatest core hydrophobicity include His 18, Pro 19, Ile 22, Leu 25, Val 41, Lys 42, Leu 43, Ser 44, and Phe 65. Figure 3B highlights these core residues according to their hydrophobicity scores. This pattern of sequence positions is conserved within all members of the α subfamily. Conservation of amino acid type within the α subfamily is found only for positions Ile 22 and Leu 43, positions found to be among the best-packed residues in the IL-8 core. Inspection of sequences in the β subfamily (listed in the lower portion of Fig. 2B) indicates that the core cluster with greatest hydrophobicity is not positioned exactly coincidental with that for the α subfamily, but staggered by 1 or 2 positions, with overlap observed at positions 22, 25, and 41. Strong hydrophobic residues

are found at each of these sequence positions, with complete conservation of amino acid type found within the β subfamily of sequences at Tyr 25 and Phe 41.

The results of Figure 2A and B indicate a strong conservation of core hydrophobic residues for sequences of both subfamilies, when applied to either the hMIP-1 β (β chemokine) or IL-8 (α chemokine) coordinates. A comparison between the most hydrophobic core cluster for the hMIP-1 β and IL-8 structures indicates that 3 equivalent positions appear on both lists, namely 25, 28, and 42 on hMIP-1 β and 22, 25, and 41 on IL-8. In both cases these residue positions are located within the core of the monomer and are occupied by strongly hydrophobic residues. Conservation of residue type at these positions for both subfamilies of sequences supports the observed structural similarity between monomers in each subfamily.

Determinants of the dimeric structures:

Monomer surface clusters

The results for clusters of solvent-exposed amino acids on the surface of the monomers with the best hydrophobicity scores are shown in Figure 4A and B for the β and α subfamilies analyzed on the hMIP-1 β and IL-8 coordinates, respectively. The alignments are identical to those shown in Figure 2. However, these figures spectrally highlight residue positions that form the surface clusters on the monomers with the best hydrophobicity scores.

Monomer surface clusters for sequences applied to the coordinates of the hMIP-1 β structure are listed in Figure 4A, with the β subfamily at the top and the α subfamily at the bottom. The sequence positions in the most hydrophobic surface cluster are found at locations 12–16 and 49–51. The second-ranked hydrophobic surface cluster involves sequence positions 3–9 and

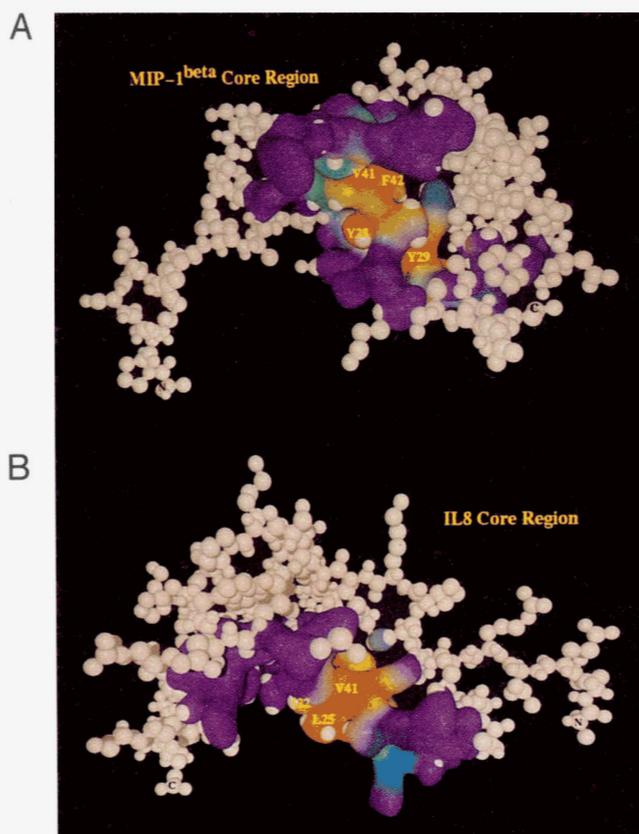


Fig. 3. Ball-and-stick model of the hMIP-1 β (A) and IL-8 (B) monomers illustrating the most hydrophobic core clusters. Atoms of residues included in the most hydrophobic core cluster are enveloped in a sheath with its surface colored spectrally. Red indicates regions of greatest hydrophobicity, whereas weaker hydrophobicity is indicated in colors approaching blue in the color spectrum. Atoms of residues not close to the most hydrophobic core cluster are shown in gray. These figures were generated using Rayshade as modified by George McGregor, Program Resources Incorporated. The coordinates of hMIP-1 β and IL-8 and from Lodi et al. (1994) and Clore et al. (1990), respectively.

32–34. The location of these clusters in the structure of hMIP-1 β is illustrated in Figure 5A. The amino acids in these clusters correspond to those amino acids buried at the interface of the hMIP-1 β dimer, as listed in Table 2. The amino acids in the 2 most hydrophobic surface clusters comprise 73% of the 907 \AA^2 of total surface buried by the hMIP-1 β monomer at the dimer interface. The most hydrophobic cluster accounts for nearly 30% of the total buried surface. The pattern of sequence positions for the most hydrophobic surface cluster is completely conserved within the β subfamily. There is, however, no conservation of amino acid type at these positions (excluding cysteines).

A similar correspondence of hydrophobic monomer surface clusters (cf. Fig. 4B, bottom) and residues buried at the dimer interface (cf. Table 2) is observed for the α subfamily when applied to the IL-8 coordinates. However, a completely different set of positions is involved in the interface of the α and β chemokines. The highest-ranked surface cluster for native IL-8 includes sequence positions 27–29, 37–39, 53, and 59. Their location in the structure is visualized in Figure 5B. These locations correspond to those found at the dimer interface of IL-8 and comprise 40% of the 751.5 \AA^2 of buried monomer surface with this

Table 2. Molecular surface area buried at the dimer interfaces of hMIP-1 β and IL-8

hMIP-1 β		IL-8	
Residue	\AA^2	Residue	\AA^2
Pro 2	50.1	Lys 23	59.1
Glu 4	37.9	Glu 24	55.3
Ser 5	61.3	Leu 25	37.6
Asp 6	99.4	Arg 26	36.9
Pro 7	39.1	Val 27	82.3
Pro 8	100.3	Ile 28	29.7
Thr 9	36.5	Glu 29	119.2
Ala 10	34.1	Ser 30	26.7
Cys 11	34.8	Asn 36	17.6
Phe 13	112.9	Thr 37	44.6
Thr 16	18.9	Gln 59	18.8
Leu 34	40.7	Phe 65	32.7
Cys 35	19.1	Leu 66	35.3
Gln 37	10.7	Arg 68	19.4
Arg 46	40.9	Ala 69	55.6
Ser 47	32.5	Glu 70	25.2
Lys 48	26.9	Ser 72	55.5
Gln 49	53.7		
Val 50	26.9		
Cys 51	30.7		
Total	907.4	Total	751.5

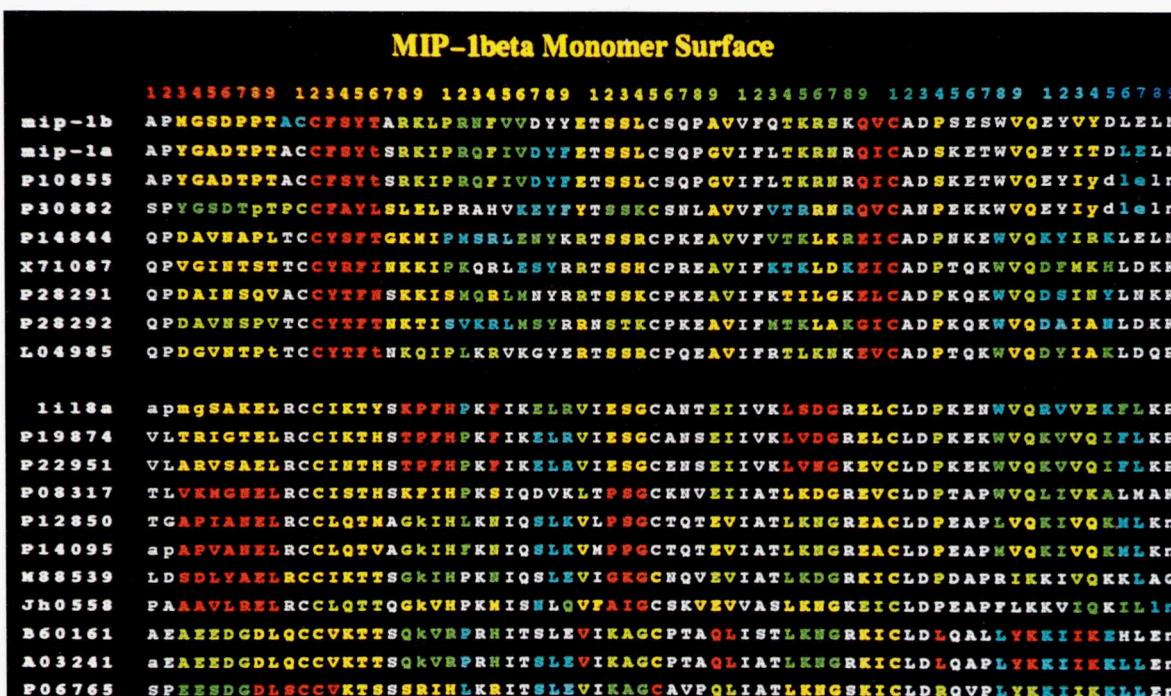
interaction. As in the case of the β subfamily, the α subfamily of sequences conserves the pattern of surface clustering found for native IL-8, albeit composed of an entirely different set of sequence positions from that found for the β subfamily of sequences. None of the amino acid types found at the positions in the strongest surface cluster are conserved within the α subfamily.

The lower portions of Figure 4A and B list the strongest surface clusters for the α subfamily of sequences applied to the hMIP-1 β coordinates and for the β subfamily of sequences applied to the IL-8 coordinates, respectively. In no case was the pattern of strongest surface cluster conserved with the sequence from the opposite subfamily. The α subfamily of sequences superimposed on the hMIP-1 β coordinates tended to shift the strongest hydrophobic surface cluster to the amino-termini and to regions flanking the positions of greatest hydrophobicity for the β subfamily (cf. compare upper and lower panels of Fig. 4A). In particular the portion of the surface identified as the second-ranked hydrophobic cluster for the β subfamily of sequences becomes the first-ranked cluster for 5 of the 11 members of the α subfamily. A similar shift was observed for the β subfamily onto the IL-8 coordinates, but to amino-termini positions slightly different than those found for the α subfamily (cf. compare upper and lower panels of Fig. 4B). Here too the same exchange of first- and second-ranked clusters occurs between the β and α subfamilies.

Potential ligand interaction sites on the surface of the dimer

The structures of the hMIP-1 β and IL-8 dimers clearly have important implications for the development of agonists and antag-

A



B

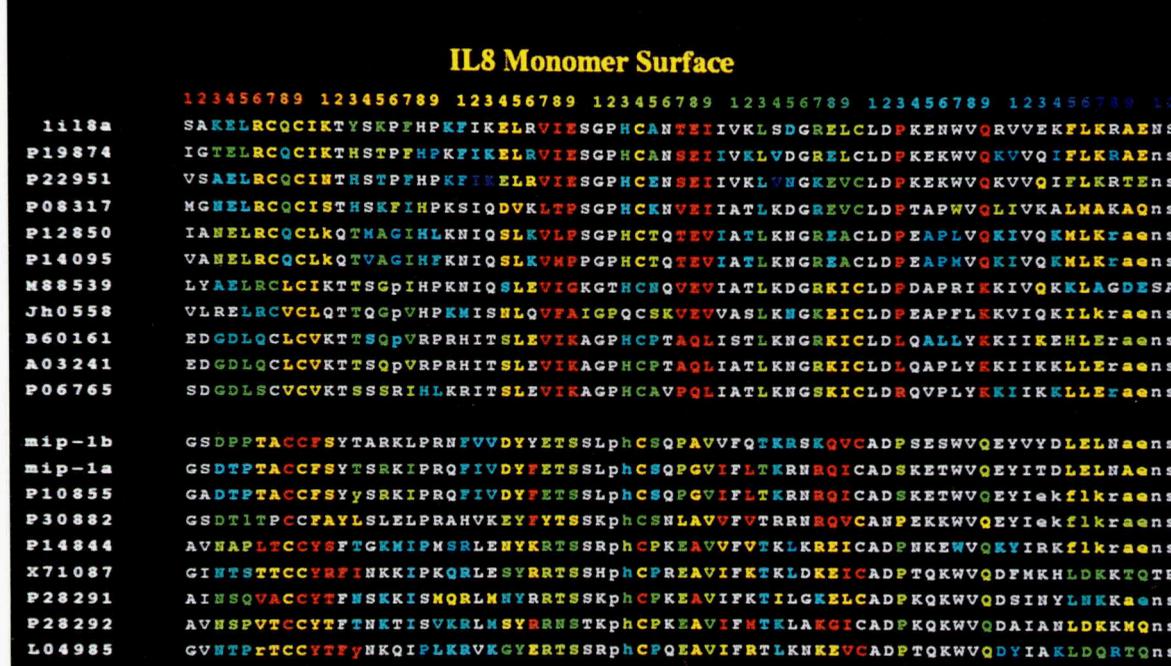


Fig. 4. Alignments of sequence members of each subfamily applied to the hMIP-1 β coordinates (A) and the IL-8 coordinates (B). Sequence positions are spectrally highlighted according to the strength of hydrophobicity of each monomer surface cluster. Colorations and orderings are as described for Figure 2.

onists. Identification of an appropriate target for ligand attachment would be an important first step in rational drug design. The procedure used here to analyze the monomer surfaces has been useful for identifying ligand binding sites in the case of enzyme-inhibitor systems (Young et al., 1994). When applied to the dimerized molecules, candidate target sites can be suggested for attachment sites on hMIP-1 β and IL-8. In other

words, an analysis identical to that performed on each monomer can be applied to the dimer structures. Such surface clusters will invariably be composed of a mixture of hydrophobic and hydrophilic residues. Hence it is important to keep in mind that the analysis seeks the most hydrophobic surface cluster from *all* surface clusters for a given molecule. It should also be remembered that these sites do not necessarily need to consti-

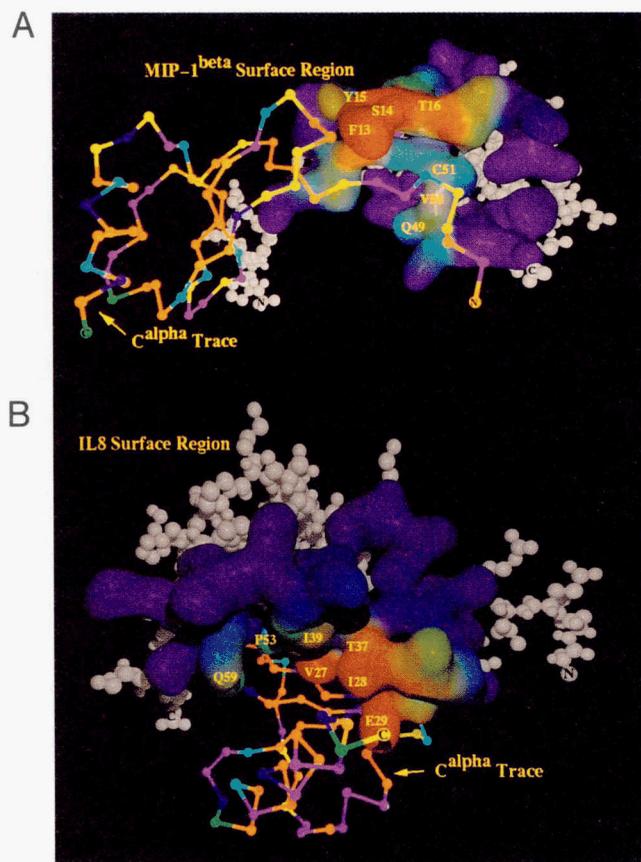


Fig. 5. Ball-and-stick model of hMIP-1 β (A) and IL-8 (B) illustrating the most hydrophobic clusters on the surface of the monomer. The color scheme is the same as that described in Figure 2. The second subunit of the dimer is shown as a C α backbone with the color of each residue indicating the residue's hydrophobicity. Each molecule is viewed from the same position as shown in Figure 3.

tute the binding surface with the cell surface receptor as this interaction need not necessarily be dominated by hydrophobic effects and may have a large contribution from electrostatic components.

The results of this analysis for the hMIP-1 β dimer indicate that the most hydrophobic surface of this structure involves the large concave surface, running at about 60° to the long axis of the dimer. This region corresponds to the lengthwise center of the complex. The most hydrophobic cluster at the surface of the dimer includes residues Ala 10, Tyr 28, Tyr 29, Glu 30, Thr 31, Val 41, Gln 43, Lys 48, Gln 49 on one monomer, and the 2 prolines at positions 7 and 8 on the other monomer. These regions are highlighted in Figure 6A and correspond to the narrow region of the dimer that separates the 2 bulkier portions of each monomer. They comprise a large number of residues because these regions involve relatively flat portions of the dimer's surface. Although these regions share overlap with some of the residues buried at the interface, their geometry and amino acid composition suggest portions of their exposed surface may also serve as target binding sites. Interestingly, this site corresponds to that proposed for the receptor binding site on the basis of sequence comparisons of hMIP-1 β , hMIP-1 α , RANTES, and MCAF (Lodi et al., 1994).

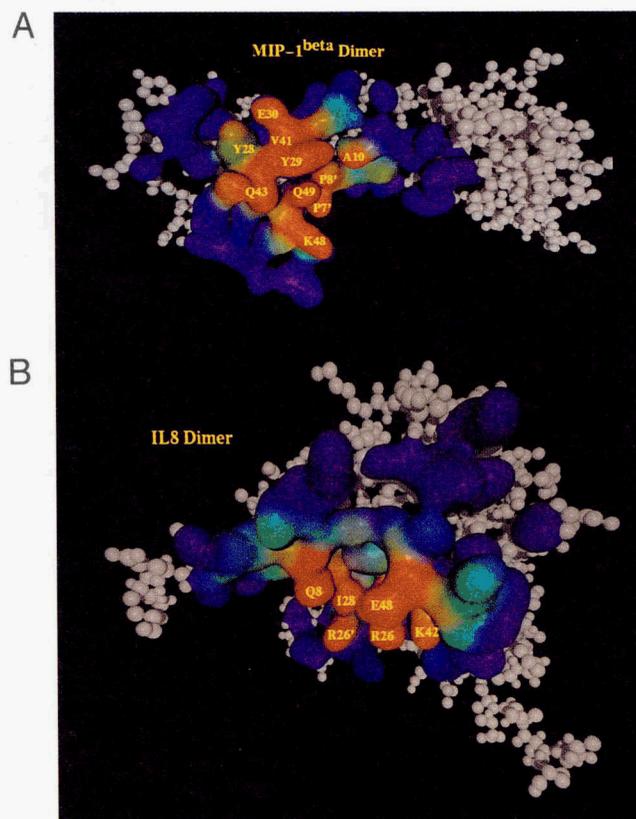


Fig. 6. Ball-and-stick model of the hMIP-1 β (A) and IL-8 (B) dimers illustrating the most hydrophobic clusters at the surface of the dimers. The color scheme is the same as that in Figure 2.

The position of greatest hydrophobicity found for the hMIP-1 β dimer is conserved for 7 of the 9 members of the β subfamily. The 2 exceptions are murine MIP-1 α (P10855) and SISD/RANTES (P30882) where a portion of the surface involving residues Phe 13, Ser/Ala 14, Cys 35, Ser 36, Gln/Asn 37, and Pro/Leu 38 on one monomer and Leu/Lys 34, Cys 35, Ser 36, Gln/Asn 37, and Pro/Leu 38 on the other monomer are more hydrophobic. These residues are also located along the lengthwise center of the molecule but on the side opposite to that of the most hydrophobic surface cluster in hMIP-1 β (i.e., the former is on the convex upper surface, whereas the latter is on the concave undersurface of the molecule).

The corresponding analysis of the IL-8 structure indicates that the most hydrophobic surface on the dimer lies at the side of the cleft formed by the 2 monomers. The residues in this cluster include Gln 8, Arg 26, Val 27, Ile 28, Ile 40, Val 41, Lys 42, and Glu 48 on one monomer and Arg 26 of the other monomer. These residues are highlighted in Figure 6B. The overlap between these residues and those found to form the most hydrophobic surface cluster on the monomer indicate that a portion of this surface is available for binding even in the dimerized molecule. It is interesting to note that this region overlaps the binding surface mapped out by NMR for an N-terminal fragment of the type I human IL-8 receptor, which comprised residues in strand β 3 (Glu 48 to Cys 50), the turn preceding β 3 (Ser 44), and the irregular N-terminal loop region (Gln 8, Thr 12, Lys 15, His 18, Lys 20, and Phe 21) (Clubb et al., 1994).

The position of greatest surface hydrophobicity found for the IL-8 dimer is conserved across the sequence members of the α subfamily with the exception of PCEF-4 (P08317). In the latter case, the most hydrophobic position includes interactions with Leu 43, Arg 47, and Val 49, instead of interactions with Ile 40, Val 41, and Lys 42, as found for IL-8. This region of greatest hydrophobicity is shifted toward the side of the cleft as opposed to the bottom of the cleft as in the case of IL-8. Both regions, however, have an adjacent border identified by the spatial proximity of positions 43, 47, and 49 with those of 40, 41, and 42.

It is interesting to compare these results with the experimental finding that the Glu 4-Leu 5-Arg 6 sequence at the N-terminus of IL-8 is a critical component of receptor binding such that IL-8 activity is lost when these 3 residues are replaced (Oppenheim et al., 1991; Baggiolini et al., 1994). The Glu-Leu-Arg region contributes to the second- and third-ranked most hydrophobic clusters. Glu 4 and Leu 5 are included in the lower-ranked surface regions of this pair. This lower-ranked cluster is comprised of many hydrophilic residues (Lys 25, Lys 30, Asp 45, and Arg 47) and suggests that important electrostatic interactions are likely to be involved with this portion of the surface.

Concluding remarks

The quantitative analysis of the hydrophobic properties of the α and β chemokines presented in this paper reveals patterns of strongly hydrophobic core and monomer surface clusters that are consistent with the similarities in tertiary structures and differences in quaternary structure. Both monomers have their strongest core hydrophobic cluster at equivalent positions. This supports the observation of structural similarity between each monomer subunit. In contrast, the pattern of surface hydrophobicity differs between hMIP-1 β and IL-8 in a manner consistent with the observed differences in their quaternary structures. The strongest hydrophobic surface clusters of each monomer involve regions on opposite sides of the monomer, regions that correlate well with the most buried residues at their respective dimer interfaces. The features of hydrophobicity that are involved in the similarities of monomeric tertiary structure and differences in quaternary structure are preserved within all the members of the β and α subfamilies. Thus, similarities are observed in the pattern of residues observed in the core cluster of greatest hydrophobicity for the entire chemokine family. Sequence members of both subfamilies demonstrate a consistent pattern of core hydrophobicity, when mapped onto the coordinates of either hMIP-1 β or IL-8. In contrast, the pattern of strongest monomer surface hydrophobicity observed within each subfamily is not preserved for sequence members of the opposite subfamily, and, moreover, no overlap is observed between the strongest hydrophobic monomer surface clusters for the hMIP-1 β and IL-8 structures. These differences provide a rational explanation for the lack of receptor crossbinding and reactivity between the β and α chemokine families.

Acknowledgments

This work was supported by the AIDS Targeted Anti-Viral Program of the Office of the Director of the National Institutes of Health (G.M.C.

and A.M.G.). The BLAST search was done at the National Center for Biotechnology and Information using the BLAST network service.

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tools. *J Mol Biol* 215:403–410.
- Baggiolini M, Dewald B, Moser B. 1994. Interleukin-8 and related chemotactic cytokines—CXC and CC chemokines. *Adv Immunol* 55:97–179.
- Baldwin ET, Weber IT, St. Charles R, Xyan JC, Appella E, Yamada M, Matsushima K, Edwards BFP, Clore GM, Gronenborn AM, Wlodawer A. 1991. Crystal structure of interleukin-8: Symbiosis of NMR and crystallography. *Proc Natl Acad Sci USA* 88:502–506.
- Carson M. 1987. Ribbon models of macromolecules. *J Mol Graphics* 5:103–106.
- Christenson HK, Claesson PM. 1988. Cavitation and the interaction between macroscopic surfaces. *Science* 239:390–392.
- Clore GM, Appella E, Yamada M, Matsushima K, Gronenborn AM. 1990. The three-dimensional structure of interleukin-8 in solution. *Biochemistry* 29:1689–1696.
- Clore GM, Gronenborn AM. 1991. Comparison of the solution nuclear magnetic resonance and crystal structures of interleukin-8: Possible implications for the mechanism of receptor binding. *J Mol Biol* 217:611–620.
- Clubb RT, Omichinski JG, Clore GM, Gronenborn AM. 1994. Mapping the binding surface of interleukin-8 complexed with an N-terminal fragment of the type 1 human interleukin-8 receptor. *FEBS Lett* 338:93–97.
- Cornette JL, Cease KB, Margalit H, Spouge JL, Berzofsky JA, DeLisi C. 1987. Hydrophobicity scales and computational techniques for detecting amphipathic structures in proteins. *J Mol Biol* 195:659–685.
- Covell DG. 1992. Folding protein α -carbon chains into compact forms by Monte Carlo methods. *Proteins Struct Funct Genet* 14:192–204.
- Covell DG. 1994. Low resolution models of polypeptide chain collapse. *J Mol Biol* 235:1032–1043.
- Covell D, Jernigan R. 1990. Conformations of folded proteins in restricted spaces. *Biochemistry* 29:3287–3294.
- Dill KA. 1990. Dominant forces in protein folding. *Biochemistry* 29:7133–7155.
- Eisenberg D, McLaglan AD. 1986. Solvation energy in protein folding and binding. *Nature* 319:199–203.
- Genetics Computer Group. 1991. *GCG database software package version 7*. Madison, Wisconsin, USA.
- Jernigan RL, Margalit H, Covell DG. 1989. Coarse graining conformations: A peptide binding example. In: Beveridge DL, Lavery R, eds. *Theoretical biochemistry and molecular biophysics*. Schenectady, New York: Adenine Press. pp 69–76.
- Kauzmann W. 1959. Some factors in the interpretation of protein denaturation. *Adv Protein Chem* 14:1–63.
- Levitt M. 1976. A simplified representation of protein conformations for rapid simulation of protein folding. *J Mol Biol* 104:59–107.
- Lodi PJ, Garrett DS, Kuszewski J, Tsang MLS, Weatherbee JA, Leonard WJ, Gronenborn AM, Clore GM. 1994. High resolution solution structure of the β chemokine hMIP-1 β by multidimensional NMR. *Science* 263:1762–1767.
- Mayo KH, Chen MJ. 1989. Human platelet factor 4 monomer-dimer-tetramer equilibria investigated by ^1H NMR spectroscopy. *Biochemistry* 28:9469–9478.
- Miyazawa S, Jernigan RL. 1985. Estimation of effective interresidue contact energies from protein crystal structures: Quasi-chemical approximation. *Macromolecules* 18:534–552.
- Nozaki Y, Tanford C. 1971. The solubility of amino acids and two glycine peptides in aqueous ethanol and dioxane solutions. *J Biol Chem* 246:2211–2217.
- Oppenheim JJ, Zachariae COC, Mukaida N, Matsushima K. 1991. Properties of the novel proinflammatory supergene “intercrine” cytokine family. *Annu Rev Immunol* 9:617–648.
- Schall TJ. 1991. Biology of the RANTES/SIS cytokine family. *Cytokine* 3:165–183.
- Skolnick J, Kolinski A. 1990. Simulations of the folding of globular proteins. *Science* 250:1121–1125.
- St. Charles R, Walz DA, Edwards BFP. 1989. The three-dimensional structure of bovine platelet factor 4 at 3.0 Å resolution. *J Biol Chem* 264:2092–2099.
- Young L, Jernigan RL, Covell DG. 1994. A role for hydrophobicity in protein-protein interactions. *Protein Sci* 3:717–729.