Molecular modeling of the GM-CSF and IL-3 receptor complexes

P.D. LYNE, P. BAMBOROUGH, D. DUNCAN, AND W.G. RICHARDS

Oxford Centre for Molecular Science, Physical and Theoretical Chemistry Laboratory,
South Parks Road, Oxford OX1 3QZ, United Kingdom

(RECEIVED June 9, 1995; ACCEPTED July 24, 1995)

Abstract

A model for the structure of the cytokine interleukin-3 (IL-3) is presented based on the structural homology of the hematopoietic cytokines and utilizing the crystal structures of interleukin-5 and granulocyte macrophage colony stimulating factor (GM-CSF). In addition, models of the receptor complexes of GM-CSF and IL-3 are presented based on the structural homology of the hematopoietic receptors to growth hormone. Several key interactions between the ligands and their receptors are discovered, some in agreement with previous mutagenesis studies and others that have not yet been the subject of mutagenesis studies. The models provide insights into the binding of GM-CSF and IL-3 to their receptors.

Keywords: cytokines; GM-CSF; homology modeling; IL-3; receptor complexes

Granulocyte macrophage colony stimulating factor (GM-CSF) (Diederichs et al., 1991; Walter et al., 1992), interleukin-3 (IL-3) (Yang et al., 1986), and interleukin-5 (IL-5) (Milburn et al., 1993) are all hematopoietic cytokines (Boulay & Paul, 1993), but they may be grouped together into a subfamily because they share a common receptor chain (Hayashida et al., 1990; Kitamura et al., 1991a; Tavernier et al., 1991). These cytokines are involved in maintaining and controlling the basophils (Lopez et al., 1990) and eosinophils (Lopez et al., 1991) in the body.

Cytokines form a large group of protein messengers with many complex functions. At present there are known to be several families of cytokines (for a review, see Sprang & Bazan 1993), including β -trefoil folds such as interleukin-1 α and β , and the β -jellyroll form of tumor necrosis factor (TNF)- α , but the largest is the four-helix bundle family. Members of this family include interleukins such as IL-3 and IL-5, colony stimulating factors such as GM-CSF, growth hormone (GH) and interferon (IFN) α , β , and γ , although the interferons contain additional long helices. The basic fold is, as the name implies, of four long α -helices arranged into two sets of parallel helices running in an "up-up-down-down" fashion (Abdel-Meguid et al., 1987), joined by loops of often irregular secondary structure and variable length. There is no strong sequence similarity within the family. Recent reviews of the four-helix bundle family include those by Boulay and Paul (1992), Wlodawer et al. (1993) and Mott and Campbell (1995).

The receptors for these ligands also form related groups. They are all membrane-spanning proteins, containing an extracellular N-terminal domain, a single transmembrane segment, and a C-terminal cytoplasmic domain. There is no known homology among the cytoplasmic domains of the receptors. The structural and sequence similarity lies in the extracellular domain (Bazan, 1990). The fundamental element of these receptors is a domain with the immunoglobulin fold, which contains seven strands making up two β -sheets. Two of these domains, each of which have their own conserved features, are linked by a short sequence between the C-terminus of the first and the N-terminus of the second, and constitute the basic receptor building block, hereafter referred to as the hematopoietic receptor (HR) unit. These domains seem to fold independently, because the addition of extra Ig-type domains both before and after the HR domains does not seem to disrupt their structure. In addition, the HR contains a highly conserved WSXWS motif. The biological role of this sequence still remains unclear (Kossiakoff et al., 1994).

The human GM-CSF, IL-3, and IL-5 receptors are heterodimeric, consisting of an α chain and a β chain (Hayashida et al., 1990; Kitamura et al., 1991b; Tavernier et al., 1991). The α chain is specific for its cognate ligand, which it can bind with low affinity, but in the absence of the β chain, no signal is transmitted to the cell. The same β chain is used as a component of all three receptors, and the $\alpha\beta$ complex shows high-affinity binding to the ligands, together with signal transduction. By itself, the β chain shows no detectable binding affinity, but there is likely to be some direct interaction between this chain and the ligands.

Here structural models of IL-3 and the IL-3 and GM-CSF receptor complexes are presented. These models are built by ho-

Reprint requests to: W.G. Richards, Oxford Centre for Molecular Science, Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, United Kingdom; e-mail: gr@vax.ox.ac.uk.

Fig. 1. Ribbon diagram of IL-3 drawn with MOLSCRIPT (Kraulis, 1991). Helix A, yellow; helix B, red; helix C, blue; helix D, green.

mology modeling techniques utilizing the crystal structures of GM-CSF (Diedrichs et al., 1991), IL-5 (Milburn et al., 1993), and the GH receptor complex (de Vos et al., 1992). These models have been reconciled with experimental data from the literature and it is hoped that they will aid understanding of the biological activities of GM-CSF and IL-3.

Results and discussion

For the purposes of the discussion, modeling of IL-3 and each of the receptor complexes will be presented separately for clarity.

IL-3

The three-dimensional structure of the model of IL-3 is shown in Figure 1. From Table 1 it may be seen that pairwise RMS differences of the helical C α of IL-3 and selected cytokines are low. All backbone angles have *trans* planar omega angles and a Ramachandran plot showed the backbone angles to fall into commonly observed regions of phi-psi space.

The 4 α -2 β fold is a characteristic feature of the cytokines. Both structural and mutagenesis studies reveal that the structural integrity and binding specificity of these proteins are associated with the four-helix bundles. The structures of GM-CSF and IL-5 contain β strands in addition to the four-helix bundle. Although not precluding the possible existence of β strands in IL-3, the modeling focussed on the structure of the four-helix bundle.

A Connolly surface calculation (Connolly, 1983) of the protein structure revealed that the model is free of large holes. Inspection of the structure shows that the hydrophobic core is

comprised of the residues Ile 7, Ile 10, Ile 11, Ile 14, Ile 17, Pro 18, Gly 29, Leu 33, Met 36, Leu 45, Phe 48, Ala 51, Ile 61, Ile 64, Leu 68, Phe 94, Leu 98. The core is free of charged residues.

The GM-CSF sequence has four cysteine residues, whereas IL-5 has only two residues. In both of these proteins, the cysteines have integral roles in determining the fold by forming disulfide bonds between the helices. The two cysteine residues in IL-3 (Cys 17 and Cys 83) form a disulfide bond to link the N terminus of helix A to helix C.

The implications of the structures of IL-3 and GM-CSF on ligand-receptor interactions will be discussed in the following sections.

GM-CSF receptor complex

A ribbon diagram of the complex is shown in Figure 2A. In the discussion below, residues in the receptor chains are numbered

Fig. 2. **A:** Ribbon diagram of the GM-CSF receptor complex. GM-CSF, red; α chain, green; β chain, blue. **B:** Ribbon diagram of the IL-3 receptor complex. IL-3, light green; α chain, pink; β chain, blue. Pictures drawn with MOLSCRIPT (Kraulis, 1991).

Table 1. RMSD scores (in Å) obtained on three-dimensional superposition of helical alpha carbons of the named structures

	IL-3	IL-5	IL-4	GH	GM-CSF	G-CSF
IL-3	—	0.91	0.83	1.33	0.88	1.92
IL-5		—	1.70	1.63	1.33	1.48
IL-4			—	2.29	1.71	2.17
GH				—	1.73	1.625
GM-CSF					—	1.91

according to those in their published sequence. Residues in GM-CSF are numbered from the pdb file, starting at residue serine 5 in the sequence SPSPSTEPW. . . .

The complex is a compact structure, with many favorable contacts between the three components, as shown in Table 2. It is not clear whether or not the receptor dimerizes in the absence of GM-CSF, and this is reflected by the small number of close contact between the two chains. There is no simple solution to the question of which residues control the preference of cytokines for their receptors, but studies on GM-CSF have singled out certain residues, forming a hydrophobic patch surrounded by polar side chains (Diedrichs et al., 1991). This region is centered on residues R24 and M79, both of which are observed to have close contacts in the model. This proposed identification site fits with current theory on protein-protein recognition (Janin & Chothia, 1976; Clackson & Wells, 1995), which suggests that hydrophobic groups are present to give strength to the interaction and hydrophilic groups to give specificity.

Because the human and mouse GM-CSF are specific to their own species, chimeric proteins can be used to identify regions of binding importance (Kaushansky et al., 1989). This study showed that hGM-CSF residues 21–31 and 78–94 are important to recognition. These are in helices A and C, which is consistent with the pattern of binding in hGH and in this model. In addition, one or both of residues R23 and R24 are important in species recognition, because the hGM-CSF mutant R23Q loses much of its activity, and because the mGM-CSF ²³LNL-RRL mutation (where the mouse tripeptide is replaced by the human sequence) is reduced to 2% of its former activity. In addition, the glycosylation site at hGM-CSF N27 increases in activity 20-fold on removal of the carbohydrate, and N27 is adjacent to R24 on the A-helix (Moonen et al., 1987). This is consistent with

the removal of steric interference from the glycosyl group on the receptor binding site.

In the model, side chain R24 lies close to the oppositely charged side chain D353, and to the oxygen atoms of the backbone of residues E350 and H351. Disruption of this residue would cause a loss of the salt bridge to D353, as well as possible hydrogen bonds to the backbone oxygens, and this would clearly cause a loss of binding affinity.

Mutation experiments have confirmed the importance of residue E21 for biological activity and for high-, but not low-affinity binding (Lopez et al., 1992a; Shanafelt et al., 1992; Hercus et al., 1994). This is consistent with the contacts made by this residue (Table 2), which are with the β -chain of the receptor and hence which only take place in the high-affinity (biologically active) complex. These studies have replaced E21 with alanine and arginine, and both mutants lose much of their biological activity, E21A 30-fold, and E21R 100-fold (Greenfield et al., 1993). The fact that both hydrophobic and hydrophilic replacements cause loss of activity is indicative that it is not merely a polar interaction being lost. Because, from the model, E21 is close to a histidine residue, this may explain why substituting arginine (which has the same charge as histidine) leads to such a large loss. Presumably, the E21A mutant has reduced activity due to a reduction in hydrophobic contacts with I352 (see Table 2). The side chain of E21 is shown in the environment of receptor residues in Figure 3.

From the model, other residues may be identified as important for binding of GM-CSF to the β chain. E14 forms a salt bridge with K263 and forms a hydrogen bond with the amide group of Q302. It may be noted also that N17 is in close proximity to the backbone oxygen of H300. It is presumed that the main source of binding between GM-CSF and the β receptor is

Table 2. Selected favorable close contacts between GM-CSF and its receptor ligands^a

Interface 1			Interface 2			Interface 3		
GM-CSF	GM-CSFR α	Distance	GM-CSF	GM-CSFR β	Distance	GM-CSFR α	GM-CSFR β	Distance
R23 NE	Y226 OH	3.0	E14 NE1	Q302 N	3.3	N250 N	N377 OD1	2.7
R23 NH2	L247 O	2.9	E14 OE1	K263 NZ	2.6	S251 N	T455 O	3.0
R23 NH2	D228 OD2	2.6	E14 OE2	K263 NZ	2.5	Y259 OH	S380 OG	2.5
N27 OD1	R275 NH2	3.0	N17 ND2	H300 O	2.9	R277 NH1	D336 OD2	2.9
N27 OD1	R275 NH2	2.7	E20 OE1	G404 N	2.9	R277 NE	D336 OD2	2.9
T98 O	N134 ND2	2.7	E20 CB	I352 CD1	3.8			
Q99 NE2	S156 OG	2.8	E21 OE2	H300 ND1	4.0			
I100 O	R167 NH2	2.6	E21 OE2	A239 O	3.4			
I100 O	R167 NH1	2.5	E21 CB	I352 CD1	4.0			
E104 OE2	K192 NZ	2.6	E21 CG	I352 CD1	3.9			
E104 OE1	K192 NZ	2.6	R24 NH1	D353 OD1	2.6			
K107 O	R275 OD1	2.6	R24 NH2	E350 O	2.6			
Q108 OE1	K192 NZ	2.7	K72 NZ	A239 O	2.4			
Q108 OE2	K192 NZ	3.1	M79 CE	Y262 OH	2.8			
K111 NZ	D254 OD1	2.6						
L115 O	K223 NZ	2.5						
V116 O	K223 NZ	2.6						
I117 O	K223 NZ	2.6						

^a Residues are given by their number in the relevant sequence. Distances are quoted in Ångstroms between the indicated atoms for each close contact.

Fig. 3. Interactions of Glu 21 on helix A of GM-CSF with the β receptor chain. The GM-CSF main chain is shown as light grey and the β main chain is shown as dark grey. Picture drawn with MOLSCRIPT (Kraulis, 1991).

the A α helix of GM-CSF (Shanafelt et al., 1992). However, from the model it is noted that there are possible strong interactions between K72 and M79 of the C α helix with A219 and Y262, respectively, of the β receptor.

Recently, Woodcock et al. (1994) have identified a region on the β receptor chain that is important for high-affinity binding to GM-CSF. Site-directed mutagenesis studies led the authors to the conclusion that the region Y365–I368 (this sequence corresponds to Y349–I352 in this work), excluding E366, is essential for recognizing the A helix of GM-CSF, in particular for binding to E21 on the A helix. The results are consistent with the report of Lock et al. (1994) that H367 is essential for binding, but conflict with the latter authors' conclusion that T365 is unimportant for binding. As may be seen from Figure 3 and Table 2, the side chains of H351 and I352 are ideally positioned to interact with E21. The side chain of E350 (corresponding to E366 in the Woodcock paper) is seen to be directed away from the binding site. It is also found from the model that Y249 is not involved in binding, as found by Lock et al. (1994). An additional histidine residue, H300 is also situated in close proximity to E21.

As yet, no mutagenesis experiments have been carried out that give any information about the binding of GM-CSF explicitly to the α chain. Table 2 shows some contacts taken from the model. Although the majority of interactions between GM-CSF

and the α receptor chain are centered on helix D of GM-CSF, it is noted that residue R23, which was assumed to take part in β -chain binding as it lies on helix A, forms a charge pair partnership with GM-CSF α residue D251, and with the backbone oxygen of GM-CSF α L270D.

The N27 residue is also seen to interact with the α receptor because it is ideally placed for forming hydrogen bonds with the side chain of R299. Thus, deglycosylation of this residue should not only enhance the activity of the protein by reducing steric interference with R23, but also by affording another method to bind the receptor.

Another point of interest that can be seen in Figure 7 is that the side chains β H300 and H351 are aligned with the two most important hGHR side chains, W104 and W169, respectively. That these appear to be so closely involved in binding to hGM-CSF side chains that are known to be important is further strong evidence in favor of the models. Mutagenesis experiments directed at these positions would be useful either to confirm the model's accuracy, or to suggest that it may be in error.

A study into the antagonist action of growth hormone mutants (Fuh et al., 1992) suggested that the antagonist action arises from prevention of sequential dimerization of the receptors. Dimerization of many receptors seems to be the initiator of, and essential for, propagation of the cell messenger process. GM-CSF is proposed to be similar to growth hormone in that receptor aggregation can only take place after GM-CSF is bound. If this is the case, then dimerization would not be possible without binding of GM-CSF to either one of the receptors. In Table 3, hGHR 1 is the equivalent to the GM-CSF β -chain because it forms interface I with GH. Consequently, it can be seen that, unlike GH, the GM-CSF complex has a greater buried surface area around the α -chain, suggesting that the binding site around interface II is of higher affinity than the β -chain receptor site. K_d for the low-affinity (α -chain only) complex is 5–8 nM; K_d for the high-affinity ($\alpha\beta$ complex) is 20–50 pM (Hoang et al., 1993).

It is also noticeable that, in both the GM-CSF complex and the hGHR crystal structure, the conserved WSXWS motif of the receptor plays no part in dimerization. The sheet preceding this motif is also highly conserved, as are the proline-rich regions, 217NPPSN (α -chain) and 325APPSL (β -chain). All of these three conserved regions are brought close in space on the individual receptor chains; they are found in the hinge region, but on the opposite side to the GM-CSF binding face, on the concave surface of the receptors. It has been shown that the WS motif of erythropoietin receptor is of great importance to cell surface expression, perhaps having an effect in assisting the folding (Yoshimura et al., 1992). Alternatively, it may recognize an unknown protein that either assists in transport to the membrane or acts as another signaling molecule.

Table 3. Connolly surface area buried between the ligand and the appropriate receptor chain on binding for GM-CSF and GH

	GM-CSFR α	GM-CSFR β	hGHR 1	hGHR 2
Buried hydrophobic surface area/ $\text{\AA}^2$	729	550	—	—
Buried hydrophilic surface area/ $\text{\AA}^2$	1,227	1,269	—	—
Total buried surface area/ $\text{\AA}^2$	1,956	1,819	1,230	900

Table 4. Selected favorable close contacts between IL-3 and its receptor ligands^a

Interface 1			Interface 2			Interface 3		
IL-3	IL-3 α	Distance	IL-3	IL-3 β	Distance	IL-3 α	IL-3 β	Distance
K28 NZ	E258 OE2	2.6	N18 ND2	N406 N	2.8	T202 CG2	P384 CG	3.5
K28 NZ	R259 O	2.5	M19 CA	V320 CG1	3.4	N246 ND2	D336 OD2	2.7
K28 NZ	V260 O	2.5	M19 CE	V320 CG1	3.7	N246 ND2	D336 OD1	3.4
D103 OD1	S159 OG	2.7	D21 OD1	Y405 OH	2.9	R259 NH1	D353 OD1	2.7
D103 OD2	S159 OG	2.7	D21 OD2	Y405 OH	2.7	R259 NH2	D353 OD2	2.7
D103 OD2	R127 NH1	2.8	D21 CB	I352 CD1	4.3			
D103 OD1	Q186 NE2	3.3	D21 CB	I303 CG2	4.0			
E106 OE1	Q186 NE2	3.4	E22 OE2	H300 CD2	3.3			
F107 CZ	L155 CD2	3.7	E22 OE2	H300 N	3.3			
R109 NH1	Y261 OH	2.8	E22 OE2	T299 CB	3.7			
T112 OG1	K217 NZ	2.5	T25 OG1	H351 ND1	2.6			
T117 O	K217 NZ	2.7	K79 NZ	H300 O	2.6			
T117 OG1	N215 ND2	2.9	K116 NZ	Y405 O	2.6			

^a Residues are given by their number in the relevant sequence. Distances are quoted in Ångstroms between the indicated atoms for each close contact.

IL-3 receptor complex

The tertiary structure of the IL-3 receptor complex is shown in Figure 2B. In Table 4, a selection of favorable close contacts between IL-3 and each receptor, and between the two receptors, are shown.

There are less mutation data relating to the formation of the IL-3 complex than for the GM-CSF complex. Nevertheless, because IL-3, GM-CSF, and IL-5 are so closely related, it is expected that the key binding features evident for GM-CSF may also apply to the IL-3 complex.

Site-directed mutagenesis studies have identified a number of residues that are important for activity (Lopez et al., 1992b). Like GM-CSF, IL-3 binds to the α receptor with low affinity, which then forms a complex with the common β chain that is essential for signaling (Kitamura et al., 1992). Lopez et al. (1992b) have demonstrated previously that the mutants D21A, E22L, and T25A lack high-affinity binding. From Table 5 and Figure 4, the main interactions between these residues and the receptor chains are shown. The D21 residue forms hydrogen bonds to Y405 on the β chain and makes hydrophobic contacts with the side chains of I352 and I302. The E22 residue, conserved in the GM-CSF, IL-3, and IL-5 subfamily, binds to H300 and T299 on the β chain (Barry et al., 1994). Another residue identified as important from mutagenesis studies, T25, is seen to interact with H351.

Table 5. Conserved helical residues used in the three-dimensional superposition and in the building of the structurally conserved core of IL-3

	Helix A	Helix B	Helix C	Helix D
GM-CSF	V16–N27	Q56–G65	K74–H87	E104–V116
IL-5	S2–S13	E22–E31	L40–Q53	V59–V71

There is a dearth of information concerning the nature of the interaction of the α receptor to IL-3. Before the analysis of the GM-CSF model, it was found that helix A also binds to the α receptor in addition to the common β chain. There is a short motif on the IL-3 α receptor, 258ERV, that is seen to bind to K28 on the A helix of IL-3. Other interactions between the β chain and the ligand include N18–N406 on the receptor and, interestingly, K79 to the β chain. The K79 residue was aligned to an

Fig. 4. Interactions of Glu 22 and Thr 25 on helix A of IL-3 with the β receptor chain. The IL-3 main chain is shown as light grey and the β main chain is shown as dark grey. Picture drawn with MOLSCRIPT (Kraulis, 1991).

analogous residue in IL-5 and one residue away from K72 of GM-CSF, which is also seen to bind to the β chain in the GM-CSF complex. The majority of interactions between the α receptor and the ligand are with helix D. Several of these contacts are shown in Table 4. As with the GM-CSF complex, there are few contacts between the receptor chains. There is a salt bridge between R259 on the α chain and D353 on β . Once again, D336 on β is seen to be involved in interreceptor interactions.

Conclusions

It is clear from the models that there are many similarities in the specific interactions that GM-CSF and IL-3 make to their respective receptors. In particular, residues on helix A bind to both receptor chains, whereas helix D binds only to the specific α chains. Many of the close contacts established from the model concur with results from previous mutagenesis studies, and some new key residues have also been discovered.

The largest potential for errors in the receptors will probably stem from the initial alignment and uncertainty in exactly where the sheets begin and end, leading to the possibility that the hydrogen bonds between paired sheet strands will be out of registry. The existence of two disulfide bonds in the N-terminal domain of each receptor constrains the strands and will reduce this problem, as will the disulfide bond proposed in the second domain of the α -chain. For GM-CSF, the disulfide bonds are between cysteines 123–133, 162–176, and 229–234 in the α -chain, whereas for IL-3, the disulfide bonds are between 94–104, 133–147, and 199–275. For the common β chain, the disulfides are between cysteines 234–244 and 274–291.

The docking of the receptors to the ligands could also have introduced errors. Although several favorable interactions exist in the model, other binding modes could exist, especially if binding does not follow the GH binding model as closely as was assumed. It was also supposed that there were no substantial changes in the conformation of the ligands on binding, and this may not be the case. Changes in conformation of the receptors should be less significant, if the GH binding model is followed, because the receptors were modeled on hGHR already bound in a complex. However, the large size of the interface makes performing an exhaustive search of all docking orientations im-

possible. The fact that several of the significant interactions produced are between residues predicted to be essential for activity lends credence to the binding model used.

Although not proven, the models appear to be reasonable in several significant ways. Firstly, the receptor is a low-energy structure with a hydrophobic core. In addition, one of the hydrophobic tryptophan residues in the WSXWS motifs in each of the two chains is at least partially exposed. Charged residues and many polar residues are exposed to solvent, except those that occur in the interface regions. More than 95% of residues have backbone angles that fall into commonly observed regions of the Ramachandran plot. The receptors seem to form a surface that is largely complementary to their respective ligands.

It is hoped that the models may be of help in understanding which residues are involved in stabilizing the binding interaction, and will help focus any future mutation studies carried out on human GM-CSF and IL-3. The coordinates are available upon request.

Methods

IL-3

Previous alignments of the cytokines have relied on the UWGCG Bestfit algorithm (Devereux et al., 1984) or similar automatic sequence alignment methods. Because of the relatively low sequence identity between the cytokines, the results of this approach do not agree with the alignments obtained by tertiary structure superposition of the known cytokines, so it must be assumed that there is a high probability that the results will also be inaccurate for alignments involving IL-3. An alignment obtained using the UWGCG Pileup program is shown in Figure 5. As can be seen, the program fails to align residues in the hGH helix D to the structurally equivalent residues in the other sequences, and inserts gaps into the regions aligned with helix A, B, and C. In addition, the cytokine receptors have 19.4% sequence identity over a 191-residue range (Bamborough et al., 1993). Therefore, multiple structural alignment was used, along with manual alignment using the graphical tool CAMELEON (Morris, 1988), to generate an alignment of IL-3 with the rest of the family. Three-dimensional structures (Bernstein et al., 1977) of GM-CSF (Diedrichs et al., 1991) and both monomers of IL-5 (Milburn et al., 1993) were aligned in QUANTA (Molecular Simulations, Inc.) by superposition of C α atoms, and the RMSD scores are given in Table 1. In CAMELEON, hydrophobic residue properties were used, rather than residue identities, because the sequences are clearly evolutionarily distant and thus have low sequence identity. Because the structure is a small helical bundle with a hydrophobic core, a more realistic alignment will be obtained by assigning helices in such a way as to bury their hydrophobic faces. The resulting alignment is shown in Figure 6.

Other published alignments of these three sequences have all differed (Boulay & Paul, 1993; Goodall et al., 1993; Rozwarski et al., 1994) and this alignment is different again. However, most of these former studies differ in their alignment of IL-5 rather than IL-3, so structural alignment of IL-5 has been used here to reinforce the sequence alignment. It agrees with all published alignments for helix A and with that of Goodall et al. (1993) for helix D: helix D was the hardest to assign because

hGH	IP[SR]L...FDN...AMLRAHRLHQLAFDTYQEFEEA[IPKEQKYSFLQNPQ
hIL5	IPTSAL...VKETLALLSTHRTLLIANETL.....RIPVPVHKN
hGMCSF	SPSPSTQPWEHVNIAIQEARRLLNLSRDTAEMNE.....
hIL3	LKTSWVNCNMDIEIITHLQKPLPLLDPNNG.....EQDILMENN
hGH	[TSLCFSEISPTPSNR]E...ETQOKS[NLELLRISL.LLIQSWLEPQF]RSVF
hIL5	HQLCTEEIPQIGITLESQTVQGGTVERLFPNL.SLIKKYID.....
hGMCSF	...TVEVISEMFDLQEPCTCLQTRLELYKQGLRGSLLTKLGLPTMMASHY
hIL3	RREPLEAFNRKAVSLQNA...SAIESILKNLLPCLPLATAAPTRHP1HI
hGH	ANSLVYGA[DSNVYDL...LKDLEEGITQLMGR]LEDGSPRTGQIFKQTYSA
hIL5	GQKKKCGEERRRVNQF...LDYLQEFGLVMNTEWIIIES.....
hGMCSF	KQHCPPTPTSCATQIITFESFENLKDPLLVIPFDCEWEPVQH.....
hIL3	KDGDWNEFRKRLTFYKLLENQAQQTTLTSLAIF.....
hGH	FDTSNHNDDA[LKNYGLLYCFRKMDKQVETFLRIVQCR]VEGSCGF
hIL5	
hGMCSF	
hIL3	

Fig. 5. Pileup alignment of human growth hormone, IL5, GM-CSF, and IL3. The hGH helices are boxed.

Fig. 6. Multiple sequence alignment of human IL-3 with human GM-CSF (Diedrichs et al., 1991) and IL-5 (Milburn et al., 1993). The alignment was performed with CAMELEON (Morris, 1988) based on the hydrophobic nature of the residues. Hydrophilic residues are shaded blue and hydrophobic residues are shaded orange. Residues without a strong hydrophobic preference are shaded grey.

there were a number of alignments possible from sequence alone. However, the alignment in Figure 6 was chosen to try and preserve, as much as possible, the positions of buried hydrophobic residues using a feature of CAMELEON that allows a three-dimensional structure to be displayed associated with a sequence and changes seen immediately on realignment. Another possible alignment approach would be to use a profile technique (Bowie et al., 1991). This technique was applied to study the helical cytokines (Bamborough, 1994) and was found not to produce results as good as the structural superposition alignment method used here.

The COMPOSER module of the Sybyl package (Sutcliffe et al., 1987a,b) was then used to build the core of helices A-D, using GM-CSF and the two copies of IL-5 as the basis. To prevent the structure being biased toward IL-5, GM-CSF was given a weighting of 2 in the model-building process, whereas each IL-5 structure was given a weight of 1.

Once the conserved core was built, the positions of the side chains were used to mutate the residues to those of IL-3. Observation showed the core to have no unfavorable interactions, thus it was accepted for the moment and loops were then added to the structure.

The loop between helices A and B in IL-3 is long and loop-searching a database would not provide reasonable structures. Consequently, a loop of the same length was cut from between the same region in IL-2, inserted into the structure, and mutated

to the residues expected of IL-3. This allowed the preservation of hematopoietic geometry in the loop. The other loops were inserted by database searching with care being taken to ensure that the loops chosen followed the correct "threaded" topology common to the other helical cytokines. Once the loops were added, the core helices were extended as shown in Figure 6. This allowed the loops to take up more favorable geometries and improved the packing in the structure.

All the structural characterizations of helical cytokines to date have shown the molecules to contain disulfide bonds. IL-3 contains only two cysteine residues and it could be seen from the model that these were close to each other in space and at an appropriate orientation for forming a disulfide bond, linking the beginning of helix A with helix C.

GM-CSF and IL-3 receptor chains

For the alignment of the cytokine receptor chains, the structure of the human growth hormone receptor was used as the basis for building the models. Unlike the case above for IL-3, there is only one structure known that can be used as the foundation for homology modeling. Nonetheless, a multiple alignment was still valuable to make areas of similarity more obvious, and where possible deletions and insertions were confined to regions that were not β -strand in the hGHR structure. The main features

of the hGHR structure that were used to aid the alignment were the WSXWS motif in domain 2 and four cysteine residues in domain 1 that form two disulfide bonds. In addition, as shown in Figure 7, the hGHR structure has a number of residues buried in the core of the structure. These are expected to be conserved throughout the family and, therefore, the hydrophobic nature of these residues was conserved during the alignment.

Traditional sequence alignment methods were found to be unsatisfactory and thus CAMELEON was used to complete the alignment. The resultant alignment is shown in Figure 7.

Although CD4 (Ryu et al., 1990) and FN3 (Potts et al., 1994) only consist of one domain, they were also aligned with the cytokine receptors because they have a similar structure to that seen in the hGHR chains. They do not share the conserved features of the hematopoietic receptors and, because they are only one domain, they were aligned twice, once with each domain of the other receptors, the two alignments being denoted as x3fn and x3fn_2, and CD4 and CD4_2, in Figure 7. Three-dimensional alignment of hGHR, CD4, and FN3 was undertaken, but using COMPOSER to model the GM-CSF and IL-3 receptor chains on the basis of these three proteins would have led to the rejection of CD4 and FN3 on the basis of insufficient similarity. Consequently, initial modeling of the receptor chains was done by mutation of the GH receptor chain into the appropriate residues using Sybyl.

Insertions and deletions were dealt with by loop searching a database of immunoglobulin structures (proteins in this database are listed in the Appendix using their Brookhaven accession codes), again using Sybyl. A specific database was constructed to increase the accuracy of the loop searches; hematopoietic receptor folding topology is very similar to the immunoglobulin fold, with end-to-end loop distances being approximately the same. Loops with unfavorable backbone angles or bad van der Waals clashes were excluded. Side chains were added automatically using a side-chain rotamer library.

After mutation of the structures, it was noted that, even after no optimization had been carried out, the two cysteines in the second domain of the GM-CSF α receptor chain were very close together in space and in a favorable conformation for a disulfide bond to be built between them. Consequent manipulation of the structure took this into account because this was seen as evidence in support of the model. Sheet regions in the chains were extended or contracted as shown in Figure 7. Extension of sheets was made where loop regions were longer than the equivalent region in hGHR or where the hydrophilic/hydrophobic pattern common to sheets was observed (see sheet F in the second domain of the common β chain for an example). Contraction of sheets was made where loops would otherwise have been too short to join the elements together or where there were disruptive residues such as glycines or prolines—the proline-rich region at the end of sheet E in domain 2 of both receptor chains illustrates this.

Receptor–ligand complexes

The specific α -chains and the common β -chain of the receptors were now built and had to be docked with GM-CSF and IL-3. Because homology with the GH/receptor complex was presumed in the construction of the receptor chains and both GM-CSF and IL-3 are thought to bind in a similar fashion (Goodall et al.,

1993), docking of the receptors and ligands was undertaken by superposition of the structures onto the GH/receptor complex. Overlay of GM-CSF onto GH would not produce a perfect match, thus, in order to maximize the accuracy of the docking, the superposition was done in two parts. GM-CSF has two proposed interfaces with its receptor, one on helices A and B, and the other on helix D; consequently, GM-CSF was first of all overlaid on GH using the C α atoms of helices A and B (the sequences were already aligned [Bamborough, 1994]). The β chain of the receptor, which is proposed to bind to helices A and B, was then overlaid onto the equivalent receptor chain in the GH complex; the resulting half complex is denoted as GM-CSF/R β . Superposition of GM-CSF/R β onto the GH/receptor complex was then performed, using the C α atoms of helix D of GM-CSF as a basis, and the α chain of the receptor, which binds to helix D, then overlaid onto the equivalent chain in the GH structure. A similar methodology was employed for docking IL-3. This meant that the resultant ligand/receptor complexes matched the GH complex most closely in the interface regions, which is the desired result given the expected homologous nature of the complexes. The two complexes were then written out as single files; this would permit optimization of the entire structures because the minimum energy conformations of the components in the complex need not be the same as the conformations in the free ligands and receptors.

Refinement

IL-3

Visual inspection and manual rotation around side-chain bonds removed unfavorable core interactions. Energy minimization was carried out using the CHARMM force-field (Brooks et al., 1983), initially on hydrogen atoms only, then just on side chains, then finally on the entire protein. Dynamics at 300 K was carried out for 20 ps on the loops and side chains, then for 20 ps on the entire structure. The protein was solvated in a water sphere, re-minimized, and a further 40 ps of dynamics was performed before a final minimization.

Receptor complexes

The unsolvated structure was energy minimized in CHARMM in the following sequence; initially the hydrogen atoms, then the side chains and loops, and finally the whole molecule. This was then followed by molecular dynamics at 300 K for 60 ps to improve side-chain packing and to increase the number of hydrogen bonds formed. After this procedure, the structure was manually modified to remove unfavorable backbone angles and close contacts, and to increase favorable interactions in the interface regions. The structure was then re-minimized and another dynamics run made. The process was continued until as many favorable interfacial contacts as possible were made; the resulting structure was then energy minimized and 60 ps of dynamics carried out with an harmonic constraint on the sheet C α atoms to preserve secondary structure.

Acknowledgments

P.D.L. and P.B. are grateful to Hoechst Roussel for a postdoctoral fellowship and predoctoral studentship, respectively.

Fig. 7. Multiple sequence alignment of hGHR (de Vos et al., 1992), FN3, CD4, IL3R α , GMCSFR α , and the first and second hematopoietic repeat of human β chain. The alignment was performed with CAMELEON (Morris, 1988) based on the hydrophobic nature of the residues. Hydrophilic residues are shaded blue and hydrophobic residues are shaded orange. Residues without a strong hydrophobic preference are shaded grey.

References

- Abdel-Meguid SS, Shieh HS, Smith WW, Dayringer HE, Violand BN, Bente A. 1987. Three dimensional structure of a genetically engineered variant of porcine growth hormone. *Proc Natl Acad Sci USA* 85:6434-6437.
- Bamborough P. 1994. Theoretical structure prediction applied to proteins of the immune system [thesis]. Oxford, UK: Oxford University.
- Bamborough P, Grant GH, Hedgecock CJR, West SP, Richards WG. 1993. A complex model of the interleukin-4/receptor complex. *Protein Struct Funct Genet* 17:11-19.
- Barry SC, Bagley CJ, Phillips J, Dottore M, Cambareri B, Moretti P, D'Andrea R, Goodall GJ, Shannon MF, Vadas MA, Lopez AF. 1994. Two contiguous residues in human interleukin-3, Asp²¹ and Glu²², selectively interact with the α - and β -chains of its receptor and participate in function. *J Biol Chem* 269:8488-8492.
- Bazan JF. 1990. Structural design and molecular evolution of a cytokine receptor superfamily. *Proc Natl Acad Sci USA* 87:6934-6938.
- Bernstein FC, Koetzle TF, Williams GJB, Meyer EF Jr, Brice MD, Rodgers JR, Kennard O, Shimanouchi T, Tasumi M. 1977. The Protein Data Bank: A computer-based archival file for macromolecular structures. *J Mol Biol* 112:535-542.
- Boulay JL, Paul WE. 1992. The interleukin-4 related lymphokines and their binding to hematopoietin receptors. *J Biol Chem* 267:20525-20528.
- Boulay JL, Paul WE. 1993. Hematopoietin sub-family classification based on size, gene organization and sequence homology. *Curr Biol* 3:573-581.
- Bowie JU, Luthy R, Eisenberg D. 1991. A method to identify protein sequences that fold into a known three dimensional structure. *Science* 253:164-179.
- Brooks BR, Brucoleri RE, Olafson BD, States DJ, Swaminathan S, Karplus M. 1983. CHARMM: A program for macromolecular energy, minimization and dynamics calculations. *J Comput Chem* 4:187-217.
- Clackson T, Wells JA. 1995. A hot-spot of binding energy in a hormone-receptor interface. *Science* 267:383-386.
- Connolly ML. 1983. Solvent-accessible surfaces of proteins and nucleic acids. *Science* 221:709-713.
- Devereux J, Haeboli P, Smithies O. 1984. A comprehensive set of sequence analysis programs for the VAX. *Nucleic Acids Res* 12:387-395.
- de Vos AM, Ultsch M, Kossiakoff AA. 1992. Human growth hormone and extracellular domain of its receptor: Crystal structure of the complex. *Science* 255:306-312.
- Diederichs K, Boone T, Karplus PA. 1991. Novel fold and putative receptor-binding site of granulocyte-macrophage colony stimulating factor. *Science* 254:1779-1782.
- Fuh G, Cunningham BC, Fukunaga R, Nagata S, Goeddel DV, Wells JA. 1992. Rational design of potent antagonists to the human growth hormone receptor. *Science* 256:1673-1677.
- Goodall GJ, Bagley CJ, Vadas MA, Lopez AF. 1993. A model for the interaction of the GM-CSF, IL-3 and IL-5 receptors with their ligands. *Growth Factors* 8:87-97.
- Greenfield RS, Braslawsky GR, Kadow KF, Spitalny GL, Chace D, Bull CO, Bursucker I. 1993. Murine granulocyte macrophage colony stimulating factor using monoclonal antibodies to synthetic peptides. *J Immunol* 150:5241-5251.
- Hayashida K, Kitamura T, Gorman DM, Arai KI, Yokota T, Miyajima A. 1990. Molecular cloning of a second subunit of the receptor for human granulocyte-macrophage colony-stimulating factor (GM-CSF): Reconstitution of a high affinity GM-CSF receptor. *Proc Natl Acad Sci USA* 87:9655-9659.
- Hercus TR, Bagley CJ, Cambareri B, Dottore M, Woodcock JM, Vadas MA, Shannon MF, Lopez AF. 1994. Specific human granulocyte-macrophage colony stimulating factor antagonists. *Proc Natl Acad Sci USA* 91:5838-5842.
- Hoang T, Delean A, Haman A, Beauchemin V, Kitamura T, Clark SC. 1993. The structure and dynamics of the granulocyte macrophage colony stimulating factor receptor defined by the ternary complex model. *J Biol Chem* 268:11881-11887.
- Janin J, Chothia C. 1976. Stability and specificity of protein-protein interactions: The case of the trypsin-trypsin inhibitor complexes. *J Mol Biol* 100:197-206.
- Kaushansky K, Shoemaker SG, Alfaro S, Brown C. 1989. Hematopoietic activity of GM-CSF is dependent upon two distinct regions of the molecule: Functional analysis based upon the activities of interspecies hybrid growth factors. *Proc Natl Acad Sci USA* 86:1213-1221.
- Kitamura T, Hayashida K, Sakamaki K, Yokota T, Arai K, Miyajima A. 1991a. Reconstitution of functional receptors for human granulocyte/macrophage colony stimulating factor (GM-CSF): Evidence that the protein encoded by the AIC2B cDNA is a subunit of the murine GM-CSF receptor. *Proc Natl Acad Sci USA* 88:5082-5086.
- Kitamura T, Miyajima A. 1992. Functional reconstitution of the human IL-3 receptor. *Blood* 80:84-90.
- Kitamura T, Sato N, Arai KI, Miyajima A. 1991b. Expression cloning of the human IL-3 receptor cDNA reveals a shared β -subunit for the human IL-3 and GM-CSF receptors. *Cell* 66:1165-1174.
- Kossiakoff AA, Somers W, Ultsch M, Andow K, Muller YA, de Vos AM. 1994. Comparison of the intermediate complexes of human growth hormone bound to the human growth hormone and prolactin receptors. *Protein Sci* 3:1697-1705.
- Kraulis P. 1991. MOLSCRIPT: A program to produce both detailed and schematic plots of protein structures. *J Appl Crystallogr* 24:946-950.
- Lock P, Metcalf D, Nicola N. 1994. Histidine 367 of the human common β chain of the receptor is critical for high affinity binding of human granulocyte macrophage colony stimulating factor. *Proc Natl Acad Sci USA* 91:252-256.
- Lopez AF, Eglinton JM, Lyons AB, Tapley PM, To LB, Park LS, Clark SC, Vadas MA. 1990. Human interleukin-3 inhibits the binding of granulocyte macrophage colony stimulating factor and interleukin-5 to basophils and strongly enhances their functional activity. *J Cell Physiol* 145:69-77.
- Lopez AF, Elliott MJ, Woodcock J, Vadas MA. 1992a. GM-CSF, IL-3 and IL-5: Cross-competition on human haemopoietic cells. *Immunol Today* 13:495-500.
- Lopez AF, Shannon MF, Barry S, Phillips JA, Cambareri B, Dottore M, Simmons P, Vadas MA. 1992b. A human interleukin-3 analog with increased biological and binding activities. *Proc Natl Acad Sci USA* 89:11842-11846.
- Lopez AF, Vadas MA, Woodcock JM, Milton SE, Lewis A, Elliott MJ, Gillis D, Ireland R, Olwell E, Park LS. 1991. Interleukin-3, interleukin-3 and granulocyte macrophage colony stimulating factor cross compete for binding to cell surface receptors on human eosinophils. *J Biol Chem* 266:24741-24747.
- Milburn MV, Hassell AM, Lambert MH, Jordan SR, Proudfoot AE, Graber P, Wells TN. 1993. A novel dimer configuration revealed by the crystal structure at 2.4 Å resolution of human interleukin-5. *Nature* 363:172-176.
- Moonen P, Mermod JJ, Ernst JF, Hirshi M, DeLamarter JF. 1987. Increased biological activity of deglycosylated recombinant human GM-CSF produced by yeast or animal cells. *Proc Natl Acad Sci USA* 84:4428-4431.
- Morris GM. 1988. The matching of protein sequences using colour intra-sequence displays. *J Mol Graph* 6:135-142.
- Mott HR, Campbell ID. 1995. Four helix bundle growth factors and their receptors: Protein-protein interactions. *Curr Opin Struct Biol* 5:114-121.
- Potts JR, Campbell ID. 1994. Fibronectin structure and assembly. *Opinion in Cell Biology* 6:648-655.
- Rozwarski DA, Gronenborn AM, Clore GM, Bazan JF, Bohm A, Wlodawer A, Hatada M, Karplus PA. 1994. Structural comparisons among the short-chain helical cytokines. *Structure* 2:159-173.
- Ryu SE, Kway PD, Truneh A, Porter TG, Arthos J, Rosenberg M, Dai X, Xuong NH, Axel R, Sweat RW, Hendrickson WA. 1990. Crystal structure of an HIV-binding fragment of human CD4. *Nature* 348:419-426.
- Shanafelt AB, Kastelein RA. 1992. High affinity ligand binding is not essential for granulocyte macrophage colony stimulating factor receptor activation. *J Biol Chem* 267:25466-25472.
- Somers W, Ultsch M, de Vos AM, Kossiakoff AA. 1994. The X-ray structure of a growth hormone prolactin receptor complex. *Nature* 372:478-481.
- Sprang SR, Bazan JF. 1993. Cytokine structural taxonomy and mechanisms of receptor engagement. *Curr Opin Struct Biol* 3:815-827.
- Sutcliffe MJ, Haneef I, Carney D, Blundell TL. 1987a. Knowledge based modelling of homologous proteins, part I: Three dimensional frameworks derived from the simultaneous superposition of multiple structures. *Protein Eng* 1:377-384.
- Sutcliffe MJ, Hayes FR, Blundell TL. 1987b. Knowledge based modelling of homologous proteins, part II: Rules for the conformations of substituted side chains. *Protein Eng* 1:385-392.
- Tavernier J, Devos R, Cornelis S, Tuypens T, Van der Heyden J, Fiers W, Plaetinck G. 1991. A human high affinity interleukin-5 receptor (IL5R) is composed of an IL5-specific α chain and a β chain shared with the receptor for GM-CSF. *Cell* 66:1175-1184.
- Walter MR, Cook WJ, Ealick SE, Nagabushan TL, Trotta PP, Bugg CE. 1992. Three dimensional structure of recombinant human granulocyte-macrophage colony stimulating factor. *J Mol Biol* 224:1075-1085.
- Wlodawer A, Pavlovsky A, Gustchina A. 1993. Hematopoietic cytokines: Similarities and differences in the structures, with implications for receptor binding. *Protein Sci* 2:1373-1382.

- Woodcock JM, Zacharakis B, Plaetinck G, Bagley CJ, Qiyu S, Hercus TR, Tavernier J, Lopez AF. 1994. Three residues in the common β chain of the human GM-CSF, IL-3 and IL-5 receptors are essential for GM-CSF and IL-5 but not IL-3 high affinity binding and interact with Glu 21 of GM-CSF. *EMBO J* 13:5176-5185.
- Yang YC, Ciarletta AB, Temple AB, Chung MP, Kovacic S, Witekgiannotti JS, Leary AC, Kriz R, Donahue RE, Wong GG, Clark SC. 1986. Human IL-3 (Multi-CSF) – Identification by expression of a novel chematopoietic growth factor related to murine IL-3. *Cell* 47:3-10.
- Yoshimura A, Zimmers T, Neumann D, Longmore G, Yoshimura Y, Lodish HF. 1992. Mutations in the Trp-Ser-X-Trp-Ser motif of the erythropoietin receptor abolish processing, ligand binding and activation of the receptor. *J Biol Chem* 267:11619-11625.

Appendix

Listed below are the Brookhaven (Bernstein et al., 1977) accession codes for the proteins used in the loop-search database for the GM-CSF and IL-3 receptors.

1fdl	2fbj	2rhe
1rei	2mcp	3hfm
1bjl	2hfl	3fab
1fai	2bjl	3cd4
1fc1	2fb4	3dpa
1fc2	2ig2	3mcg
1igf	2igf	4fab
1pfc		