The EMBO Journal Vol.17 No.6 pp.1665–1674, 1998

Crystal structure of a cytokine-binding region of gp130

Jerónimo Bravo[1], David Staunton[2], John K.Heath[3] and E.Yvonne Jones[1,2,4]

[1]Laboratory of Molecular Biophysics, The Rex Richards Building, South Parks Road, Oxford OX1 3QU, [2]Oxford Centre for Molecular Sciences, New Chemistry Building, South Parks Road, Oxford OX1 3QT and [3]Cancer Research Campaign Growth Factor Group, School of Biochemistry, The University of Birmingham, Edgbaston, Birmingham B15 2TT, UK

[4]Corresponding author
e-mail: Yvon@biop.ox.ac.uk

The structure of the cytokine-binding homology region of the cell surface receptor gp130 has been determined by X-ray crystallography at 2.0 Å resolution. The β sandwich structure of the two domains conforms to the topology of the cytokine receptor superfamily. This first structure of an uncomplexed receptor exhibits a similar L-shaped quaternary structure to that of ligand-bound family members and suggests a limited flexibility in relative domain orientation of some 3°. The putative ligand-binding loops are relatively rigid, with a phenylalanine side chain similarly positioned to exposed aromatic residues implicated in ligand binding for other such receptors. The positioning and structure of the N-terminal portion of the polypeptide chain have implications for the structure and function of cytokine receptors, such as gp130, which contain an additional N-terminal immunoglobulin-like domain.
Keywords: cytokine receptor/gp130/multiple anomalous diffraction/X-ray crystallography

Introduction

Cytokines, generally in the form of secreted molecules, mediate intercellular signalling by high affinity interaction (K_D ~10^{-10} M) with the extracellular regions of specific cell surface receptors. This promotes receptor oligomerization which, in turn, triggers intracellular signalling cascades within the target cell. The ability of a given cytokine to elicit biological responses in a target cell is therefore dictated by the specificity of interaction between ligand and receptor.

Gp130 is a transmembrane receptor which is required for signal transduction by a set of cytokines, the gp130 family, which have many significant biological functions of potential therapeutic interest (reviewed in Kishimoto *et al.*, 1995). Gp130-mediated signalling has been implicated in the regulation of a wide variety of adult tissue systems, including haemopoesis, nervous system, bone, heart, adipose tissue, testes, liver and muscle (reviewed in Kishimoto *et al.*, 1995). Targeted inactivation of the gp130 gene results in a complex pre-natal lethal phenotype including defects in cardiac and haematological function (Yoshida *et al.*, 1996). In addition, chronic activation of gp130 signalling in a transgenic mouse model results in cardiac hypertrophy (Hirota *et al.*, 1995). The gp130 family of ligands currently comprises interleukin-6 (IL-6), IL-11, herpes virus IL-6 (HSVIL-6), leukaemia inhibitory factor (LIF), oncostatin (OSM), cardiotrophin (CT-1) and ciliary neurotrophic factor (CNTF). The three-dimensional structures of three members of the gp130 family, LIF (Robinson *et al.*, 1994), CNTF (McDonald *et al.*, 1995) and IL-6 (Somers *et al.*, 1997; Xu *et al.*, 1997) have been defined by crystallographic or solution NMR techniques. This reveals that these cytokines share a common topology, being composed of four regions of α helix (helices A–D) linked by polypeptide loops in the 'up-up-down-down' conformation typical of the 'long chain' family of cytokines (Boulton *et al.*, 1994).

The signalling functions of gp130 are initiated by the ligand-mediated formation of oligomeric complexes with other specific partner receptors. Gp130 initially was cloned as an essential transmembrane component for signalling mediated by IL-6 (Hibi *et al.*, 1990). This occurs via dimerization of gp130 (Murakami *et al.*, 1993) following the formation of a hexameric complex containing two molecules of gp130, two molecules of IL-6 and two molecules of a soluble specific IL-6 receptor (Ward *et al.*, 1994; Paonessa *et al.*, 1995). A similar mechanism pertains to the case of IL-11, where homodimerization of gp130, and subsequent execution of signalling functions, is brought about by association with a complex of IL-11 and specific IL-11 receptors (Hilton *et al.*, 1994; Karow *et al.*, 1996).

Gp130 was also cloned as a receptor required for signalling mediated by cytokines which associate with a second transmembrane receptor of the cytokine type–LIF-R (Gearing *et al.*, 1991). These include OSM (Gearing *et al.*, 1992; Liu *et al.*, 1992), LIF and CNTF (Ip *et al.*, 1992), and CT-1 (Pennica *et al.*, 1995). In this case, signal transduction is initiated by ligand-mediated heterodimerization of gp130 and LIF-R (facilitated, in the case of CNTF, by association with a third non-signalling receptor component CNTF-R). Recently it has been discovered that OSM can also mediate signalling by heterodimerization of gp130 with a novel transmembrane signalling receptor of the cytokine type, OSM-R (Mosley *et al.*, 1996). The intracellular signalling pathways activated by ligand-mediated homo- or heterodimerization of gp130 include activation of the receptor-associated JAK/Tyk tyrosine kinases (Boulton *et al.*, 1994; Stahl *et al.*, 1994), the STAT family of transcription factors (Stahl *et al.*, 1995) and src-family tyrosine kinase pathways (Ernst *et al.*, 1994).

The sequence of the extracellular ligand-binding region of gp130 reveals that it is a member of the 'cytokine' superfamily of receptors characterized by a canonical cytokine-binding homology region (CHR) containing the

Table I. gp130-CHR data collection statistics

	Wild-type	Seλ_1	Seλ_2	Seλ_3	Seλ_4	Seλ_5
Wavelength (Å)	1.030	9790	0.9791	0.9535	0.9793	0.9789
Unique	48 189	20 549	20 280	22 400	20 264	19 303
$<I/\sigma(I)>$	19.4 (5.0)	23.2	27.2	24.8	18.8	28.6
R_{merge} (%)	5.8 (20.6)	5.1	4.3	4.2	4.2	5.0
Resolution (Å)	30–2.0	30–2.55	30–2.55	30–2.50	30–2.55	30–2.55
Completeness (%)	98.8 (96.6)	93.6	93.5	92	93.5	89.4

$R_{merge} = \Sigma|I - <I>| / \Sigma<I>$.
Values in parentheses correspond to the highest resolution shell (2.07–2.00 Å).

'WSXWS' motif, a proline-rich 'hinge' region and a characteristic spacing of cysteine residues (reviewed in Cosman, 1993). In addition, the extracellular region of gp130 contains an N-terminal module predicted to adopt a seven-stranded immunoglobulin-like conformation and, C-terminal to the CHR, three fibronectin type III (FN III) domains. Deletion studies have revealed that the gp130-CHR is sufficient for interaction with ligand (Horsten *et al.*, 1995). Mutation studies of both IL-6 (Paonessa *et al.*, 1995) and LIF (Hudson *et al*., 1996) have shown that this interaction involves topologically analogous receptor recognition epitopes (site II) in both ligands. A second, physically discrete, gp130 ligand recognition epitope (site III) has also been described for the interaction with IL-6 (Paonessa *et al.*, 1995). This interaction requires regions of gp130 outside the CHR (Simpson *et al.*, 1997; D.Staunton, K.R.Hudson and J.K.Heath, unpublished observations). The function of the CHR of gp130 is therefore the association with partner ligands (alone or complexed with receptor) via their site II recognition epitopes.

Crystal structures are available for four cytokine receptors containing a CHR; the growth hormone receptor (GHR; De Vos *et al.*, 1992), the prolactin receptor (PRLR; Somers *et al*., 1994), the erythropoietin receptor (EPOR; Livnah *et al*., 1996) and the interferon-γ receptor (IFNγ-R; Walter *et al.*, 1995). These four prototypes undergo an exclusively homodimerization mode of action with a restricted range of ligands; little is known currently of the detailed structural features of receptors which undergo heterodimerization in the presence of ligand or interact with multiple ligand and receptor partners.

We report here the high resolution crystal structure of the cytokine-binding homology region of gp130. This structure provides the first detailed three-dimensional information for a receptor component crucial to the signalling complexes of a large family of growth factors (IL-6, IL-11, LIF, CNTF, CT-1 and OSM) allowing assessment of the molecular basis of specific recognition and ligand engagement.

Results and discussion

Expression and structure determination

A soluble form of the gp130 cytokine-binding homology region (gp130-CHR) was expressed in *Escherichia coli* as a maltose-binding protein (MBP) fusion (D.Staunton, K.R.Hudson and J.K.Heath, in preparation). The final purified protein comprised four residues from the fusion linker, residues 100–303 of human gp130 (the CHR) and a further 14 residues corresponding to a three-alanine linker and c-myc tag. Surface plasmon resonance studies confirmed that this recombinant form of gp130-CHR binds OSM with an affinity (K_d ~6.5×10^{-8} M) equivalent to the complete gp130 extracellular domain expressed in eukaryotic cells (D.Staunton, K.R.Hudson and J.K.Heath, in preparation). Gp130-CHR therefore retains structural features required for ligand engagement via site II.

Crystallization trials yielded highly ordered crystals of space group $C222_1$ (unit cell dimensions a = 84.5 Å, b = 132.3 Å, c = 121.9 Å) which contained two gp130-CHR molecules per crystallographic asymmetric unit. The non-crystallographic symmetry does not reveal any possible mode of receptor dimerization. The structure was determined by multiple anomalous dispersion (MAD) phasing techniques using X-ray diffraction data collected on BM14 at the European Synchrotron Radiation Facility (ESRF) from crystals of a selenomethionyl form of the protein. The structure has been refined to a crystallographic *R*-value of 21.5% for all data between 30 and 2.0 Å resolution. Crystallographic statistics are reported in Tables I and II. All 204 residues of the CHR are well ordered for one of the two copies in the crystallographic asymmetric unit, but residue 100 at the N-terminus and residues 212–213 of a loop region are disordered in the second copy; the final model also includes certain of the residues which derive from the expression construct (Figure 1A, see Materials and methods). Domain-wise superpositions of the two molecules in the crystallographic asymmetric unit show essentially identical structures (r.m.s. deviation for equivalent C_α atoms between 90 residues of the CHR N-terminal domain is 0.44 Å and between 99 residues of the CHR C-terminal domain is 0.32 Å). The following text focuses exclusively on the structure of the human gp130-CHR and refers to these residues by the intact gp130 numbering, 100–303.

Structure description and comparison with other CHR structures

As anticipated from sequence analysis, the topology of gp130-CHR is similar to those of the three other class 1 receptors of the cytokine superfamily (Cosman, 1993) for which structures have been determined (hGHR, de Vos *et al.*, 1992; hPRLR, Somers *et al.*, 1994; EPOR, Livnah *et al.*, 1996). The CHR comprises two specialized FN III domains. The basic structural scaffold for each domain consists of a β sandwich primarily formed from a three-strand (A, B, E) and four-strand (C, C′, F, G) β sheet. These domains are connected by a short 3_{10} helix and are oriented such that the whole molecule has an approximate L shape (Figure 1).

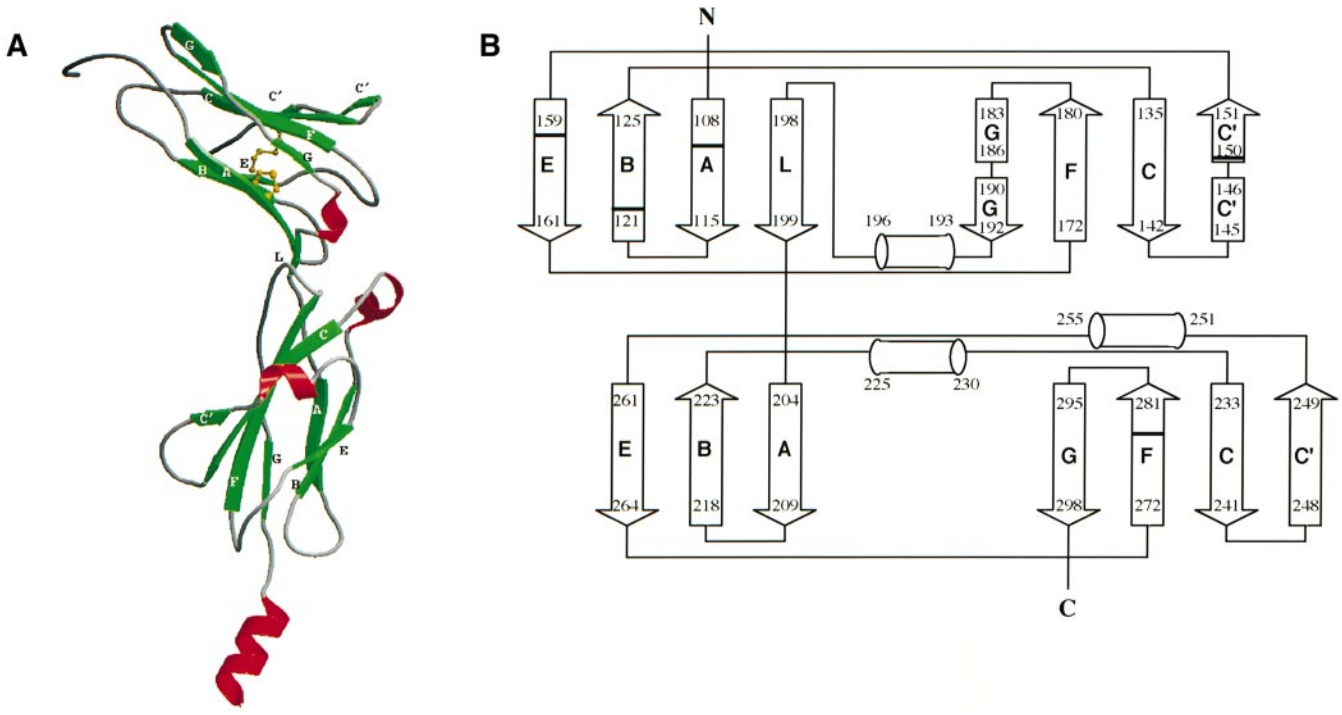

Fig. 1. The structure of gp130-CHR. (**A**) Ribbon representation of the structure of gp130-CHR. Helical segments are shown in red and β strands in green. The full crystal structure is illustrated which includes, in addition to the structure of residues 100–303 of gp130, an extra three N-terminal residues and eight C-terminal residues which derive from the expression construct. The C-terminal α helix results from these latter residues. (**B**) Topology diagram of the two domains of gp130-CHR. Helices are represented by cylinders and β strands by arrows. The positions of the five cysteine residues are marked by black bars. (A) and all components of Figures 2, 3 and 5 were drawn using programs MOLSCRIPT (Kraulis, 1991), with modifications by R.Esnouf (Esnouf, 1997), and RASTER3D (Merrit and Murphy, 1994).

Gp130-CHR D1

The N-terminal domain (D1; residues 103–192) has the standard arrangement of A, B, C, C′, E, F and G β strands (Figures 1 and 2). One notable feature, unique to gp130-CHR D1, is the division of strand G into two approximately equal portions. This is the result of a β bulge and, as discussed below, shows similarity to the WSXWS motif of the second domain. In common with other examples of this fold, gp130-CHR D1 contains two disulfides which link the A and B strands (Cys112–Cys122) and the C′ and E strands (Cys150–Cys160). The lengths of the interstrand loops are similar to those in the EPOR D1 (with the exception of the shorter region between strand C′ and E, and a longer FG loop) and show little significant increase in flexibility relative to the core of the domain (as judged from crystallographic *B*-factors).

Given this general level of topological equivalence, it is somewhat surprising to find that structural superpositions with other CHR D1 structures (Figure 2A) show relatively poor agreement within the core framework (for example, 1.0 Å r.m.s. deviation for 53 structurally equivalent Cα atoms with EPOR D1). As illustrated in Figure 2B, this discrepancy arises from the distinctive angle of β strands C, F and G in the upper half of the GFCC′ sheet. A superposition on gp130-CHR D2 shows a more extensive match over these main secondary structure elements (1.2 Å r.m.s. deviation for 64 structurally equivalent Cα atoms). The key feature, common to both gp130-CHR D1 and D2, is the region of extended polypeptide chain (residues 102–107 of D1, residues 198–204 of D2) which packs tightly against the edge of the β sandwich between strands B and G before starting strand A at Asn109 in D1 and Asn205 in D2. The tight packing of this part of the polypeptide chain against the core of the β sandwich is mediated by the insertion of proline residues Pro103 and Pro107 in D1, and Pro200 and Pro203 in D2. The incorporation of this feature necessitates the shift in orientation of the upper part of the GFC sheet. In gp130-CHR D1, this is achieved through the distinctive β bulge in strand G at residues 187–189 stabilized by the hydrogen bonding of the Ser187 hydroxyl to the main-chain nitrogen of Val176 in strand F (Figure 2C). An identical function is performed by the two serine residues in the canonical WSXWS motif of CHR D2 structures, with the superposition of the gp130-CHR D1 and D2 domains indicating that Ser187 and Ser292 are structurally equivalent. None of the other class 1 CHR D1 structures contain the equivalent length of polypeptide tightly clamped between the B and G β strands or the β bulge in strand G.

Gp130-CHR D2

The C-terminal domain (D2; residues 200–300) conforms to the standard A,B,E and GFCC′ β sheet arrangement. Of the two domains of the CHR, the second appears to be generally the more structurally conserved within the cytokine receptor family. Structural superpositions indicate closest similarity to the EPOR domain (1.05 Å r.m.s. deviation for 81 structurally equivalent Cα atoms). This arises primarily from the shorter length of the β strands in these two molecules compared with those in other members of the superfamily. The only secondary structure element to not correspond closely in position between the gp130-CHR and EPOR D2 structures is the C′ strand. In common with the other class-1 CHR D2 structures, the

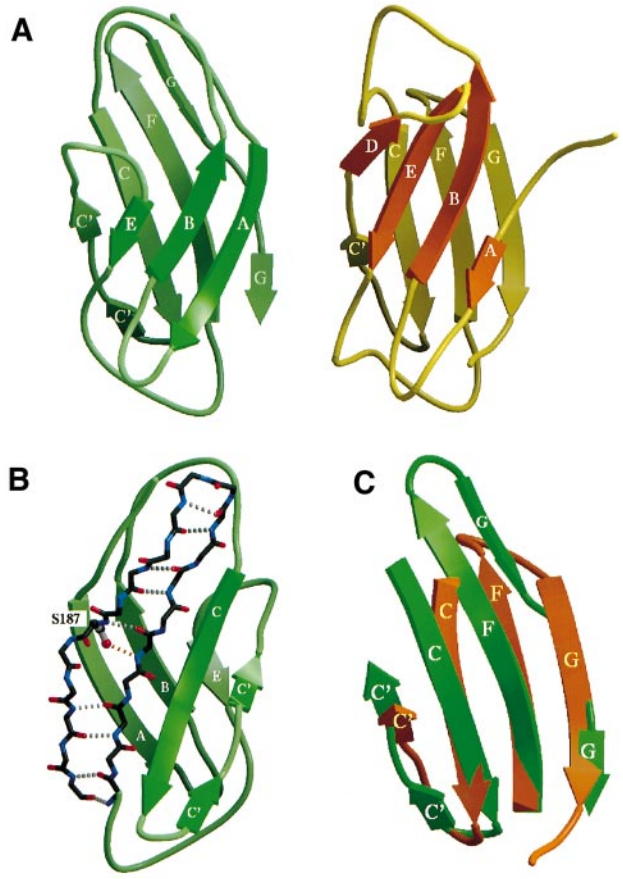

Fig. 2. Gp130-CHR domain 1. (**A**) Comparison of D1 of gp130 (green) and D1 of EPOR (brown). The C′ and G strands are broken into two portions in gp130, and there is no D strand. (**B**) The β bulge in strand G. The main-chain atoms of β strands F and G are shown in stick representation, Ser187 is shown in ball and stick representation, and hydrogen bonds are denoted by broken lines. (**C**) The distinctive C′,C,F,G sheet in gp130. The C′,C,F,G strands are shown for gp130 (green) and EPOR (brown) positioned on the basis of a whole domain superposition (performed using the program SHP; Stuart *et al.*, 1979) to illustrate the novel nature of the top half of the gp130 β sheet.

interstrand loops in gp130-CHR D2 show relatively limited mobility (as judged from crystallographic *B*-factors), with the exception of the AB loop. This loop is stabilized by a crystal contact in one copy of the gp130-CHR but in the other copy the AB loop is exposed to solvent and is disordered in the electron density map (residues 212 and 213). The domain contains one free cysteine residue (Cys279) which is buried within the core of the β sandwich. This precludes it from involvement in disulfide-linked homodimerization of gp130 during IL-6-related cytokine signalling (Murakami *et al.*, 1993).

The WSXWSX sequence (gp130 residues 288–293), a defining feature of this receptor superfamily, is situated in the N-terminal portion of strand G (Figure 3A and B) and has an essentially identical double β bulge structure to that of the homologous region in EPOR. The two successive β bulges are stabilized by hydrogen bonds from the hydroxyls of Ser289 and Ser292 to the main-chain nitrogens of Cys279 and Ile277 respectively. As in the other class-1 CHR D2 structures, the side chains of Trp288 and Trp291 participate in an extended π-cation system, stacking between the side chains of strand F residues Arg276, Arg278 and Met280. This feature is considerably more extended in gp130-CHR D2 than in EPOR D2, however, since it also involves side chains from residues Arg240 and Trp247 in strand C and at the start of strand C′ respectively. The absence of these latter interactions in EPOR may underlie the difference in the position of its C′ strand relative to that in gp130-CHR D2. The extended π-cation system in gp130 is most closely matched by that in PRLR.

The interdomain region and relative domain orientation

As observed for the other superfamily members, residues of the interdomain linker region (gp130 residues 193–197) form a 3_{10} helix. Additionally, residues 198 and 199 in gp130-CHR form a short β strand. This novel structural element hydrogen-bonds to both strand A of D1 and the WSXWS region at the N-terminus of strand G in D2, providing an extra constraint on the relative orientation and positioning of the two domains (Figure 3C). The resultant juxtaposition of D1 and D2 produces a tight interdomain interface which, excluding the contribution of the linker, buries ~350 Å^2 of solvent-accessible surface. This is contributed mainly by residues in strand A and the EF loop of D1 (I113, E116 and Y168) and residues from the BC loop and strand G of D2 (I227, V230, I231 and Y287). Superposition of the two copies of gp130-CHR in the crystallographic asymmetric unit reveals a 3° difference in the relative orientation of their domains. This appears to originate from very slight changes in the main-chain torsion angles for linker residues V198–K199. The structural constraints imposed by the linker 3_{10} helix and β strand plus the hydrophobic, close-packed nature of the interface, argues against any more substantial degree of interdomain orientational freedom than the observed 3° range.

The overall shape of the molecule can be quantified in terms of a tilt angle (defined as the angle between the long axes, running approximately parallel to the β strands, in the two domains; Bork *et al.*, 1996). With a tilt angle of 78°, the relative domain orientation in gp130-CHR corresponds most closely to the general ‘L-shaped’ (~90° tilt angle) arrangement characteristic of the other class 1 members of the superfamily (hGHR, hPRLR and EPOR) rather than the more upright (~50° tilt angle) arrangement of the more distantly related class 2 members IFNγ-R and TF or the more extreme 120° or so of tilt very recently observed in the natural killer inhibitory receptor (Fan *et al.*, 1997). A detailed comparison (combining the effect of tilt angle and twist in the orientation of D1 relative to the D2 axis) reveals that the precise interdomain orientation in gp130-CHR differs by some 23° from that of the most closely related quaternary structure, that of EPOR.

Sequence comparisons with other species

The human gp130-CHR sequence employed in this study was aligned with the homologous regions from mouse, rat and *Xenopus* gp130 (Figure 4). This reveals that 74/204 (36%) of residues in this region are conserved amongst all versions of the gp130-CHR. It is notable that the majority (64) of these shared residues are in a relatively buried location in the human gp130-CHR structure. This indicates that these conserved residues most probably play a role in determining the structural framework of gp130

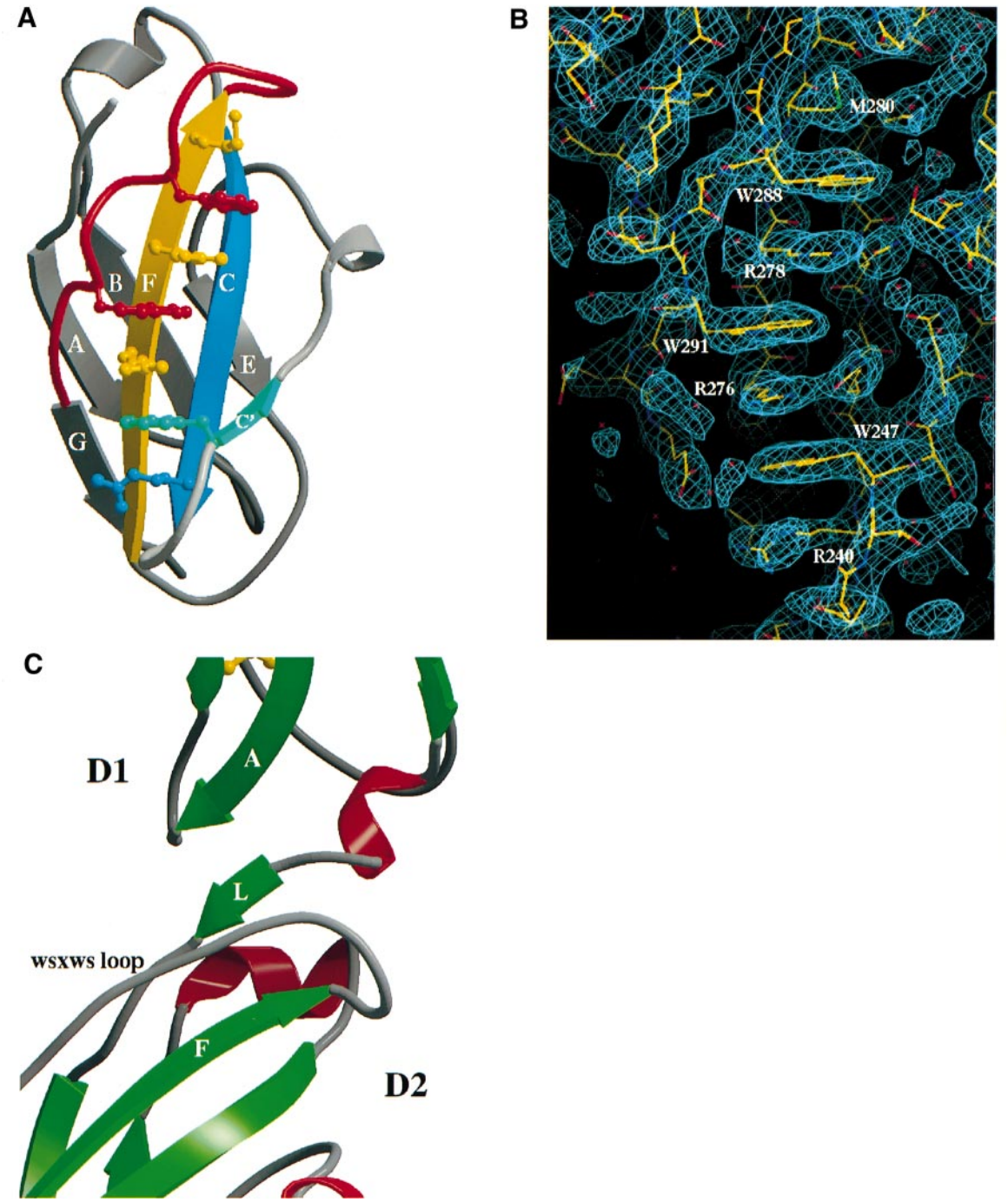

Fig. 3. Gp130-CHR domain 2 and interdomain linker. (**A**) Schematic diagram of the WSXWS box. The side chains of residues contributing to this structural feature are colour coded according to the secondary structure element from which they originate. (**B**) Atomic coordinates and electron density map for the WSXWS box. The refined coordinates are displayed with the original 2.9 Å resolution electron density map calculated using MAD phases followed by density modification (program DM, see Materials and methods). The map is contoured at 1σ in program O. (**C**) Schematic diagram of the D1–D2 linker region. Strand L in the linker region hydrogen-bonds to strand A in DL and the WSXWS box region of the polypeptide chain prior to the start of strand G in D2.

rather than ligand recognition. It follows that the putative ligand recognition epitopes of gp130 may exhibit variation between species.

Implications for ligand recognition

The topological similarity of gp130 to hGHR and hPRLR, systems for which ligand binding has been structurally and functionally well characterized (Cunningham and Wells, 1989; reviewed in Sprang and Bazan, 1993; Wells *et al.*, 1993), permits the identification of candidate structural features of gp130 that mediate ligand recognition via site II. The cognate site II gp130 recognition epitopes have been defined for two ligands, LIF (Hudson *et al.*, 1996) and IL-6 (Savino *et al.*, 1994). In both cases, these consist of a small number (4–6) of solvent-exposed residues located in the adjacent helices A and C of the ligand. This suggests that the site II ligand recognition site of gp130 will also be formed from relatively few

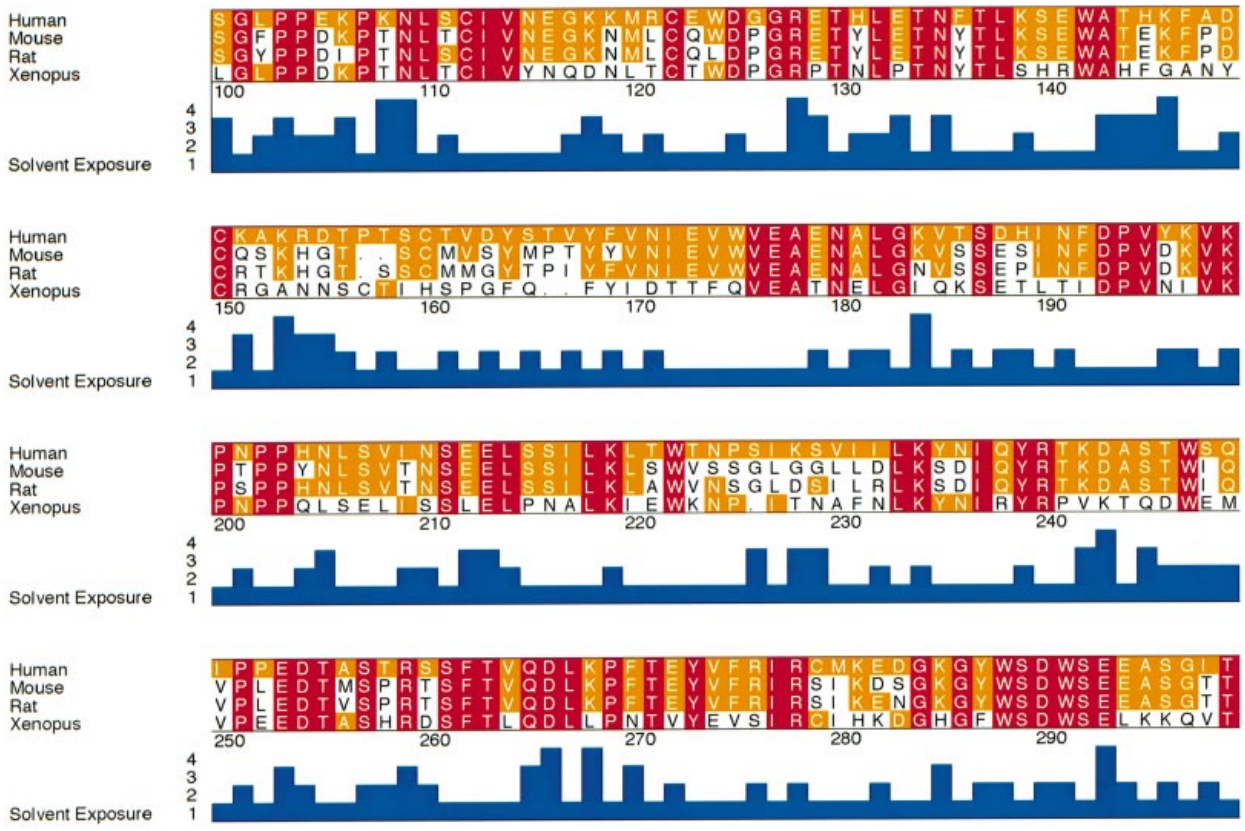

Fig. 4. Sequence alignment and solvent accessibility of gp130 sequences. The CHR region of human (Swissprot accession No. P40189, Hibi *et al*., 1990), murine (Q00560, Saito *et al*., 1992), rat (P40190, Wang *et al*., 1992) and *Xenopus* (K.Chien, A.Grace and J.Chen, personal communication) gp130 sequences were aligned using the progressive pairwise algorithm of Feng and Doolittle (1987) implemented in the Pileup programme of the GCG package followed by minor manual editing. Solvent exposure scores were calculated using DSSP implemented in the program Turbo-6 (Roussel and Cambillau, 1989). Scores 0–50 were assigned a value of 1, 51–100:2, 101–150:3 and 151–200:4. Conserved residues are coloured in orange and residues which are identical to human gp130 are coloured in red. The figure was generated using the programme Alscript (Barton, 1993).

solvent-exposed residues forming a complementary binding site.

The hGH–hGHR complex (De Vos *et al.*, 1992) and the related hGH–hPRLR complex (Somers *et al.*, 1994) reveal that recognition of the ligand via site II involves solvent-exposed residues located in three loops linking the main β strand elements. The first of these is a prominent aromatic residue located in the loop between strands E and F of D1 (Trp104 in both hGHR and hPRLR). In the human gp130-CHR structure, the analogous EF loop contains a similar prominent, solvent-exposed residue Phe169 (Figure 5). Sequence alignment of gp130 sequences (Figure 4) suggests that the analogous residue is present in rat gp130 but is replaced by the conservative substitution of a tyrosine residue in the mouse and *Xenopus* proteins. Sequence alignment of the equivalent region in other receptors for members of the long chain cytokine family reveals that a non-polar residue (Phe, Tyr or Trp) in the predicted EF loop of D1 is a common feature (data not shown).

The other two potential binding sites are located in D2 in the form of the loops linking strands BC (gp130 residues 226–230) and FG (gp130 residues 281–285). A non-conservative substitution mutant V230D, which is located in the BC loop, results in loss of affinity for IL-6–IL-6R (Horsten *et al.*, 1997). V230 is, however, relatively buried in the three-dimensional structure (DSSP score of relative exposure 23%), and the effect of this mutation on ligand recognition may, therefore, be indirect. The FG loop has also been implicated recently in the recognition of granulocyte colony-stimulating factor (GCSF) by the D2 domain of the GCSFR CHR (Yamasaki *et al.*, 1997). In addition, non-conservative substitution mutants of two exposed residues (G286W and K285E) in this region of gp130 result in loss of affinity for IL-6–IL-6R (Horsten *et al.*, 1997), suggesting that the FG loop does indeed play a significant role in ligand recognition. The equivalent regions of gp130 proteins from other species exhibit conservation, but not identity, of residues in these regions.

A key feature of gp130 is its ability to interact with a range of ligands in the context of a number of other partner receptors. In cases such as LIF and OSM (Hudson *et al.*, 1996), the ligand is able to interact with gp130 with high affinity on its own. In other cases such as IL-6 and IL-11, high affinity interaction between gp130 and the ligand requires that ligand is associated with a partner receptor (IL-6R and IL-11R respectively). This suggests that there may exist additional sites on gp130 which are involved in interaction with partner receptors. Mutagenesis of the IL-6R (Yawata *et al.*, 1993) and the IL-11R (M.A.Hall, P.Bilinski, A.Gossler and J.K.Heath, unpublished observations) have revealed that specific non-conservative substitutions of residues in the membrane-proximal region of the predicted D2 in these receptors can block the ability of the ligand–receptor complex to interact with gp130. Inspection of the dimeric hGHR

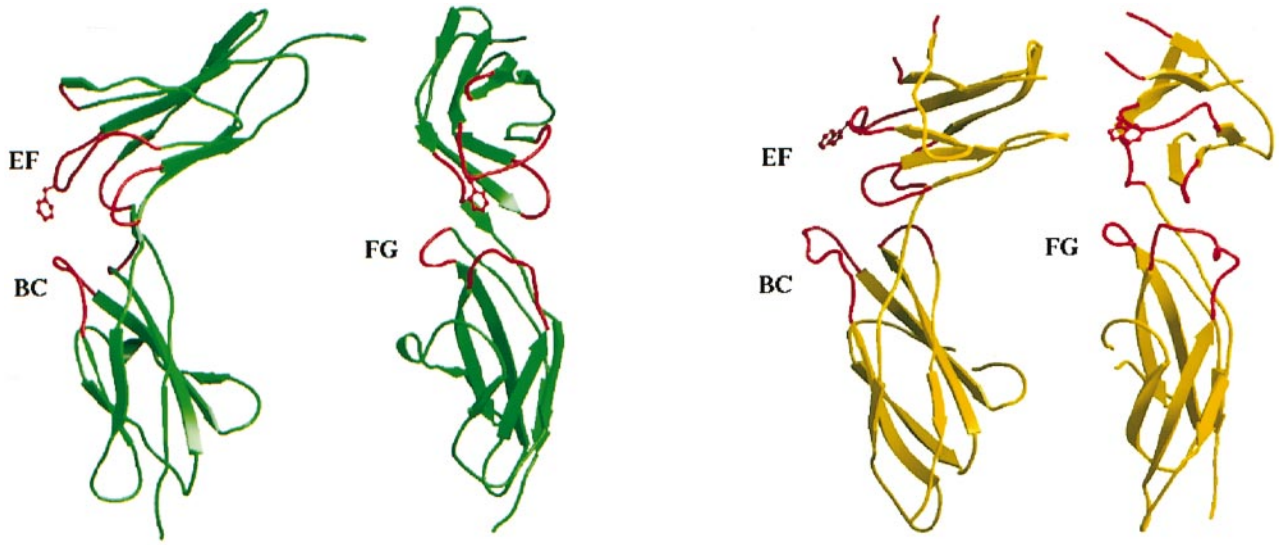

Fig. 5. The putative ligand-binding region in gp130-CHR compared with hGHR. Two orthogonal views are shown for gp130-CHR (green) and hGHR (yellow). Loops implicated in ligand binding for the hGH–hGHR complex, and their equivalents in the gp130-CHR structure, are denoted in red. The side chains of the structurally equivalent residues Phe169 and Trp104 are shown in ball and stick representation.

complex (de Vos *et al.*, 1992) shows that D2 of each hGHR partner receptor is in apposition, forming a receptor dimer interface. It is likely, therefore, that non-conservative mutations in the dimer interface could disrupt the formation of a high affinity complex. It may therefore be anticipated that the analogous dimer interface region of gp130 (formed from the AB loop and strand E of D2) may also include sites of receptor–receptor recognition. It is of interest, therefore, that amongst the few species-conserved residues exhibiting significant access to solvent, five (Glu213, Leu214, Lys219, Gln265 and Asp266) are located in this region (Figure 4), suggesting that they may be involved in the formation of the receptor dimer interface.

Conclusions

The gp130-CHR structure reported here is the first for an unliganded receptor in this superfamily. As discussed previously, the observed, L-shaped, quaternary structure and the nature of the interdomain linker region imply that there is little domain reorientation on ligand binding. Gp130 is competent to bind several different ligands; the current structure implies that it does so with essentially identical global structure. Conformational variability in the ligand-binding loops could be envisaged as one potential mechanism for such adaptability in ligand recognition. Indeed, the BC loop of D2 in hGHR adopts a different conformation on binding site I or site II of hGH (de Vos *et al.*, 1992). The unliganded structure of gp130, however, provides little indication of any potential for such a mechanism, with the D2 BC loop showing a single rigid conformation. Clearly the structure of unliganded gp130-CHR represents only the first stage in studies of these recognition events. A comparison of the conformations of various interdomain linkers with that of the N-terminal portion of the polypeptide in gp130-CHR does, however, provide some pointers to the structure/function of cytokine receptors containing an additional Ig-like domain. The gp130 Ig-D1 linker bears closest similarity in relative position and overall conformation to that of the linker between variable and constant domains in the light chain of Fab NEW (Saul and Poljak, 1992, comparison not shown). The implied positioning of the gp130 Ig-like domain is fully consistent with previously proposed models for this domain's role in multimeric recognition complexes (e.g. Simpson *et al.*, 1997), but a detailed analysis must await further structural studies.

Materials and methods

Native and selenomethionyl protein production

Full details of the expression, purification and functional characterization of gp130-CHR are reported elsewhere (D.Staunton, K.R.Hudson and J.K.Heath, in preparation). Briefly, the 25 kDa CHR of human gp130 (residues 100–303) was expressed as a folded fusion protein secreted into the periplasmic space of *E.coli* by the pMALp2 expression vector (New England Biolabs). Purification of the MBP–gp130-CHR protein fusion was achieved by ion exchange chromatography. The fusion protein was cut with rhinovirus 3c protease and the gp130-CHR separated from the MBP by high resolution ion exchange chromatography. The activity of the purified gp130-CHR was determined by its ability to bind the cytokine OSM as followed by surface plasmon resonance analysis (BIAcore).

The selenomethionyl form (SeMet-gp130-CHR) was produced by similar methods but in the presence of selenomethionine-enriched media. As for the native protein, the gp130-CHR fragment fused with MBP was expressed in *E.coli* strain HW1110 using the pMALp2 vector. The bacteria were grown in defined medium prepared as described by LeMaster and Richards (1985). A 50 ml pre-culture grown in LeMaster's medium supplemented with 1 mg/ml thiamine and 50 μg/ml carbenicillin was used to innoculate 4 l of media additionally supplemented with 50 mg/l seleno-L-methionine (Sigma). The bacteria grew at 37°C with a doubling time of ~150 min. At an absorbance of 1.5/cm (at 600 nm), the culture was induced with 0.1 mM isopropyl-β-D-thiogalactopyranoside (IPTG) and the incubation temperature reduced to 25°C. The cultures were harvested after a further 3 h of growth when the absorbance had reached 3.0/cm. The purification was performed as for the native protein except that an additional step of amylose resin (New England Biolabs) affinity chromatography was introduced after the initial ion exchange to remove contaminating proteins from the fusion protein and all buffers contained 1 mM dithiothreitol and 1 mM EDTA to prevent oxidation of the selenomethionine.

Mass determinations by electrospray ionization mass spectrometry were performed on a VG BioQ mass spectrometer. Analysis of the gp130-CHR was consistent with a polypeptide of the expected sequence (calculated mass 25 210 Da, observed mass 25 209 ± 5 Da). Gp130-CHR contains two methionine residues, and a comparison of the

measured mass of the selenoderivative indicated that the SeMet-gp130-CHR had 95% seleno-L-methionine incorporation in place of normal methionine. Forms corresponding to mono- and non-substituted gp130-CHR were not detected in the purified sample. The SeMet-gp130-CHR had an affinity for OSM identical to that of the native gp130-CHR, with a dissociation constant of 50 nM as followed by surface plasmon resonance analysis (BIAcore).

Crystallization

Both wild-type and SeMet-gp130-CHR were crystallized at 4°C using the hanging drop vapour diffusion method over a reservoir solution containing 1.8–2.1 M ammonium sulfate, 0.1 M Tris pH 8.0. Typically, 1.5 µl of the protein solution (12 mg/ml in 20 mM Tris pH 8.0) was mixed with an equal volume of the reservoir solution. SeMet-gp130-CHR crystals were slightly smaller than wild-type ($0.7\times0.4\times0.2$ mm^3) but, when exposed to synchrotron radiation, showed ordered Bragg diffraction to about the same resolution (2 Å). The crystals are of the $C222_1$ space group with unit cell dimensions $a = 84.48$ Å, $b = 132.29$ Å and $c = 121.93$ Å. There are two molecules per asymmetric unit and the crystal solvent content is ~62%.

Data collection

All data sets were collected using synchrotron radiation. Crystals used for cryo-crystallographic data collection were transfered to mother liquor containing 25% glycerol, flash-cooled in liquid propane and stored at –170°C in liquid nitrogen. Data for MAD-based phase determination were collected from one cryo-cooled SeMet-gp130-CHR crystal at the BM14 beam line of the ESRF. The characteristics of the Se absorption edge for this crystal were determined by a fluorescence scan, and five wavelengths were selected for data collection at points corresponding to the peak (maximum f′, $\lambda = 0.9790$ Å), inflection point (minimum f′, $\lambda = 0.9791$ Å), and two remote energies, below at $\lambda = 0.9793$ Å and above the absorption edge at $\lambda = 0.9535$ Å. A fifth data set was collected just above the maximum f′ at $\lambda = 0.9789$ Å. Diffraction data sets were collected, at each wavelength in turn, over the same 95° ϕ range as 0.5° oscillation images using a CCD detector (Hammersley *et al.*, 1994; Moy, 1994). A single wavelength ($\lambda = 1.03$ Å) high resolution (2 Å) data set subsequently was collected as 1° oscillation images from a cryo-cooled wild-type gp130-CHR crystal at the same beam line using an imaging plate detector (MarResearch 30 cm diameter).

The diffraction data were autoindexed and integrated using program DENZO. Data at each wavelength were scaled separately and merged (preserving Bjivoet pairs for the MAD data) using SCALEPACK (Otwinowski and Minor, 1997). Crystallographic data collection statistics are reported in Table I.

Structure determination

The Se positions were determined from the SeMet-gp130-CHR data sets by manual inspection of difference and anomalous Patterson syntheses and confirmed using the program SHELX (Sheldrick *et al*., 1993). The four Se sites were refined in program MLPHARE, and MAD phases were calculated to 2.9 Å resolution (figure of merit 0.48). An initial map based on these phases and F_{obs} from the $\lambda = 0.9791$ Å data set showed clear electron density for both copies of gp130-CHR in the asymmetric unit. Density modification using solvent flattening and histogram matching procedures as implemented in the program DM (Cowtan, 1994; CCP4 suite of programs) yielded a high quality electron density map which was used in the interactive graphics program O (Jones *et al.*, 1991) to trace the gp130-CHR structure (Figures 1 and 3B).

Refinement and structural analysis

Refinement was carried out using standard protocols in the program X-PLOR (Brünger, 1992). A model for copy 1 of the two molecules in the asymmetric unit was built using program O and a model for copy 2 generated from these coordinates by application of non-crystallographic symmetry operators. The model was refined initially at 2.5 Å resolution against the SeMet-gp130-CHR $\lambda = 0.9535$ data set. Tight non-crystallographic symmetry restraints were maintained between the two molecules, and 7.5% of the data were reserved for R_{free} monitoring. Once available, the 2.0 Å resolution native data set was used for all subsequent refinement with an extension of the test array for the cross-validation. Inclusion of a bulk solvent correction allowed all measured data for 30.0–2.0 Å resolution to be used. In the final stages of refinement, non-crystallographic symmetry restraints were released and ordered water molecules were modelled.

For one of the copies in the asymmetric unit, significant electron density is present for all 204 residues of the human gp130-CHR plus all three N-terminal residues and eight of the C-terminal residues derived from the expression construct. The second copy is less well ordered, lacking clear electron density for the N-terminal residues up to and including gp130 residue Ser100 and gp130 residues 212–213 in a loop region of the second domain; seven residues from the expression construct are visible at the C-terminus of this molecule. For this final model, no non-glycine residues fall in the disallowed regions of the Ramachandran plot. Refinement and model statistics are reported in Table II.

Table II. gp130-CHR structural refinement statistics

Resolution range (Å)	30–2.0
Completeness (%)	95.0
No. of reflections ($F>0$)	44 627
R_{cryst} (%)	21.5
R_{free} (%)	25.0
No. of non-hydrogen atoms	
Protein	3359
Water	287
Sulfate	30
R.m.s.d. from ideality	
Bond lengths (Å)	0.006
Bond angles (°)	1.40
Dihedrals (°)	25.36
Improper (°)	1.11
Average *B*-factor (Å^2)	
Main chain	19.5
Side chain	20.3
Water	19.1

$R_{cryst} = \Sigma||F_{obs}| - |F_{calc}||/\Sigma|F_{obs}|$.
R_{free} is as for R_{cryst} but calculated for a test set comprising reflections not used in the refinement (7.5%).

Atomic coordinates

Atomic coordinates for gp130-CHR have been deposited with the Protein Data Bank, Brookhaven National Laboratory, USA.

Acknowledgements

We thank R.Bryan, K.Measures and R.Esnouf for computing facilities and programs, M.Deller, K.Harlos, J.Tormo and the staff of the ESRF BM14 and EMBL outstation in Grenoble for assistance with X-ray data collection, S.Lee for help in the preparation of figures, and I.Campbell and D.Stuart for discussion. This work was funded by the Cancer Research Campaign. The Oxford Centre for Molecular Sciences is supported by the BBSRC, EPSRC and MRC. J.B. is supported by an EC Fellowship and E.Y.J. by the Royal Society.

References

Barton,G.J. (1993) ALSCRIPT: a tool to format multiple sequence alignments. *Protein Eng.*, **6**, 37–40.

Bork,P., Downing,A.K., Kieffer,B. and Campbell,I.D. (1996) Structure and distribution of modules in extracellular proteins. *Q. Rev. Biophys.*, **29**, 119–167.

Boulton,T.G., Stahl,N. and Yancopoulos,G.D. (1994) Ciliary neurotrophic factor/leukemia inhibitory factor/interleukin 6/oncostatin M family of cytokines induces tyrosine phosphorylation of a common set of proteins overlapping those induced by other cytokines and growth factors. *J. Biol. Chem.*, **269**, 11648–11655.

Brünger,A.T. (1992) *XPLOR Version 3.1: A system for X-Ray Crystallography and NMR.* Yale University Press, New Haven, CT.

Cosman,D. (1993) The hematopoietin receptor superfamily. *Cytokine*, **5**, 95–106.

Cowtan,K. (1994) DM, an automated procedure for phase improvement by density modification. *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography*, **31**, 24–28.

Cunningham,B.C. and Wells,J.A. (1989) High-resolution epitope mapping of hGH–receptor interactions by alanine-scanning mutagenesis. *Science*, **244**, 1081–1085.

De Vos,A.M., Ultsch,M. and Kossiakoff,A.A. (1992) Human growth hormone and extracellular domain of its receptor: crystal structure of the complex. *Science*, **255**, 306–312.

Ernst,M., Gearing,D.P. and Dunn,A.R. (1994) Functional and biochemical association of Hck with the LIF/IL-6 receptor signal transducing subunit gp130 in embryonic stem cells. *EMBO J.*, **13**, 1574–1584.

Esnouf,R.M. (1997) An extensively modified version of Molscript that includes greatly enhanced colouring capabilities. *J. Mol. Graphics*, **15**, 133–138.

Fan,Q.R., Mosyak,L., Winter,C.C., Wagtmann,N., Long,E.O. and Wiley,D.C. (1997) Structure of the inhibitory receptor for human natural killer cells resembles haematopoietic receptors. *Nature*, **389**, 96–100.

Feng,D.F. and Doolittle,R.F. (1987) Progressive sequence alignment as a prerequisite to correct phylogenetic trees. *J. Mol. Evol.*, **25**, 351–360.

Gearing,D.P., Thut,C.J., VandeBos,T., Gimpel,S.D., Delaney,P.B., King,J., Price,V., Cosman,D. and Beckmann,M.P. (1991) Leukemia inhibitory factor receptor is structurally related to the IL-6 signal transducer, gp130. *EMBO J.*, **10**, 2839–2848.

Gearing,D.P. *et al.* (1992) The IL-6 signal transducer, gp130: an oncostatin M receptor and affinity converter for the LIF receptor. *Science*, **255**, 1434–1437.

Hammersley,A.P., Svensson,S.O. and Thompson,A. (1994) Calibration and correction of spatial distortions in 2D detector systems. *Nucl. Instrum. Methods*, **A346**, 312–321.

Hibi,M., Murakami,M., Saito,M., Hirano,T., Taga,T. and Kishimoto,T. (1990) Molecular cloning and expression of an IL-6 signal transducer, gp130. *Cell*, **63**, 1149–1157.

Hilton,D.J. *et al*. (1994) Cloning of a murine IL-11 receptor alpha-chain; requirement for gp130 for high affinity binding and signal transduction. *EMBO J.*, **13**, 4765–4775.

Hirota,H., Yoshida,K., Kishimoto,T. and Taga,T. (1995) Continuous activation of gp130, a signal-transducing receptor component for interleukin 6-related cytokines, causes myocardial hypertrophy in mice. *Proc. Natl Acad. Sci. USA*, **92**, 4862–4866.

Horsten,U., Schmitz Van de Leur,H., Mullberg,J., Heinrich,P.C. and Rose John,S. (1995) The membrane distal half of gp130 is responsible for the formation of a ternary complex with IL-6 and the IL-6 receptor. *FEBS Lett.*, **360**, 43–46.

Horsten,U., Muller Newen,G., Gerhartz,C., Wollmer,A., Wijdenes,J., Heinrich,P.C. and Grotzinger,J. (1997) Molecular modeling-guided mutagenesis of the extracellular part of gp130 leads to the identification of contact sites in the interleukin-6 (IL-6)·IL-6 receptor·gp130 complex. *J. Biol. Chem.*, **272**, 23748–23757.

Hudson,K.R., Vernallis,A.B. and Heath,J.K. (1996) Characterization of the receptor binding sites of human leukemia inhibitory factor and creation of antagonists. *J. Biol. Chem.*, **271**, 11971–11978.

Ip,N.Y. *et al.* (1992) CNTF and LIF act on neuronal cells via shared signaling pathways that involve the IL-6 signal transducing receptor component gp130. *Cell*, **69**, 1121–1132.

Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr.*, **A47**, 110–119.

Karow,J., Hudson,K.R., Hall,M.A., Vernallis,A.B., Taylor,J.A., Gossler,A. and Heath,J.K. (1996) Mediation of interleukin-11-dependent biological responses by a soluble form of the interleukin-11 receptor. *Biochem. J.*, **318**, 489–495.

Kishimoto,T., Tanaka,T., Yoshida,K., Akira,S. and Taga,T. (1995) Cytokine signal transduction through a homo- or heterodimer of gp130. *Ann. NY Acad. Sci.*, **766**, 224–234.

Kraulis,P.J. (1991) MOLSCRIPT: a program to produce both detailed and schematic plots of protein structures. *J. Appl. Crystallogr.*, **24**, 946–950.

LeMaster,D.M. and Richards,F.M. (1985) 1H-15N heteronuclear NMR studies of *Escherichia coli* thioredoxin in samples isotopically labeled by residue type. *Biochemistry*, **24**, 7263–7268.

Liu,J., Modrell,B., Aruffo,A., Marken,J.S., Taga,T., Yasukawa,K., Murakami,M., Kishimoto,T. and Shoyab,M. (1992) Interleukin-6 signal transducer gp130 mediates oncostatin M signaling. *J. Biol. Chem.*, **267**, 16763–16766.

Livnah,O., Stura,E.A., Johnson,D.L., Middleton,S.A., Mulcahy,L.S., Wrighton,N.C., Dower,W.J., Jolliffe,L.K. and Wilson,I.A. (1996) Functional mimicry of a protein hormone by a peptide agonist: the EPO receptor complex at 2.8 A. *Science*, **273**, 464–471.

McDonald,N.Q., Panayotatos,N. and Hendrickson,W.A. (1995) Crystal structure of dimeric human ciliary neurotrophic factor determined by MAD phasing. *EMBO J.*, **14**, 2689–2699.

Merritt,E.A. and Murphy,M.E.P. (1994) Raster3D version 2.0. A program for photorealistic molecular graphics. *Acta Crystallogr.*, **D50**, 869–873.

Mosley,B., De Imus,C., Friend,D., Boiani,N., Thoma,B., Park,L.S. and Cosman,D. (1996) Dual oncostatin M (OSM) receptors. Cloning and characterization of an alternative signaling subunit conferring OSM-specific receptor activation. *J. Biol. Chem.*, **271**, 32635–32643.

Moy,J.P. (1994) A 200 mm input field, 5–80 keV detector based on an X-ray image intensifier and CCD camera. *Nucl. Instrum. Methods*, **A384**, 641–644.

Murakami,M., Hibi,M., Nakagawa,N., Nakagawa,T., Yasukawa,K., Yamanishi,K., Taga,T. and Kishimoto,T. (1993) IL-6-induced homodimerization of gp130 and associated activation of a tyrosine kinase. *Science*, **260**, 1808–1810.

Otwinowski,Z. and Minor,W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.*, **276**, 307–326.

Paonessa,G., Graziani,R., De Serio,A., Savino,R., Ciapponi,L., Lahm,A., Salvati,A.L., Toniatti,C. and Ciliberto,G. (1995) Two distinct and independent sites on IL-6 trigger gp 130 dimer formation and signalling. *EMBO J.*, **14**, 1942–1951.

Pennica,D. *et al.* (1995) Cardiotrophin-1. Biological activities and binding to the leukemia inhibitory factor receptor/gp130 signaling complex. *J. Biol. Chem.*, **270**, 10915–10922.

Robinson,R.C., Grey,L.M., Staunton,D., Vankelecom,H., Vernallis,A.B., Moreau,J.F., Stuart,D.I., Heath,J.K. and Jones,E.Y. (1994) The crystal structure and biological function of leukemia inhibitory factor: implications for receptor binding. *Cell*, **77**, 1101–1116.

Roussel,A. and Cambillau,C. (1989) TURBO-FRODO. In *Silicon Graphics Geometry Partner Directory.* Silicon Graphics, Mountain View, CA, pp. 77–78.

Saito,M., Yoshida,K., Hibi,M., Taga,T. and Kishimoto,T. (1992) Molecular cloning of a murine IL-6 receptor-associated signal transducer, gp130, and its regulated expression *in vivo*. *J. Immunol.*, **148**, 4066–4071.

Saul,F.A. and Poljak,R.J. (1992) Crystal structure of human immunoglobulin fragment Fab New refined at 2.0 Å resolution. *Proteins*, **14**, 363–371.

Savino,R., Lahm,A., Salvati,A.L., Ciapponi,L., Sporeno,E., Altamura,S., Paonessa,G., Toniatti,C. and Ciliberto,G. (1994) Generation of interleukin-6 receptor antagonists by molecular-modeling guided mutagenesis of residues important for gp130 activation. *EMBO J.*, **13**, 1357–1367.

Sheldrick,G.M., Dauter,Z., Wilson,K.S., Hope,H. and Sieker,L.C. (1993) The application of direct methods and Patterson interpretation to high-resolution native protein data. *Acta Crystallogr.*, **D49**, 18–23.

Simpson,R.J., Hammacher,A., Smith,D.K., Matthews,J.M. and Ward,L.D. (1997) Interleukin-6: structure–function relationships. *Protein Sci.*, **6**, 929–955.

Somers,W., Ultsch,M., De Vos,A.M. and Kossiakoff,A.A. (1994) The X-ray structure of a growth hormone–prolactin receptor complex. *Nature*, **372**, 478–481.

Somers,W., Stahl,M. and Seehra,J.S. (1997) 1.9 Å crystal structure of interleukin 6: implications for a novel mode of receptor dimerization and signaling. *EMBO J.*, **16**, 989–997.

Sprang,S.R. and Bazan,J.F. (1993) Cytokine structural taxonomy and mechanisms of receptor engagement. *Curr. Opin. Struct. Biol.*, **3**, 815–827.

Stahl,N. *et al.* (1994) Association and activation of Jak-Tyk kinases by CNTF-LIF-OSM-IL-6 beta receptor components. *Science*, **263**, 92–95.

Stahl,N., Farruggella,T.J., Boulton,T.G., Zhong,Z., Darnell,J.E.,Jr and Yancopoulos,G.D. (1995) Choice of STATs and other substrates specified by modular tyrosine-based motifs in cytokine receptors. *Science*, **267**, 1349–1353.

Stuart,D.I., Levine,M., Muirhead,H. and Stammers,D.K. (1979) The crystal structure of a cat pyruvate kinase at a resolution of 2.6 Å. *J. Mol. Biol.*, **134**, 109–142.

Walter,M.R., Windsor,W.T., Nagabhushan,T.L., Lundell,D.J., Lunn,C.A., Zauodny,P.J. and Narula,S.K. (1995) Crystal structure of a complex between interferon-gamma and its soluble high-affinity receptor. *Nature*, **376**, 230–235.

Wang,Y., Nesbitt,J.E., Fuentes,N.L. and Fuller,G.M. (1992) Molecular cloning and characterization of the rat liver IL-6 signal transducing molecule, gp130. *Genomics*, **14**, 666–672.

Ward,L.D., Howlett,G.J., Discolo,G., Yasukawa,K., Hammacher,A., Moritz,R.L. and Simpson,R.J. (1994) High affinity interleukin-6 receptor is a hexameric complex consisting of two molecules each of interleukin-6, interleukin-6 receptor, and gp-130. *J. Biol. Chem.*, **269**, 23286–23289.

Wells,J.A., Cunningham,B.C., Fuh,G., Lowman,H.B., Bass,S.H., Mulkerrin,M.G., Ultsch,M. and deVos,A.M. (1993) The molecular basis for growth hormone–receptor interactions. *Recent Prog. Horm. Res.*, **48**, 253–275.

Xu,G.Y., Yu,H.A., Hong,J., Stahl,M., McDonagh,T., Kay,L.E. and Cumming,D.A. (1997) Solution structure of recombinant human interleukin-6. *J. Mol. Biol.*, **268**, 468–481.

Yamasaki,K., Naito,S., Anaguchi,H., Ohkubo,T. and Ota,Y. (1997) Solution structure of an extracellular domain containing the WSxWS motif of the granulocyte colony-stimulating factor receptor and its interaction with ligand. *Nature Struct. Biol.*, **4**, 498–504.

Yawata,H., Yasukawa,K., Natsuka,S., Murakami,M., Yamasaki,K., Hibi,M., Taga,T. and Kishimoto,T. (1993) Structure–function analysis of human IL-6 receptor: dissociation of amino acid residues required for IL-6-binding and for IL-6 signal transduction through gp130. *EMBO J.*, **12**, 1705–1712.

Yoshida,K. *et al*. (1996) Targeted disruption of gp130, a common signal transducer for the interleukin 6 family of cytokines, leads to myocardial and hematological disorders. *Proc. Natl Acad. Sci. USA*, **93**, 407–411.

Received December 5, 1997; revised January 8, 1998;
accepted January 15, 1998