Disulfide Bond Structure and N-Glycosylation Sites of the Extracellular Domain of the Human Interleukin-6 Receptor

(Received for publication, September 11, 1998, and in revised form, December 7, 1998)

Adam R. Cole, Nathan E. Hall‡, Herbert R. Treutlein‡, James S. Eddes, Gavin E. Reid, Robert L. Moritz, and Richard J. Simpson§

From the Joint Protein Structure Laboratory, Ludwig Institute for Cancer Research (Melbourne Tumour Biology Branch) and The Walter and Eliza Hall Institute of Medical Research, Parkville, Victoria 3050, Australia and The Ludwig Institute for Cancer Research (Melbourne Tumour Biology Branch) and The Cooperative Research Centre for Cellular Growth Factors, Parkville, Victoria 3050, Australia

The high affinity interleukin-6 (IL-6) receptor is a hexameric complex consisting of two molecules each of IL-6, IL-6 receptor (IL-6R), and the high affinity converter and signaling molecule, gp130. The extracellular “soluble” part of the IL-6R (sIL-6R) consists of three domains: an amino-terminal Ig-like domain and two fibronectin-type III (FN III) domains. The two FN III domains comprise the cytokine-binding domain defined by a set of 4 conserved cysteine residues and a WSWS sequence motif. Here, we have determined the disulfide structure of the human sIL-6R by peptide mapping in the absence and presence of reducing agent. Mass spectrometric analysis of these peptides revealed four disulfide bonds and two free cysteines. The disulfides Cys102-Cys113 and Cys146-Cys157 are consistent with known cytokine-binding domain motifs, and Cys28-Cys77 with known Ig superfamily domains. An unusual cysteine connectivity between Cys8-Cys174, which links the Ig-like and NH2-terminal FN III domains causing them to fold back onto each other, has not previously been observed among cytokine receptors. The two free cysteines (Cys192 and Cys258) were detected as cysteinylation-cysteines, although a small proportion of Cys258 was reactive with the alkylating agent 4-vinylpyridine. Of the four potential N-glycosylation sites, carbohydrate moieties were identified on Asn36, Asn74, and Asn202, but not on Asn226.

Interleukin-6 (IL-6) is a multifunctional cytokine that plays a central role in host defense due to its wide range of immune and hematopoietic activities, as well as its potent ability to induce the acute phase response (1–3). Since overexpression of IL-6 has been implicated in the pathology of a number of diseases (for reviews, see Refs. 4 and 5), it is anticipated that selective antagonists of IL-6 action may offer therapeutic benefits in the treatment of IL-6-related diseases.

The biological activities of IL-6 are mediated by the IL-6 receptor system which comprises two receptor proteins: the specific ligand-binding α-subunit receptor (IL-6R) and the signal transducing β-subunit, gp130. gp130 also forms part of the receptor complexes of leukemia inhibitory factor, ciliary neutrophoric factor, oncostatin M, cardiotrophin-1, and IL-11 (6) which, in part, provides a molecular basis for the functional redundancy of these cytokines. IL-6 first binds the IL-6R with an affinity of ~1 nM and the IL-6-IL-6R complex then binds gp130 with a resulting affinity of ~10 pM. The ternary complex of the IL-6 receptor system is a hexamer, comprising two molecules each of IL-6, IL-6R, and gp130 (8, 9).

The cDNA of the human IL-6R encodes a protein of 468 amino acids (7), including a signal peptide of 19 amino acids, an extracellular region of 339 amino acids, a transmembrane domain of 28 amino acids, and a short cytoplasmic domain of 82 amino acids. This sequence shows 54 and 57% overall amino acid identity with the cDNA sequences for mouse (10) and rat (11) IL-6R, respectively. The mature 80-kDa IL-6R is a glycosylated form of the predicted 50-kDa precursor (12) and contains six potential N-linked glycosylation sites. The extracellular region has a modular structure, consisting of three domains of approximately 100 amino acids. The amino acid sequence of the NH2-terminal domain is characteristic of the immunoglobulin superfamily (Ig-like) (13, 14). Members of this family share a common β-sheet folding topology called a Greek Key (15), whereby neighboring β-strands form hydrogen bonds in an anti-parallel fashion to form a β-pleated sheet. Two β-sheets are then packed against each other to produce a hydrophobic core. Similarly, the two COOH-terminal domains of the IL-6R are classified as fibronectin type III-like (FN III) modules, a subclass of the β-sandwich fold (14). The topology of these domains is similar to those of Ig-like modules, with the notable exception of the “sheet switching” of β-strand D from the first β-sheet of an Ig-like domain to form β-strand C’ on the second β-sheet of FN III domains. Together, the two FN III domains form a cytokine-binding domain (CBD) which is characteristic of class I cytokine receptors (16) (e.g. receptors for interleukins-3, -5, -6, -11, gp130, erythropoietin (EPO), ciliary neutrophoric factor, granulocyte-colony stimulating factor, growth hormone (GH), and prolactin (PRL)). Generally, these receptors are characterized by two conserved disulfide bonds located in the NH2-terminal FN III domain and a conserved WSXWS motif located in the COOH-terminal FN III domain.

The cytoplasmic and transmembrane domains of the IL-6R are not required for IL-6 signaling (17) and biologically active soluble forms of IL-6R (sIL-6R) are naturally found in low
concentrations in human urine (18) and serum (19, 20) of healthy individuals. In contrast to many other soluble cytokine receptors that act as inhibitors by competing for ligand binding with cellular receptors (e.g. tumor necrosis factor, IL-1, -2, -4, interferon-γ, nerve growth factor, leukemia inhibitory factor, granulocyte-stimulating factor and granulocyte macrophage-colony stimulating factor) (for reviews, see Refs. 21 and 22), the sIL-6R acts as an agonist of IL-6 activity (17). It is not clear whether sIL-6R is generated by proteolytic shedding of membrane-bound IL-6R (23), or from an alternatively spliced mRNA species (24, 25), or both. In certain disease states, for example, patients with human immunodeficiency virus infection or multiple myeloma, increased levels of sIL-6R have been reported (26, 27). Therefore, inhibition of the IL-6/sIL-6R complex has been labeled as a key target to antagonize the in vivo action of IL-6 (28).

To elucidate the tertiary structure of the IL-6R extracellular region, we have purified human sIL-6R using a Chinese hamster ovary (CHO) cell expression system (29, 30). This form of the sIL-6R contains four potential N-linked glycosylation sites and 10 cysteine residues. Three cysteines are located in the Ig-like domain, six in the NH₂-terminal FN III domain, and one in the COOH-terminal FN III domain. Previously, we have shown that the ligand affinity purified sIL-6R bound IL-6 and gp130 with a 2:2:2 stoichiometry of IL-6, sIL-6R, and sgp180 (8) and was bioactive as determined by the ability of the IL-6/sIL-6R complex to prevent the differentiation of embryonic stem cells (31). Here, reversed-phase HPLC peptide mapping under nonreducing and reducing conditions, in combination with mass spectrometric and NH₂-terminal sequence analysis, was used to determine the disulfide structure and carbohydrate attachment sites of sIL-6R. On the basis of these results, we have created a model that depicts the topology of the extracellular region of the IL-6R and predicts its interactions with IL-6 and gp130.

EXPERIMENTAL PROCEDURES

Materials—Trypsin (sequencing grade) and neuraminidase (EC 3.2.1.18) were obtained from Boehringer-Mannheim. An ent-dolichol-lactose preparation obtained from Flavobacterium meningosepticum (32) containing three β-N-acetylglucosidase F (endo F) activities (F1, F2, and F3) as well as peptide-N-glycosidase (33, 34) was a kind gift from Dr. G. E. Norris (Massey University, New Zealand). Tris-(2-carboxyethyl)-phosphine (TCEP) was obtained from Pierce. 4-Vinylpyridine was purchased from Sigma. All other chemicals were HPLC grade.

Purification of the Extracellular or "Soluble" Domain of the Human IL-6 Receptor (sIL-6R)—sIL-6R was purified from the conditioned media of CHO cells transfected with an expression vector (pECEdhfr344) which encodes the extracellular binding domain of the IL-6R (truncated at residue 345) (29). The sIL-6R was concentrated from CHO cell conditioned media using a Saratoc Miniapparatus (Saratoc, Goettingen, Germany) equipped with a 30,000 molecular weight cut-off membrane and purified by ligand affinity chromatography using an IL-6-Sepharose column (30).

SDS-PAGE and Isoelectrofocusing—SDS-PAGE analysis of sIL-6R samples was performed using pre-cast 4–20% polyacrylamide/SDS gels (Novex) according to the method of Laemmli (35). Isoelectrofocusing (IEF-PAGE) was performed on pre-cast linear pH gradient (pH 3–10), (Novex) according to the method of Laemmli (35). Isoelectrofocusing (IEF-PAGE) was performed on pre-cast linear pH gradient (pH 3–10), and hydrophobic patterns of the β-sheets.

Peptide Mapping of the Tryptic Digest of sIL-6R—Deglycosylated sIL-6R (200 μg, 4 nmol) was digested with trypsin (1:20 w/w, 0.05 M NaH₂PO₄, pH 6.0) overnight at 37 °C. The trypptic peptide mixture was fractionated by reversed-phase HPLC using a Hewlett-Packard Liquid Chromatograph (model HP 1090A) and a Vydac C18 column (250 × 4.6-mm inner diameter). The column was developed at 0.5 ml/min using a 60-min linear gradient of 0–100% B, where solvent A was aqueous 0.1% trifluoroacetic acid and solvent B was 60% acetonitrile in aqueous 0.09% trifluoroacetic acid (45 °C). The column eluent was split (1:160), post-detector, using a stainless steel Tee-union (Upchurch catalog number U428, Upchurch, Oak Harbor, WA) directing 0.6% of the total flow at 3 μl/min to the mass spectrometer, while the remainder (99.4%) was collected into polypropylene microcentrifuge tubes (Eppendorf) for further analysis. To identify disulfide-containing fractions, 25% of the digest (1 nmol) was reduced with an equal volume of 10 mTCEP in 0.2 n sodium citrate buffer, pH 6.0 (5 °C, 10 min (36), and then re-chromatographed under identical conditions.

Ion-trap Mass Spectrometry—On-line MS analysis of peptide fractions was performed on a Finnigan-MAT LCQ quadrupole ion trap mass spectrometer equipped with an ESI source (San Jose, CA). “Triple play” experiments, consisting of MS/MS zip scan/MS/MS, were performed as described elsewhere (37). Source CID/single ion monitoring (sCID/ SIM) was employed to identify S-pyridylethyl cysteine-containing peptides (38). For sCID/SIM, the relative collision energy in the source region was set at 70% (arbitrary value) and the mass range was scanned from m/z 104.5–107.5 to detect S-pyridylethyl fragment ions (M–2H) of the peptides were identified using the Finnigan PEPMAP® program, and from their CID product ion spectra using the MS-Tag and MS-Product algorithms (Prospector pacific rim mirror site, http://jpsl.ludwig.edu.au).

NH₂-terminal Sequence Analysis—Automated Edman degradation of proteins and peptides was performed using a HP-G1005A bipheric NH₂-terminal protein sequencer (Hewlett-Packard, version 3.0 chemistry) (38).

Homology Modeling—The Ig-like domain of the IL-6R was modeled using the structure of the mouse monoclonal antibody FAB D44.1 V₁ domain (39) which showed 23% sequence identity, the highest for all known Ig three-dimensional structures. The coordinates for the template were taken from the Protein Data Bank (40), entry 1MLB (chain A, residues 1–109). Template structures for the CBD of gp130 (41), Protein Data Bank entry 1BQO; GHR, chain B (first binding receptor) of Protein Data Bank entry 3HHR (42); EPOR, Protein Data Bank entry 1EBP; and PRLR, Protein Data Bank entry 1BFP (44). The sequence alignment was prepared in two parts. The Ig-like domain of IL-6R was manually aligned with the template FAB D44.1 V₁ structure. For the CBD, a structure-based multiple sequence alignment was prepared. The β-sheets of the known CBD structures were superimposed to provide the basis of the alignment. The remaining sequences of unknown structure were manually aligned with these structures, conserving the disulfide patterns, WSXWS motif, and hydrophobic patterns of the β-sheets.

The MODELLER program (45) was used to generate separate models of the Ig-like domain and CBD. The quality of the models was assessed as described previously (46); in particular using the ProsaII program (47). MODELLER was also used to determine the relative orientations of the Ig-like domain and CBD. A disulfide restraint between Cysα and Cysα’ of the Ig-like domain and FN III domain, respectively, was introduced in accordance with our experimental results. Fifty models were generated with a range of orientations between the Ig-like domain and the CBD. The final two models were chosen on the basis of the quality checks described above and agreement with experimental data. The model of the sIL-6R, complexed with the crystal structures of IL-6 (48) (Protein Data Bank entry 1ALU) and the gp130 CBD (41), was constructed by superimposing these moieties over the human GH receptor complex (42).

RESULTS

Initial Characterization of Recombinant sIL-6 Receptor—NH₂-terminal sequence analysis of the first 20 residues of the purified sIL-6R in agreement with the published sequence (7) (data not shown). The purified sIL-6R yielded a single broad band on SDS-PAGE with an apparent molecular mass of ~52,000 (Fig. 1A, lane 1), a value significantly higher than 36,368 Da calculated from the amino acid composition (7). Upon treatment with either neuraminidase or a combination of neuraminidase and an endoglycosidase mixture, the M₀ of the sIL-6R was reduced to ~50,000 and ~40,000, respectively (Fig. 1A, lanes 2 and 3). These data suggest that the increased M₀ of

7208 Extracellular Domain of Human IL-6 Receptor
sIL-6R (−12,000) is due to glycosylation of the CHO cell-derived protein. Pronounced charge heterogeneity of the mature sIL-6R was observed upon IEF (Fig. 1B, lane 1). Treatment with neuraminidase (Fig. 1B, lane 2) and neuraminidase plus an endoglycosidase mixture (Fig. 1B, lane 3) reduced this complexity to one or two major bands, respectively, indicating that the heterogeneity of the sIL-6R preparation is primarily due to differential N-linked glycosylation.

**Labeling of Free Cysteine Residues**—Free cysteine residues in sIL-6R (−200 μg) were modified with 4-vinylpyridine at pH 8.5. Following enzymatic deglycosylation, the treated sIL-6R was digested with trypsin at pH 6.0, and subjected to RP-HPLC/ESI-IT-MS analysis as described under “Experimental Procedures.” The total ion current profiles of nonreduced and reduced tryptic digest of sIL-6R are shown in Fig. 2 (panels A and B, respectively). sCID/SIM of the nonreduced digest for S-pyridylethyl ions (m/z 106) revealed a single peak at retention time 31.42 min (Fig. 2C). The mass spectrum of the peak at retention time 31.42 min (Fig. 2D) revealed the presence of two peptides: peptides T1a and T1b with calculated masses of 2007.3 Da and 1992.5 Da, respectively. Automated CID MS/MS of the doubly charged ion (m/z 1003.8) of peptide T1a (Fig. 2E) identified this peptide as residues 253–268 of the sIL-6R (DLQHHCVIHDAWSGLR) containing an additional mass of 118.9 Da located at Cys258. Upon reduction with TCEP, the mass of peptide T1a decreased by 119.3 Da and was located at 32.20 min (peptide T9, Fig. 2B; see Table I), consistent with cysteinylation of Cys258.

Sequence data of peptide T1b was not available due to the low abundance of this ion (−20% compared with peptide T1a). However, the 14.8-Da difference observed between peptides T1a and T1b is consistent with the difference in mass between cysteinylated (+119 Da) and S-pyridylethylated (+105 Da) Cys258 of peptide Asp253–Arg268. These data suggest that the peak observed in the sCID/SIM profile (Fig. 2C) emanates from an S-pyridylethylated form (Cys258) of peptide T1b. Peptide T1b was also observed (in low abundance) at approximately the same retention time in the TCEP-reduced total ion current profile (Fig. 2B). Taken together, these data indicate that the majority of sIL-6R purified from CHO cell conditioned medium contain a modified Cys258 (cysteinylated), and only a small portion (−20%) remains as the unmodified Cys258 (free sulphydryl).

**Determination of Disulfide Linkages of sIL-6R by Tryptic Peptide Mapping**—A portion (25%) of the tryptic digest of deglycosylated/S-pyridylethylated sIL-6R was reduced with 10 mM TCEP at pH 6.0 and subjected to on-line RP-HPLC/ESI-IT-MS analysis (Fig. 2B) using the same chromatographic conditions described for the nonreduced digest. Inspection of the nonreduced and reduced tryptic peptide maps of sIL-6R (Fig. 2, panels A and B) revealed that upon reduction of the digest, the retention times of five peptide fractions (T1-T5) in the nonreduced tryptic map (Fig. 2A) changed with the concomitant appearance of nine peptide fractions (T6-T14) in the reduced tryptic map (Fig. 2B). A summary of peptide masses found in fractions T1-T14 is shown in Table I.

As mentioned above, MS analysis of peptide fraction T1 revealed tryptic peptide Asp253–Arg268 (Fig. 2, A, D, and E) containing a cysteinyl-cysteine at position 258 in the sequence (T1a), as well as the small proportion of the peptide that had
been reacted with 4-vinylpyridine (T1b). Similarly, MS/MS analysis of nonreduced T2 (data not shown) identified this peptide as residues Arg5-Arg13 of the sIL-6R with a cysteinylated cysteine residue at position 192. The reduced form of this peptide (T1 in Table I) was also confirmed by MS/MS analysis (data not shown). Modification of Cys192 with 4-vinylpyridine (T1b). Similarly, MS/MS analysis of nonreduced T2 (data not shown) identified this peptide as residues Arg5-Arg13 of the sIL-6R with a cysteinylated cysteine residue at position 192. The reduced form of this peptide (T1 in Table I) was also confirmed by MS/MS analysis (data not shown). Modification of Cys192 with 4-vinylpyridine (T1b).

The mass of peptide fraction T5 (4259.9 Da, Table I) was consistent with the summation of the masses of peptides T10 (1571.8 Da) and T12 (2689.8 Da), where the observed 1.7 Da decrease in the mass of T5 corresponded to the formation of a disulfide link between peptides T10 and T12. The identities of peptides T10 (Lys105-Arg118) and T12 (Ala80-Arg104) in Fig. 2B were confirmed by MS/MS analysis (data not shown). Analysis of the CID fragment ion data (m/z 1436.4 and 958.3) from the +4 charged ion of T5 (m/z 1065.9) was consistent with fragmentation at Pro104 resulting in residues Pro104-Arg104 from peptide T12 and tryptic peptide T10 (Lys105-Arg118) linked through Cys102-Cys113 (Fig. 3C). Automated Edman degradation of tryptic peptide fractions T1-T5 (Fig. 2A) confirmed the NH2-terminal sequence of the peptide components in these fractions (Table I).

N-Linked Glycosylation Sites in sIL-6R—The N-linked glycosylation sites in sIL-6R were determined by the non-appearance of asparagine residues during automated Edman degrada-
tion of tryptic and Asp-N endopeptidase peptides containing Asn-Xaa-(Ser/Thr) motifs (data not shown). Of the four potential N-glycosylation sites in the extracellular domain of the IL-6R, NAT (residues 36–38), NIT (residues 202–204), and NYS (residues 74–76) were found to be glycosylated, whereas one potential N-glycosylation site, NSS (residues 226–228), was not modified. Further confirmation of N-linked glycosylated asparagine residues was provided by Asn/Asp conversion following endoglycosidase F treatment (49); see Asn/Asp conversion at position 74 in peptide T7 (Fig. 4).

Homology Modeling—Two separate models of the sIL-6R were created that depict different orientations between the

---

### Table I

| Peptide fraction* | Sequenceb | Residue numberc | Nonreduced massc | Reduced massc | Δ massc | Cysteine connectivity |
|------------------|-----------|----------------|-----------------|--------------|--------|---------------------|
| T1 a             | DLQHKVHDAWSSLR (T9) | 253–268 | 2007.3 | 1888.0 (1887.1) | −119.3 |           |
| Cys              |           |               |                 |              |        |                     |
| T1 b             | DLQHKVHDAWSSLR (T1b) | 253–268 | 1992.5 | 1992.5 (1887.1) | +0.0   |                     |
| T2               | TQTFCCGILQPDPANNIVTAVAR (T11) | 186–210 | 2718.8 | 2599.9 (2599.0) | −118.9 |           |
| Cys              |           |               |                 |              |        |                     |
| T3               | KFCGLAVPEGDSFYIVSMCVASSGK (T14) | 155–182 | 6544.0 (6543.3) | 2898.5 (2899.3) | +2.2   | 6–174, 146–157     |
| KFONSPFDEQFCQYSESQR (T8) | 133–154 | 2618.7 (2618.8) |               |              |        |                     |
| GVITSLPESVLTCPGVEPDNATVHWL (T13) | 14–44 | 3264.8 (3263.7) |               |              |        |                     |
| T4               | SVQLHDGNSCYR (T7) | 66–79 | 1629.6 (1628.8) | 28–77 |        |                     |
| AGFAGTVHLLDVPPFPQPLCSFR (T12) | 80–104 | 2689.8 (2689.1) |               |              |        |                     |
| T5               | KPSLHHVCEWGR (T10) | 105–118 | 1571.8 (1571.8) | 102–113 | +1.7   |                     |

---

* Tryptic peptide fractions are labeled according to their order of retention as shown in Fig. 2A.

b Peptide sequences identified from MS/MS spectral data using the programs MS-Tag, MS-Product, and manual assignment. Underlined sequences were confirmed by NH2-terminal sequencing. One-letter abbreviations were used for amino acids. Tryptic peptides shown in parentheses refer to peptides produced upon reduction of the tryptic digest prior to RP-HPLC (Fig. 2B).

c Numbers denote amino acid positions in the sequence of the mature protein (Fig. 6).

d Numbers refer to observed mass while those in parentheses are calculated from the amino acid sequence minus 2 daltons for each disulfide bond.

e Numbers refer to observed mass while those in parentheses are calculated from the amino acid sequence.

f Difference in mass between the sum of observed reduced masses and the observed nonreduced mass.

g The difference in mass between the nonreduced and reduced peptide is consistent with S-cysteinylated cysteine.

h The difference in mass between the nonreduced and reduced peptide is consistent with S-pyridylethylated cysteine.
Extracellular Domain of Human IL-6 Receptor

Knowledge of the cysteine connectivity pattern and carbohydrate attachment sites of sIL-6R provides a foundation for the determination of the three-dimensional structure of this molecule and, ultimately, of the IL-6-IL-6R complex. Mass spectrometric and NH2-terminal sequence analysis of a tryptic digest of the sIL-6R identified four disulfide linkages (Cys192-Cys258, Cys77-Cys102, Cys146-Cys157, and Cys6-Cys174), two “free” cysteines (Cys192 and Cys174), and three N-linked glycosylation sites (Asn36, Asn74, and Asn202). The second and third domains of the IL-6R are classified as FN III domains. Sequence analysis of the sIL-6R revealed 25% sequence identity to the CBD of gp130, which is higher than any other template in the database (Fig. 6B).

The MODELLER program works in such a way that the calculated model follows most closely the best template (50). Therefore, our models for the tandem FN III modules of sIL-6R are most closely aligned to the recently published crystal structure for the CBD of gp130 (41). Models A and B of the sIL-6R were shown to be of good quality as assessed by the structural quality checks described under “Experimental Procedures.”

The ProsaiI Z-scores of the CBD of models A and B were -4.8 and -5.3, respectively, comparing favorably with a score of -4.8 for the CBD of gp130. The ProsaiI plot of the IL-6R Ig-like domain is less favorable than the profile of the template, FAB D44.1, a general feature observed in homology model building with low sequence identity.

**DISCUSSION**

The template used for the Ig-like domain was the Vδ domain of FAB D44.1; a V-type Ig-like fold comprised of nine β-strands (39). The amino acid sequence of FAB D44.1 aligned well with the IL-6R Ig-like domain in β-strands B, C, D, and F (Fig. 6A). These strands form the hydrophobic core of the domain and contain a number of conserved hydrophobic and aromatic residues, including the disulfide bond linking β-strands B and F. In contrast, the peripheral strands A, E, and G give less favorable alignments. The absence of β-strands C′ and C″ of the FAB D44.1 Vδ domain in the IL-6R suggests that the Ig-modeled domain forms a C1-type rather than a V-type fold (for review, see Ref. 14). Regardless of the assignment of the particular Ig-like fold, the overall β-sandwich structure is retained. In particular, the relative position of Cys6 will not vary. The second and third domains of the IL-6R are classified as FN III domains. Sequence analysis of the sIL-6R revealed 25% sequence identity to the CBD of gp130, which is higher than any other template in the data base (Fig. 6B).

The two consecutive disulfide linkages in the NH2-terminal FN III domain of the sIL-6R (Cys192-Cys258 and Cys146-Cys157) correspond with those reported for other class I cytokine receptors, namely, granulocyte-colony stimulating factor receptor (51, 52), gp130 (41), and PRLR (44). Similarly, the Cys6-Cys7 connectivity in the Ig-like domain that links the two β-sheets, thus forming the β-sandwich fold, is characteristic of the Ig superfamily (13, 14).

The Cys6-Cys174 disulfide is novel among the cytokine receptors. Sequence alignment of the IL-6R with GHR, PRLR, and gp130, whose CBD structures are known (Fig. 6), predicts that Cys174 is located on β-strand F of the NH2-terminal FN III domain, while Cys6 is situated at the NH2-terminal end of β-strand A of the Ig-like domain. Therefore, the disulfide bond formed between these two residues provides an additional link between the Ig-like and FN III domains. Cys6 is conserved in the IL-11R (54, 55), however, the IL-11R has no corresponding cysteine partner in the NH2-terminal FN III domain with which to form an inter-domain disulfide bond. No other known class I cytokine receptors contain a cysteine arrangement which could facilitate the same disulfide bond as observed for the IL-6R. Interestingly, substitution of Cys174 with Asp in a
The soluble form of the CBD resulted in a complete loss of ligand binding (17). The significance of this observation is not yet known. Truncation experiments of the IL-6R have shown that the Ig-like module is not required for IL-6 binding (17, 56), although the affinity of the truncated receptor for IL-6 was not determined. Recently, it has been suggested that the Ig-like domain may be required for receptor internalization (57). Together, these observations indicate that further studies are needed.

Fig. 4. CID spectrum of the +2 charged parent ion (m/z 815.9) of tryptic peptide T7 (observed mass, 1629.6 Da) from Fig. 2B. The m/z difference of 115.2 between ions y₅ and y₆ and not 114.1 as expected from the amino acid sequence, is indicative of the conversion Asn⁷⁴ to Asp caused by endoglycosidase F removal of the N-linked carbohydrate chain. The m/z difference between the ions either side of y₅ and y₆ (y₆-y₅, 163.0 (Tyr), and y₅-y₇, 56.8 (Gly)) supports this conclusion.

Fig. 5. Homology models of the three-dimensional structure of the sIL-6R. The Cys²⁶-Cys¹⁷⁷ connectivity causes the Ig-like domain to fold back onto the NH₂-terminal FN III domain either in the plane (A) or across the plane (B) of the CBD. Cysteine residues are shown in yellow. Disulfide bonds link Cys²⁶-Cys¹⁷⁷, Cys²⁶-Cys⁷⁷, Cys¹⁰²-Cys¹¹³, and Cys¹⁴⁶-Cys¹⁵⁷. Cys¹⁹² and Cys²⁵⁸ do not form intramolecular disulfide bonds. Glycosylated asparagine residues are shown in pink. Asn²²⁶ (non-glycosylated) is colored blue. The Ig-like domain is modeled on a mouse antibody structure; the CBD follows most closely the structure of the gp130 CBD. Model coordinates are available at http://www.liba.ludwig.edu.au.
needed to elucidate the role of the inter-domain disulfide bond and Ig-like module in IL-6 receptor complex formation.

Cys192 and Cys258 do not form intramolecular disulfide bonds and occur as either free or in a cysteinylated form. These residues are predicted to be located at the COOH-terminal end of \( \beta \)-strand G of the NH2-terminal FN III domain and \( \beta \)-strand E of the COOH-terminal FN III domain, respectively. The three-dimensional models of the sIL-6R shown in Fig. 5 predict that Cys192 and Cys258 are spatially distant from each other and are unable to form a disulfide bond. Cys 258 is conserved among human, mouse, and rat IL-6 receptor species, while Cys192 is replaced by a leucine in the mouse and rat sequences (Fig. 6). Cys 258 is also conserved in the EPOR (58) and cytokine-like factor-1, a recently cloned soluble member of the class I cytokine receptor family whose function is presently unknown (59). The presence of free cysteine residues in the membrane-proximal region of the IL-6R extracellular domain suggests that the IL-6R may homodimerize via intermolecular disulfide bond formation, in a manner similar to that of gpl30 (60), or EPOR (61). Moreover, disulfide linkages between specific \( \beta \)-strand B and F of the Ig-like domain are highlighted pink, and conserved disulfides of class I cytokine receptors in the NH2-terminal FN III domain are blue. The disulfide between the Ig-like and FN III domain is highlighted green and the free cysteines are yellow. Potential glycosylation sites are red, although Asn226 is not glycosylated.

Implications of the cysteine connectivities and carbohydrate attachments on the overall topology of the IL-6R extracellular region are modeled in Fig. 5. The inter-domain disulfide bond which links the Ig-like and NH2-terminal FN III domains is distant from the hinge region connecting these two modules. This causes the Ig-like domain to fold back onto the NH2-terminal FN III domain either in the plane (Fig. 5A) or across the plane of the CBD (Fig. 5B). A different type of inter-domain disulfide bond has been reported for the extracellular region of the interferon-\( \gamma \) receptor, a class II cytokine receptor (67, 68). Here, the inter-domain disulfide bond is located in the hinge region of the CBD, close to the linker sequence connecting the two FN III domains. This maintains a hinge angle of approximately 120°, whereas the inter-domain disulfide bond found in the IL-6R imposes a hinge angle approaching 0° between the Ig-like and NH2-terminal FN III domains.

A model of the sIL-6R (Fig. 5A) complexed with the crystal
structure of IL-6 (48) and CBD of gp130 (41), superimposed over the crystal structure of the growth hormone receptor complex (42), is shown in Fig. 7. This model is consistent with our previously proposed hexameric IL-6 receptor complex model which is based upon the association of two GH/GHR-like trimers (2). While other IL-6/IL-6R/sgp130 trimer models have been reported (69–71), these were generated prior to the publication of the IL-6 and gp130 CBD three-dimensional structures and most likely contain significant errors in the sequence alignments for gp130 and IL-6R. For instance, Kalai et al. (71) align Cys174 of the IL-6R with Ile109 of the GHR, thereby directing the side chain of Cys174 toward the core of the NH2-terminal FN III domain which would prevent the formation of an inter-domain disulfide bond with Cys6. In our model, Cys174 is aligned with Thr112 of the GHR and Glu177 of gp130 (Fig. 7). This orients the side chain of Cys174 outwards from the core of the protein, which is consistent with a Cys6-Cys174 connectivity. The formation of a disulfide bond between Cys6 and Cys174 restrains the Ig-like domain away from the IL-6-binding site (Fig. 7). This suggests that the Ig-like domain is unlikely to be involved in ligand binding.

Our model predicts that cysteinylated Cys192 is located within the IL-6-binding site of the IL-6R. However, this modification seemingly does not interfere with ligand binding since the modified receptor was purified by affinity chromatography using IL-6-Sepharose. This finding is in accord with a previous
study, which demonstrated that mutation of Cys to alanine did not inhibit IL-6 binding (17). It is unlikely that the IL-6R forms an inter-molecular disulfide bond with IL-6 upon binding, since biophysical experiments on both human (72) and mouse (73) IL-6 have shown that it does not contain free cysteines. These observations suggest that despite its location, Cys266 is also located outside the predicted ligand-binding site of the IL-6R which is consistent with this residue not being glycosylated, since carbohydrate chains in this region would be expected to sterically inhibit ligand binding. N-Linked carbohydrate chains located on Asn266 and Asn202 do not interfere with any of the protein-protein interaction sites of the receptor complex predicted in our model. The role of the Ig-like domain and the free cysteine residues in the extracellular domain of the IL-6R must await further studies.

Acknowledgments—We thank J. Bravo and W. Somers for the coordinates of the CDD of gp130 and PRLR, respectively, prior to release into the Protein Data Bank.

REFERENCES
1. Kishimoto, T., Akira, S., Narazaki, M., and Taga, T. (1990) J. Biol. Chem. 265, 2833–2838.
2. Simpson, R. J., Hammacher, A., Smith, D. K., Matthews, J. M., and Ward, L. D. (1997) J. Biol. Chem. 272, 13040–13046.
3. Akira, S., Taga, T., and Kishimoto, T. (1993) Adv. Immunol. 54, 1–78.
4. Novick, D., Engelmann, H., Wallach, D., and Rubinstein, M. (1989) J. Biol. Chem. 264, 1471–1482.
5. Kishimoto, T., Akira, S., and Taga, T. (1998) J. Biol. Chem. 273, 3111–3115.
6. Kishimoto, T., Taga, T., and Akira, S. (1994) Cell 76, 253–262.
7. Taniguchi, T., Hirano, T., and Kishimoto, T. (1988) J. Immunol. 141, 297–314.
8. Ward, L. D., Howlett, G. J., Discolo, G., Yasukawa, K., Hammacher, A., Moritz, R. L., and Simpson, R. J. (1994) J. Biol. Chem. 269, 23286–23289.
9. Sasaki, T., Sueoka, S., Mamasaki, K., Kato, T., Hirata, T., and Kishimoto, T. (1998) J. Immunol. 160, 1371–1379.
10. Naesens, G., Graziani, R., De Serio, A., Savino, R., Ciapponi, L., Lahm, A., Salvati, A., Toniatti, C., and Ciliberto, G. (1995) EMBO J. 14, 1942–1951.
11. Sugita, T., Totsuka, T., Saito, M., Yamasaki, K., Taga, T., and Kishimoto, T. (1990) J. Exp. Med. 171, 2001–2009.
12. Baumann, M., Baumann, H., and Fey, G. H. (1990) J. Biol. Chem. 265, 18653–18662.
13. Hirata, Y., Taga, T., Hibi, M., Nakano, N., Hirano, T., and Kishimoto, T. (1989) J. Immunol. 143, 2900–2906.
14. Bork, P., Holm, L., and Sander, C. (1994) J. Mol. Biol. 242, 309–320.
15. de Vos, A. M., Ultsch, M., and Kossiakoff, A. A. (1994) Science 263, 989–997.
16. Bazan, J. F. (1990) Proc. Natl. Acad. Sci. U. S. A. 87, 6934–6938.
17. Clogston, C. L., Boone, T. C., Crandall, B. C., Mendiaz, E. A., and Lu, H. S. (1996) J. Biol. Chem. 271, 820–824.
18. Simpson, R. J., Hammacher, A., Smith, D. K., Matthews, J. M., and Ward, L. D. (1997) J. Biol. Chem. 272, 13040–13046.
19. Yasukawa, K., and Simpson, R. J. (1995) Biochemistry 34, 2901–2907.
20. Simpson, R. J., and Winzor, D. J. (1995) Biochemistry 34, 2901–2907.