Human Leptin Receptor

DETERMINATION OF DISULFIDE STRUCTURE AND N-GLYCOSYLATION SITES OF THE EXTRACELLULAR DOMAIN

Mitsuru Haniu‡, Tsutomu Arakawa, Edward J. Bures, Yunjen Young, John O. Hui, Michael F. Rohde, Andrew A. Welcher, and Thomas Horan

From the Departments of Protein Structure, Protein Chemistry, and Immunology, Amgen Inc., Thousand Oaks, California 91320-1789

The leptin receptor (OB-R) is a member of the class I cytokine receptor family and mediates the weight regulatory effects of its ligand through interaction with cytoplasmic kinases. The extracellular domain of this receptor is comprised of two immunoglobulin-like and cytokine-receptor homology domains each and type III fibronectin domains. The extracellular domain of human leptin receptor was expressed in and purified from Chinese hamster ovary cells and was found to contain extensive N-glycosylation (approximately 36% of the total protein). The purified protein had a molecular weight of approximately 145,000 and exhibited ligand binding ability as evidenced by formation of ligand-receptor complex, followed by chemical cross-linking. The determined disulfide motif of the soluble leptin receptor contained several distinct cystine knots as well as 10 free cysteines. The N-glycosylation analysis revealed that Asn

The \( \text{ob} \) gene product, leptin, is an important circulating signal for regulation of body weight (1–4). Weight reducing effects of recombinant leptin were observed in both normal mice and mice with diet-induced obesity (5). The biological function of leptin is mediated through its membrane-associated receptor, leptin receptor (OB-R). Because of the biological and therapeutic importance of the leptin/OB-R system, numerous studies have been carried out to understand how this interaction is related to body weight regulation (6–9)

Human OB-R is a membrane-spanning glycoprotein consisting of a signal sequence, two immunoglobulin domains, two cytokine receptor homology (CRH) domains each containing a WSXS motif, fibronectin type III domains, a transmembrane region, and an intracellular domain. The predicted extracellular domain consists of 839 amino acid residues and shows significant similarity to members of the class I cytokine receptor family (6), in particular the gp-130 signal-transducing component of various cytokine receptors as well as the granulocyte-colony-stimulating factor receptor (G-CSFR) (10). Although overall sequence identity between OB-R and gp-130 is low (approximately 24%), several key regions are conserved, particularly the second cytokine homology domain (CRH)-2 (6). In this paper, we report structural information including the disulfide motif and N-glycosylation map of the extracellular domain of OB-R.

EXPERIMENTAL PROCEDURES

Materials—Pepsin was obtained from Sigma. Trypsin and thermolysin were purchased from Boehringer Mannheim. Cross-linkers BS and DSS were obtained from Pierce and 4-HCCA from Sigma. The N- and O-glycansases were purchased from Genzyme (Cambridge, MA), and sialidase was from Boehringer Mannheim. Recombinant human leptin was prepared as described previously (11). All chemicals used were of reagent grade or analytical grade.

Antisera Preparation—New Zealand White rabbits were injected subcutaneously on day 1 with 0.2 mg of soluble OB-R and an equal volume of Freund’s complete adjuvant. Further boosters (days 7, 21, 35, and 56) were given with the substitution of Freund’s incomplete adjuvant. Antibody titers were monitored by enzyme-linked immunosassay. After the third booster, 20 ml of blood was obtained from each animal.

The antisera for leptin was affinity purified on a leptin-Sepharose column and then conjugated to horseradish peroxidase using a Freezyme Kit (Pierce).

Western Blotting—Samples were electrophoresed under reducing conditions on 4–20% SDS-PAGE gels (Novex, San Diego) and then transferred to BA83 nitrocellulose (Schleicher & Schuell) overnight at 10 V in 25 mM Tris, 192 mM glycine, 20% v/v methanol. After transfer, blots were treated for 1 h in 1% Tween 20 (Bio-Rad) containing 5% w/v non-fat dry milk before adding antisera (1/2000 dilution) or antibody (0.1 µg/ml) for an additional hour. The blots were then washed with PBS/Tween 20 four times, 10 min each. For detection of receptor, a secondary anti-rabbit Ig (Amersham Pharmacia Biotech) was diluted 1/6000 in PBS/Tween 20 and reacted with the blot for 30 min, afterward repeating the washes. Blots were then treated with enhanced chemiluminescence detection reagents (Amersham Corp.) and exposed to Kodak AR film.

Expression and Purification of an Extracellular Domain of OB Receptor—DNA encoding the extracellular domain and signal sequence of the human leptin receptor (residues 1–839) (6) was cloned into the mammalian expression vector pDSRα (12) and then transfected into CHO D cells (13). Individual cell colonies were selected based upon expression of a dihydrofolate reductase gene in the vector. The highest expressing cells were adapted to 30 µm methotrexate to stimulate amplification of soluble OB-R expression. After expansion in spinner flasks, roller bottles were inoculated with 2 × 10⁶ cells in 200 ml of Dulbecco’s modified Eagle’s medium:Ham’s F12 supplemented with nonessential amino acids (Life Technologies, Inc.) and 5% fetal bovine serum (HyClone, Logan, UT). After the cells reached confluency (3–4 days), the medium was replaced with 200 ml of growth medium with no serum. Conditioned media were harvested after 7 days and immediately chilled on ice. CHO cell-conditioned media containing soluble OB-R were concentrated 10-fold with a Pellicon tangential flow ultra-

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ To whom correspondence should be addressed. Tel.: 805-447-3117; Fax: 805-499-7464; E-mail: mhaniu@amgen.com.

1 The abbreviations used are: R, receptor; CRH, cytokine receptor homology; BS, bis-disuccinimidyl suberate; CHO, Chinese hamster ovary; Me₂SO, dimethyl sulfoxide; DSS, disuccinimidyl suberate; G-CSFR, granulocyte colony-stimulating factor receptor; 4-HCCA, 4-α-hydroxy-
cyano-cinnamic acid; MALDI, matrix-assisted laser desorption ioniza-
tion; OB-R, leptin receptor; NEM, N-ethylmaleimide; PTH, phenylthio-
hydantoin; PAGE, polyacrylamide gel electrophoresis; HPLC, high pressure liquid chromatography; PBS, phosphate-buffered saline.
filtration device (Amicon, Danvers, MA) fitted with a 100,000 molecular weight cut-off filter (Filtron, Bearfoot, MA). The concentrate was dialyzed with 4 volumes of 25 mM Tris-HCl, pH 7.8, at 4 °C and then loaded onto a Q-Sepharose column (80 x 140 mm) equilibrated with 25 mM Tris-HCl, 30 mM NaCl, pH 7.8. After eluting the column with a linear NaCl gradient from 30 to 350 mM in 20 column volumes, aliquots of column fractions were applied to 4–20% polyacrylamide gels (Novex, San Diego) (14) and run under reducing SDS-PAGE conditions. Protein bands were visualized by silver staining and a pool made of fractions containing soluble OB-R. This pool was reacted with hydroxylysylatite for 4 h at 4 °C (Calbiochem) which bound contaminating proteins but very little soluble receptor. The hydroxylatate unbound material was concentrated and applied to a Superose 12 column (35 x 600 mm) (Amersham Pharmacia Biotech) equilibrated with 10 mM sodium phosphate, 200 mM NaCl, pH 7.6. Column fractions were analyzed by SDS-PAGE as described above, except using Comassie Blue to stain, and a pool consisting of more than 95% pure receptor was made.

Chemical Cross-linking Experiments—A stable complex was made by incubating soluble OB-R (100 µg) with a 2-fold molar excess of recombinant human leptin at 24 °C for 4 h. The ligand-receptor complex was purified by Superose 12 (10 x 300 mm) equilibrated with 1x PBS, pH 7.8. Cross-linking was performed by treating the complex with 1 mM DSS (dissolved in Me2SO) or BS (dissolved in water) for 1 h at room temperature. The reaction was terminated by addition of 0.1 M ammonium acetate. The reactions were analyzed by SDS-PAGE under reducing conditions. Cross-linked products were detected by both silver staining and Western blotting using an affinity purified, horseradish peroxidase-conjugated, polyclonal antibody raised against human leptin.

Proteolytic and Chemical Cleavages of Soluble OB-R—In order to obtain Cys-containing peptides, soluble OB-R (100 µg) was digested with pepsin (2 µg) in 0.02 N HCl, pH 2, for 20 h at 37 °C. Under acidic conditions, disulfide rearrangements can be avoided. The digestion was terminated by direct injection into an HPLC column. Tryptic digestion of soluble OB-R was performed in the presence of 5 mM iodoacetate in 0.1 M Tris-HCl, pH 7.5. The digestion was carried out at 37 °C for 24 h with an enzyme:substrate ratio of 1:50 (w/w). Thermolytic or chymotryptic digestions of peptide fragments were performed in 0.1 M Tris-HCl, pH 7.5. The digestion was carried out at 37 °C for 24 h with an enzyme:substrate ratio of approximately 1:20 (w/w). Cyanogen bromide cleavage of OB-R (100 µg) was performed with 2 µg of cyanogen bromide dissolved in 100 µl of 70% formic acid. The reaction was allowed to proceed for 24 h at 25 °C in the dark and was stopped by direct HPLC injection. The cyanogen bromide peptides were treated with N-, O-glycanase (0.5 units) in 0.1 M Tris-HCl buffer, pH 7.5, and then digested with endoproteinase Asp-N, followed by endoproteinase Glu-C in the same buffer with an enzyme:substrate ratio of 1:50 (w/w).

Deglycosylation of Soluble OB-R—Soluble OB-R was reduced and denatured by adding SDS to 0.4% and β-mercaptoethanol to 50 mM and then heating at 95 °C for 5 min. Nonidet P-40 was then added to 1.8% with an enzyme:substrate ratio of 1:50 (w/w). Thermolytic or chymotryptic digestions of peptide fragments were performed in 0.1 M Tris-HCl buffer, pH 7.5, and then digested with endoproteinase Asp-N, followed by endoproteinase Glu-C in the same buffer with an enzyme:substrate ratio of 1:50 (w/w).

**FIG. 1.** SDS-polyacrylamide gel electrophoresis of soluble OB-R. A, lane 1, standard markers; lane 2, intact OB-R; lane 3, intact OB-R treated with buffer; lane 4, OB-R treated with N-glycanase; lane 5, OB-R treated with sialidase; lane 6, OB-R treated with sialidase and O-glycanase; and lane 7, OB-R treated with sialidase and N-, O-glycanases. B, Western blotting using antibody raised against human soluble OB-R. See “Experimental Procedures.”

**FIG. 2.** Amino acid sequence of soluble OB-R. The sequence was deduced from cDNA sequence (6). The double lined sequence shows a signal peptide.

|   | MECKCEPVVL  | LEMPFLVIT | APNLGSPITP | WRFKLSCMPF | NSTVDFFLLP 50 |
|---|-------------|-----------|------------|-------------|----------------|
| 51 | AGISMHNTNSNGYDEYRAP | KNNSGQGS | NLYSGIPTCF | FGSRQHDCRNCS 100 |
| 101 | LCAICDUGEGFPYTVNLSLVF | QUGNWQINTQ | ICYVSISLEL | 150 |
| 151 | LIPSNYKBNNYLLVYVPDLEVYQ | DPVLQDGKQ | MPNHCWNCY | HCCCECMLRV 200 |
| 201 | PTAKLUYNYLCNLCSQGCV | LPKCTWSMY | RFTNNPCW | LKLMNICEO 250 |
| 251 | GNLKYSWSPSLVEYPFLQYQ | VXK REV | TREADKHYSA | TSLYDVLDSL 300 |
| 301 | GSSVCKVQVGRLQDGQNSHS | DWSTPRFVPT | QDQVYTPPPK | LTYSVGVNMP 350 |
| 351 | HCTYKKEKIKVESEK1999VN | NLSBPKFQ | YVDINSSVVR | VTFNLNSNP 400 |
| 401 | PROKYDYYAV| YCNMCRBCY | RVEAYLVIDV | NINISCPETQ | YLUGMCNCRS 450 |
| 451 | QGISTQOLASES | TLOQRYVSSHSS | LCSDPFSEH | FLSEREPDCY | QQDSQFPKFT 500 |
| 501 | QFTPLSGGY | MGRINHISLGL | SLSDSPEPVTVCY | ESYSVKLP | SGRVASKRIN 550 |
| 551 | IGLKQSWK | VFPENNMLQ | VTRLQVSLQ | VXQWMZVYD | AKSTWVLPSV 600 |
| 601 | PDLCAVAYAVQ | VRECRDGQGGL | YWNSWNSYAP | TVVMKDFK | RGGPRRIN 650 |
| 651 | QDPIMKERNVLTLMHKPMRN | DLSGCWQYVR | INBPASCTON | WEVDWGNHTK 700 |
| 701 | PTFLSTQSAH | TPVVLAISN | GASVANFNLT | FSPMMVKN | VQSLSSAPL 750 |
| 751 | SSCVIVSWIPv | SPSDRLMYVF | IREDNMLQND | GEIKHRMRSS | SVKSYMIDH 800 |
| 801 | FIPFPEQFOF | LYPIFMRQV | PKPIKINGTFQ | DDERRHQSD 839 |
The reduced peptide was alkylated with 20 mM N-ethylmaleimide for 30 min at room temperature after adjusting to pH 6.0 with 1 M Tris base as described previously (15). The modified peptides were sequenced after purification by reversed phase HPLC using a Zorbax CN column (2.1 × 150 mm) as described above.

Mass Spectrometry—Matrix-assisted laser desorption ionization (MALDI) mass spectrometry of the peptides was performed using either a Kratos IV (Kratos Analytical) or Voyager mass spectrometer (PerSeptive Biosystems). Samples were dissolved in 0.1% trifluoroacetic acid, 50% acetonitrile and then spotted on the sample cartridge with a 4-HCCA matrix. Cys-containing peptides were also analyzed using a Sciex API triple quadrupole mass spectrometer with an ion-spray interface using a Michrom Biosource Ultrafast Microprotein Analyzer. The carrier solvent was 50% acetonitrile/water with 0.1% trifluoroacetic acid flowing at 5 μl/min. The scan range was 300–2400 atomic mass unit with a step of 0.5 atomic mass unit. The mass units and standard deviation were calculated using Sciex hypermass software.

RESULTS

Molecular Properties of the Soluble OB-R—Human soluble OB-R was expressed in CHO cells and purified from conditioned medium by Q-Sepharose, hydroxylapatite, and gel filtration. The purified soluble OB-R exhibited a single, broad band on SDS-PAGE, corresponding to approximately 145,000 daltons (Fig. 1). Deglycosylation of the purified soluble OB-R with N-glycanase reduced the molecular mass to approximately 90,000 daltons, indicating extensive N-glycosylation al-

**Fig. 3.** Ligand-receptor complex. A, HPLC analysis of OB-R-leptin (1:1) complex. The purified complex was subjected to reversed phase HPLC using a Zorbax CN column as described under “Experimental Procedures.” Protein peaks I and II correspond to OB-R and leptin, respectively. B, control proteins; OB-R and leptin (L).

**Fig. 4.** SDS-PAGE of chemical cross-linking between soluble OB-R and leptin. A (left), cross-linking between OB-R and leptin with 1 mM BS and DSS. Lane 1, standard markers; lane 2, OB-R; lane 3, leptin; lane 4, cross-linked with 1 mM DSS; lane 5, cross-linked with 1 mM BS. The 4–20% gel was used for these experiments. Right, Western blotting using human leptin antibody (see “Experimental Procedures”). B, electrophoresis under the same conditions as in A except for 8% gel used.

**Fig. 5.** Pepsin-generated peptide map of soluble OB-R. The protein sample (100 μg) was digested with pepsin (2 μg) and directly injected to a Vydac C18 column (2.1 × 150 mm) as described under “Experimental Procedures.”
though treatment with sialidase and O-glycanase slightly increased its mobility, suggesting that O-glycosylation is present but not significant. The N-terminal sequence of the purified protein was identified to be FXLSYPTP (residues 22–30), whereas the C-terminal sequence of DIEKHKQSD (residues 832–839) was identified from a pepsin-generated fragment.

The observed N- and C-terminal sequences were identical to those predicted from the cDNA-deduced sequence reported by Tartaglia et al. (6), which is shown in Fig. 2. Thus, the soluble OB-R purified here contained 818 amino acid residues (from residues 22 to 839) with a calculated molecular weight of 93,501, indicating that the OB-R contained approximately 36% carbohydrates of its total mass.

**Ligand Binding Activity and Preparation of Ligand-Receptor Complex**—The leptin-OB-R complex was prepared by mixing the receptor with a 2-fold molar excess of leptin. After incubation at 4 °C for 24 h, the mixture was subjected to size exclusion chromatography using a Superose 12 column (10 × 300 mm) equilibrated with PBS, pH 7.8. Two peaks were obtained, corresponding to complex and free leptin. From the elution position in Superose 12 column, it is likely that the complex contained only one molecule of OB-R. The complex peak was subjected to reversed phase HPLC using a Zorbax CN (2.1 × 150 mm), which indicated that it contained a 1:1 ratio of leptin and OB-R (Fig. 3).

**Chemical Cross-linking of the Ligand-Receptor Complex**—In order to examine ligand binding specificity of OB-R, the purified complex was cross-linked with 1 mM BS and DSS, followed by SDS-PAGE analysis. One gel was silver-stained and other electroblotted to nitrocellulose and detected with an antibody. TABLE I, Table II.

![Fig. 6. MALDI mass spectrometry of peptic peptide P21. A, peptide P21 was analyzed by MALDI mass spectrometry using 4-HCCA. Signal 3573.3 corresponds to the intact peptide. B, increasing laser power shows individual peptide mass 1761.7 from the cleavage of disulfide bond.](image)
purified, polyclonal antibody raised against recombinant human leptin. In the silver-stained gel (Fig. 4A), a protein band just above that corresponding to free receptor is observed. Fig. 4B shows the Western blot of this gel using an 8% gel, verifying the presence of leptin (cross-linked to the OB-R) in this band.

Pepsin-generated Peptide Map of the Soluble OB-R for Determining Disulfide Linkages—The soluble OB-R was fragmented by pepsin digestion, resulting in numerous peptide peaks separated by reversed phase HPLC (Fig. 5). Therefore, sequence analysis was performed on all peptic peptides and indicated that more than 15 Cys-containing peptides were isolated. Table I shows the summary of the sequence analyses of the Cys-containing peptides including both disulfides and free cysteines. Peak 8 contained two peptides, VRX (residues 611–613) (P8a) and LTKMTXRW (residues 442–449) (P8b), in which X was expected to be a cysteine residue according to the cDNA sequence. MALDI mass spectrometry showed that these peptides (P8a and P8b) were not linked together through a disulfide bond but existed as free cysteines, Cys^{447} and Cys^{613}, respectively. Although the two peptides are different in size, they co-eluted from the HPLC column under the conditions used.

Peptic peptide P16 contained the single sequence MKND-SLS (residue 668–674), in which X is an unidentified residue but is a Cys residue based on the cDNA sequence. Since this peptide was solely isolated, Cys^{674} was assigned to be in the

### Table III

| No. | Peptide               | Sequence* |
|-----|-----------------------|-----------|
| 1   | T1                    | LSCMPPNSTYDFFLPGILSK (35–55) |
| 2   | T1-Th1                | LSCMPNNSTVDY (35–46) |
| 3   | T2                    | ILTSVGNSVSPHCYK (340–355) |
| 4   | T2-CT1                | HC1Y (351–354) |
| 5   | T2-CT2                | CCNEHECHR (412–421) |
| 6   | T3                    | GSFOQMVHNCSCVHECCECLPVPTAK (179–204) |
| 7   | T4                    | VNIVQSLSYPLNSSCVIVSWILPSDYK (738–766) |
| 8   | T5a                   | LNDTDMCMCLK (205–214) |

*N denotes the glycosylation sites. Parentheses show the residue number.

### Table IV

**Sequence analysis of peptide T1-E1**

| No. | Peptide 1 (PTH) | Peptide 2 (PTH) | Peptide 3 (PTH) |
|-----|----------------|----------------|----------------|
| 1   | Leu 3.6        | Thr 3.0        | Gln 3.6        |
| 2   | Ser 2.0        | Thr 2.4        | Asp 2.0        |
| 3   | Cys^{37}       | Phe 1.6        | Arg +          |
| 4   | Met +          | His 0.5        | Asn (N)*       |
| 5   | Pro 1.9        | Cys^{89} (0.94)* | Cys^{89} (0.94)* |
| 6   | Pro 2.1        | Cys^{89}       | Ser 0.2        |
| 7   | Asn (N)*       | Phe 0.5        | Leu 1.5        |
| 8   | Ser 0.3        | Arg +          | Cys^{102}      |
| 9   | Thr 0.2        | Ser 0.2        | Ala 0.2        |
| 10  | Tyr 0.5        | Gru 0.3        | Asp            |

* Numbers show PTH recovery (picomoles). Only 10 cycles were analyzed.
* Not detected due to N-glycosylation.
* Detected as diPTH-Cys.

### Table V

**Sequence analysis of peptide T2-CT2**

| No. | Peptide 1 (PTH) | Peptide 2 (PTH) |
|-----|----------------|----------------|
| 1   | Cys^{352} (13.9)* | Cys^{342} (13.9)* |
| 2   | Ile 12.5        | Cys^{413}      |
| 3   | Tyr 7.5         | Asn 7.9        |
| 4   |                | Glu 6.5        |
| 5   |                | His 2.6        |
| 6   |                | Glu 7.3        |
| 7   |                | Cys^{413}      |
| 8   |                | His 2.7        |
| 9   |                | His 4.3        |
| 10  |                | Arg 0.9        |

* Numbers show PTH recovery (picomoles).
* Detected as diPTH-Cys.

Mass observed (calculated): 1662.4 (1660.4)

---

**Human Leptin Receptor**

**Figure 7. Tryptic peptide map of soluble OB-R.** The receptor (100 µg) was digested with trypsin (2 µg) in 0.1 M Tris-HCl buffer, pH 7.5, in the presence of 5 mM iodoacetate. The numbered peaks (T1–T3) contained cysteine clusters. Due to carbohydrate heterogeneity, some peptides showed their multiplicities despite having the same peptide sequence.
free state. Mass spectrometry of this peptide revealed it not to be glycosylated although it contained an N-glycosylation sequence, NXS. Two peptides, P18 and P19, had the same sequence of RYVINHHTSXNGTW (residues 678–691) where N denotes an N-glycosylation site. Mass spectrometry of the intact peptides showed a larger mass than expected. After deglycosylation with N-glycanase, these peptides showed the same mass of 1689.0, which is nearly identical to the expected mass 1687.9. The observed difference in HPLC retention time is evidently due to carbohydrate heterogeneity, and the data indicated that Cys687 was in a free state.

Peptide P21 contained the single sequence ED-SPLVPFQKGFQVMHXXSVNXEXXL (residues 170–197). Since this peptide contains five cysteines according to the DNA sequence, at least one must be in the free state. Disulfide linkages in this peptide were determined after cyanogen bromide cleavage, as discussed later. Additionally, Asn187 was glycosylated since no PTH-Asn was detected at this cycle. Meanwhile, peptide 27b contained two sequences, KLSXMPPNSTY (residues 34–44) and SKTTFHXXFRSEQDRNXLXAD (residues 83–104), containing one of three cystine knots in the receptor. However, disulfide linkages in these peptides could not be determined from this digestion. Further digestion of this peptide with thermolysin gave the three sequences, LSXMPPNSTY (residues 35–44), FHXX (residues 87–90), and RSEQDRNXLXAD (residues 92–104). Asn41 and Asn98 were assigned to be N-glycosylation sites, as determined by sequence analysis. Table II shows sequence analysis of peptide P27b. Upon PTH analysis of this peptide, the third cycle did not show significant diPTH-Cys, indirectly suggesting that Cys37 is not disulfide-linked to Cys89 but probably to Cys90. The complete disulfide structure of this Cys cluster was eventually determined by tryptic digestion as discussed later.

Peptide P28a corresponded to the same sequence as peptide P21 but contained another seven N-terminal amino acids. Peptide P28b showed a single peptide MCLK (residues 211–214) with one free cysteine, Cys212. Sequence analysis of peptide P29 showed two sequences, PKDXX (residues 485–489) and YEIXFQ (residues 496–501), clearly demonstrating the presence of a disulfide linkage between Cys488 and Cys498. This result was confirmed by mass spectrometry, experimental mass 1425.0 versus the theoretic mass 1423.0. Similarly, peptide P33 showed two sequences, NIQXW (residues 123–132) and LYX (residues 141–143), again using mass spectrometry to confirm that Cys131 and Cys142 are linked together. Peptide P34b contained the two sequences, YXSDDIPIHPSGI (residues 472–484) and DSPPTXVLPDS (residues 523–533), and similarly the molecular mass of this peptide was 2583.4, close to the
theoretic mass of 2587.2, clearly demonstrating a disulfide linkage of Cys \(^{477}\)-Cys \(^{528}\). A glycopeptide P36, SAYPLNSSCNSE (residues 434–458), confirmed by mass spectrometry, it contained a free cysteine Cys 753, although limited amount. Since peptide P37 had the single sequence VSLVPVLAXYA (residues 596–608), confirmed by mass spectrometry, it contained a free cysteine Cys 604. Meanwhile, peptide P38 contained two cysteines in a single peptide IS-NEHE-SE (residues 745–753) contained a free cysteine, Cys 753, although.

**Determination of the Disulfide Linkages of Three Cystine Knots**—In order to determine the disulfide linkages of three cystine knots present in OB-R, intact receptor was digested with trypsin in the presence of 5 mM iodoacetate. The peptide map shown in Fig. 7 gave several key Cys peptides for disulfide linkages determined by PTH analysis after further digestion of this peptide with trypsin in the presence of 5 mM iodoacetate. The peptide CB1 was deglycosylated and digested with endoproteinase Glu-C, resulting in a new fragment 16, 17). From the chymotryptic digest, two Cys-containing peptides were isolated as follows: T2-CT1 containing cysteine residues located at the same position (16, 17). To confirm this result, a thermolytic digest of peptide T1 provided two sequences LSCMPNStYDFILFLAPGLSK (residues 35–55), TTPHC-CFRSE (residues 85–94), and QDRNCSDCQDNE (residues 95–107), respectively. When sequenced, the first cycle corresponding to Cys\(^ {37}\) and Cys\(^ {90}\) showed significant yields (–5%–10%) of diPTH-Cys (Table IV) (16, 17). To confirm this result, a thermolytic digest of peptide T1 provided two sequences LSCMPNStYDFILFLAPGLSK (residues 35–46) and HCPPRSEQDRNCQDNE (residues 88–100) which also showed significant recovery of diPTH-Cys at cycle 3 corresponding to Cys\(^ {37}\) and Cys\(^ {90}\). Subsequently, these results indicated that Cys\(^ {89}\) was linked to Cys\(^ {90}\), supporting another linkage Cys\(^ {37}\)-Cys\(^ {90}\). Cys\(^ {102}\) was assigned to be a free state.

![image](https://example.com/image1.png)

**Fig. 10.** Purification of partially reduced and alkylated CB-1 and isolation of Cys-containing peptides from peptide CB-1. A, partial reduction and alkylation of CB-1. B, peptide map of CB1. The peptide CB1 was deglycosylated and digested with endoproteinases Asp-N and Glu-C as described under “Experimental Procedures.” Peaks 1, 2, and 3 are denoted as CB1-D.E.1, -2, and -3, respectively.

**TABLE VI**

Sequence analysis of peptides CB1, CB1-NEM1, CB1-NEM2, and CB1-NEM3

| No. | CB1 (PTH)a | CB1-NEM1 (PTH) | CB1-NEM2 (PTH) | CB1-NEM3 (PTH) |
|-----|------------|----------------|----------------|----------------|
| 1   | Val 29.5   | 3.5            | 6.7            | 11.2           |
| 2   | His 14.9   | 2.7            | 5.7            | 5.8            |
| 3   | Cys\(^ {196}\) b | 1.5d         | 2.9d           |                |
| 4   | Aan (N) e |                |                |                |
| 5   | Cys\(^ {188}\) b | 0.4d          | 1.1d           | 3.2d           |
| 6   | Ser 10     | 2.5            | 3.6            | 5.0            |
| 7   | Val 14.2   | 3.3            | 6.9            | 8.9            |
| 8   | His 9.0    | 1.5            | 2.6            | 3.1            |
| 9   | Glu 8.3    | 1.9            | 2.9            | 3.1            |
| 10  | Cys\(^ {193}\) b | 1.5d          |                | 2.4d           |
| 11  | Cys\(^ {194}\) b | 1.5d          | 2.4d           |                |
| 12  | Glu 5.5    | 1.2            | 2.7            | 4.7            |
| 13  | Cys\(^ {196}\) b | 0.5d          | 1.5d           |                |
| 14  | Leu 5.5    | 0.9            | 1.1            | 2.4            |
| 15  | Val        | 3.7            | 1.2            | 1.8            | 3.7            |
| 16  | Pro        | 3.5            |                |                |
| 17  | Val        | 3.9            |                |                |
| 18  | Pro        | 4.4            |                |                |
| 19  | Thr        | 1.6            |                |                |
| 20  | Ala        | 1.9            |                |                |

Mass no. (calculated) 3346 (3047 + 305 = 3352)

a PTH recovery (picomoles). Data from 20-cycle or 15-cycle runs are presented here.

b Not detected.

c Not detected due to N-glycosylation.

d Detected as PTH-N-ethylsuccinimidocysteine.

e Detected as PTH-N-ethylsuccinimidocysteine.
etry of peptide T1-CT2 showed a signal of 1662.4, which is identical to the theoretical mass (Fig. 8). When peptide T2-CT1 was sequenced, no significant diPTH-Cys was detected at any cycle upon PTH analysis, whereas peptide T2-CT2 showed a significant diPTH-Cys at the first cycle corresponding to Cys^352_Cys^412. Recovery of diPTH-Cys was approximately 95%, demonstrating the presence of Cys^352_Cys^412. Since mass spectrometry indicated that peptide T2-CT2 did not contain any carboxymethylcysteine, the remaining cysteine linkage was assigned to be Cys^413-Cys^418.

Peptide T3 has a sequence GSFQMVHCNCSVHECC (residues 179–204) in which an N-glycosylated asparagine is underlined (Table III). This single peptide contained five cysteine residues and, as before, one must be in the free state. Since the peptide was recovered in low yield, we isolated a corresponding peptide after CNBr cleavage of intact soluble OB-R. Fig. 9 shows the HPLC peptide map of the CNBr fragments. The peptide CB1 contained a sequence VHCNCSVHECC (residues 184–211). Mass spectral analysis revealed that deglycosylated CB1 has a mass of 3346, corresponding to a single glutathione adduct (Table VI). To determine the two disulfide bonds, the CB1 was partially reduced with 10 mM Tris-(2-carboxyethyl)phosphine at 25 °C for 5 min, followed by alkylation with NEM in order to identify a preferentially reduced disulfide bond. Fig. 10A shows an HPLC profile after partial reduction and alkylation of peptide CB1. Sequence analysis of the peptide CB1-NEM1 revealed that only Cys^186 and Cys^193 were detected as PTH-N-ethylsuccinimidocysteines appearing as double peaks between PTH-Pro and PTH-Met. Peptide CB1-NEM2 contained four NEM-modified cysteines, Cys^186_Cys^193 and Cys^196, respectively. Finally, the peptide CB1-NEM3 contained five NEM-modified cysteines including additional Cys^194 (Table VI). These results suggested the presence of one disulfide linkage, Cys^186_Cys^193, and thus suggested another linkage, Cys^186_Cys^193. Cys^194 was therefore assigned to be glutathione adduct. This assignment was confirmed by direct isolation of Cys peptides from peptide CB1. The CB1 peptide was deglycosylated with N-glycanase treatment and was digested with endoproteinase Asp-N, followed by endoproteinase Glu-C. The peptides were purified by reversed phase HPLC as shown in Fig. 10B. Peptide CB1-D.E-1 contained the sequence VHCNCSVHECC (residues 187–194), showing a small amount of diPTH-Cys at cycle 7 corresponding to Cys^193 (Table VII). This suggests a presence of diPTH linkage Cys^188-Cys^193. Mass spectrometry indicated one of the cysteines (presumably Cys^194) was modified probably with γ-EC (mass 250), a breakdown product of glutathione. Meanwhile, peptide peak CB1-D.E-2 contained two peptides VHC (184–186) and CLVPVPTAK (196–202), indicating a disulfide linkage Cys^186-Cys^196, as confirmed by mass spectrometry (Table VII).

**Determination of N-Glycosylation Sites**—Sequence analyses of both peptic and tryptic fragments provided information on the extensive N-glycosylation of the OB-R. Determination of the glycosylation sites relied upon no detection of Asn, obtained in sequence analysis. As indicated in Table I, the Cys-containing peptides have eight N-glycosylation sites. The remaining 10 N-glycosylation sites were identified by sequence analysis and mass spectrometry of other peptic or tryptic peptides. Some peptides were re-treated with N-glycanase in order to obtain

![H. R. L. P.](image)

**Fig. 11. Disulfide structure and N-glycosylation map of the extracellular domain of OB-R.** Three cystine clusters were schematically drawn as well as free cysteine residues. N-Glycosylation sites were indicated as rhombic marks. CRH, cytokine receptor homology domain; WSXWS, a common motif in cytokine receptor.

**Table VIII**

| No. | Residue no. | Consensus sequence (Murine) | Glycosylated or non-glycosylated a |
|-----|-------------|-----------------------------|-----------------------------------|
| 1   | Asn^23      | NLS (NLA) b                 | G                                 |
| 2   | Asn^41      | NTS (NAS)                   | G                                 |
| 3   | Asn^56      | NTS (NAS)                   | G                                 |
| 4   | Asn^73      | NTS (NAS)                   | G                                 |
| 5   | Asn^81      | NTS (NAS)                   | G                                 |
| 6   | Asn^98      | NTS (NAS)                   | G                                 |
| 7   | Asn^106     | NTS (NAS)                   | G                                 |
| 8   | Asn^208     | NTS (NAS)                   | G                                 |
| 9   | Asn^270     | NTS (NAS)                   | G                                 |
| 10  | Asn^347     | NTS (NAS)                   | G                                 |
| 11  | Asn^397     | NTS (NAS)                   | G                                 |
| 12  | Asn^432     | NTS (NAS)                   | G                                 |
| 13  | Asn^511     | NTS (NAS)                   | G                                 |
| 14  | Asn^624     | NTS (NAS)                   | G                                 |
| 15  | Asn^659     | NTS (NAS)                   | G                                 |
| 16  | Asn^670     | NTS (NAS)                   | G                                 |
| 17  | Asn^697     | NTS (NAS)                   | G                                 |
| 18  | Asn^728     | NTS (NAS)                   | G                                 |
| 19  | Asn^750     | NTS (NAS)                   | G                                 |

a G shows the peptide in which asparaginyl residue was N-glycosylated. N denotes non-glycosylation.
b Underlined sequence denotes non-glycosylation site in murine OB-R.
their exact mass. Table VIII shows all the possible N-glycosylation sites of the human receptor and a comparison with the corresponding murine sequence. Almost every N-glycosylation sequence was found to be modified by carbohydrates and only two sites were not, Asn\(^{433}\) and Asn\(^{670}\). Interestingly, residue Asn\(^{624}\) located at the second WS motif was found to be N-glycosylated. Since the WSXS motif has been shown to be important in the folding of interleukin-2 receptor or at least its ligand binding region (19), N-glycosylation at Asn\(^{624}\) may affect the folded structure of OB-R. From peptide mapping, 18 of 20 N-glycosylation sequences might be occupied by N-glycans. An overall structural diagram of the extracellular domain of human OB-R detailing 9 disulfide linkages, 10 free cysteines, and \(\text{XWS}^{18}\) motif is shown in Fig. 11. These disulfide and N-glycosylation motifs are not homologous with any known cytokine receptors including G-CSFR (18).

**DISCUSSION**

In comparing the sequences of the human and murine OB-R, the extracellular domain of the murine receptor lacks several cysteines (Cys\(^{102}\), Cys\(^{104}\), Cys\(^{212}\), and Cys\(^{567}\)) present in the human receptor (Fig. 11). This leads us to believe that these four cysteines may be in the free state. This study revealed that all of these cysteine residues are indeed present in the free state and, additionally, that Cys\(^{536}\), Cys\(^{647}\), Cys\(^{650}\), Cys\(^{613}\), Cys\(^{674}\), and Cys\(^{723}\) are also in the free state. Although 10 free cysteine residues are present in the human OB-R, most of them are largely unreactive to alkylating reagents, suggesting that they are buried within the protein molecule or otherwise protected from alkylation by O- or N-glycosylation. Possible functional roles of these free cysteines were not addressed in this study.

OB-R contains an unusual N-glycosylation motif, indicated as an NCS sequence, where the cysteine residue is involved in the disulfide formation. This sequence was found in several proteins including protein C and G-CSFR (18, 20), and there are two sites in OB-R. Several other unusual N-glycosylation sites containing cysteine have been reported for proteins including CD69, immunoglobulin \(\mu\) chain, and von Willebrand factor in which the atypical sequences NXC or NGGT are also found to be N-glycosylated (20–23). However, these sequences contained a Ser or Thr residue in the middle residue, for instance NSCX or NXCX, whereas X denotes variable residues.

Both the G-CSF receptor and gp-130 have significant sequence homology to the OB-R (6). The CRH domains of these and other receptors include a WSXS motif, which is thought to play an important role in folding and is located between two \(\beta\) sheet strands F’ and G’ (24, 25). Furthermore, another sequence motif, SSFY, has been postulated to be important in the ligand-binding region of interleukin-6 receptor (26). A similar sequence was also found in OB-R, SSLY (residues 469–472). Although OB-R contains two repeating Ig-CRH domains, the disulfide structures of these two domains were found to be slightly different from each other (Fig. 11) in that the F’–G’ loop of the C-terminal CRH contained two free cysteines, suggesting a different conformation from the N-terminal CRH which did not have any such free cysteines. These findings suggest that domains CRH I and II may have different mechanisms for ligand interaction or perhaps other implications for biologic activity. To study further this structure-function relationship site-directed or deletion mutagenesis of the receptor would be useful.

**Acknowledgment**—We thank Joan Bennett for preparation of the manuscript.

**REFERENCES**

1. Zhang, Y., Proenca, R., Maffei, M., Barone, M., Leopold, L., and Friedman, J. M. (1994) *Nature* 372, 425–431
2. Gainford, T., Wilson, T., Metcalfe, D., Handman, E., McFarlane, C., Ng, A., Nicola, N. A., Alexander, W. S., and Hilton, D. J. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 14564–14568
3. White, D. W., and Tartaglia, L. A. (1996) *Cytokine Growth Factor Rev.* 7, 303–309
4. White, D. B., and Martin, R. J. (1997) *Proc. Soc. Exp. Biol. Med.* 214, 222–232
5. Coleman, D. L. (1995) *Diabetologia* 9, 294–298
6. Tartaglia, L. A., Dembski, M., Wang, J., Deng, N., Culpepper, J., Devos, R., Richards, G. J., Campfield, L. A., Clark, F. T., Deeds, J., Muir, C., Sanker, S., Moriarty, A., Moore, K. J., Smutko, J. S., Mays, G. G., Woolf, E. A., Monroe, C. A., and Tepper, R. L. (1996) *Cell* 83, 1263–1271
7. Baumann, H., Morella, K. K., White, D. W., Dembski, M., Bailon, P. S., Kim, H., Lafi, C.-F., and Tartaglia, L. A. (1996) *Proc. Natl. Acad. Sci. U. S. A.* 93, 8374–8378
8. Tartaglia, L. A. (1997) *J. Biol. Chem.* 272, 6993–6996
9. White, D. W., Kurapatwinski, K. K., Devos, R., Baumann, H., and Tartaglia, L. A. (1997) *J. Biol. Chem.* 272, 4065–4071
10. Fukunaga, R., Ishizaka-Ikeda, E., Pan, C.-X., Seto, Y., and Nagata, S. (1991) *EMBO J.* 10, 2855–2865
11. Pellemounter, M. A., Cullen, M. J., Baker, M. B., Hecht, R., Winters, D., Boone, T., and Collins, F. (1995) *Science* 269, 540–543
12. DeClerck, Y. A., Yean, T.-D., Ho, S. H., Ting, J., and Langley, L. (1991) *J. Biol. Chem.* 266, 3893–3899
13. Wigler, M., Silverstein, S., Lee, L. S., Pellicer, A., Chang, Y. C., and Axel, R. (1977) *Cell* 11, 223–232
14. Laemmli, U. K. (1970) *Nature* 227, 680–685
15. Agnes, H. (1986) in *Advanced Methods in Microsequence Analysis* (Wittmann-Liebold, B., ed.) pp. 244–255, Springer-Verlag, Berlin Heidelberg
16. Marti, T., Rosselet, S. J., Titani, K., and Walsh, K. A. (1987) *Biochemistry* 26, 8099–8109
17. Haniu, M., Horan, T., Arakawa, T., Le, J., Katta, V., Hara, S., and Rohde, M. F. (1996) *Biochemistry* 35, 13040–13046
18. Miyaizaki, T., Maruyama, Y., Yamada, G., Hatakeyama, M., and Taniguchi, T. (1991) *EMBO J.* 10, 3191–3197
19. Stenflo, J., and Fernlund, P. (1982) *J. Biol. Chem.* 257, 12180–12190
20. Keoh, M., Sibley, C., Fuhrman, J., Schilling, J., and Houd, L. E. (1979) *Proc. Natl. Acad. Sci. U. S. A.* 76, 2932–2936
21. Vance, B. A., Wu, W., Risbood, K. K., Segal, D. M., and Karsee, K. P. (1997) *J. Biol. Chem.* 272, 23117–23122
22. Titani, K., Kumar, S., Takio, K., Ericsson, L. H., Wade, R. D., Ashida, K., Walsh, K. A., Chepko, M. W., Sadler, J. E., and Fujikawa, K. (1986) *Biochemistry* 25, 3171–3184
23. Gavel, Y., and von Heijne, G. (1990) *Protein Eng.* 3, 433–442
24. Rajotte, D., Cadieux, C., Haman, A., Wilkes, B. C., Clark, S. C., Heroe, T., Woodcock, J. A., Lopes, A., and Hoang, T. (1997) *J. Exp. Med.* 185, 1939–1950
25. Kala, M., Montero-Julian, F. A., Grontzinger, J., Weilmm, A., Morelle, D., Brochier, J., Rose-John, S., Heinrich, P. C., Brailly, H., and Content, J. (1996) *Eur. J. Biochem.* 238, 714–723