Proc. Natl. Acad. Sci. USA
Vol. 81, pp. 6486–6490, October 1984
Immunology

Amino acid sequence and post-translational modification of human interleukin 2

(T-cell growth factor/glycosylation/lymphokines)

RICHARD J. ROBB*, RUSTY M. KUTNY†, MARIA PANICO‡, HOWARD R. MORRIS‡, AND VINAY CHOWDHRY†

*Central Research and Development Department, E. I. du Pont de Nemours and Company, Glenolden Laboratory, Glenolden, PA 19036; †Central Research and Development Department, E. I. du Pont de Nemours and Company, Experimental Station, Wilmington, DE 19898; and ‡Department of Biochemistry, Imperial College, London SW7 2AZ, England

Communicated by H. E. Simmons, June 22, 1984

ABSTRACT **Human interleukin 2 was separated into multiple molecular forms by selective immunoaffinity chromatography and chromatofocusing. For the most part, this heterogeneity was attributed to variations in glycosylation of the threonine residue in position 3 of the polypeptide chain. The various molecular forms of interleukin 2 had nearly identical specific activities in the *in vitro* proliferation assay, indicating that the glycosylation had no significant effect on this response. The entire primary sequence of interleukin 2, including the location of the intramolecular disulfide bridge, was determined by a combination of peptide mapping and protein sequencing. This information should aid in the determination of the active site(s) of the molecule.**

Proliferation of T lymphocytes after stimulation by mitogens or antigens involves a hormone-like growth factor termed interleukin 2 (IL2) or T-cell growth factor (1, 2). Based on dose–response curves and inhibition studies employing monoclonal antibodies, the proliferative response appears to be mediated by interaction of the factor with a high-affinity membrane receptor (2). In addition to T-cell proliferation, IL2 has been implicated in other immunological responses (2), and impairment of its secretion is associated with certain diseases (3). The factor was recently purified to homogeneity from several sources (4, 5), making possible a detailed analysis of the relationship between its structure and function.

Biochemical characterization of human IL2 indicated that the molecule is heterogeneous with respect to size and charge (6). On the basis of enzymatic digestion, these differences were attributed to a variable carbohydrate component (6). Microheterogeneity in the amino acid sequence could not be ruled out, however, until the isolation of cDNA coding for IL2 (7). Studies using mRNA derived from normal cells and cell lines indicated a single amino acid sequence as well as a single copy of the structural gene (7, 8). Although such studies provided the primary sequence of the factor, a complete understanding of IL2 structure and function requires knowledge of the nature and location of post-translational modification as well as of the three-dimensional folding of the polypeptide chain. In this paper, we provide a complete characterization of the covalent structure of human IL2, including the position of the intramolecular disulfide bridge.

MATERIALS AND METHODS

Production of IL2. IL2 was prepared from the cell supernatant of a high-producer subclone (J6.8.9.15.32) of the human T-cell line, JURKAT, after induction with phytohemagglutinin and phorbol 12-myristate 13-acetate (5).

IL2 Assay. The bioactivity of IL2 was determined by the IL2-dependent stimulation of proliferation (measured by [^{3}H]thymidine incorporation) of a cloned murine T-lymphocyte line (CTLL-2, subclone 15H). Units of activity were calculated as described (5).

Fractionation of IL2. Crude supernatant from induced JURKAT cells was first passed through a PTHK cassette (100,000 M_r retention, Millipore Pellicon system) to remove large molecular complexes and then concentrated 60-fold using a PTGC cassette (10,000 M_r retention). The concentrated material was passed through a column of CNBr-activated Sepharose 4B coupled to either murine monoclonal IgG2a antibody 1H11-1A5 or murine monoclonal IgG1 antibody 46C8-A2, both of which have the same binding selectivity. About 60% of the IL2 bioactivity bound to the affinity column (Aff-IL2) and could be quantitatively recovered by elution at low pH.

The material that passed through the immunoaffinity column (FT-IL2) was further purified by gel filtration on Sephadex G-100. The pooled bioactivity from the G-100 column (1.4×10^6 units of IL2 activity/155 mg of protein) was divided into aliquots of 20 mg of protein and each aliquot was fractionated by chromatofocusing on a Mono P column (Pharmacia). The starting buffer was 25 mM Tris·HCl, pH 8.8, and the elution buffer was 60 ml of Polybuffer 96 (1:20 dilution) adjusted to pH 7.0. The fractions were tested for bioactivity and three regions (A, B, and C) of activity were pooled (total recovery of activity was ≈90%). Aliquots (0.3–2 mg of protein) of each pool were injected directly onto a 4.6 × 75 mm Ultrapore RPSC column (Altex, Berkeley, CA) and eluted with a 0–60% (vol/vol) gradient of acetonitrile in 0.1% aqueous trifluoroacetic acid.

Solutions of the four forms of IL2 were judged to be free of phytohemagglutinin and phorbol 12-myristate 13-acetate (<10 pM) based on gel electrophoresis data and radioactive tracer analysis (5).

Reduction and Alkylation. Purified IL2 (dissolved in 8M guanidine·HCl) was reduced with dithiothreitol and alkylated with iodoacetic acid as described (5).

Fractionation of Tryptic Peptides. Reduced and alkylated IL2 (0.1–1 mg) was dissolved in 200 μl of 0.1 M ammonium bicarbonate and digested with L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin (United States Biochemicals, Cleveland, OH) for 6 hr at 37°C, using a 1:100 (wt/wt) ratio of trypsin to IL2. The mixture was injected directly onto a 4.6 × 250 mm Ultrasphere-ODS column (Altex) and eluted with a gradient of acetonitrile in 0.1% trifluoroacetic acid.

Sequence Analysis. Intact IL2 and tryptic peptides were sequenced using the gas-vapor sequencer model 470A (Applied Biosystems).

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

Abbreviations: IL2, interleukin 2; Aff-IL2, IL2 that bound to an immunoaffinity column; FT-IL2, IL2 species that passed unretarded through such a column; GalNAc, *N*-acetyl-D-galactosamine; >PhNCS, phenylthiohydantoin.

Table 1. Purification of JURKAT-derived IL2

Sample	Volume, ml	Protein, mg	IL2 activity,* units × 10^{-6}	Recovery, %	Specific activity, units/mg
Cell supernatant	70,500	4900	4.15	100	850
Immunoaffinity					
Bound (Aff-IL2)	30	7.6	2.45	59.0	322,000
Unbound (FT-IL2)	2,000	4300	1.60	38.6	370
CF and RP-HPLC					
FT-IL2-A	29	2.7	0.85	20.6	316,000
FT-IL2-B	11	0.32	0.10	2.5	315,000
FT-IL2-C	16	0.68	0.21	5.1	313,000

CF, chromatofocusing; RP, reversed-phase.
*These values should be multiplied by a factor of 41.5 in order to express the bioactivity in terms of the BRMP IL2 Reference Reagent supplied by the Biological Response Modifiers Program of the National Cancer Institute.

Mass Spectrometry. The fast atom bombardment mass spectrum of tryptic peptides was obtained on a V.G. ZAB high field instrument equipped with an M-Scan fast atom bombardment ion gun. The gun was operated using xenon gas at a beam current of 15 μA at 10 kV. Peptides were dissolved in 5% (vol/vol) acetic acid and 2-μl aliquots (1–5 pmol) were loaded onto the fast atom bombardment target, which had been coated with 3 μl of glycerol. An addition of 0.5 μl of thioglycerol was made just prior to obtaining spectra. Mass spectra were recorded up to a mass of 3200 Da at full accelerating voltage.

RESULTS

Fractionation of JURKAT-Derived IL2. Previous studies had shown that JURKAT-derived IL2 could be separated into multiple molecular forms based on differences in isoelectric focusing behavior (6) and reactivity with a murine monoclonal antibody (5). To determine the nature of this heterogeneity, crude JURKAT-derived IL2 was first fractionated on the immunoaffinity column. This was followed by chromatofocusing of the IL2 remaining in the column effluent.

When the crude JURKAT cell supernatant was passed over the immunoaffinity column, about 60% of the bioactivity was bound. This bound material (designated Aff-IL2) was recovered by elution at low pH (Table 1). It appeared homogeneous on two-dimensional gel electrophoresis and reversed-phase HPLC (5). Chromatofocusing of the material that did not bind to the immunoaffinity column (FT-IL2) separated the remaining bioactivity into one major and two minor peaks (Fig. 1*a*). IL2 in each of these peaks (designated FT-IL2-A, -B, and -C) was purified to homogeneity by reversed-phase HPLC (Fig. 1*b*).

$NaDodSO_4$/PAGE of Aff-IL2 and HPLC-purified FT-IL2-A, -B, and -C showed that, although each preparation exhibited a single silver-stained band, the relative mobilities of the bands varied (data not shown). The M_r values for Aff-IL2 and FT-IL2-A, -B, and -C were 15,900, 15,600, 15,400, and 17,200, respectively. Treatment of FT-IL2-C with neuraminidase reduced its M_r to ≈16,100, indicating that much of its apparent size disparity was due to covalently linked sialic acid.

The structural heterogeneity evident upon electrophoresis appears to play little, if any, role in the molecule's bioactivity since the specific activities of the four forms of IL2 in the *in vitro* proliferation assay (using murine or human T cells as targets) differed by <5% (Table 1). All of the four forms of IL2 were also equally effective on a weight basis in maintaining the continuous proliferation of the IL-2-dependent murine CTLL-2 cell line and all had similar affinities (to within 10%) for the high-affinity cellular receptor (2) on human lymphoblasts (data not shown).

Variations in Post-translational Modification. Fig. 2*a* shows the elution profile obtained by reversed-phase HPLC of the tryptic peptides of Aff-IL2. Identical profiles were obtained for the peptides of FT-IL2-A, -B, and-C, with the exception that the peptide designated T-2 varied slightly in its elution position (Fig. 2*b*). Amino acid compositions of T-2 from Aff-IL2, FT-IL2-A, and FT-IL2-C were identical and indicated that T-2 is the NH_2-terminal octapeptide (9). T-2 from FT-IL2-B, in contrast, contained very little alanine and proline, which are normally found at the NH_2 terminus of the molecule (5). The coincidence of all but the T-2 peaks on the HPLC profiles indicates that the differences between the four forms of IL2 are confined to the first eight amino acid positions in the molecule.

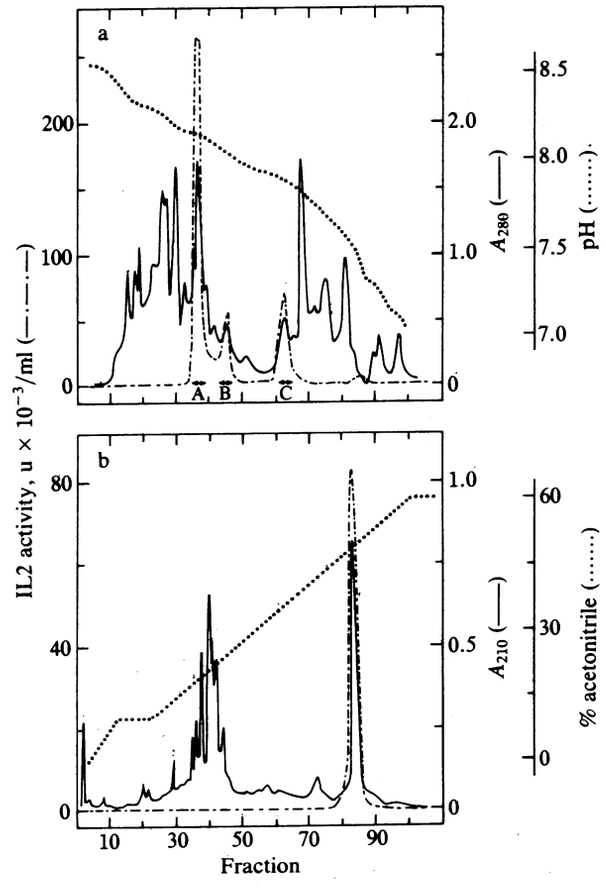

FIG. 1. (*a*) Separation of FT-IL2-A, -B, and -C by chromatofocusing on a Mono P column. Peak fractions pooled are indicated by the double-headed arrows. (*b*) Typical HPLC purification of an FT-IL2 component obtained by chromatofocusing (in this case, FT-IL2-A). u, Units.

Sequencing of the four forms of intact IL2 revealed variability in the HPLC mobility of the phenylthiohydantoin (>PhNCS) derivative of the amino acid at position 3 of the polypeptide chain (data not shown). The cycle 3 >PhNCS derivative from FT-IL2-A corresponded to an unmodified threonine while that from cycle 3 of Aff-IL2 migrated in a position previously attributed to the *N*-acetylgalactosamine (GalNAc) derivative of threonine (9). The cycle 3 >PhNCS-amino acid from FT-IL2-C did not correspond to either of these positions. It therefore was some other derivative of threonine since the FT-IL2-C NH_2-terminal octapeptide (T-2) contained two threonines (positions 3 and 7). The amino acids in the other seven NH_2-proximal positions were the same for these three forms of IL2 and confirmed our previous result obtained with Aff-IL2 (refs. 5 and 9; see Fig. 4). In contrast, the NH_2-terminal sequence of FT-IL2-B began with an unmodified threonine, followed by Ser-Ser-Ser-Thr-Lys.

During fast atom bombardment mass spectrometry of the NH_2-terminal octapeptides from the various forms of IL2, an abundant quasi-molecular ion signal was observed at *m/e* 778 for FT-IL2-A and at *m/e* 981 for Aff-IL2. These values

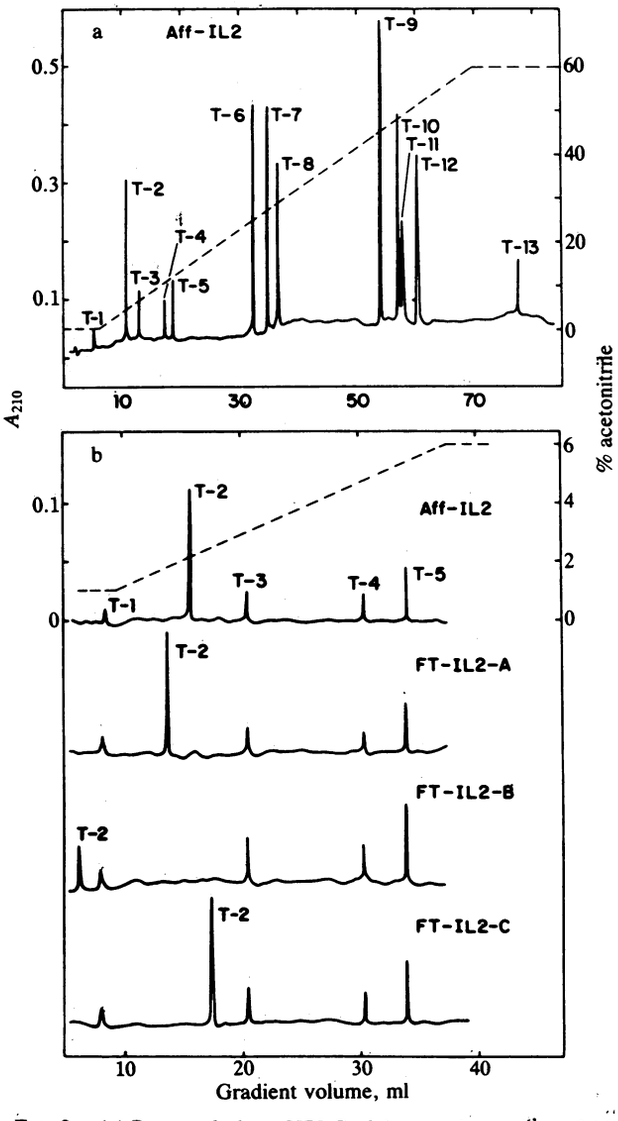

FIG. 2. (*a*) Reversed-phase HPLC of the tryptic peptides of Aff-IL2. (*b*) Comparison of the elution positions of peptide T-2 (residues 1–8) from the various forms of IL2. The acetonitrile gradient used for elution is shown by the broken line.

corresponded to those expected for an unmodified and GalNAc-modified octapeptide, respectively. The spectrum for the T-2 peptide of FT-IL2-C corresponded to that expected from an octapeptide containing *N*-acetylhexosaminylhexosyl-(*N*-acetyl)neuraminic acid (Fig. 3). The hexosamine was identified as galactosamine by procedures previously described (9) and the hexose was tentatively identified as galactose by retention time on gas-liquid chromatography. The mass spectrometric and sequence data suggest that the neuraminic acid is attached to the galactose, which is in turn linked to the GalNAc on the threonine at position 3 of FT-IL2-C. The latter linkage is most likely β-D-Gal(1→3)-D-GalNAc since ≈90% of neuraminidase-treated FT-IL2-C bound to peanut agglutinin-coupled beads, which are quite selective for this configuration.

An earlier study suggested that glycosylation of Thr-3 was involved in recognition of IL2 by the 1H11-1A5 antibody (9). Indeed, the Aff-IL2 NH_2-terminal octapeptide with the GalNAc modification was as effective on a molar basis as intact IL2 in competing with the binding of radiolabeled IL2 to antibody-coupled Sepharose beads. In contrast, the NH_2-terminal octapeptides derived from FT-IL2-A and -C were at least 300-fold less effective on a molar basis in blocking the binding of intact IL2. This result confirms the indication that an unmasked GalNAc residue on threonine at position 3 contributes to high-affinity binding to the 1H11-1A5 antibody.

Amino Acid Sequence Analysis. The complete sequence (Fig. 4) of IL2 was derived from peptide fragments of Aff-IL2 (Fig. 2*a*). The peptides were aligned by comparison with the sequence expected from the cDNA data (7) and no discrepancy with the nucleotide-derived information was found. Examination of the amino sugar content of the intact protein molecule indicated a single galactosamine residue, suggesting that position 3 was the only site of such glycosylation in Aff-IL2. This conclusion was substantiated by mass spectrometric analyses of tryptic peptides T-3, T-5, T-6, T-7, T-8, T-9, T-10, T-11, and T-12 in which the quasi-molecular ion signals were in agreement with the masses expected for unmodified or carboxymethylated (T-9, T-10, and T-12) peptides.

Location of the Intramolecular Disulfide Bridge. JURKAT-derived IL2 contains three cysteine residues, two of which form an intramolecular disulfide bond (5). Treatment of IL2 with 5 mM dithiothreitol resulted in a 70% loss of bioactivity in a physiological buffer (10 mM Tris·HCl, pH 7.5/0.14 M NaCl) and a >95% loss in a denaturing solution (0.2% $NaDodSO_4$). In contrast, alkylation of the free cysteine with iodoacetic acid resulted in <10% loss of activity. To determine which of the cysteines formed the critical disulfide bond, the incorporation of iodo[^{3}H]acetic acid into the tryptic peptides of Aff-IL2 was determined with and without prior reduction. The results showed that alkylation of Cys-125 (tryptic peptide T-9, Fig. 2*a*) was unaffected by reduction while alkylation of Cys-58 and Cys-105 (tryptic peptides T-12 and T-10) was markedly diminished in the absence of the reducing agent (Fig. 5). These results are consistent with an intramolecular disulfide bridge connecting Cys-58 and Cys-105.

DISCUSSION

In this paper we describe the complete primary structure of human IL2. The sequence of amino acids was that expected from the cDNA data (7). In addition, the intramolecular disulfide bridge was found to link the cysteines in positions 58 and 105. Post-translational modification of the JURKAT-derived molecule was largely restricted to variations in the *O*-linked glycosylation of threonine in position 3. Such modifications had little effect on the *in vitro* specific activity of the

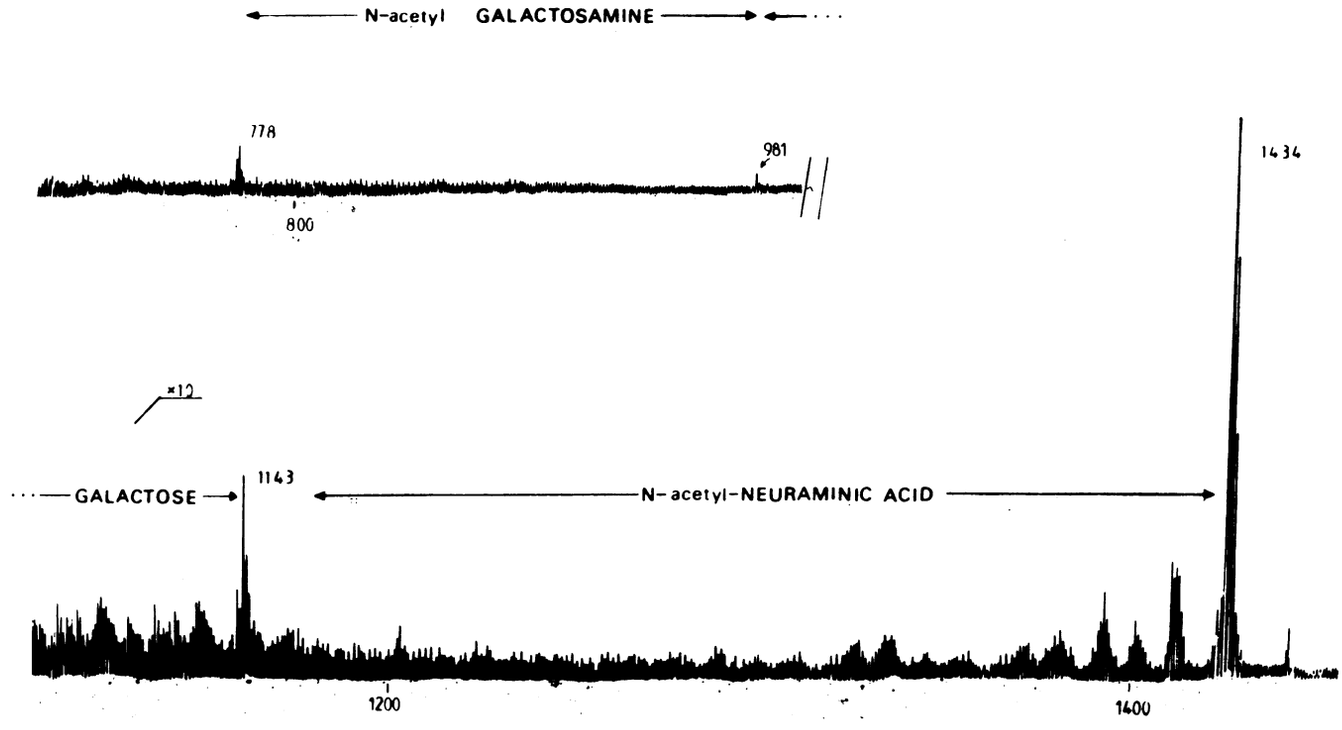

FIG. 3. Mass spectrum of the NH_2-terminal octapeptide (T-2) of FT-IL2-C indicating the quasi-molecular ion (*m/e*, 1434) and several fragments. The successive attachment of GalNAc, galactose, and *N*-acetylneuraminic acid to the basic octapeptide structure (*m/e*, 778) is consistent with the mass differences between the progressively larger fragments, as indicated by the double-headed arrows. Note that no signal was recorded corresponding to a fragment composed of the octapeptide, GalNAc, and *N*-acetylneuraminic acid (*m/e*, 1272), even though that region of the spectrum was recorded at a greater (10×) sensitivity. This finding suggests that the neuraminic acid is linked to the galactosyl-*N*-acetylgalactosaminyl-peptide through galactose rather than through GalNAc.

factor and the binding of IL2 to its cellular receptor and therefore appear to be unimportant for bioactivity. Two lines of evidence indicate an absence of additional modifications to the molecule. First, no aberrant PhNCS-derivative other than that of Thr-3 was encountered, and second, mass spectrometry of tryptic peptides failed to detect evidence of any other alteration. Thus, the marked hydrophobicity of IL2 appears to be an intrinsic characteristic of its amino acid sequence and does not involve modifying agents such as fatty acids as was proposed (4).

Although variations in glycosylation were the principal source of heterogeneity for JURKAT-derived IL2, a small fraction of the growth factor (FT-IL2-B) lacked the alanine and proline residues normally found at positions 1 and 2 of the polypeptide chain. Moreover, during chromatofocusing, another minor form (<2%) of the molecule was occasionally observed as a shoulder on the high pI side of the FT-IL2-A peak. This form of IL2 contained proline at position 1, followed by Thr-Ser-Ser-Ser-Thr. These minor, truncated forms of the molecule may result from variations in the removal of the signal peptide or from nonspecific proteolytic degradation.

The fact that only a single structural gene for human IL2 has been found in various cell lines as well as normal cells (7, 8) indicates that, apart from possible allelic differences, post-translational modification is the source of the observed

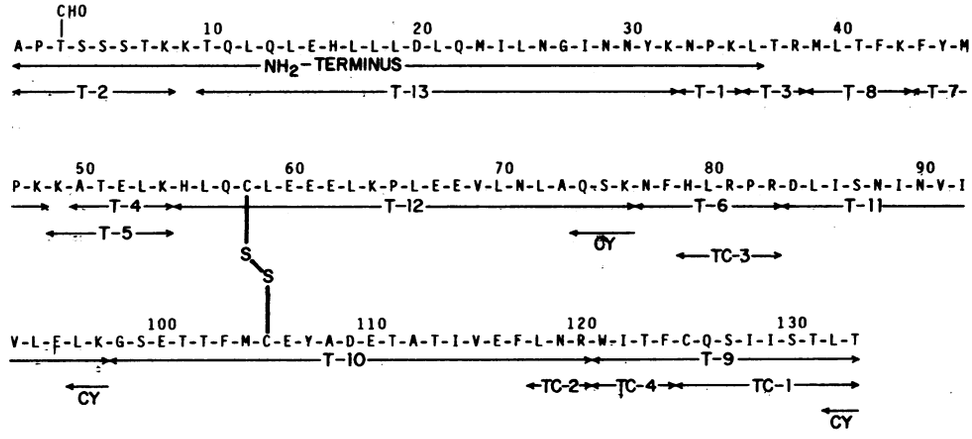

FIG. 4. Amino acid sequence of Aff-IL2. The double-headed arrows designate peptides used during sequencing. The T series (see Fig. 2) denotes tryptic peptides while the TC series denotes peptides that were fortuitously derived using a preparation of trypsin that apparently retained some chymotryptic activity. The designation CY and the single-headed arrows denote information derived from carboxypeptidase Y digestion. Each of the 133 positions in the sequence was determined using the intact molecule or one or more of the relevant peptides.

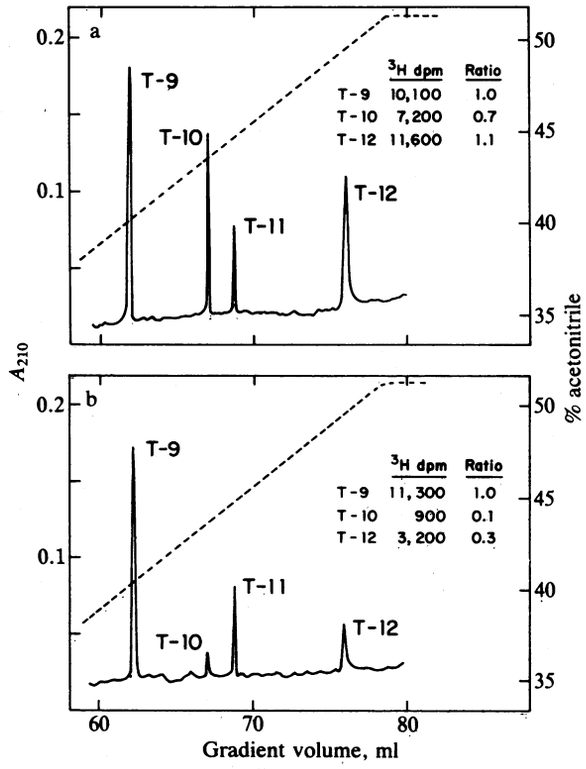

FIG. 5. A portion of the HPLC profile of tryptic fragments of Aff-IL2 alkylated by iodo[^{3}H]acetic acid with (*a*) and without (*b*) prior reduction with dithiothreitol. Ratios of the radioactivity recovered in peptides T-9, T-10, and T-12 are given.

charge and size heterogeneity of the molecule (6). The present results thus provide an explanation of such heterogeneity at the molecular level. JURKAT-derived IL2 consisted mainly of a mixture of unmodified polypeptide (FT-IL2-A) and the GalNAc-derivative (Aff-IL2). In addition, a small amount of a hexosyl-sialylated form (FT-IL2-C) was isolated. Such sialylated forms of IL2 were much more prominent when the factor was derived from human tonsil cells. Tonsil-derived IL2 was separated into three major and one minor peaks of bioactivity by isoelectric focusing (6). The position of one major peak corresponded to that of both JURKAT-derived FT-IL2-A and Aff-IL2. A second major peak, which consisted of IL2 containing one residue of sialic acid, migrated in a position corresponding to JURKAT FT-IL2-C, while the third major peak, which consisted of IL2 with two residues of sialic acid, was rarely observed in the JURKAT-derived preparation (6). The minor peak of tonsil-derived material migrated in a position corresponding to JURKAT FT-IL2-B. In all likelihood, the heterogeneity in the major forms of tonsil-derived IL2 is also attributable to variations in glycosylation at Thr-3.

The signal for *O*-linked glycosylation is contained in the secondary and tertiary structure of a polypeptide chain and is often associated with clustered, exposed regions of hydroxyl amino acids containing adjacent proline residues (10). The threonine in position 3 of human IL2 fits these characteristics. Synthesis of the *O*-linked chains occurs in the Golgi apparatus (10) and begins with the enzymatic transfer of GalNAc from UDP-GalNAc. Following this reaction, galactose and sialic acid are frequently added to GalNAc. FT-IL2-A, Aff-IL2, and FT-IL2-C thus represent the starting material, intermediate, and final product of this reaction sequence. Although the variability in the extent of glycosylation of IL2 may reflect a failure of some molecules to complete the sequence of modification, it is possible that some carbohydrate is enzymatically removed during the period of IL2 secretion or during storage and processing of the cell supernatants. In this regard, a significant portion (25–35%) of the cells die during stimulation, a circumstance that could result in either incomplete synthesis of the carbohydrate chain or its partial enzymatic removal.

The presence of carbohydrate on human IL2 appears to have no effect on its ability to promote T-cell proliferation. Nevertheless, the extent and nature of the glycosylation may affect the rate of clearance of the molecule *in vivo*. Another potential problem associated with variable glycosylation is that patients receiving huge quantities of such IL2 forms as Aff-IL2 or nonglycosylated, bacteria-derived IL2 may make antibodies to the molecules. The fact that T lymphocytes normally release a mixture of glycosylated forms *in vitro* (6), however, may mitigate this particular complication. Unfortunately, it is not possible with the sensitivity of present techniques to ask whether endogenous IL2 released *in vivo* also consists of such a mixture.

The structural characterization of IL2 presented here, including the positions of the disulfide bridge and of the attachment site for carbohydrate, should prove useful in determining the three-dimensional structure of the molecule. Such data will provide a basis for designing experiments to determine the region(s) of IL2 responsible for interaction with its cellular receptor.

The authors acknowledge the excellent technical assistance of L. Gehman, B. Krieger, P. Webber, T. Miller, and F. Kitzen and are grateful to the Medical Research Council (Great Britain) for support of the mass spectrometry facility.

1. Smith, K. A. (1980) *Immunol. Rev.* **51,** 337–357.
2. Robb, R. J. (1984) *Immunol. Today* **5,** 203–209.
3. Altman, A., Theofilopoulos, A. N., Weiner, R., Katz, D. H. & Dixon, F. J. (1981) *J. Exp. Med.* **154,** 791–808.
4. Henderson, L. E., Hewetson, J. F., Hopkins, R. F., III, Sowder, R. C., Neubauer, R. H. & Rabin, H. (1983) *J. Immunol.* **131,** 810–815.
5. Robb, R. J., Kutny, R. M. & Chowdhry, V. (1983) *Proc. Natl. Acad. Sci. USA* **80,** 5990–5994.
6. Robb, R. J. & Smith, K. A. (1981) *Mol. Immunol.* **18,** 1087–1094.
7. Taniguchi, T., Matsui, H., Fujita, T., Takaoka, C., Kashima, N., Yoshimoto, R. & Hamuro, J. (1983) *Nature (London)* **302,** 305–310.
8. Devos, R., Plaetinek, G., Cheroutre, H., Simons, G., Degrave, W., Travernier, J., Remaut, E. & Fiers, W. (1983) *Nucleic Acids Res.* **11,** 4307–4323.
9. Robb, R. J., Kutny, R. M., Panico, M., Morris, H., DeGrado, W. F. & Chowdhry, V. (1983) *Biochem. Biophys. Res. Commun.* **116,** 1049–1055.
10. Kornfeld, R. & Kornfeld, S. (1976) *Annu. Rev. Biochem.* **45,** 217–237.