Journal of Chromatography, 562 (1991) 403–419
Biomedical Applications
Elsevier Science Publishers B.V., Amsterdam

CHROMBIO. 5569

Mass spectrometric determination of glycosylation sites and oligosaccharide composition of insect-expressed mouse interleukin-3

MICHAL SVOBODA and MICHAEL PRZYBYLSKI*

Fakultät für Chemie, Universität Konstanz, Postfach 5560, D-7750 Konstanz (F.R.G.)

JOLANDA SCHREURS and ATSUSHI MIYAJIMA

Department of Molecular Biology, DNAX Research Institute of Molecular and Cellular Biology, 901 California Avenue, Palo Alto, CA 94304 (U.S.A.)

and

KENNETH HOGELAND and MAX DEINZER*

Department of Agricultural Chemistry, Oregon State University, Corvallis, OR 97331 (U.S.A.)

ABSTRACT

The primary structure of Baculovirus-expressed mouse interleukin-3 produced in infected *Bombyx mori* larvae was characterized by liquid secondary ion mass spectrometry and ^{252}Cf-plasma desorption mass spectrometry in combination with selected protein microchemical reactions. Interleukin-3 was found to consist of at least two glycoprotein species of *ca.* 17 000 dalton. Characterization of tryptic and *S. aureus* V8 protease peptides by Edman degradation combined with plasma desorption mass spectrometry showed that two N-glycosylation sites, Asn-16 and Asn-86, were present. N-Glycan residues were shown by liquid secondary ion mass spectrometry and high-performance liquid chromatography to consist of mannose, fucose, and glucosamine. The presence of galactosamine indicated that O-glycosylated residues were present, in addition to the N-glycosylated residues. Glucose was also present, which indicated incomplete processing of the insect-expressed N-linked oligosaccharides.

INTRODUCTION

Interleukin-3 (IL-3), a T-cell- and mast-cell-derived factor [1,2], supports the growth and differentiation of a diverse array of hemopoietic and lymphoid cells, including mast cells, macrophages, megakaryocytes, eosinophils, erythrocytes, pre-B cells, and possibly pre-T cells [1,3,4]. Its ability to promote the proliferation of progenitor cells and multipotential stem cells makes it a primary regulator of the early stages of hemopoiesis and lymphopoiesis, particularly during the inflammatory process [5].

IL-3 was initially purified from WEHI-3B conditioned cell supernatants [6,7] as a heterogeneous species with an apparent molecular mass of 28–33 kilodalton

0378-4347/91/$03.50

(kDa), as shown by sodium dodecyl sulphate–polyacrylamide gel electrophoresis (SDS-PAGE). Characterization of murine IL-3 [8] expressed from a cDNA clone in monkey cells demonstrated that a 26-amino-acid N-terminal signal peptide is removed, resulting in a secreted protein (Fig. 1) with a molecular mass of 15 674 Da. Post-translational glycosylation appears to differ between various expression systems. Baculovirus-expressed IL-3 produced in silkworms had an apparent molecular mass of *ca.* 20 kDa [5], whereas L cell and C-127 fibroblast-derived IL-3 had an apparent molecular mass of 23–27 kDa [9]. In spite of these differences the biological activities *in vitro* of the different recombinant and natural IL-3s so far appear to be qualitatively and quantitatively similar. However, no comparative studies of the different forms have yet been made *in vivo*.

The recent application of a variety of lymphokines, including IL-3, GM-CSF, and IL-2, as therapeutic agents in humans requires a detailed knowledge of lymphokine structure and post-translational modifications. Structures different from the natural product may result in different pharmacokinetics, cause neutralizing antibody production, or change the receptor binding characteristics [10,11].

The recent development of efficient ionization and desorption methods in mass spectrometry has enabled the direct determination of molecular mass and structural analysis of oligo- and polypeptides [12]. Two of the most successful methods are liquid secondary ion mass spectrometry (LSIMS), *i.e.* bombardment of a homogeneous sample solution with primary ions [13], and ^{252}Cf-plasma desorption mass spectrometry (PDMS) with MeV primary particle beams [14]. Molecular ions for compounds as large as small proteins have been obtained [14,15]. The particular capability of both PDMS and LSIMS to analyse complex peptide

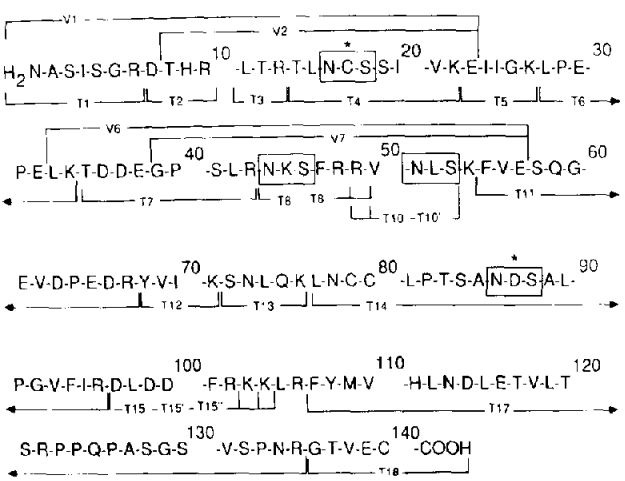

Fig. 1. Interleukin-3 protein sequence (without signal sequence), tryptic (T) and *S. aureus* (V) peptides and potential N-glycosylation sites.

mixtures has been used in the specific molecular mass determinations of proteolytic digest mixtures (peptide mapping) [16]. Information to be obtained by peptide-mapping techniques on protein structures include the specific molecular mass of glycopeptide fragments, glycosylation sites and mass of carbohydrate residues, and the presence and type of other potential post-translational modifications, such as the identification of disulphide linkages. In this study, we have applied LSIMS and PDMS in combination with peptide-mapping procedures to study the amino acid sequence, glycosylation pattern and carbohydrate structure of IL-3 produced by the silkworm, *Bombyx mori*, infected with a recombinant baculovirus.

EXPERIMENTAL

Chemicals and reagents

Trypsin from bovine pancreas, TPCK-treated, was obtained from Sigma (Steinheim, F.R.G.). Nanograde trifluoroacetic acid (TFA) was obtained from Merck (Darmstadt, F.R.G.). Phenylisothiocyanate (PITC) was obtained from Aldrich (Milwaukee, WI, U.S.A.) and pyridine was obtained from Baker (Deventer, The Netherlands).

Preparation, isolation and purification of interleukin-3

The silkworm-derived IL-3 was produced by infection of silkworm larvae with a recombinant baculovirus as previously described [17]. The hemolymph was aspirated and frozen at −70°C.

IL-3 was purified by a two-step process of affinity purification and C_8 reversed-phase chromatography. In the first step, the monoclonal anti-IL-3 antibody, 8F8.1 [18], was conjugated to Affi-Gel Hz Hydrazide (Bio-Rad), according to the manufacturer's directions. Briefly, 10 mg in 5 ml of the monoclonal antibody was dialysed against 0.1 *M* sodium acetate (pH 5.5)–0.15 *M* NaCl (buffer 1) and then treated with sodium periodate at a final concentration of 10 mg/5 ml for 1 h in the dark. The reaction was stopped by adding enough glycerol to achieve a final concentration of 10 m*M*. The antibody was then dialysed overnight against buffer 1 and incubated on a rotator for 10 h at room temperature with 2 ml of the hydrazide resin. The resin was washed with one column volume of 0.5 *M* NaCl (pH 7.0) followed by several washes with phosphate-buffered saline (PBS). Then 1 ml of the silkworm hemolymph was added to the resin, incubated overnight at 4°C on a rotator, then transferred to a small Bio-Rad open column and washed with 40 ml of PBS. The effluent was monitored for protein content by a mini Bio-Rad protein assay, and when no further detectable protein came off the column, a series of acid washes was used. The column was washed sequentially with 10 ml of the following citrate–phosphate buffers: pH 6.0, pH 5.0, and pH 4.0. IL-3 was eluted with 1-ml aliquots of citrate–phosphate buffer (pH 3.0). Biological activity was monitored using the MC/9 cells in the MTT assay devel-

oped by Mosmann [19]. Samples were pooled, concentrated with a Centricon 10 (Amicon), treated with 0.1% TFA, and then applied to a 10 × 0.46 cm I.D. C_8 reversed-phase column (Pharmacia). The sample was eluted using a 0.1% TFA, shallow 29–37% acetonitrile gradient [17], and was shown by SDS-PAGE to consist of two predominant bands of *ca.* 21 and 20 kDA. Several minor species were observed at 17–19 kDa. The aliquots of purified IL-3 were lyophilized and frozen at −70°C until further analysis.

Digestion and isolation of IL-3 and preparation of peptides

Trypsin digestion of IL-3. To 23 μg of IL-3 in 4.5 μl of 50 m*M* NH_4HCO_3 buffer, 0.5 μl of a solution of trypsin (1 μg/μl) in 0.1% aqueous TFA was added. The enzyme-to-substrate ratio was 1:50 (w/w). The solution was incubated at 30°C, and 1-μl aliquots were taken after 0.25, 0.75, 2, and 20 h for PDMS analysis. The peptide composition did not vary with time; only the relative concentrations changed.

Reduction and carboxymethylation of IL-3. IL-3 (96 μg) was dissolved in 0.1 ml of reduction and denaturation buffer solution [6 *M* GuHCl, 50 m*M* ammonium bicarbonate, 1 m*M* EDTA, 10 mg of dithiothreitol (DTT)]. The alkylation reaction was carried out with 4.9 mg of iodoacetamide at 25°C in the absence of light and under a nitrogen atmosphere. After 30 min the reaction was terminated by addition of 3 μl of 2-mercaptoethanol. The reaction mixture was desalted by depositing the solution onto a microdialysis membrane (Millepore VS, 0.025 μm), which had been equilibrated for 1 h with 0.1% TFA, then transferred to an Eppendorf tube and lyophilized.

Digestion of carboxymethylated IL-3 with trypsin. The carboxymethylated IL-3 (48 μg) was dissolved in 50 m*M* ammonium bicarbonate (pH 8) solution. After the addition of 0.1% aqueous TFA (1:10 enzyme-to-substrate ratio) the solution was incubated with trypsin for 4 h at 37°C and then lyophilized. The residue was taken up in 24 μl of 0.1% TFA, and 1 μl of this solution was added to a nitrocellulose target for analysis by PDMS.

Digestion of IL-3 with S. aureus V8-protease. IL-3 (15 μg) was dissolved in 15 μl of ammonium bicarbonate buffer (pH 8). The V8-protease was dissolved in Milli-Q water, and an amount of this solution was added to the IL-3 solution sufficient to give a 1:10 enzyme-to-substrate ratio. The digestion was carried out at 37°C for 48 h. The digest was lyophilized, the residue was dissolved in 7 μl of 0.1% aqueous TFA, and 1 μl of the solution was deposited on a nitrocellulose target for analysis by PDMS.

Combined Edman–PDMS sequence analysis

A fresh 5% solution of phenylisothiocynate (PITC) in pyridine–water (1:1, pH 8.5) was used for each Edman coupling–hydrolysis cycle. Proteolytic peptides isolated by HPLC fractionation were dissolved in 40 μl of PITC solution. After the addition of 40 μl of water, the coupling reaction was carried out at 25°C for 45

min and the sample was lyophilized. The residue was dissolved in 30 μl of 0.1% TFA, and 2 μl of this solution were placed on a nitrocellulose target for analysis by PDMS. The remaining sample was lyophilized. To the residue were added 40 μl of 100% TFA. After 15 min, the TFA was carefully removed using a stream of nitrogen. The residue was again dissolved in 30 μl of 0.1% TFA solution and, without carrying out the conversion reaction, 2 μl of this solution were placed on a nitrocellulose target and analysed directly by PDMS. This cycle was repeated for sequential analyses of the N-terminal amino acids.

Isolation of tryptic peptides

The HPLC system used in the isolation of tryptic peptides consisted of a modified Waters Model 244 Chromatograph with binary gradient Waters 6000A and Waters M-45 high pressure pumps. The detector was a multiwavelength Waters 490E system set at wavelengths of 220 nm, 250 nm, and 280 nm. A Waters Maxima 820 data system was used for data reduction. A reversed-phase 25 cm × 4 mm I.D. column, packed with C_{18}-Nucleosil (7 μm particle size), was used. The solvents consisted of (A) 0.04% aqueous TFA (pH 2) and (B) acetonitrile (Merck) with 0.03% aqueous TFA. The tryptic digest solution was membrane-filtered (Schleicher & Schuell, 0.4 μm pore size) and then degassed by sonication. The elution gradient was as follows: 0–10 min, 100% A; 10–60 min, 100% A to 30% A; 60–70 min, 30% A to 0% A. The peptides were collected with an automatic fraction collector (Gilson) in micro-Eppendorf cups. The sample fractions were lyophilized, and the residue was dissolved in 10 μl of ammonium bicarbonate (pH 8) solution. From each sample, 1 μl of solution was deposited onto a nitrocellulose target and analysed by PDMS.

Californium-252 PDMS of IL-3

Spectra were obtained on a Bioion 20 PDMS from Bio-Ion Nordic (Uppsala, Sweden). The desorption source was a ^{252}Cf foil, which produced fission fragments equivalent to *ca.* 1000 start pulses per second [20]. The targets consisted of metal rings of 13-mm diameter covered with an aluminium mylar foil, onto which a nitrocellulose surface was electrosprayed as previously described [21].

A sample of IL-3 for molecular mass determination was prepared by dissolving 23 μg (*ca.* 14 nmol) of IL-3 in 10 μl of 0.1% aqueous TFA. A 2-μl aliquot of this solution was deposited on the PDMS target. The acceleration voltage was set at 17 kV, and the analysis was carried out for $30 \cdot 10^6$ counts.

Carbohydrate determination

Glassware was silanized by treatment with a 2% solution of dimethyldichlorosilane in toluene and subsequently rinsed with methanol, warm water and deionized water. IL-3 (120 μg) was treated with a solution of glacial acetic acid–acetic anhydride–sulphuric acid (10:2:1) in a 5-ml conical vial. This was mixed and heated at 60°C in an oven for 6 h. The reaction mixture was cooled and quenched

with 1 ml of water and then extracted with two 1-ml portions of chloroform. The organic layer was transferred to a 5 ml conical vial and washed with water. The organic layer was carefully concentrated under nitrogen, and the concentrate then was taken up in methanol for analysis of the acetylated sugar fragments by LSIMS.

LSIMS analysis was carried out on a Kratos MS-50 mass spectrometer. Xenon was used to generate the primary ionizing beam from an Ion-Tech gun operated at 8 keV. Samples in a glycerol–thioglycerol (1:1, v/v) matrix were loaded on the FAB target.

IL-3 (30 μg) in 300 μl of water was treated with 200 μl of 2 *M* TFA. The tube was purged with nitrogen, sealed, and placed in an oven at 105°C for 4.5 h. The sample was cooled, concentrated, washed with water, and then analysed by HPLC. The HPLC system consisted of a Dionex BioLC with an anion exchange/pulsed amperometric detector (AE-PAD) and a 250 × 4 mm I.D. Carbo-Pac PA1 column and 25 × 3 mm I.D. Carbo-Pac PA guard column. The carbohydrates were eluted isocratically with water–100 m*M* NaOH (90:10), at a flow-rate of 1 ml/min. Post-column addition of 300 m*M* NaOH at 1 ml/min was used to meet detector requirements.

RESULTS

Molecular mass determination

Analysis of IL-3 by PDMS showed two peaks at 5498 and 5830 Da for the triply charged molecular ions, two at 8270 and 8688 Da for the doubly charged ions, and two at 16 772 and 17 394 Da for the protonated parent ions (Fig. 2). The isotopically averaged masses derived from these molecular ions yielded molecular masses of 16 601 and 17 419 Da, respectively. Based on a core protein of

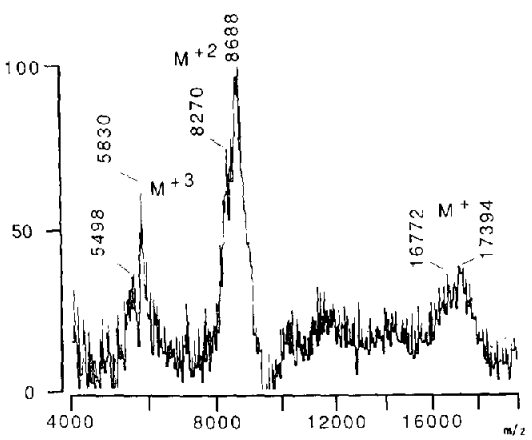

Fig. 2. ^{252}Cf-Plasma desorption mass spectrum of interleukin-3 (4.6 μg; 270 pmol); start counts, 30 · 10^6; accelerating voltage, 17 kV.

mass 15 674 Da [8], molecular masses for two carbohydrates of 928 and 1746 Da were estimated. Thus, the molecular mass of IL-3 as determined by PDMS differs from previous estimates made by SDS-PAGE, and reinforces the use of absolute measurements as a key in obtaining molecular masses closer to those predicted.

Mass spectrometric peptide-mapping of tryptic digests

The digestion of non-reduced IL-3 by trypsin yielded a mixture that could be analysed directly by PDMS. Various peaks were present in the spectra in the range 400–3279 Da (Fig. 3). Eighteen possible tryptic peptides with at least three amino acids (Table I) were expected from complete digestion. Several cumulative digestion sites leading to multiple cleavages were observed in the protein sequence. The corresponding peptides are indicated by T′ and T″ (Table I), and contain one or two adjacent arginine or lysine residues.

A high mass peak with high intensity is observed with *m/z* 3215. This peak is due to the protonated T17 peptide, whose calculated mass is 3214 Da. Prominent peaks in the spectrum with *m/z* 416 (T16′) (not shown in Fig. 3), 522 (T12), 528 (T2), 545 (T16), 566 (T9′), 782 (T15), 991 (T7) and 1507 (T11) were assigned to the protonated forms of several predicted peptides (Table I). Similarly, protonated tryptic peptides T15 (*m/z* 782) and T15′ (*m/z* 910) appear to be present in the digest. In several cases, however, unambiguous assignment could not be made between two possible peptides because of identical nominal masses. Peptides T1 and T13, for instance, have calculated masses of 590 Da and 589 Da, respectively. The peak with *m/z* 591 would better fit protonated T1, but a difference of only 1 amu between the two peptides prevents unambiguous assignment. T1 is part of the silkworm IL-3 core, as the amino terminal sequence analysis showed it to be Ala-Ser-Ile [5,22]. Earlier results of WEHI-3-derived IL-3 [23] indicated that the first six amino acids, *i.e.* T1, were missing from the core protein, but it was not

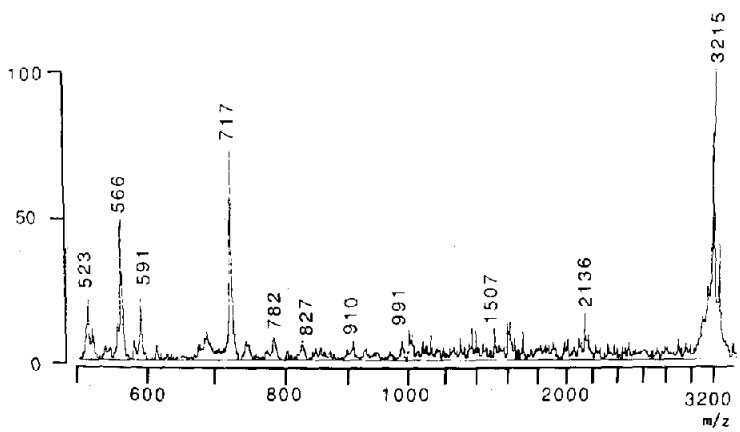

Fig. 3. ^{252}Cf-Plasma desorption mass spectrum of a tryptic digest of interleukin-3 (4.6 μg; 270 pmol) after 2-h digestion. Accelerating voltage, 15 kV, 10 · 10^6 counts. (Masses assigned by PDMS data system.)

TABLE I

PREDICTED AND EXPERIMENTAL MASSES OF TRYPTIC PEPTIDES OF INTERLEUKIN-3

Peptide	Sequence	Molecular mass	MH^+ in PD spectra
T1	1–6	590	591
T2	7–10	527	528
T3	11–13	388	Target background
T4	14–22	964	[a]
T5	23–27	559	561
T6	28–34	825	827
T7	35–43	989	991
T8	44–48	651	Mult. digestion sites
T8′	44–49	807	Mult. digestion sites
T9	46–48	408	Target background
T9′	46–49	565	566
T10	49–54	716	717
T10′	50–54	560	561
T11	55–67	1507	1507
T12	68–71	522	523
T13	72–76	589	591
T14	77–96	2091	[a]
T15	97 102	780	782
T15′	97–103	908	910
T15″	97–104	1036	–
T16	103–106	544	545
T16′	104–106	415	416
T17	107–135	3214	3215
T18	136–140	508	–
Other	49–76	3279	3279
	120–135	1639	1640[b]

[a] Glycosylated.

[b] Presence shown by Edman degradation in F6 (Fig. 5).

clear whether further processing had occurred physiologically or whether T1 had been removed during purification [22].

Another pair of peptides that cannot be readily distinguished in the protonated forms are T5 (calculated mass of 559 Da) and T10′ (calculated mass of 560 Da). A peak in the spectrum with m/z 561 fits protonated T10′ exactly. Peptides T8 and T8′ with calculated masses of 651 Da and 807 Da were not observed. These peptides would be expected only on incomplete digestion, as there is overlap with the expected T9 and T9′ peptides.

Peaks in the spectrum in the range m/z 390–410 cannot be assigned to tryptic peptides, as this is the region in which nitrocellulose background peaks appear. Thus, if T3 (388 Da) had been present in the digest it could not have been confirmed by these studies. Likewise, T9 with a calculated mass of 408 Da would fall almost exactly at the mass of the most intense peak (m/z 409) for nitrocellulose.

TABLE II

RELEVANT CALCULATED AND EXPERIMENTAL[a] MASSES OF *S. aureus* V8-PROTEASE PEPTIDES OF INTERLEUKIN-3[b]

Peptide	Sequence	Molecular mass	MH^+ in PD spectra
V1	1–23	2545	[c]
V2	8–23	1858	[c]
V3	8–30	2609	[c]
V4	8–32	2835	[c]
V5	24–32	995	997
V6	33–57	2936	2939[d]
V7	38–57	2364	2365
V8	62–87	2936	2939[d]
V9	100–116	2225	2226
V10	117–140	2441	2440

[a] ^{252}Cf-PDMS, 15 kV accelerating voltage, $4 \cdot 10^6$ counts

[b] 2 μg (120 pmol) IL-3.

[c] Probable glycosylated peptide.

[d] V6 and V8 have the same nominal mass, but T14 appears to be glycosylated (see Fig. 6), indicating that V8 may be a glycopeptide and that mass 2939 corresponds to V6.

Mass spectrometric peptide-mapping of S. aureus V8-protease digests

V8-protease digests of IL-3 produced a number of major peptides (Table II). The most intense peaks were due to peptides V5 with *m/z* 997, and V6 (or V8) with *m/z* 2939. There were also smaller peaks due to peptides V7 (*m/z* 2365), V9 (*m/z* 2226), and V 10 (*m/z* 2440) (Fig. 4). Peaks for several predicted peptides, V1–V4, were not observed. These peptides contain the Asn-16 N-glycosylation consensus sequence, and the results suggested glycosylation at this site.

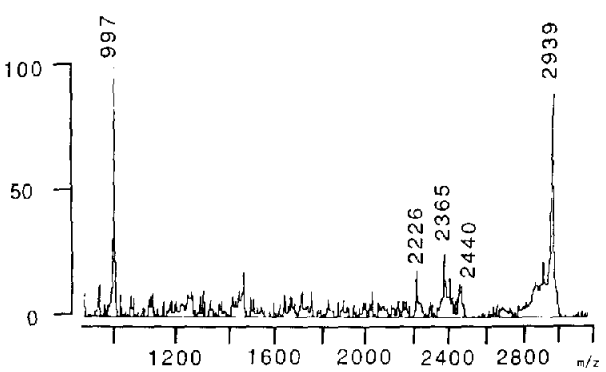

Fig. 4. ^{252}Cf-Plasma desorption mass spectrum of an *S. aureus* V8-protease (pH 4.0) digest of interleukin-3 (2 μg; 118 pmol) after 19-h digestion; start counts, $4 \cdot 10^6$; accelerating voltage, 15 kV.

Glycosylation sites

IL-3 has consensus sequences for four potential N-glycosylation sites (Asn-X-Ser) at the following positions: Asn-16, Asn-44, Asn-51, Asn-86. The presence in the tryptic digest of the peptide T10, with an observed *m/z* 717, which is identical with that predicted from the amino acid sequence, rules out glycosylation at Asn-51. Peptides T8 (651 Da) and T8′ (807 Da), which contain Asn-44, were not observed, but these peptides could have been digested further and their absence did not necessarily indicate glycosylation. The results from the *S. aureus* digestion, however, are informative. Peptide V7, which includes amino acids 38–57 (Table II), was observed in the digest (Fig. 4). Moreover, the peak due to protonated peptide V6 (*m/z* 2936), which contains amino acids 33—57, is fairly intense in the PDMS spectrum. These results not only conclusively confirm the absence of glycosylation at Asn-51, but also strongly indicate that Asn-44 is not glycosylated either.

A small peak in the spectrum of the tryptic digests at *m/z* 2091 could be due to peptide T14 (77–96). The peak intensity is very weak, however, and T14 was assumed to be glycosylated at Asn-86. Similarly, there were no peaks in the tryptic peptide mapping analysis corresponding to T4 (14–22). Moreover, in the V8-protease digests, no peaks were observed for the expected sequences 1–23 (V1), 8–23 (V2), 8–30 (V3), and 8–32 (V4); sequences which contain the N-glycosylation site Asn-16. It is very likely, therefore, that Asn-16 is glycosylated.

Reduction and carboxymethylation of cysteine residues

Peptides connected by disulphide linkages may not be recognizable directly by peptide-mapping analysis. Moreover, digestion of the protein under these conditions may be incomplete. Three of the four cysteine residues of IL-3 are located in tryptic peptides that contain N-glycosylation sites, *i.e.* Asn-16 and Asn-86, and these peptides were not detected in the tryptic digests. To determine whether these two sites are glycosylated or the disulphide bridges lead to production of unrecognizable peptide fragments, reduction of the disulphide linkages by dithiothreitol (DTT) followed by carboxymethylation (CM) and digestion by trypsin was carried out. This protocol was expected to produce three modified peptides in the mixture: T4 with one CM-Cys (1022 Da), T14 with two CM-Cys (2206 Da) and T18 with one CM-Cys (565 Da). An intense peak with *m/z* 566 in the PDMS spectrum could be due to carboxymethylated T18, but it could also be due to T9′, which has the same nominal mass (Table I). There was no peak at *m/z* 1022 and this result strongly indicated that T4 is modified by glycosylation. A small peak in the spectrum at *m/z* 2207 suggested the presence of some protonated T14. Several new peaks appeared that had not been observed in the spectrum of unreduced IL-3 tryptic digest. Of particular interest are three intense ions at *m/z* 1551, 1595 and 3086 (Fig. 5).

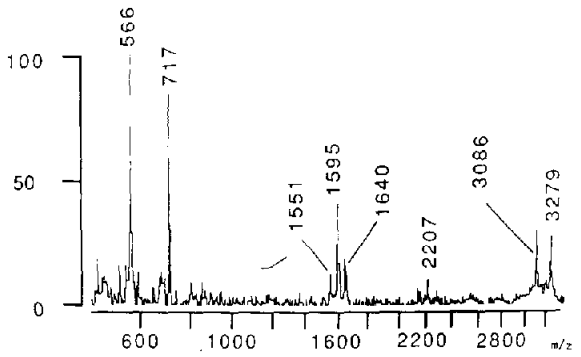

Fig. 5. ^{252}Cf-PDMS spectrum of reduced and carboxymethylated IL-3 (2 μg; 120 pmol) tryptic peptides. Peaks at *m/z* 1551 and *m/z* 1595 are probably due to glycosylated T4, and the peak at *m/z* 3086 is glycosylated T14.

Edman degradation combined with PDMS analysis of glycosylated peptides

With a combination of Edman degradation and molecular mass determination of truncated peptides by PDMS [24], peptides with modified amino acids can be identified directly (Fig. 6). The carboxymethylated peptide with *m/z* 3086, for example, after two cycles of coupling with PITC and hydrolysis, showed molecular mass reductions of 113 Da (Leu) and 114 Da (Asn) (Fig. 6b), respectively. Two additional Edman cycles showed losses of two CM-cysteines. Clearly this is the tryptic peptide, T14, whose unmodified mass with the two CM-cys residues should be 2206 Da. The difference in mass between the carboxymethylated peptide observed and that calculated is 878 Da, which, thus, is determined to be the mass of a carbohydrate residue attached at Asn-86. The mass of the N-glycan core sequence [Manα1→3((Manα1→6)Manβ1)→4GlcNAcβ1→4GlcNAc-] of complex oligosaccharides is 893 Da. Although we cannot yet explain the low mass of the carbohydrate residue observed, it may result from a modification of the carbohydrate from the purification or the analytical procedure.

Tryptic peptides were initially isolated by fractionation of the peptide mixture by HPLC on a reversed-phase C_{18} column (Fig. 7). Each fraction contained one or two peptides. The masses of these isolated peptides were determined by PDMS, and their identities were thereby confirmed (Table I). The late-eluting fraction (F14) contained the glycosylated peptides and those peptides interconnected with disulphide linkages. Peptides T4, T14, and T18 are believed to have been connected via disulphide bridges to Cys-17, Cys-79, Cys-80, and Cys-140. Following reduction of the disulphide linkages and carboxymethylation of the peptides isolated in F-14 (Fig. 7), the identities of T4, T14, and T18 were established by gas-phase Edman sequencing.

Edman degradation coupled with PDMS was not carried out on T4, but if it is assumed that the molecular ions, *viz. m/z* 1551 and *m/z* 1595, observed for the

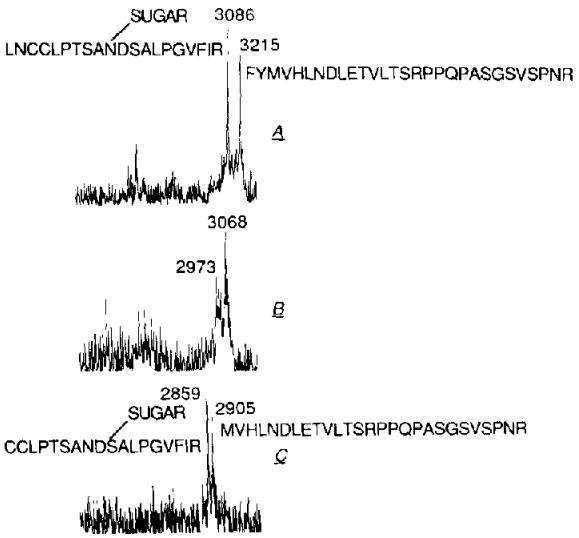

Fig. 6. ^{252}Cf-PDMS measurements of (A) tryptic peptides T17 and glycosylated T14 from reduced and carboxymethylated IL-3; (B) after one Edman cycle; (C) after two Edman cycles.

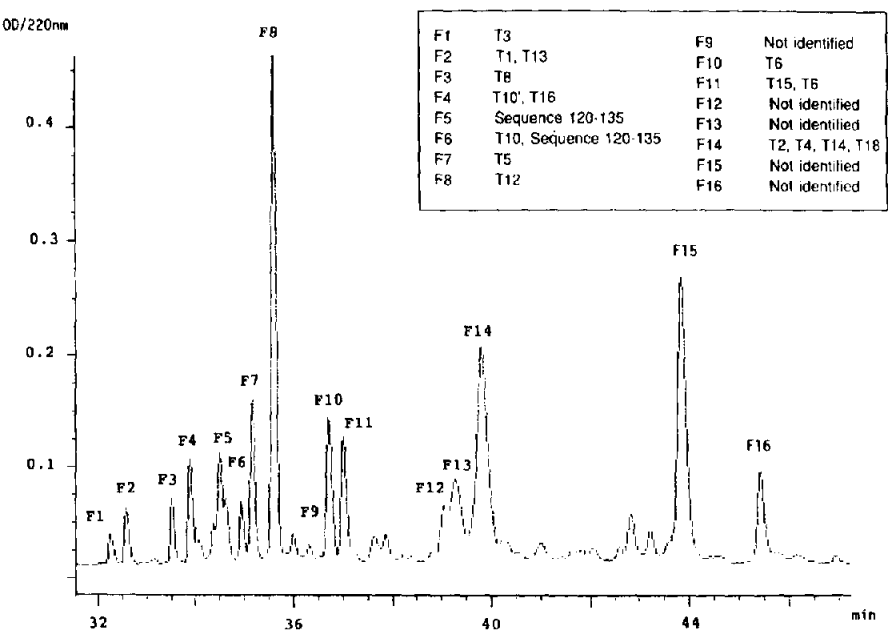

Fig. 7. High-performance liquid chromatogram of tryptic peptides from IL-3 on 25 cm × 4 mm I.D. C_{18} Nucleosil (7 μm particle size); UV adsorption at 220 nm; gradient elution (see Experimental).

carboxymethylated tryptic peptides are both due to glycosylated T4, then carbohydrate residues with masses of 528 and 572 Da are associated with T4.

The presence of T2 in F14 (Fig. 7), which was identified by gas-phase Edman sequencing, suggests modification of this peptide. This relatively small peptide has a hydrophilicity not much different from that of some of the early eluting peptides, which would not place it in this fraction. Therefore, T2 may be modified. One would expect a glycosylated peptide to be very hydrophilic and elute early. However, the size of the determinant may play a role in the relative elution order. PDMS measurements indicated that T-2 was also present in the tryptic peptide mixture, suggesting it is partially modified (Table I).

Several other late-eluting fractions, *i.e.* F15 and F16 (Fig. 7) were not identified. These likewise apear to contain hydrophobic linkages. A tryptic peptide with *m/z* 1639 in the spectrum of an HPLC fraction (F5) was shown to contain the N-terminal sequence Thr-Ser-Arg; this peptide was thus recognized as 120–135, arising by non-specific trypsin cleavage at the C-terminal end of Leu-119.

It was not possible to observe a peak for T-3 (Table I) at the appropriate mass (389 Da) by PDMS because interference from nitrocellulose peaks obscure this region. But the identity of this peptide was confirmed by Edman sequencing and glycosylation is ruled out. Peptides T-1 and T-13 have the same nominal mass, but although a peak appeared in the tryptic digest spectrum with the correct mass (591 Da), O-glycosylation of one or the other could not be ruled out from the proteolysis experiments alone. The V8-protease digestion experiment was not informative, since both V1 and V8, which has the same nominal mass as V6, contain N-glycosylation sites, Asn-16 and Asn-86.

Carbohydrate residues

With the aid of a hydrolysis and peracetylation procedure [25] followed by chromatographic and LSIMS analyses, both high mannose and complex-type oligosaccharides were identified. The ions in the LSIMS spectrum associated with complex-type oligosaccharides [23] were observed at *m/z* 408, *m/z* 619, *m/z* 636, *m/z* 924, and *m/z* 966 (Table III). A peak with *m/z* 1094 in the spectrum could be due to NeuNAc-Hex-HexNAc-OAcH$^+$ (1095 Da), and provides some evidence for the presence of sialic acid. Fucose was indicated by the presence of an ion at *m/z* 273. The presence of fucose as well as of glucosamine, mannose and glucose was also ascertained by polyanion chromatography.

There is good evidence for the presence of oligomannosidic oligosaccharides in the LSIMS spectrum as well. A small peak at *m/z* 1272 was observed, which could be Hex_4OAc/NH_4^+ from the protonated molecular-ion series. However, the A-type fragment ions, *i.e.* oxonium ions from cleavage of the glycosidic linkage, were more intense and clearly visible. These were observed at *m/z* 1195, *m/z* 907, *m/z* 619 and *m/z* 331. Both the HPLC and LSIMS results indicate a large percentage of mannose in the carbohydrate residue.

TABLE III

MOLECULAR IONS OBSERVED BY POSITIVE-ION LSIMS[a] IN MIXTURE OF PERACETYLATED CARBOHYDRATES FROM THE ACETOLYSIS OF IL-3 [25]

m/z	Probable carbohydrate composition
331	Hex^+
408	$Hex\text{-}OAc/NH_4^+$ [b]
619	$Hex\text{-}HexNac^+/H$ or Hex_2^+
636	$Hex\text{-}HexNAc\text{-}OH/H^+$
678	$Hex\text{-}HexNac\text{-}OAc/H^+$
966	$Hex\text{-}Hex\text{-}HexNAc\text{-}OAc/H^+$
1094	$NeuNAc\text{-}Hex\text{-}HexNAc\text{-}OAc^+$
1136[c]	Hex-Hex-HexNAc-Fuc

[a] 1:1 thioglycerol–glycerol matrix; xenon atom beam (6–8 KeV).

[b] Ammonium chloride added to matrix.

[c] *m/z* 1094 can arise by loss of ketene from *m/z* 1136.

DISCUSSION

It is well established that different molecules with IL-3 activity show considerable microheterogeneity [22]. Although populations of IL-3 differing in the amino acid start position [22,23] have been reported, the extensive microheterogeneity is most likely related to the differences in glycosylation. It has been demonstrated, for example, that treatment of WEHI-3 IL-3 with neuraminidase and addition of tunicamycin to IL-3-producing cells reduced the charge heterogeneity, which strongly indicates the presence of carbohydrates [22]. In addition, a stably transformed L-cell line expressing mouse IL-3 has produced a glycoprotein that has an apparent molecular mass on SDS-PAGE of 26 kDa [9]. In contrast, the apparent molecular mass of silkworm-expressed IL-3 on SDS-PAGE is only 22 kDa [5]. Our estimated molecular masses of 16 601 and 17 419 Da show that the silkworm-expressed IL-3 has an apparent molecular mass 928–1746 Da greater than would be predicted from the amino acid sequence of IL-3 (15 674). This mass difference could be accounted for only by glycosylation of the IL-3 preparation. The binding of IL-3 to concanavalin A also indicates that it contains complex biantennary N-glycans [26].

We have shown two N-glycosylation sites (Asn-16 and Asn-86) of IL-3 contain carbohydrate residues, and two other potential N-glycosylation sites (Asn-44 and Asn-51) were shown not to be glycosylated. In addition, O-linked oligosaccharides involving linkages to serine and threonine could be present. IL-3 has fifteen serines and eight threonines, which are potential O-linked glycosylation sites (Table IV). The hydrolysis mixture contained a small amount of galactosamine, which is evidence for O-glycosylation. The tryptic and V8-protease digests in-

TABLE IV

O-LINKED GLYCOSYLATION SITES OF INTERLEUKIN-3

Peptide	Potential glycosylation site		Glycosylation
	S	T	
T1	2, 4		No
T2		8	[a]
T3		12,14	[b]
T4	18,19		[c]
T7		35	No
V6	41, 46, 53	35	No
V7	41, 46, 53		No
T11	58		No
T13	72		No
T14	84, 88	83	[d]
T17	121, 128, 130, 132	117, 120	No

[a] A tryptic peptide was observed (see Table I), but F14 (Fig. 7) contains T2 peptide, indicating it may be partially modified.

[b] Expected mass for T3 is 388 Da, which would be obscured in PDMS by chemical and background noise. T3 appeared in F1 (Fig. 7), indicating it is not glycosylated.

[c] T4 is glycosylated at Asn-16, but glycosylation at Ser-18 and Ser-19 cannot be ruled out.

[d] T14 is glycosylated, probably at Asn-86. However, glycosylation at Thr-83, Ser-84, Ser-88 cannot be ruled out.

dicate that most of the peptides are not likely to be O-glycosylated, but four serines and five threonines could not be shown unambiguously to be non-glycosylated, based on peptide-mapping analysis. Of these, Ser-18 and Ser-88 are part of peptides that are N-glycosylated.

Recombinant baculoviruses are increasingly being used as expression vectors for a wide variety of genes because of their ability to efficiently produce large amounts of proteins. It has been suggested that the determinants of processing reside in the polypeptide, and that they are universally recognized despite differences in the oligosaccharide processing pathways in insect and vertebrate cells [27]. A comparison of glycoproteins from the mosquito cell line (*Aedes albopictus* $C_6/36$) with those from vertebrate cells, for example, shows the oligosaccharide synthesis is similar. But the oligosaccharides of glycoproteins from *Sindbis* virus, derived from Chinese hamster ovary (CHO) cells and expressed in mosquito cells, differ from the vertebrate-derived oligosaccharides in that the $Man_3GlcNAc_2$ structure does not progress to a complex type in the *Aedes* cells [28]. Thus, it was concluded that the low activity of the necessary glycosyltransferases is consistent with the immature N-glycans of mosquito membrane glycoproteins [28]. But it was shown that *Spodoptera frugiperda* cells express a human tissue plasminogen

activator preparation that is partially resistant to *endo*-N-acetyl-D-glucosaminidase H, an enzyme that removes immature high-mannose type oligosaccharides [29]. These studies suggest that insect cells are able to carry out the trimming reactions but processing does not take place such that the trimmed oligosaccharides progress to complex side-chains that contain fucose, galactose and neuraminic acid.

More recently, however, it was shown that the oligosaccharides of influenza virus, hemagglutinin expressed by *S. frugiperda* (SF) using a baculovirus vector, though severely truncated are, nevertheless, extensively fucosylated [30]. This suggests that, in SF cells at least, fucosyltransferases are present. In the case of IL-3, the L-cells-expressed protein has an apparent maximum mass by SDS-PAGE of 26 kDa whereas the *Bombyx-mori*-expressed protein gives an apparent mass by SDS-PAGE of 22 kDa. These results alone suggest that processing is incomplete in the insect-expressed protein, and this is confirmed by the relatively large amount of glucose [24] found in this study. Despite clear evidence for incompletely processed IL-3 from *B. mori* cells, fucose is present and there is even some evidence for neuraminic acid. Thus, these insect cells also must possess some of the appropriate glycosyltransferase activity to act on trimmed oligosaccharides to synthesize complex carbohydrates.

The differences and possible divergence of oligosaccharide processing pathways can only be understood if methods are available to determine the chemical structures of oligosaccharides and glycopeptides directly and efficiently. In this study it is demonstrated that PDMS, LSIMS and peptide mapping analysis are highly suitable for (1) determining molecular masses accurately where other methods, *e.g.* SDS-PAGE, fail badly, (2) confirming the consensus protein sequence, (3) locating N- and O-linked glycans on the protein, and (4) rapidly identifying the carbohydrate classes that are present.

ACKNOWLEDGEMENTS

This work was supported in part by grants from the National Institutes of Health, NIEHS ES001210 and from the Deutsche Forschungsgemeinschaft, Bonn, F.R.G. (Pr-175/2). We gratefully acknowledge many helpful suggestions by Professor George Rohrmann, Oregon State University. We also thank Professor S. Maeda, University of California, Davis, for helpful discussion.

REFERENCES

1 D. M. Rennick, F. D. Lee, T. Yokota, K. Arai, H. Cantor and G. J. Nabel, *J. Immunol.*, 134 (1985) 910–914.
2 A. Wodnar-Filipowicz, C. Heusser and C. Moroni, *Nature (London)*, 339 (1989) 150–152.
3 A. D. Whetton and T. M. Dexter, *Trends Biochem. Sci.*, 11 (1986) 207.
4 J. N. Ihle and Y. Weinstein, *Adv. Immunol.*, 39 (1986) 1.
5 A. Miyajima, T. Yokota, T. Otsuka, S. Miyatake, Y. Takebe, H. Hagiwara, J.Schreurs, K. Kaibuchi, F. Lee, N. Arai, T. Mosmann and K. Arai, *Immunol. Ser.*, 40 (1988) 103–148.

6 J. N. Ihle, J. Keller, L. Henderson, F. Klein and E. Palaszynski, *J. Immunol.*, 129 (1982) 2431–2436.
7 I. Clark-Lewis, S. B. H. Kent and J. W. Schrader, *J. Biol. Chem.*, 259 (1984) 7488.
8 T. Yokota, F. Lee, D. Rennick, C. Hall, N. Arai, T. Mosmann, G. Nabel, H. Cantor and K. Arai, *Proc. Natl. Acad. Sci. U.S.A.*, 81 (1984) 1070–1074.
9 T. R. Mosmann, J. S. Abrams, K. Arai, N. Arai, M. W. Bond, F. D. Lee, A. Miyajima, S. Miyatake, D. Rennick, J. Schreurs, C. A. Smith, Y. Takebe, T. Yokota, G. Zurawski and S. M. Zurawski, in S. Gilles (Editor), *Recombinant Lymphokines and Their Receptors*, Marcel Dekker, New York, 1987.
10 J. J. Delente, *Trends Biotech.*, 3 (1985) 218.
11 P. W. Berman and L. A. Lasky, *Trends Biotech.*, 3 (1985) 51–53.
12 C. J. McNeal (Editor), *Mass Spectrometry in the Analysis of Large Molecules*, Wiley, New York, 1986.
13 M. Barber, R. S. Bardoli, R. D. Sedgwick and R. W. Tyler, *J. Chem. Soc., Chem. Commun.*, (1981) 325–327.
14 B. R. Sundquist and R. D. MacFarlane, *Mass Spectrom. Rev.*, 4 (1985) 421–460.
15 I. Kamensky and A. G. Craig, *Anal. Instr.*, 16 (1987) 71–91.
16 B. W. Gibson and K. Biemann, *Proc. Natl. Acad. Sci. U.S.A.*, 81 (1984) 1956–1960.
17 A. Miyajima, J. Schreurs, K. Otsu, A. Kondo, K. Arai and S. Maeda, *Gene*, 58 (1987) 273–281.
18 J. F. Abrams and M. K. Pearce, *J. Immunol.*, 140 (1988) 131–137.
19 T. Mosmann, *J. Immunol. Methods*, 65 (1983) 55–63.
20 P. Roepstorff, P. F. Nielsen and P. Hojrup, *Biochem. Soc. Trans.*, 15 (1987) 155–158.
21 G. P. Jonsson, A. B. Hedin, A. B. Hakansson, B. R.Sundquist, G. S. Save, P. F. Nielsen, P. Roepstorff, K. E. Johansson, I. Kamensky and M. S. K. Lindberg, *Anal. Chem.*, 58 (1986) 1084–1087.
22 I. Clark-Lewis and J. W. Schrader, *Lymphokines*, 15 (1988) 1–37.
23 J. N. Ihle, J. Keller, S. Oroszlan, L. E. Henderson, T. D. Copeland, F. Fitch, M. B. Prystowsky, E. Goldwasser, J. W. Schrader, E. Palaszynski, M. Dy and B. Lebel, *J. Immunol.*, 131 (1983) 282–287.
24 P. F. Nielsen, M. Svoboda, B. Landis, K. Schneider and M. Przybylski, *Anal. Biochem.*, (1990) in press.
25 S. Naik, J. E. Oates, A. Dell, G. W. Taylor, P. M. Dey and J. B. Pridham, *Biochem. Biophys. Res. Commun.*, 132 (1985) 1–7.
26 R. K. Merkle and R. D. Cummings, *Methods Enzymol.*, 138 (1987) 232–259.
27 P. Hsieh and P. W. Robbins, *J. Biol. Chem.*, 259 (1984) 2375–2382.
28 T. D. Butters, R. C. Hughes and P. Vischer, *Biochim. Biophys. Acta*, 640 (1981) 627–686.
29 D. L. Jarvis and M. D. Summers, *Molec. Cell Biol.*, 9 (1989) 214–223.
30 K. Kuroda, H. Geyer, R. Geyer, W. Doerfler and H-D. Klenk, *Virology*, 174 (1990) 418–429.
31 T. D. Butters and R. C. Hughes, *Biochim. Biophys. Acta*, 640 (1981) 655–671.