Glycosylation pattern of human inter-α-inhibitor heavy chains

Christophe FLAHAUT*†, Calliope CAPON*†, Malika BALDUYCK*, Guy RICART‡, Pierre SAUTIERE§ and Jacques MIZON*[1]

*Laboratoire de Biochimie, Faculté de Pharmacie, Université de Lille II, Avenue du Professeur Laguesse, B.P. 83, F-59006 Lille, France, †Laboratoire de Chimie Biologique, Université des Sciences et Technologies de Lille I, F-59655, Villeneuve d'Ascq, France, ‡Laboratoire de Spectrométrie de Masse, Centre Commun de Mesures, Université des Sciences et Technologies de Lille I, F-59655, Villeneuve d'Ascq, France, and §Département V, URA 1309 CNRS, Institut de Biologie et Institut Pasteur de Lille, 1 rue du Professeur Calmette, B.P. 447, F-59021 Lille, France

Human inter-α-inhibitor (IαI) is a plasma serine-proteinase inhibitor. It consists of three polypeptide chains covalently linked by a glycosaminoglycan chain: a light chain named bikunin carrying the anti-proteinase activity and two heavy chains, H1 and H2, which exhibit specific properties, e.g. they interact with hyaluronan thus stabilizing the extracellular matrix. In this study, using matrix-assisted laser desorption ionization–time-of-flight MS and amino acid sequencing of tryptic peptides, we provide a detailed analysis of the glycosylation pattern of both heavy chains. H1 carries two complex-type N-glycans of predominantly biantennary structure linked to asparagine residues at positions 256 and 559 respectively. In contrast, the oligosaccharides attached to H2 are a complex-type N-glycan in the N-terminal region of the protein (Asn^{64}) and three to four type-1 core-structure O-glycans mono- or di-sialylated, clustered in the C-terminal region. We propose that these O-glycans might function as a recognition signal for the H2 heavy chain. The biological implications of this hypothesis, notably for the biosynthetic pathway of IαI, are discussed.

INTRODUCTION

Inter-α-inhibitor (IαI) is the leading member of a family of human plasma serine-proteinase inhibitors. IαI consists of three polypeptide chains, covalently linked by a glycosaminoglycan chain: two heavy chains H1 and H2 (M_r approx. 80000) and a light chain (M_r 26000) named bikunin, because it is structurally related to Künitz-type proteinase inhibitors (recently reviewed in [1]). For many years, IαI has been considered to be a precursor for smaller anti-proteinases carrying bikunin, able to diffuse more easily and thus participate in tissue protection [2]. Furthermore, besides its proteinase-inhibiting activity, bikunin may play additional roles, e.g. modulation of cellular proliferation, anti-inflammatory activity, etc.

It is now evident that the IαI heavy chains also exert specific functions, e.g. they are responsible for the hyaluronan (HA)-binding capacity of IαI reported previously [3–5]. In this respect, IαI may help to stabilize the extracellular matrix. On the other hand, it has been proposed that the heavy chains mediate the binding of IαI to the cell surface [6]; thus they would serve as specific carriers for bikunin delivery in various tissues.

However, the possible mechanisms for HA binding to IαI remain open to discussion and therefore a thorough structural characterization of the heavy chains might assist in a better understanding of their physiological role. The amino acid sequences of the two heavy chains H1 and H2 are highly similar, with 40% identity [7]. With regard to glycosylation, two asparagine residues were identified as carbohydrate-binding sites in each of the heavy chains [8]. However, we have reported the presence of *N*-acetylgalactosamine (GalNAc) in the H2 heavy chain, thus suggesting the probable occurrence of O-glycosylation [9], and more recently [10] it has been demonstrated that Thr^{637} in H2 is effectively O-glycosylated. Here, we provide a detailed analysis of the glycosylation pattern of both H1 and H2 of IαI.

MATERIALS AND METHODS

Materials

Human IαI was isolated as described recently [11] and supplied by the Laboratoire Français du Fractionnement et des Biotechnologies (Lille, France). Tosylphenylalanylchloromethane (Tos-Phe-CH_2Cl)-treated trypsin from bovine pancreas and endoproteinase Glu-C from *Staphylococcus aureus* were from Sigma and Boehringer respectively. Recombinant peptide-N-glycosidase F (PNGase F) from *Flavobacterium meningosepticum* was from Oxford GlycoSystems (Abingdon, Oxon., U.K.). All the solvents and reagents used for gas-phase sequencing were from Perkin–Elmer (Norwalk, CT, U.S.A.).

Isolation and characterization of the IαI heavy chains

The H1 and H2 heavy chains were isolated from IαI by anion-exchange chromatography following dissociation by treatment with NH_2OH. Briefly, IαI (80 mg) in 100 ml of water was mixed with an equal vol. of 2 M NH_2OH adjusted to pH 9.2 and stirred for 1 h at 37 °C before dialysis overnight at 4 °C against a 20 mM Tris/HCl buffer, pH 7.6, containing 4 mM EDTA (buffer A). The efficacy of the dissociation was checked by SDS/PAGE (7.5% gels) with Coomassie Brilliant Blue staining. The dissociated IαI was then loaded on to a Q-Sepharose fast-flow column (16 cm × 1.5 cm) equilibrated in buffer A at a flow rate of 100 ml/h. After washing the column for 30 min with the same

Abbreviations used: Iαl, inter-α-inhibitor; GalNAc, *N*-acetylgalactosamine; Tos-Phe-CH_2Cl, tosylphenylalanylchloromethane; PNGase F, peptide-N-glycosidase F; i.d., internal diameter; RP, reverse-phase; TFA, trifluoroacetic acid; MALDI–TOF-MS, matrix-assisted laser desorption ionization–time-of-flight MS; PTH, phenylthiohydantoin; 2-AB, 2-aminobenzamide; HA, hyaluronan; NeuAc, *N*-acetylneuraminic acid.

[1] To whom correspondence should be addressed (e-mail jmizon@phare.univ-lille2.fr).

buffer, the elution was performed by using a linear gradient from 0 to 0.5 M NaCl in buffer A for 4 h and then 1 M NaCl in buffer A was applied to the column. The A_{280} was monitored and the eluted fractions further analysed by SDS/PAGE (10 % gels). The fractions containing each heavy chain were collected and desalted if necessary by HPLC using a Synchropak RP1 column [100 mm × 4.6 mm internal diameter (i.d.); Alltech, Deerfield, IL, U.S.A.].

For PNGase F digestion, each pure heavy chain (60 μg) in 60 μl of a 10 mM phosphate buffer, pH 7.5, containing 25 mM EDTA, 0.5 % (w/v) SDS and 5 % (v/v) 2-mercaptoethanol was heated for 2 min at 100 °C. After cooling the mixture, 20 μl of 10 % (v/v) Triton X-100 and 10 μl of PNGase F (5 units) were successively added. For H1, the reaction mixture was incubated for 3 h at 25 °C in a water bath.

In order to enhance the deglycosylation of H2, an additional 5 units of PNGase F was added after 65 h and the incubation was prolonged up to 72 h. Samples were taken at different incubation times for SDS/PAGE analysis.

Reduction, alkylation and tryptic hydrolysis of the heavy chains

5 mg of each heavy chain in 5 ml of 100 mM Tris/HCl, pH 7.25, were denatured by 6 M guanidine hydrochloride. The pH was then adjusted to 8.0. The samples were flushed with nitrogen and incubated for 30 min at 50 °C. After addition of 12 mM dithiothreitol, the samples were again flushed with nitrogen and incubated at 50 °C for 4 h. 300 mM iodoacetamide was then added. The reaction was carried out in the dark at room temperature for 30 min under a N_2 atmosphere. After extensive dialysis of the reaction mixtures against a 100 mM Tris/HCl buffer, pH 8.0, Tos-Phe-CH_2Cl-treated trypsin was added with an enzyme-to-substrate ratio of 1:50 (by mass) and the mixture incubated at 37 °C for 24 h. The hydrolysis was stopped by addition of 4-(2-aminoethyl)-benzenesulphonylfluoride (AEBSF) at 1 mM final concentration. The fragments obtained were separated either by reverse-phase (RP)-HPLC or affinity chromatography using lectin-coupled agarose.

RP-HPLC separation of trypsin-digested peptides of the H2 heavy chain

The tryptic digest of H2 was applied to a C18 RP 5 μm Zorbax column (250 mm × 4.6 mm i.d.) equipped with a guard column (50 mm × 4.6 mm i.d.) of the same phase (Interchim, Montluçon, France) and equilibrated with 0.1 % (v/v) trifluoroacetic acid (TFA) in water. For elution, isocratic conditions with the same solvent were applied for 5 min, followed by a linear gradient of acetonitrile from 0 to 80 % in 0.1 % (v/v) TFA over 120 min at a flow rate of 0.7 ml/min. Eluates were monitored by A_{215}. Fractions 21–23 containing glycopeptides were detected by molar carbohydrate composition and matrix-assisted laser desorption ionization–time-of-flight MS (MALDI–TOF-MS) and further purified on the same column by using a linear gradient of acetonitrile from 20–70 % over 100 min.

Separation of glycopeptides by affinity chromatography on immobilized lectins

The peptide digests corresponding to 5 mg (in 5 ml buffer) of H1 or H2, after reduction, alkylation and tryptic hydrolysis, were loaded on to a Concanavalin-A–Sepharose column (5 ml) equilibrated in 50 mM Tris/HCl/100 mM NaCl buffer, pH 7.4, containing 1 mM $CaCl_2$, 1 mM $MnCl_2$, 1 mM $MgCl_2$ and 0.02 % (w/v) NaN_3 (buffer B). After washing the column with 10 ml of buffer B, the elution was performed by 10 ml of 0.5 M α-methylglucoside in buffer B at a flow rate of 20 ml/h, with detection by A_{280}. For the H2 digest, the washed fractions were pooled and loaded at the same flow rate on to a Jacalin–agarose column (2 ml) equilibrated in buffer B. After washing the column (6 ml buffer B), glycopeptides were eluted by 0.5 M α-methylgalactoside in buffer B (6 ml). The different separated fractions were concentrated into 2 ml aliquots in a vacuum centrifuge (Jouan, Saint-Nazaire, France) before desalting on a Sep-Pak C18 cartridge (Waters U.K. Ltd., Watford, Herts., U.K.) according to the manufacturer's instructions. They were then freeze-dried before analysis by MALDI–TOF-MS. In preparation for amino acid sequencing, the glycopeptide fractions were purified using an ABI 130A MicroBore HPLC system (Perkin–Elmer). The freeze-dried fractions were dissolved in 0.05 % (v/v) TFA and separated on a C18 Vydac 218 TP column (250 mm × 1 mm i.d., 10 μm particle size, 300 nm pore size), using a linear gradient from 0 to 65 % acetonitrile in 0.05 % (v/v) TFA over 120 min at a flow rate of 50 μl/min. The A_{215} was monitored.

Analysis of oligosaccharides

Glycans were released by hydrazinolysis from 2 mg of heavy chains, using the GlycoPrep 1000 (Oxford GlycoSystems) with N-mode. The resulting oligosaccharides were desalted by gel-filtration chromatography on a Bio-Gel P2 column (90 cm × 1.6 cm) with deionized H_2O. Desalted glycans were analysed with respect to molar carbohydrate composition and labelled with 2-aminobenzamide (2-AB) [12]. AB-labelled glycans were then separated by HPLC on a GlycoSep H column (100 mm × 3 mm i.d., Oxford GlycoSystems), using a gradient of acetonitrile from 0 to 40 % in 0.05 % (v/v) TFA over 30 min, followed by 5 min at 40 % and 15 min from 40–60 % at a flow rate of 0.5 ml/min. The A_{254} was monitored. All fractions were further analysed by MALDI–TOF-MS.

Analytical methods

PAGE was performed using 7.5 % and 10 % polyacrylamide slab gels under reducing or non-reducing conditions [13] with Coomassie Blue staining [14].

Carbohydrate analysis was performed using gas chromatography with a silicone OV 101 capillary column (0.32 mm × 25 m). Samples were analysed after methanolysis (0.5 M HCl in methanol for 24 h at 80 °C), followed by N-reacetylation and trimethylsilylation, as described in [15] with slight modifications [16].

N-terminal sequencing was carried out on a gas-phase sequencer (Procise 492, Perkin–Elmer) using the pulsed-liquid programme. Phenylthiohydantoin (PTH) derivatives of amino acids were identified on-line on a 120 A amino acid analyser (Perkin–Elmer).The yields expressed in picomol of PTH-amino acids correspond to raw data.

MALDI–TOF-MS was performed on a 'Vision 2000' time-of-flight instrument (Finnigan MAT, Bremen, Germany) equipped with a 337 nm UV Laser. The mass spectra were acquired in linear or reflection modes with positive or negative detection and with or without 10 kV accelerating voltage. Aliquots (1 μl) of the analyte solution (15 pmol) were mixed with an equal vol. of the matrix solution, 2,5-dihydroxybenzoic acid [10 mg/ml dissolved in acetonitrile–H_2O (30:70, by vol.)]. External calibration was performed using either BSA (M_r 66431; Sigma) for mass determination of the heavy chains or angiotensin I (M_r 1296.7; Sigma) as standards, and bovine insulin (M_r 5733.6; Sigma) for

mass determination of the glycopeptides and the released oligosaccharides respectively.

For endoproteinase Glu-C hydrolysis of O-glycopeptides of H2, pure freeze-dried peptides (6 nmol) were dissolved in 40 μl of 50 mM ammonium acetate buffer, pH 4.0, and digested at 37 °C for 15 h using an enzyme-to-peptide ratio of 1:50. Digestion was then stopped by the addition of 5 μl of 10% (v/v) formic acid before freeze-drying and MALDI–TOF-MS analysis.

Chemical desialylation was carried out by heating peptides at 80 °C in a closed reaction vial in the presence of 5% (v/v) formic acid at pH 2.0.

RESULTS

Isolation and characterization of the IαI heavy chains

IαI was dissociated by NH_2OH treatment, which selectively cleaves the ester bonds between an internal GalNAc molecule in the glycosaminoglycan chain and the C-terminal aspartate of each heavy chain, namely H1 and H2. They were then isolated by anion-exchange chromatography and analysed by SDS/PAGE with prior reduction. H1 and H2 moved as a unique protein band; their relative molecular masses were 78000 and 85000 respectively, as described previously [17].

The purity of the preparations was also assessed by MALDI–TOF-MS (Figure 1). The mass of H1, determined from the singly charged ion, was 76258. Subtraction of the protein mass of 71892, calculated from the amino acid sequence, resulted in a mass of 4366 for the carbohydrate moiety. By employing the same method, the molecular mass on the basis of the amino acid sequence of H2 (72432) was compared with the mass 76508 determined by MALDI–TOF-MS. The difference (4076) indicates an increase in mass due to a post-translational modification, presumably glycosylation. For each IαI heavy chain, MS analysis showed a broader peak than for BSA used as external calibrator; the broadening of the peak was related to the heterogeneity of glycosylation. Thus the glycosylation patterns of H1 and H2 were successively studied.

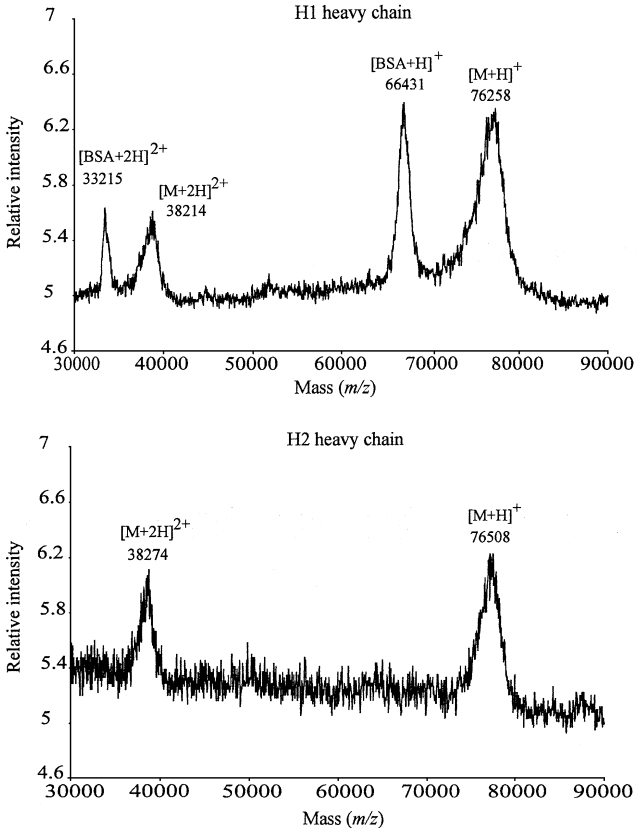


Figure 1 Linear time-of-flight mass spectrum for the H1 and H2 heavy chains

The singly charged ion of H1 (top panel) appeared at *m/z* 76258; the same signal for H2 (bottom panel) appeared at *m/z* 76508. The double-charged ions of the two heavy chains were also detected (*m/z* 38214 for H1; *m/z* 38274 for H2). BSA was used as an external calibration standard in both cases.

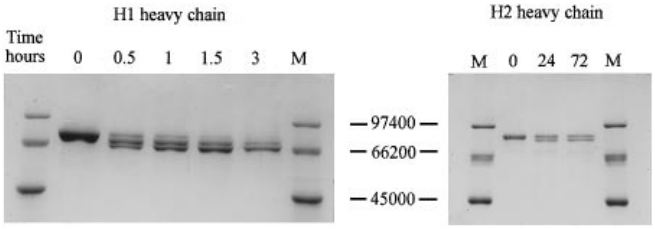


Figure 2 PNGase F digestion patterns of the H1 and H2 heavy chains

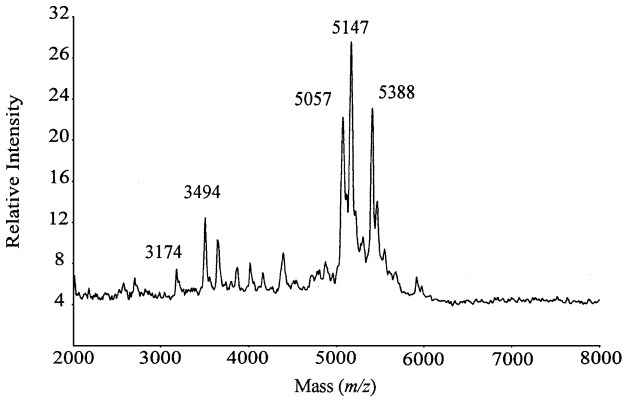
Samples (1 μg), after different incubation times as indicated, were analysed by SDS/PAGE (10% gels) under reducing conditions with Coomassie Blue staining. Left panel: H1; right panel: H2. M, low-molecular-mass markers from Bio-Rad: phosphorylase B, 97400 Da; BSA, 66200 Da; egg white albumin, 45000 Da.

Glycosylation pattern of the H1 heavy chain

Carbohydrate composition of H1 (Table 1) was in agreement with the presence of N-linked glycans, mainly of biantennary type. The presence of two N-linked glycans was established by PNGase F digestion of H1 and subsequent SDS/PAGE analysis. Indeed, two bands (M_r approx. 76000 and 74000) successively appeared during the incubation time (Figure 2).

To identify the glycosylation sites, glycopeptides obtained by tryptic digestion of reduced and carboxamidomethylated H1 heavy chain were extracted by affinity chromatography on Concanavalin-A–Sepharose. The glycopeptides eluted by α-methylglucoside were desalted on Sep-Pak cartridges and analysed by MALDI–TOF-MS. The spectrum (Figure 3) showed three strong ion signals at *m/z* 5388, 5147 and 5057. To determine the structure of these components, the glycopeptides were further purified by RP-HPLC and subjected to N-terminal sequencing. The intense signals 5388 and 5147 were assigned to peptides of Asp^{236}–Lys^{262} and Ile^{238}–Lys^{262} respectively, which both comprise the potential N-glycosylation acceptor site Asn^{256}-Leu-Thr and a di-sialylated biantennary glycan. The signal observed at *m/z* 5057 was consistent with an identical di-sialylated biantennary structure linked to the second potential N-glycosylation site Asn^{559}-Leu-Ser located in the peptide Ala^{558}–Arg^{583}. As indicated in Figure 3, we also assigned a mass composition to other signals (3494 and 3174). Interestingly, the signal at *m/z* 3494 corresponds to the peptide Phe^{252}–Lys^{262}, which results from cleavage at the C-terminal side of Phe^{251}. Such a cleavage by trypsin has been previously observed [18], even when chymotryptic activity had been chemically inhibited. Therefore our results confirm that the two potential N-glycosylation sites of H1 (namely Asn^{256} and Asn^{559}) are effectively occupied.

Furthermore, total glycans were released from H1 by hydrazinolysis, labelled with 2-AB and analysed on a GlycoSep H column. The chromatographic elution profile was followed by



Glycopeptides			
Observed Mass	N-terminal Sequence	Identification or position in sequence	Theoretical Mass
5388	D K I C D	D^{236}-$\underline{N}^{256}$-K^{262}	5378
5147	I C D L L	I^{238}-$\underline{N}^{256}$-K^{262}	5135
5057	A - L S S	A^{558}-$\underline{N}^{559}$-R^{583}	5061
3494	F A P Q -	F^{252}-$\underline{N}^{256}$-K^{262}	3481
3174	A - L S S	A^{558}-$\underline{N}^{559}$-R^{566}	3163

Figure 3 Linear time-of-flight MS and assignment of signals to the N-glycopeptides of H1 obtained by affinity chromatography on Concanavalin-A–Sepharose

The conditions used are as described in the Materials and methods section. Asn at positions 256 and 559 correspond to potential sites of N-glycosylation.

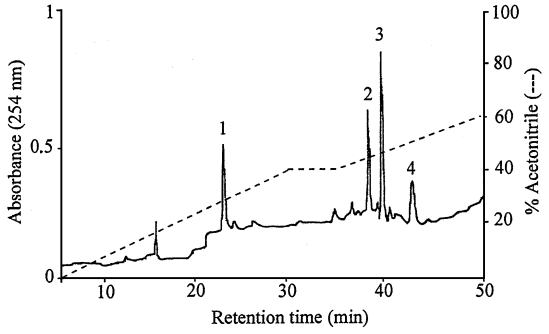


Figure 4 Fractionation of 2-AB-labelled oligosaccharides from the H1 heavy chain

The labelled oligosaccharides were fractionated on a GlycoSep H column (100 mm × 3 mm, i.d.) developed with an acetonitrile gradient, as described in the Materials and methods section. According to MALDI–TOF-MS analysis, peak 1 corresponds to 2-AB in excess; peaks 2 and 3 contain mono- and di-sialylated biantennary structures respectively; peak 4 corresponds to a triantennary structure.

monitoring A_{254} (Figure 4). The major fractions corresponding to peaks 2 and 3 were identified by MALDI–TOF-MS as mono- and di-sialylated biantennary structures respectively. They represent about 28% and 47% of the released glycans respectively. The minor fraction corresponding to peak 4 represents 23% of these glycans and was characterized as a triantennary structure.

Table 1 Carbohydrate composition of the glycans linked to the IαI heavy chains

The molar ratios were calculated on the basis of three mannose units. Since the GlcNAc–Asn linkage is only partially cleaved by methanolysis [15,16], the number of GlcNAc is underestimated (about one residue for three mannose units).

	Molar ratios						
Heavy chain	Fuc	Gal	Man	Glc	GalNAc	GlcNAc	NeuAc
H1	–	3	3	–	–	2.5	1.7
H2	–	5.6	3	–	3.1	3	7

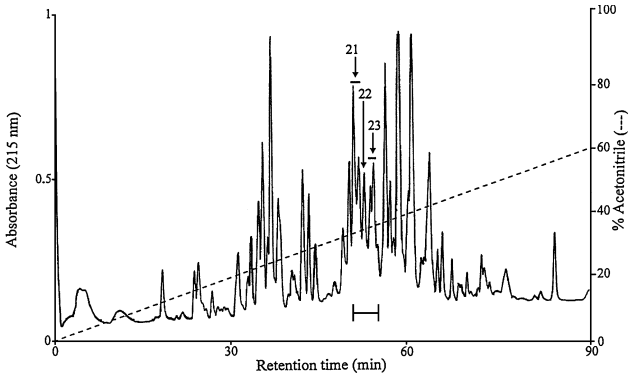


Figure 5 RP-HPLC elution profile of the tryptic digest from H2

The peptides generated by tryptic hydrolysis were injected on to a C18 RP 5 μm Zorbax column (250 mm × 4.6 mm i.d.) equipped with a guard column (50 mm × 4.6 mm i.d.) of the same phase and equilibrated with 0.1% (v/v) TFA in H_2O. The elution was performed as described in the Materials and methods section.

Glycosylation pattern of the H2 heavy chain

Carbohydrate composition of H2 (Table 1) appeared more complex compared with that of the H1 heavy chain. The increased contents of sialic acid and galactose and the presence of GalNAc suggest the co-existence of mucin-type O-linked glycans and complex-type N-glycans.

Characterization of the N-glycopeptide of H2

H2 contains three N-glycosylation sites that are candidates for modification by addition of carbohydrates: Asn^{42}-Asn-Ser, Asn^{64}-Phe-Ser and Asn^{391}-Ile-Ser. SDS/PAGE analysis showed that when H2 was incubated in the presence of PNGase F, its relative molecular mass shifted slowly from 85000 (native form) to 83000 (deglycosylated form) (Figure 2). Thus only one N-linked carbohydrate chain per mol of H2 would be present.

To identify attachment sites and structures of glycans on H2, the glycoprotein was reduced, carboxamidomethylated and then digested by trypsin. The resulting peptide fragments were separated by RP-HPLC (Figure 5). The purified fractions were analysed by MALDI–TOF-MS.

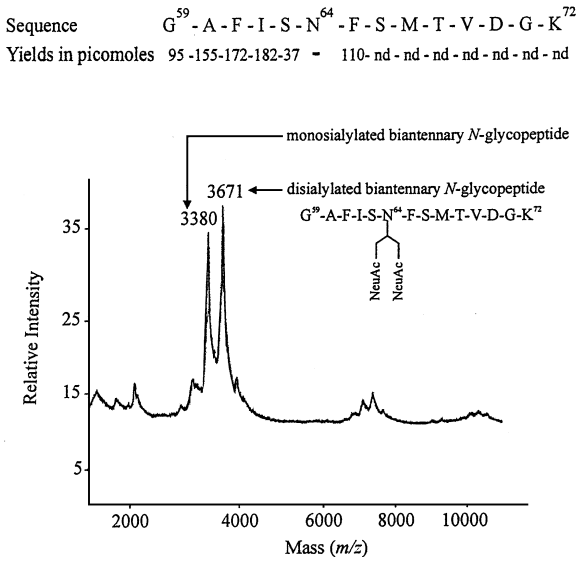


Figure 6 MALDI–TOF-MS analysis and N-terminal sequencing of the N-glycopeptide Gly[59]–Lys[72] from the H2 heavy chain

The glycopeptide isolated by RP-HPLC was analysed by MALDI–TOF-MS using a linear negative mode. Ions at *m/z* 3671 and 3380 correspond to the theoretical mass of the peptide Gly[59]–Lys[72] (1473 Da) with an additional N-glycan biantennary structure, either di-sialylated (2204 Da) or monosialylated (1913 Da). nd, not determined.

Two peptides with masses calculated theoretically from the known sequences around the potential N-glycosylation site were found: Val[40]–Asn[42]–Lys[58] and Glu[386]–Asn[391]–Lys[409]. However, the naked peptide Gly[59]–Asn[64]–Lys[72] was not detected: the predicted mass of this peptide, including the biantennary complex-type N-glycan, would be approx. 3500 Da. Glycopeptides with masses of 3500 Da were indeed characterized in fractions 21, 22 and 23 from HPLC (Figure 5). These fractions were pooled and rechromatographed as indicated in the Materials and methods section. The expected glycopeptides were clearly identified by Edman degradation (Figure 6); Asn[64] at the glycosylation site was, of course, not detected.

The structure of the N-linked carbohydrate chain was mainly of the biantennary type, as shown by MS data (Figure 6); the major signals at *m/z* 3671 and 3380 account for the presence of biantennary structure, di- and mono-sialylated respectively, attached to the peptide. Secondly, the structure was also shown by chromatographic analysis on a GlycoSep H column of oligosaccharides released from H2 by hydrazinolysis and then labelled by reductive amination with 2-AB; biantennary structure without, or with, one or two sialic acid residues represented 15%, 52% and 32% of N-glycans respectively. Thirdly, the glycopeptide Gly[59]–Lys[72] was easily purified from a tryptic digest of H2 by affinity chromatography on Concanavalin-A–Sepharose.

Characterization of the O-glycopeptides of H2

The carbohydrate composition (presence of GalNAc) and MS analysis (family of peaks with *m/z* approx. 6000) of the fractions 21, 22 and 23, obtained by RP chromatography of the tryptic digest of H2 (Figure 5), indicate the presence of O-glycopeptides in addition to the N-glycopeptide characterized previously. The pooled fractions were again fractionated by RP-HPLC. Several fractions in low amounts were obtained that, as indicated by amino acid sequencing, correspond to the C-terminal region of the H2 heavy chain, beginning at Val[607]. It was previously demonstrated that Thr[637] is O-glycosylated and MS data suggested that the O-linked carbohydrate chain has a type-1 core structure Galβ1–3-GalNAc-Thr recognized by Jacalin [10]. Indeed, by affinity chromatography on Jacalin–agarose followed by RP-HPLC on a microbore column, we separated a glycopeptide

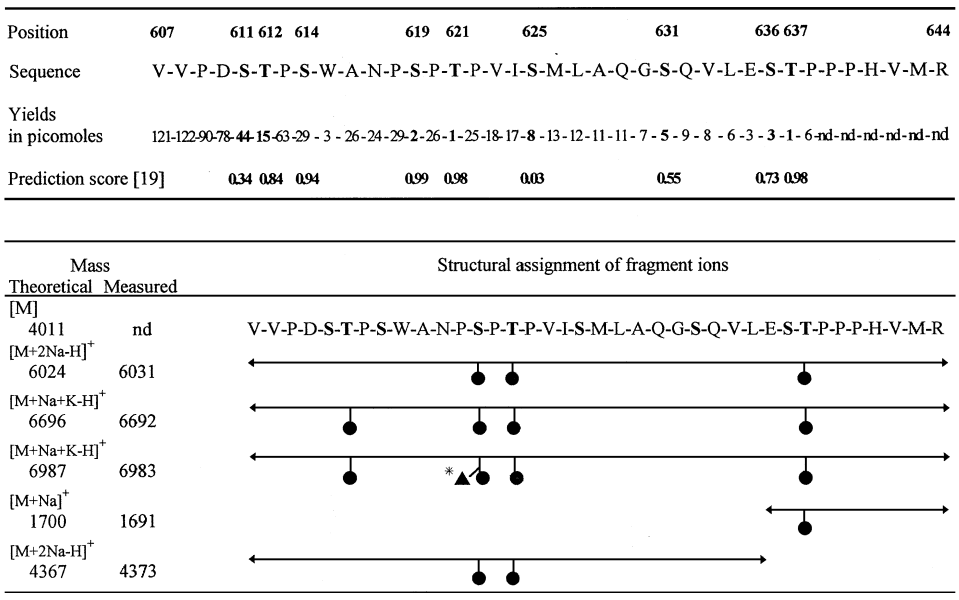


Figure 7 Structural characterization and assignment of the O-glycopeptide Val[607]–Arg[644] from the H2 heavy chain

Schematic representation showing the location of O-linked glycans on the O-glycopeptide Val[607]–Arg[644] from H2. The N-terminal sequencing data expressed in picomol of PTH-amino acids correspond to raw data. The O-linked oligosaccharides are indicated with symbols: (●), NeuAc-Gal-GalNAc; (▲), additional NeuAc residue carried by NeuAc-Gal-GalNAc; *, undefined residue position; nd, not determined.

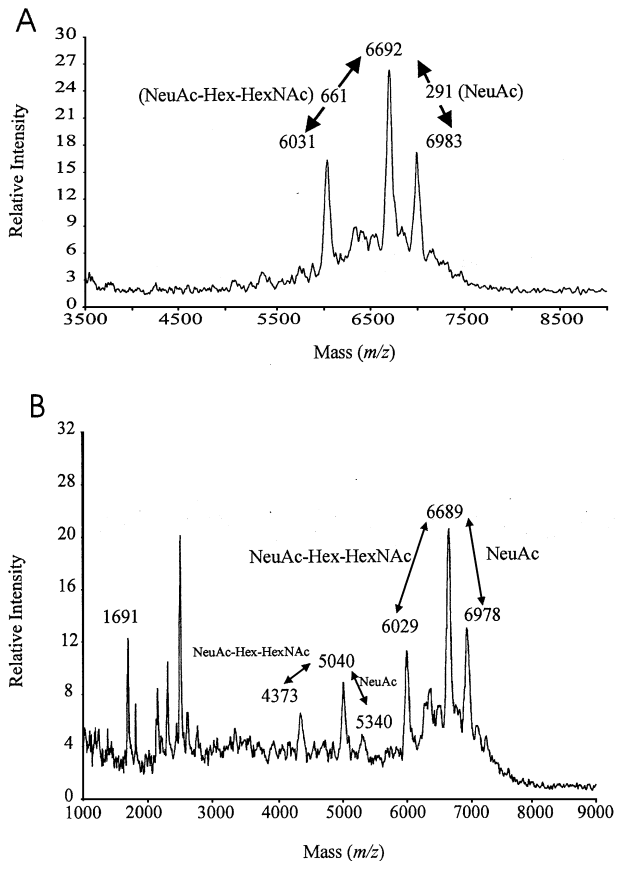


Figure 8 Linear time-of-flight spectrum of the O-glycopeptidic fraction isolated by Jacalin–agarose chromatography of the tryptic digest of H2

(**A**) Mass spectrum of the native O-glycopeptides; (**B**) mass spectrum of the same fraction after partial endoproteinase Glu-C digestion.

identified by N-terminal sequencing as the C-terminal part of H2, starting at Val^{607} (Figure 7). The predicted molecular mass of the peptide Val^{607}–Arg^{644} is 4011 Da. MALDI–TOF-MS analysis of the fraction eluted from Jacalin revealed three products with masses of 6983, 6692 and 6031 Da respectively (Figure 8A). Subtraction of the theoretical mass of the peptide (4011 Da) from the ion at m/z 6031 suggested the presence of three mucin-type O-glycans: *N*-acetylneuraminic acid (NeuAc)-Gal-GalNAc (656 Da). In the same manner, the ion at m/z 6692 was identified as the same peptide Val^{607}–Arg^{644} carrying four trisaccharides, whereas the ion at m/z 6983 was the same peptide carrying three trisaccharides and one di-sialylated tetrasaccharide. Indeed, by mass analysis after acid-catalysed desialylation of the glycopeptide, two components were characterized simultaneously at m/z 5152 and 5518 (results not shown). They correspond to the same peptide (4011 Da) carrying three or four O-glycans Gal-GalNAc (365 Da) respectively. This model is in good agreement with amino acid sequence analysis. The first 32 residues were positively identified by Edman degradation except for Ser^{619}, Thr^{621} and Thr^{637} thus suggesting that these three residues are fully glycosylated. Moreover, the low yield of PTH-Thr^{612} indicates a partial glycosylation of this residue (Figure 7).

The sequence of the glycopeptide was submitted to the Net O-Glyc Server [19] for prediction of O-glycosylation sites. All the O-glycosylated residues were correctly identified. However, Ser^{614} and Ser^{636} also exhibited a high potential for being glycosylated.

To confirm the structure of the O-glycopeptide Val^{607}–Arg^{644} eluted from Jacalin, it was digested with endoproteinase Glu-C and the digest analysed by MALDI–TOF-MS (Figure 8B). Fragment ions at m/z 1691 were assigned to be the peptide Ser^{636}–Arg^{644} with Thr^{637} fully glycosylated by the trisaccharide NeuAc-Gal-GalNAc. The three signals with m/z 4373, 5040 and 5340 corresponded to the mass of the peptide Val^{607}–Glu^{635} carrying two or three O-linked carbohydrate chains.

Taken together, our results demonstrate that the H2 heavy chain of IαI contains three or four mucin-type O-glycans clustered in the C-terminal region of the protein and one N-glycan, in a biantennary structure, linked to the Asn^{64} residue, whereas the H1 heavy chain carries two N-glycans possessing mainly the same biantennary structure. The different post-translational modifications concerning the human IαI molecule are depicted in Figure 9.

DISCUSSION

On the basis of protein sequence data, the human IαI heavy chains H1 and H2, which contain 644 and 648 amino acid residues respectively, are recognized as being highly homologous [7]. Indeed, their masses estimated in this work by MALDI–TOF-MS are very similar: 76258 for H1 and 76508 for H2. In the same way, with regard to their glycosylation patterns, two N-glycans were considered to be present on each chain [8].

However, as early as 1979, IαI was classified among the human serum proteins carrying alkali-labile carbohydrate chains able to be recognized by Jacalin [22]. Because carbohydrate moieties may play important physiological roles, we undertook a detailed analysis of the glycans bound to each heavy chain. We confirm that the two potential N-glycosylation sites of H1 are effectively fully occupied by complex-type N-glycans. They are predominantly of biantennary type, as demonstrated by MALDI–TOF-MS analysis of N-glycans released by hydrazinolysis.

However, we have clearly established that the H2 heavy chain carries only one complex-type N-glycan attached to Asn^{64}. Indeed, in the tryptic digest of H2, we detected by MALDI–TOF-MS analysis two naked peptides containing the two other potent glycosylation sites, Asn^{42}-Asn-Ser and Asn^{391}-Ile-Ser. Likewise, we demonstrated by amino acid sequencing that Asn^{617}, which belongs to a rarely glycosylated site (Asn-Pro-Ser), is effectively not modified (Figure 7).

The main result of our study is to show that, in addition to the O-glycan carried by Thr^{637} [10], two or three additional O-linked carbohydrate chains are present on H2. All these O-glycans consist of a type-1 core structure with one or two NeuAc moieties and are clustered in the C-terminal part of the peptide chain. Our results are on the basis of MS analysis and amino acid sequencing.

As suggested, the high frequency of proline residues in the C-terminal region of the different heavy chains (namely H1, H2 and H3), all belonging to the IαI family, may promote their maturation processing and their covalent linkage to the glycosaminoglycan carried by bikunin [23]. However, mucin-type O-linked glycosylation also occurs in regions with high contents of serine, threonine and proline. Thus to estimate the relative propensity of hydroxy amino acids to be glycosylated, we submitted the C-terminal sequences of the three heavy chains to the predictive method described in [19]. Our results show that the O-glycosylated sites on H2 were correctly identified (Figure 7). However, other predicted O-glycosylation sites were identified in

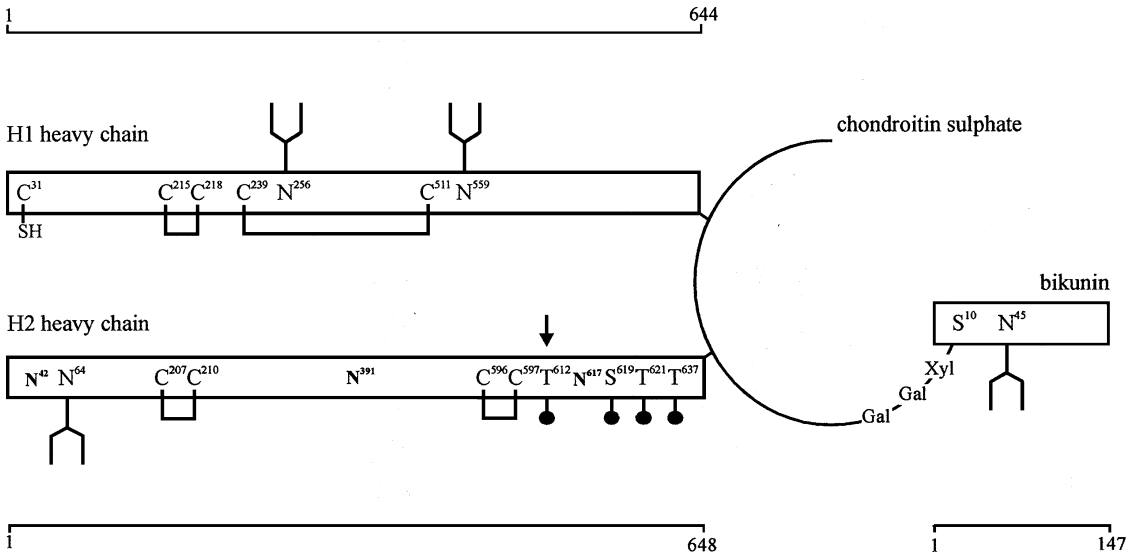


Figure 9 Schematic representation (not to scale) of the human IαI molecule showing post-translational modifications

The IαI molecule consists of three polypeptide chains covalently linked by a chondroitin sulphate chain. H1 and H2 are linked to the chondroitin sulphate chain via ester bonds between their C-terminal Asp residues and an internal *N*-acetylgalactosamine of the glycosaminoglycan chain [20]. The oligosaccharide structures are depicted with following symbols: (♈), N-glycan carrying either zero, one or two NeuAc residue(s) at the non-reducing terminal position. The biantennary N-glycan on bikunin has been described previously [21]. (●), O-linked glycan: NeuAc-Gal-GalNAc. The black arrow indicates the only site partially O-glycosylated. The disulphide bridges of human IαI [36] are depicted with solid lines. The disulphide bridges in bikunin are omitted for simplicity.

each heavy chain, whereas we demonstrated that, as assessed by their GalNAc content, H1 (this study) as well as H3 [24] did not contain O-glycans. Therefore we conclude that the glycosylation pattern of H2 (one N-glycan and a cluster of O-glycans) is clearly different from that of H1 or H3 (two N-glycans). These differences might be conserved between species; indeed, we note that the three fully O-glycosylated residues and the N-glycosylated Asn of human H2 are strictly conserved in the amino acid sequence of mouse H2 [25], as well as Syrian hamster H2 [26]. Furthermore, recent findings indicate that the segment including the cluster of O-glycans on H2 corresponds to the region in which the amino acid sequences of the three heavy chains (H1, H2 and H3) diverged during evolution, before the emergence of mammals [26]. To investigate this further, the structural study of the IαI heavy chains in different animal species might be rewarding.

Concerning the physiological implications of the presence of the O-glycan domain, several functions may be assigned. As originally described [27,28], the two heavy chains H1 and H2 are associated within the IαI structure and we have previously demonstrated by selective immunoadsorption [17] that each IαI molecule contains one H1 and one H2 heavy chain, both linked to bikunin. Obviously, O-glycosylation does not interfere in the formation of the glycosaminoglycan–protein cross-link. However, the mechanism that permits the control in hepatocytes of IαI biosynthesis by preferential assembly of these peptide chains remains unknown. We suggest that O-glycans located in the C-terminal part of H2 in an area exposed on the glycoprotein surface might carry the requisite information for a specific reconnaissance of the partners concerned. Indeed, two types of IαI have been described: a human type designated as IαIH_1,H_2,B and a bovine type that consists of H2 and H3 linked to bikunin [1]. Thus in each case, a unique highly O-glycosylated chain (H2) is associated with another peptide chain (H1 or H3). Interestingly, it has been reported that both types of IαI (human and bovine) co-exist in Syrian hamster [29]. However, the IαIH1,H3,B IαI type is still unknown.

Our hypothesis is supported by the fact that the chain assembly occurs in the latter part of the biosynthetic pathway, most likely in the *trans*-Golgi network or in secretory vesicles [30,31], whereas O-linked glycosylation is initiated in the *cis*-Golgi compartment, where polypeptide *N*-acetylgalactosaminyltransferases are located.

With regard to HA-binding capacity of the heavy chains, it has been suggested that the C-terminal half of the heavy chains participates in this binding [32]. Because O-linked glycosylation of CD44 decreases its ability to adhere to HA [33], it would be worthwhile to study further the influence of O-linked glycosylation of H2 on its affinity for HA.

At last, the metabolism of IαI might also be influenced by the presence of O-glycans linked on H2. Therefore during inflammatory processes such as septic shock, where the harmful role of leucocyte elastase has been suggested, a large consumption of IαI has been reported [34]. We have previously demonstrated that this proteinase preferentially cleaves the H2 heavy chain in its C-terminal part [35] and we hypothesized that charge-mediated interactions between the glycosaminoglycan chain of IαI and elastase, which is cationic, might explain this particular sensitivity to proteolysis. This one would be obviously enhanced by the O-glycan chains located in the C-terminal part of H2.

Multivalent O-glycosylation is largely implicated in glycoprotein functions. The O-glycosylation of H2 we describe here has to be taken into account in further studies concerning IαI.

This work was supported in part by the Ministère de l'Enseignement Supérieur et de la Recherche (DRED E.A. 1052) and by the Université de Lille II.

REFERENCES

1 Salier, J. P., Rouet, P., Raguenez, G. and Daveau, M. (1996) Biochem. J. **315**, 1–9
2 Gebhard, W. and Hochstrasser, K. (1986) in Proteinase Inhibitors (Barrett, A. J. and Salvesen, G., eds.), pp. 389–401, Elsevier Science Publishers BV
3 Sandson, J., Hamerman, D. and Schwick, G. (1965) Trans. Ass. Am. Physicians **28**, 304–313

4 Chen, L., Simon, J. T. and Larsen, W. J. (1992) J. Biol. Chem. **267**, 12380–12386
5 Blom, A., Pertoft, H. and Fries, E. (1995) J. Biol. Chem. **270**, 9698–9701
6 Kobayashi, H., Gotoh, J., Hirashima, Y. and Terao, T. (1996) J. Biol. Chem. **271**, 11362–11367
7 Diarra-Mehrpour, M., Bourguignon, J., Bost, F., Sesboüé, R., Muschio, F., Sarafan, N. and Martin, J. P. (1992) Biochim. Biophys. Acta **1132**, 114–118
8 Gebhard, W., Schreitmüller, T., Hochstrasser, K. and Wachter, E. (1989) Eur. J. Biochem. **181**, 571–576
9 Malki, N., Balduyck, M., Maes, P., Capon, C., Mizon, C., Han, K. K., Tartar, A., Fournet, B. and Mizon, J. (1992) Biol. Chem. Hoppe-Seyler **373**, 1009–1018
10 Enghild, J. J., Salvesen, G., Thogersen, I. B., Valnickova, Z., Pizzo, S. and Hefta, S. A. (1993) J. Biol. Chem. **268**, 8711–8716
11 Michalski, C., Piva, F., Balduyck, M., Mizon, C., Burnouf, T., Huart, J. J. and Mizon, J. (1994) Vox Sang. **67**, 329–336
12 Bigge, J. C., Patel, T., Bruce, J. A., Goulding, P. N., Charles, S. M. and Parekh, R. B. (1995) Anal. Biochem. **230**, 229–238
13 Laemmli, U.K. (1970) Nature (London) **227**, 680–685
14 Stephano, J. L., Gould, M. and Rojas-Galicia, L. (1986) Anal. Biochem. **152**, 308–313
15 Kamerling, J. P., Gerwig, G. J., Vliegenthart, J. F. G. and Clamp, J. R. (1975) Biochem. J. **151**, 491–495
16 Montreuil, J., Bouquelet, S., Debray, H., Fournet, B., Spik, G. and Strecker, G. (1986) in Carbohydrate Analysis: a Practical Approach (Chaplin, M. F. and Kennedy, J. F., eds.), pp. 143–204, IRL Press, Oxford
17 Mizon, C., Balduyck, M., Albani, D., Michalski, C., Burnouf, T. and Mizon, J. (1996) J. Immunol. Methods **190**, 61–70
18 Gohlke, M., Baude, G., Nuck, R., Grunow, D., Kannick, C., Bringmann, P., Donne, P. and Reutter, W. (1996) J. Biol. Chem. **271**, 7381–7386
19 Hansen, J. E., Lund, O., Tolstrup, N., Gooley, A. A., Williams, K. L. and Brunak, S. (1998) Glycoconjugate J. **15**, 115–130
20 Morelle, W., Capon, C., Balduyck, M., Sautière, P., Kouach, M., Michalski, C., Fournet, B. and Mizon, J. (1994) Eur. J. Biochem. **221**, 881–888
21 Toyada, H., Ikey, T., Demachi, Y., Toida, T. and Imanari, T. (1992) Chem. Pharm. Bull. **40**, 2882–2884
22 Chatterjee, B., Vaith, P., Chatterjee, S., Karduck, D. and Uhlenbruck, G. (1979) Int. J. Biochem. **10**, 321–327
23 Bourguignon, J., Diarra-Mehrpour, M., Thiberville, L., Bost, F., Sesboüé, R. and Martin, J. P. (1993) Eur. J. Biochem. **212**, 771–776
24 Mizon, C., Héron, A., Capon, C., Sautiére, P., Michalski, C., Sesboüé, R. and Mizon, J. (1997) J. Chromatogr. B **692**, 281–291
25 Chan, P., Risler, J. L., Raguenez, G. and Salier, J. P. (1995) Biochem. J. **306**, 505–512
26 Nakatani, T., Suzuki, Y., Yamamoto, T. and Sinohara, H. (1997) J. Biochem. **122**, 71–82
27 Balduyck, M., Laroui, S., Mizon, C. and Mizon, J. (1989) Biol. Chem. Hoppe-Seyler **370**, 329–336
28 Enghild, J. J., Thogersen, I. B., Pizzo, S. V. and Salvesen, G. (1989) J. Biol. Chem. **264**, 15975–15981
29 Yamamoto, T., Yamamoto, K. and Sinohara, H. (1996) J. Biochem. **120**, 145–152
30 Héron, A., Bourguignon, J., Callé, A., Borghi, H., Sesboüé, R., Diarra-Mehrpour, M. and Martin, J. P. (1994) Biochem. J. **302**, 573–580
31 Thogersen, I. B. and Enghild, J. J. (1995) J. Biol. Chem. **270**, 18700–18709
32 Huang, L., Yoneda, M. and Kimata, K. (1993) J. Biol. Chem. **268**, 26725–26730
33 Dasgupta, A., Takahashi, K., Cutler, M. and Tanabé, K. (1996) Biochem. Biophys. Res. Commun. **227**, 110–117
34 Witte, J., Jochum, M., Scherrer, R., Schramm, W., Hochstrasser, K. and Fritz, H. (1982) Int. Care Med. **8**, 215–222
35 Balduyck, M., Piva, F., Mizon, C., Maes, P., Malki, N., Gressier, B., Michalski, C. and Mizon, J. (1993) Biol. Chem. Hoppe-Seyler **374**, 895–901
36 Flahaut, C., Mizon, C., Aumercier-Maes, P., Colson, P., Bailly, C., Sautiere, P. and Mizon, J. (1998) Eur. J. Biochem, in the press

Received 5 February 1998/24 April 1998; accepted 14 May 1998