Aah VI, a novel, *N*-glycosylated anti-insect toxin from *Androctonus australis hector* scorpion venom: isolation, characterisation, and glycan structure determination

O. Hassani[a], D. Loew[b], A. Van Dorsselaer[1,b], M.J. Papandréou[a], O. Sorokine[b], H. Rochat[a], F. Sampieri[a], P. Mansuelle[a]

[a] *UMR 6560, Université de la Méditerranée CNRS, Laboratoire de Biochimie, Ingénierie des Protéines, IFR Jean Roche, Faculté de Médecine Secteur Nord, Bd P. Dramard, 13916 Marseille Cedex 20, France*
[b] *UMR 7509, Laboratoire de Spectrométrie de Masse Bio-Organique associé au CNRS, Université Louis Pasteur, 1 rue B. Pascal, 67008 Strasbourg, France*

Received 21 October 1998; received in revised form 10 December 1998

Abstract Aah VI was isolated from the venom of the North African scorpion, *Androctonus australis hector*. It is the first glycosylated neurotoxin from scorpion venom to be described. It was not toxic to mice, when injected intracerebroventricularly at a dose of 1.2 μg per animal. However, it had typical activity in *Blatella germanica* cockroaches resulting in gradual paralysis and very low toxicity (LD_{50} = 8.5 μg/g of animal). It consists of 66 amino acid residues and is heterogeneously *N*-glycosylated at a single site, on asparagine 9, of the Asn-Gly-Thr sequence. The potential *N*-glycosylation site was deduced from automatic Edman degradation and amino acid analysis, and glycan heterogeneity was evidenced by ESMS. Determination of the *N*-glycan structures (dHex, Hex and HexNAc) was assessed by nanoESMS/MS with picomolar amounts of sample. Current knowledge of *N*-glycan structure and composition suggests that the glycan structures are derived from a common core.

Key words: Scorpion toxin; *N*-Glycosylation; Mass spectrometry; High-performance liquid chromatography

1. Introduction

Na^+ channels generating nerve impulses are the target of the small polypeptide neurotoxins that are the very active components of scorpion venoms. These toxins have been useful as probes in Na^+ channel structure and function studies. They form a homologous family of mostly basic 60–70 residue proteins [1,2] and consist of a single chain cross-linked by four disulphide bridges [3,4]. Alignment of the eight half-cystine residues of these toxins demonstrates their sequence similarities and has made it possible to classify the toxins into several structural groups of maximal similarity. The toxins may also be classified as mammal-selective or insect-selective [5,6]. Glycosylated scorpion neurotoxins have not previously been reported.

We report here the purification and characterisation of the first glycosylated scorpion neurotoxin to be described: Aah VI. It is an anti-insect toxin, isolated from the venom of the Old World scorpion species *Androctonus australis hector*, and is weakly active on insects. It bears some unusual glycan structures and the possible function of these new structural features is discussed. We also demonstrate the efficiency of analytical mass spectrometry, and particularly nanoESMS/MS, for detecting protein heterogeneity and determining glycan structure in picomolar amounts of sample.

2. Materials and methods

2.1. Isolation and protein purification

Venom was obtained from *Androctonus australis hector* scorpions by electrical stimulation of animals collected in the Hazoua area (Tunisia). Partial purification was carried out in part as previously described [5,7]. Then, the lyophilised RIT fraction [5] was loaded onto a (7.5×75 mm) TosoHaas CM-5PW TSK column equilibrated with A (20 mM NH_4CH_3COO, pH 4.7), and eluted at a flow rate of 1 ml/min with a gradient of 0–50% of B (500 mM NH_4CH_3COO, pH 6.8) in A over 50 min. A sample of the fraction eluted at 22 min and was loaded onto a LiChrospher 100 RP-8, 125×4 mm column (Merck), equilibrated with 10% B (0.1% TFA in CH_3CN), and eluted with a gradient of 10–60% B in A (0.1% TFA in H_2O) over 25 min.

2.2. Determination of toxicity in vivo

For all injections, the solution used was 0.15 M NaCl containing 2 mg bovine serum albumin (BSA) per ml.

Toxicity in male C57BL/6 black mice (20±2 g) was tested by ICV injection of 5 μl of the solutions, and the LD_{50} was determined by a lethality test 24 h after injection. Toxicity in male cockroaches (90–110 mg) *Blatella germanica* was tested by injection of 1 μl of the solutions into the abdominal segment and the LD_{50} was determined 6 h after injection.

2.3. Enzymatic cleavage separation conditions, amino acid analyses and automated degradation of protein or peptides

These procedures were essentially performed as previously described [8]. Briefly, Aah VI was reduced and either *S*-carboxymethylated (RCM) or *S*-pyridyl-ethylated (RPE), then desalted on a C8 RP column and lyophilised. 3 nmol of RPE Aah VI was digested with trypsin (3% (w/w), in 125 mM HCO_3NH_4, pH 8.5, 20 h, 37°C). 5 nmol of RCM protein was digested with chymotrypsin (5% (w/w), in 100 mM Tris-HCl, 10 mM $CaCl_2$, pH 7.8, 20 h, 37°C). 200 pmol of RPE Aah VI was digested with Lys-C endoproteinase (4% (w/w), in 100 mM Tris-HCl, 10 mM $CaCl_2$, pH 8.3, 4 h, 37°C). The elution gradient for tryptic or chymotryptic peptide separation was from 2%

[1]Corresponding authors. A. Van Dorsselaer (Fax: (33) 388 60 46 87. E-mail: vandors@chimie.u-strasbg.fr) and P. Mansuelle (Fax: (33) 491 65 75 95. E-mail: mansuelle.p@jean-roche.univ-mrs.fr).

Abbreviations: Hex, hexose; dHex, deoxyhexose; HexNAc, *N*-acetylhexosamine; NeuAc, *N*-acetylneuraminic acid; ESMS, electrospray mass spectrometry; ESMS/MS, electrospray tandem mass spectrometry; ICV, intracerebroventricular; PEC, pyridylethyl-cysteine; RCM, reduced and *S*-carboxymethylated; RPE, reduced and *S*-pyridylethylated; PLA_2, phospholipase A_2; RP-HPLC, reverse-phase high-performance liquid chromatography; LC-MS, liquid chromatography mass spectrometry; TFA, trifluoroacetic acid; Tris, tris-(hydroxymethyl)-aminomethane; RT, retention time; SIR, single ion recording; LD_{50}, lethal dose for 50% of the tested animals

PII: S0014-5793(98)01710-4

to 88% of solvent B (0.1% TFA in CH_3CN/H_2O, 70/30) in A (0.1% TFA in H_2O) over 60 min at 200 μl/min on a 140 HPLC system equipped with a C18 PTH column (2.1×220 mm, 5 μm) from Applied Biosystems, a Division of Perkin Elmer (Foster City, CA). For Lys-C digest separation conditions, see Section 2.5. Detection was at 215 nm. Purified peptides and protein were lyophilised. Aliquots of resuspended lyophilisate were used for sequencing with an Applied Biosystems 476 A sequencer. Other aliquots were used for amino acid analysis on a Beckman 6300 apparatus [8]. Amino acid composition was calculated as follows: experimental values are means of four analyses, two after 20 h of hydrolysis, and the other two after 70 h, except for Ser, Thr, Tyr and Cys, for which the values were obtained by extrapolation to zero time hydrolysis, and for hydrophobic residues, for which the values after 70 h of hydrolysis were used. Trp was determined spectrophotometrically.

2.4. Electrospray (ES) and nanoelectrospray mass spectrometry (nanoESMS)

ES and nanoES mass spectra were obtained on a Quattro II mass spectrometer (Micromass Ltd UK, Altrincham, UK). The nanoES and ES drying gases were heated to 50°C and 80°C, respectively. NanoES was performed with a modified probe as described [9] and with no counter electrode. Samples were dissolved in H_2O/CH_3CN (1/1) containing 1% formic acid, and then introduced into the electrospray ion source. The very low flow rate (20 nmol/min) used in nanoESMS experiments allowed long analysis times (about 3 h on a sample of 4 μl) and then multiple measurements, i.e optimisation of fragmentation conditions, and adjustment of resolution.

2.5. Liquid chromatography mass spectrometry (LC-MS) and specific glycoside detection by single ion recording (SIR) of the Lys-C digest of RPE Aah VI

LC-MS was performed as previously described [10] by coupling the HPLC system described above to the ES mass spectrometer. The separation on the narrow-bore column (2 mm×125 mm) was flow-split via a stainless steel Valco tee (Supelco, Saint-Germain-en-Laye, France) with approximately 1/15 of column effluent being directed towards the mass spectrometer via a fused silica capillary. The samples were separated on a Macherey Nagel Nucleosil 300-5 C18 column (2 mm×125 mm) with a gradient of 0–60% of B (0.08% TFA in CH_3CN) in A (0.1% TFA in H_2O) over 60 min, at 250 μl/min. Total ion current (TIC) trace (*m/z* 300–1800) and single ion recording (SIR) were carried out by the stepped cone voltage technique from 50 to 90 V with an interscan delay of 0.1 s. Specific oligosaccharides were detected by scanning the oxonium ions and their water loss peaks [11] (168 for $(HexNAc\text{-}(H_2O)_2)^+$, 204 for $(HexNAc)^+$, 274 for $(NeuAc\text{-}H_2O)^+$, 292 for $(NeuAc)^+$ and 366 for $(Hex+HexNAc)^+$), during acquisition at a cone voltage of 90 V and 0.08 s duration for each specific mass.

2.6. Tandem mass spectrometry (MS/MS)

Native Aah VI, and the glycosylated peptide obtained by action of Lys-C endoproteinase, were subjected to nanoES tandem mass spectrometry. The parent ions were selected in the first quadrupole mass analyser and fragmented by collision-induced dissociation (CID) with argon. The collision energy was adjusted for each selected ion so that fragmentation was just apparent.

3. Results

3.1. Purification of Aah VI

Aah VI was purified by two chromatography steps. The first, which involved recycling gel filtration on Sephadex G-50, gave a similar pattern to that previously described [7], with three major fractions (RIT, RI and RII) obtained (not shown). Fraction RIT contains the insect-selective toxin Aah IT [5]. Fraction RI contains the anti-mammal toxins Aah I and Aah III and fraction RII contains the anti-mammal toxin Aah II and the anti-crustacean toxin, Aah CT [12]. The RIT fraction was further purified by cation-exchange chromatography using a CM-TSK column, giving several fractions. The fraction eluted at 22 min (Aah VI, Fig. 1A) produced a single peak in analytical C8 HPLC (Fig. 1B) and accounted for 16% of the absorbance at 280 nm of the insect toxin fraction.

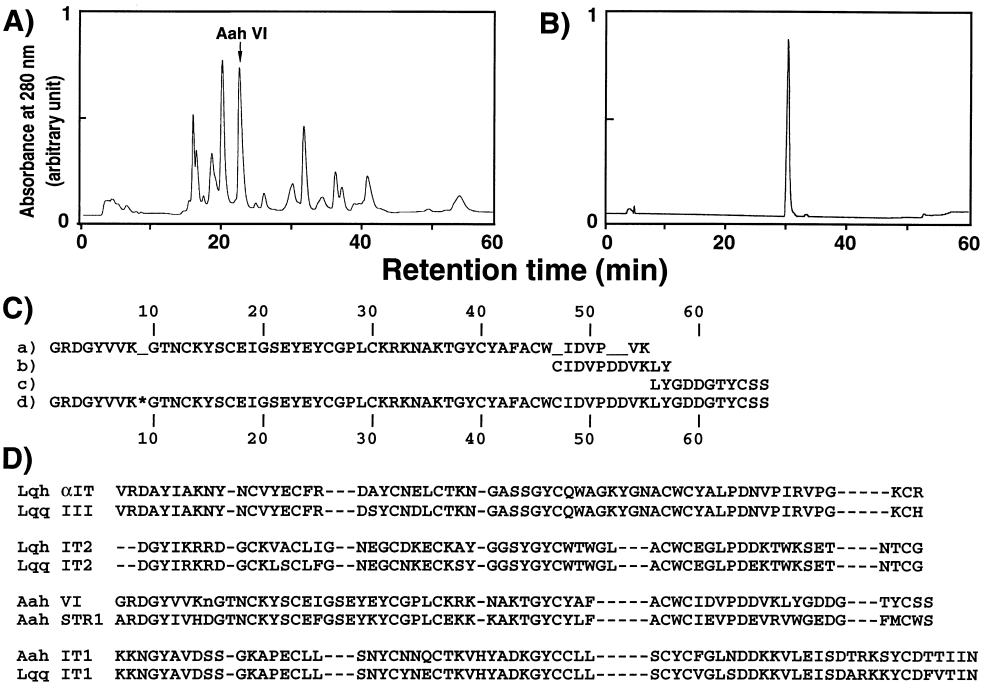

Fig. 1. Purification steps, and sequence determination and comparison. A: CM MemSep chromatography of the RIT fraction from recycling chromatography. B: RP-HPLC of the Aah VI fraction from CM MemSep. Details are given in Section 2. C: Amino acid sequencing of Aah VI. (a) Results of Edman degradation of RPE-Aah VI (_: unidentified residues); (b) chymotryptic peptide; (c) tryptic peptide; (d) complete sequence of Aah VI as deduced from a–c. The star indicates the glycosylated Asn residue. D: Sequence comparison with alignment of the half-cystine residues. The sequences were taken from references [4,16–20]. n: glycosylated asparagine residue.

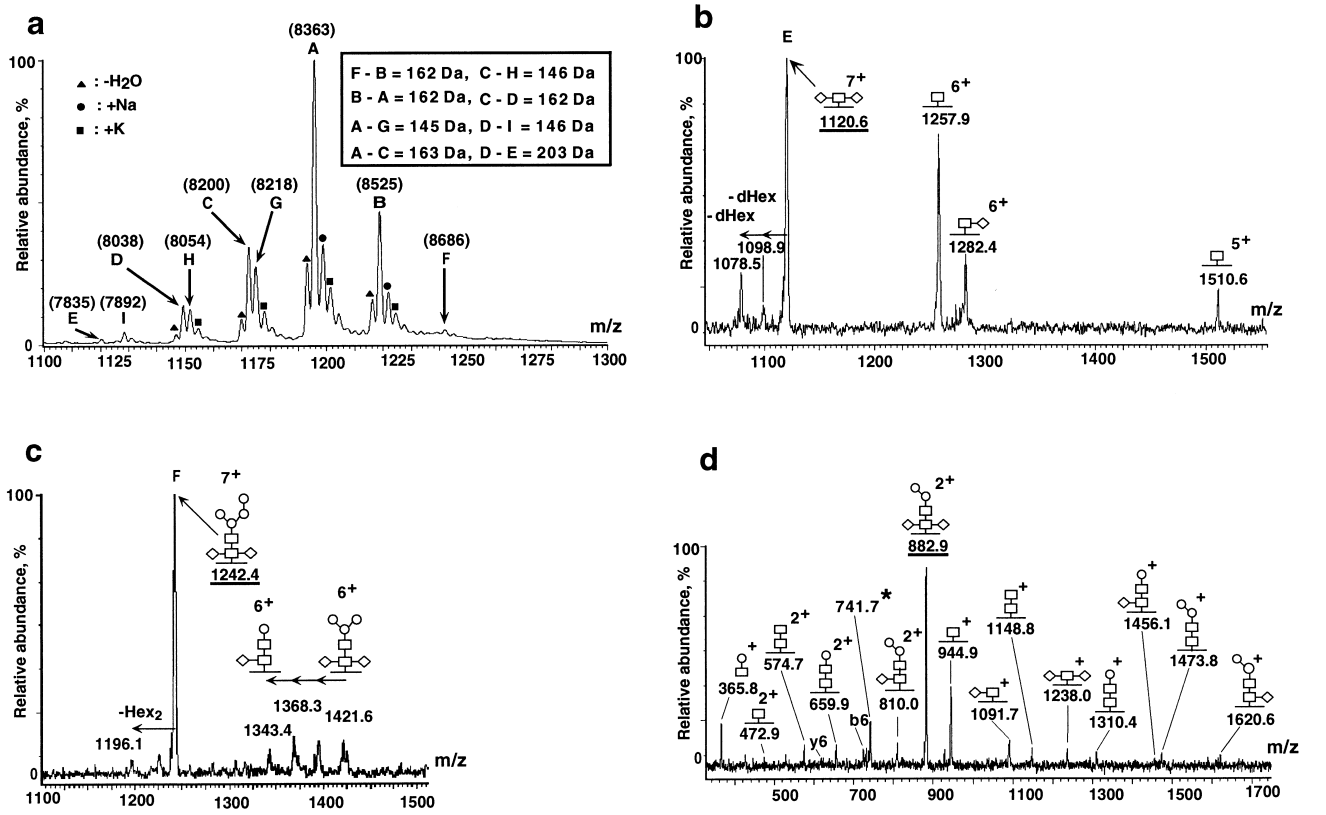

Fig. 2. Nanoelectrospray and/or nanoelectrospray tandem mass spectra obtained on intact Aah VI or peptide 9–14. a: NanoESMS of the seven charged states of AahVI. Adducts of Na and K were observed. Numbers in parentheses indicate real masses. Inset: mass differences between the indicated forms. b: NanoESMS/MS of seven charged ion at $m/z = 1120$ (underlined), i.e. form E in panel a. Two deoxyhexoses have clearly been lost, in succession, from the first *N*-acetylhexosamine, which is directly attached to the peptide backbone. c: NanoESMS/MS of seven charged ion at $m/z = 1242$ (underlined), i.e. form F in panel a. The successive loss of hexoses and deoxyhexoses is observed. The observed fragments are consistent with a trimannosylated core and two deoxyhexoses. d: NanoESMS/MS of the double charged ion at $m/z = 882.9$ (underlined), a glycoform of the Lys-C peptide (9–14), i.e. measured mass = 1763.6 (Table 1). Sequence ions are designated according to the usual nomenclature [32,33]. The complete structure of the glycans can be deduced from the fragments observed. The ion at $m/z = 741.7$ (labelled with a star) corresponds to the $(M+H)^+$ of the core peptide (9–14).

3.2. Homogeneity tests

The amino acid composition of Aah VI was calculated from amino acid analysis and spectrophotometric measurements. Values for most of the residues, other than hydrolysis time-sensitive residues, were close to integers: Asp 8.7; Thr 3.1; Ser 4.1; Glu 3.1; Pro 2.4; Gly 7.9; Ala 3.2; Cys 7.5; Val 3.7; Ile 2.0; Leu 2.3; Tyr 7.8; Phe 1.2; Lys 5.8; Trp 1; Arg 2.1. The homogeneity of the protein was verified by C8 RP HPLC, which gave a single peak (Fig. 1B), and by Edman degradation. We sequenced 400 pmol of native protein, and detected a single PTH in each of the first 14 cycles, except in cycles 9 and 13 which gave no PTH (not shown). Position 13 may be occupied by a conserved half-cystine residue, but position 9 has never been reported to correspond to such a residue in other toxins.

The nanoES mass spectrum of native Aah VI showed several series of multiply charged ions, demonstrating protein heterogeneity. Fig. 2a shows the ions with seven charges. The molecular masses of the nine main compounds (A–I) were separated by ±146 Da, ±162 or ±203 Da, typically corresponding to dHex, Hex and HexNAc moieties, respectively (Fig. 2a, inset).

3.3. Biological activity of Aah VI

Aah VI was completely inactive against mice, even when 1.2 μg was injected ICV. In cockroaches, however, Aah VI caused a very slow, gradual paralysis, but was not very toxic ($LD_{50} = 8.5$ μg/g of animal).

3.4. Sequence determination

We sequenced 700 pmol of RPE-Aah VI [8], and the N-terminal amino acid residues were assigned, up to residue 55, except for positions 9, 47, 52 and 53 (Fig. 1C, a). Two sets of peptides were generated by enzymatic digestion to fill these gaps and complete the sequence. Chymotryptic peptide 47–57 (Fig. 1C, b) and the C-terminal tryptic peptide 56–66 (Fig. 1C, c) were both useful. We therefore obtained the required level of peptide overlap (Fig. 1C, a–c) to assign the sequence of the alkylated protein entirely, except for the residue in position 9, as shown in Fig. 1C, d.

After reduction and alkylation, the measured mass of the major component A was 9212.3 Da (not shown), i.e. an increase of 849.4 Da, in good agreement with the addition of pyridylethyl moieties on the half-cystine residues of a protein with four disulphide bridges (expected increase: 849.17 Da).

3.5. A single glycosylation site: assumption and demonstration

We compared the Aah VI amino acid composition deduced by peptide sequencing with that obtained by acidic hydrolysis. There was clearly an Asp residue missing in the sequence data,

Table 1
Data from LC-MS, and suggested assignment of the peptides obtained from LC of the RPE Aah VI, digested with Lys-C endoproteinase

Data from LC-MS of Lys-C-digested Aah VI		Data from sequence	Mass difference	Proposed components of fractions from LC-MS of Lys-C-digested Aah VI
RT (min)	Meas. mass (Da)	Calc. mass (Da)		
24.88	1763.6	740.8 (9–14)	1022.8[a]	9–14+$(HexNAc)_2(dHex)_2(Hex)_2$
24.88	1617.4	740.8 (9–14)	876.6[b]	9–14+$(HexNAc)_2(dHex)_1(Hex)_2$
24.88	1925.8	740.8 (9–14)	1185.0[c]	9–14+$(HexNAc)_2(dHex)_2(Hex)_3$
28.62	893.0	893.0 (1–8)	0.0	1–8
29.83	1285.6	1285.3 (56–66)	0.3	56–66
33.57	2259.5	2259.6 (15–31)	0.3	15–31
38.74	2484.7	2484.8 (37–55)	0.1	37–55

Two expected short peptides (32–33, 34–36) were not detected and were presumably eluted in the void volume.
First column: retention time for each individual compound; second column: measured mass of each compound; third column: calculated mass for the corresponding expected core peptide (indicated in parentheses), deduced from sequence data (Fig. 1C) and Lys-C endoproteinase specificity; fourth column: difference between the measured and calculated masses. The differences between the measured masses of glycoforms A, G and B, and the calculated mass of the core protein were the same as in a, b and c, respectively; fifth column: proposed assignment of the compounds (peptide residue range and glycan composition).

suggesting that the unidentified residue at position 9 was a derivative of Asp or Asn converted back to Asp by acidic hydrolysis. If residue 9 were Asn, the sequence NGT (positions 9–11) would be a genuine glycosylation site [13,14]. This suggested that the heterogeneity was due to heterogeneous *N*-glycan structures, linked to the Asn-9 residue only. Assuming that residue 9 is indeed an Asn residue, the differences between the theoretical mass of the polypeptide (7340.2 Da) and the masses observed in the spectrum for the native protein can be used to suggest a glycan composition for each glycoform of Aah VI, taking into account the oligosaccharide structures of known *N*-linked glycans [15]. Using the two lightest forms I and E as bases (excess mass = 552 and 495 Da, respectively), a composition may be suggested for form E ($HexNAc(dHex)_2$), and then for each form A–I, with $(HexNAc)_2(dHex)_2(Hex)_4$ for form F.

LC-MS with specific detection of glycan residues (SIR) of alkylated Aah VI digested with the Lys-C endoproteinase (not shown) unambiguously showed that only one peptide (RT = 24.88 min) was glycosylated. Data are shown in Table 1 (columns 1 and 2). The mass spectrum of the glycosylated peptide fraction had signals for three compounds (column 2, first three rows), whereas the other spectra had only one. The fourth column indicates that the measured (column 2) and calculated (column 3) masses of the Lys-C peptides were similar, except for the 9–14 peptide. The three compounds in the peak at RT = 24.88 min are glycosylated forms of the core peptide 9–14 with an Asn residue in position 9, thus the mass differences (column 4, a, b, c) correspond to *N*-glycan moieties. These values are perfectly consistent with the excess mass of three of the main glycosylated forms of the native protein: A, G and B (Fig. 2a). The suggested glycan compositions are given in column 5. Thus, these results clearly demonstrate that Asn 9 was t-e only site of heterogeneous carbohydrate chain attachment.

3.6. Determination of the N-glycan structures by nanoES tandem mass spectrometry

A series of MS/MS experiments was performed for the most intense series of multiply charged ions of the various glycoforms (A–I) observed on the native protein and also for a series of very minor peaks. The spectra obtained all had regular and successive losses of 162, 146 or 204 Da from the parent ions, due to the losses of a Hex, a dHex or a $(HexNAc)^+$, respectively. Fig. 2b,c shows the fragmentation pattern and glycoform assignment based on the correlation of the ions with the smallest and the largest values of *m/z*, i.e. forms E and F. The analyses of these ions at *m/z* 1120.6 and *m/z* 1242.4 located the two dHex residues on the inner HexNAc (Fig. 2b) and the hexoses on the outer HexNAc (Fig. 2c). MS/MS analysis was also carried out for the major doubly charged ion at *m/z* = 882.9 (Fig. 2d) of the glycopeptide 9–14 (see Table 1, first line). The nanoESMS/MS analysis of this ion showed that the dHex moieties were linked to the HexNAc which was directly attached to the polypeptide chain and, that the two HexNAc are bound together. The mass of the core peptide was also detected at *m/z* = 741.7 corresponding to $(M+H)^+$ of the peptide. The structures of the various glycoforms, A–I, are shown in Fig. 3a.

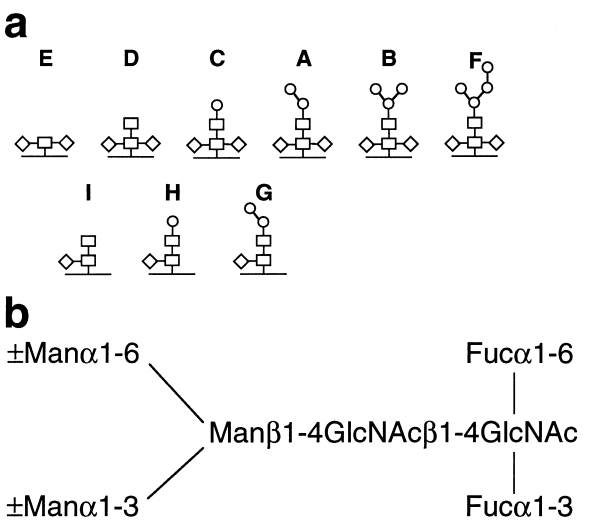

Fig. 3. Glycoforms carried by Aah VI. a: Structures determined by nanoESMS and nanoESMS/MS experiments, in order of increasing molecular weight from left to right. ◇: dHex, ○: Hex and □: HexNAc. b: Suggested probable common glycan structure of three of the main glycoforms, A, B and C. This speculation is based on (i) the determined linkages between hexoses, deoxyhexoses and *N*-acetylhexosamines for the various forms, as shown in panel a, (ii) information obtained from a series of attempts at deglycosylation, which demonstrated the presence of structures with high mannose content, and (iii) knowledge of the glycan biosynthetic pathways.

4. Discussion

HPLC techniques are continually improving and it is now possible to isolate minor fractions, with little or no pharmacological activity, from scorpion venoms. The Aah VI fraction was not identified by its anti-insect activity, but was isolated because it was eluted in the same gel filtration fraction as the contraction-causing insect-selective toxin Aah IT1 [5]. Amino acid composition determinations and automatic N-terminal sequencing of the native protein are effective for testing the core protein homogeneity, but are not appropriate for detecting glycan heterogeneity because isoforms have identical chromatographic properties. We used analytical mass spectrometry to show glycan heterogeneity of Aah VI. With this technique, we also demonstrated that Aah VI was reticulated by four disulphide bridges.

Toxicity tests on cockroaches strongly suggest that the very low anti-insect activity of Aah VI was not due to contaminating Aah IT1. If the contamination of Aah VI were 1/1000 (Aah IT1 $LD_{50} = 10$ ng/g, i.e. 850 times less than Aah VI), the amount of Aah IT1 injected in the Aah VI LD_{50} dose would have been close to the Aah IT1 LD_{50} and should have led to an immediate paralysis (5–10 s for Aah IT1). If there was less contamination, a lower lethality, but still an immediate paralysis effect, should have been observed, whereas the Aah VI-induced paralysis took about 1 h.

Aah VI has amino acid sequence motifs that are conserved in mammal-selective and insect-selective toxins, but its overall sequence is significantly different from both these well-defined groups. Its sequence differs extensively (Fig. 1D) from the sequences of the excitatory anti-insect toxins Aah IT1 [4] and Lqq IT1 [16]. It also differs from Lqh αIT [17] and Lqq III [18] (anti-insect α-toxin group), which are active against both mammals and insects. It differs to a lesser extent from the sequences of two depressant anti-insect toxins Lqq IT2 [16] and Lqh IT2 [19]. Remarkably, Aah VI is very similar in sequence to Aah STR1 (Fig. 1D), a non-glycosylated anatoxin purified from the venom of the same Old World species (Mansuelle et al., unpublished results), the overall folding of which is typical of toxins from the New World [20].

By use of NanoESMS and nanoESMS/MS, we succeeded in identifying the glycan motifs of the isoforms from only a few picomoles of sample. We used a very low flow rate (< 20 nmol/min), which made it possible to optimise fragmentation for nanoESMS/MS using very small volumes (3 μl) and amounts of material (concentration of glycopeptide 9–14 was about 1 pmol/μl). This is very much less than the amount of sample needed for ^{1}H-NMR analysis, which is about 50 μg of glycans, i.e about 70 nmol of Aah toxin VI, although ^{1}H-NMR gives much more complete information. Indeed, the most striking feature of Aah VI is that various glycan moieties are attached to it, increasing its molecular mass to 8500 Da and resulting in eluting from gel filtration among the long (70 residues) insect-selective neurotoxins. It cannot be excluded that some of the observed multiple glycosylated forms may result from the purification process, but in a recent study to identify the origin of glycan heterogeneity of an antibacterial peptide [21], we demonstrated that this heterogeneity was already present during the biosynthesis of the peptide.

Mass spectrometry detected a diversity of short *N*-glycan structures from $Hex_4(HexNAc)_2dHex_2$ to $HexNAc(dHex)_2$. These data and current knowledge of glycosylation biosynthesis suggest that the common glycan core structure of Aah VI is Asn-GlcNAc(α1-6Fuc)(α1-3Fuc)(β1-4GlcNAc). This sequence is consistent with (i) Aah VI glycans being partially sensitive to endoglycosidase F treatment (not shown), which cleaves the β1-4 linkage between the two innermost *N*-acetylglucosamine residues [22] and (ii) reports of such difucosylated structures for a phospholipase from honey bee venom [23]. Manα1-6Manβ1-4GlcNAcβ1-4GlcNAc has been isolated from honey bee venom glycoprotein [23] and from normal human blood mononuclear cell interleukin-6 [24]. Given this and what is known of *N*-glycan structures, the hexose residues of Aah VI glycans are probably mannose residues, which typically form a branched antenna on the di-*N*-acetylglucosaminyl core next to Asn (Fig. 3b). Thus, the glycosylation pathway in scorpion cells is very different from that in mammals, insects and plants [25–27]. There are at least two possible reasons for the presence of such unusually short *N*-glycan structures: (i) an original glycosylation pathway involving the sequential addition of various carbohydrates to the conserved glycosylation site (AsnXSer/Thr), as in *O*-linked glycan biosynthesis, or (ii) high glycosidase activities extensively shortening the usual *N*-glycans to a $GlcNAcFuc_2$ structure. A similar degradation of mature *N*-linked structures has been described in insect cells, in which high levels of β-*N*-acetylgalactosaminidase and β-*N*-acetylglucosaminidase activity may be responsible for the extensive shortening of complex glycans to structures with high mannose content [28].

Aah VI is the first glycosylated scorpion neurotoxin to be characterised. The characteristic consensus sequence NX(T/S) for a glycosylation site has, however, already been found in other toxins from Old World scorpions. For example, the NCT 11–13 sequence is present in the much studied Aah II toxin [29]. However, the involvement of Cys-12 in a disulphide bridge probably prevents glycosylation from occurring. The consensus sequences N(K/H/L)S at positions 7–9 and/or NK(S/T) at positions 61–63 are found in some New World toxins, but these toxins have never been reported to be glycosylated. It would be of value to systematically measure the precise molecular mass of newly characterised toxins, and also of other not so recently discovered toxins, especially those containing a potential glycosylation site, because other recently purified scorpion toxins have been found, by mass spectrometry, to be heterogeneously glycosylated (M. Alami, personal communication).

As far as we know, no glycosylated animal toxin with a size similar to that of Aah VI has been reported in the literature. The PLA_2 from honey bee venom is not toxic, whereas the toxic PLA_2 isolated from the Australian tiger snake (*Notechis scutatus scutatus*) was reported to be not glycosylated [30]. The toxic protein physalitoxin, from the venom of *Physalia physalis*, appears to be responsible for both the haemolytic and lethal activities of the venom [31], but its molecular mass is far higher than that of Aah VI. Thus, there is no actual glycosylated animal toxic protein to which Aah VI can be compared.

The resistance of Aah VI to complete deglycosylation should be further investigated to check whether this glycosylation motif contributes to the toxin activity. Note that Aah STR1, which has a sequence which is 88% identical to that of Aah VI, has no known activity. Aah VI may be the first

member of a new class of toxins with original mode of action and biological effect.

Acknowledgements: D.L., A.V.D. and O.S. would like to thank the TMR network for peptide and protein structure elucidation by mass spectrometry: ERBCHRXCT940425. We would also like to thank R. Oughideni for help with sequencing, Dr E. Fenouillet and Dr C. Kopeyan for helpful advice and fruitful discussion.

References

[1] Rochat, H., Rochat, C., Kupeyan, C., Miranda, F., Lissitzky, S. and Edman, P. (1970) FEBS Lett. 10, 349–351.
[2] Rochat, H., Bernard, P. and Couraud, F. (1979) in: Advances in Cytopharmacology (Cecarelli, B. and Clementi, F., Eds.), Vol. 3, pp. 325–334, Raven Press, New York.
[3] Kopeyan, C., Martinez, G., Lissitzky, S., Miranda, F. and Rochat, H. (1974) Eur. J. Biochem. 47, 483–489.
[4] Darbon, H., Zlotkin, E., Kopeyan, C., Van Rietschoten, J. and Rochat, H. (1982) Int. J. Pept. Protein Res. 20, 320–330.
[5] Zlotkin, E., Rochat, H., Kopeyan, C., Miranda, F. and Lissitzky, S. (1971) Biochimie 53, 1073–1078.
[6] De Lima, M.E., Martin, M.F., Diniz, C.R. and Rochat, H. (1986) Biochem. Biophys. Res. Commun. 139, 296–302.
[7] Miranda, F., Kupeyan, C., Rochat, H., Rochat, C. and Lissitzky, S. (1970) Eur. J. Biochem. 16, 514–523.
[8] Kharrat, R., Mansuelle, P., Sampieri, F., Crest, M., Oughideni, R., Van Rietschoten, J., Martin-Eauclaire, M.F., Rochat, H. and El Ayeb, M. (1997) FEBS Lett. 406, 284–290.
[9] Roecklin, D., Klarskov, K., Cavallini, B., Sabatié, J., Bouchon, B., Loew, D., Van Dorsselaer, A. and Bischoff, R. (1997) Eur. J. Biochem. 245, 589–599.
[10] Strub, J.M., Sorokine, O., Van Dorsselaer, A., Aunis, D. and Metz-Boutigue, M.H. (1997) J. Biol. Chem. 272, 11928–11936.
[11] Huddleston M.J., Bean, M.F., Barr, J.R. and Carr, S.A. (1991) Proceedings of the 39th ASMS Conference on Mass Spectrometry and allied Topics, Nashville, TN, 19–24 May, pp. 280–281.
[12] Zlotkin, E., Martinez, G., Rochat, H. and Miranda, F. (1975) Insect Biochem. 5, 243–250.
[13] Marshall, R.D. (1972) Annu. Rev. Biochem. 41, 433–442.
[14] Gavel, Y. and Von Heijne, G. (1990) Protein Eng. 3, 433–442.
[15] Snider, M.D. (1984) Biology of Carbohydrates, Vol. 2, pp. 163–198, John Wiley and Sons, New York.
[16] Kopeyan, C., Mansuelle, P., Sampieri, F., Brando, T., Bahraoui, E.M., Rochat, H. and Granier, C. (1990) FEBS Lett. 261, 423–426.
[17] Eitan, M., Fowler, E., Herrmann, R., Duval, A., Pehlate, M. and Zlotkin, E. (1990) Biochemistry 29, 5941–5947.
[18] Kopeyan, C., Mansuelle, P., Martin-Eauclaire, M.F., Rochat, H. and Miranda, F. (1993) Nat. Toxins 1, 423–426.
[19] Zlotkin, E., Eitan, M., Bindokas, V.P., Adams, M.E., Moyer, M., Burkhart, W. and Fowler, E. (1991) Biochemistry 30, 4814–4821.
[20] Blanc, E., Hassani, O., Meunier, S., Mansuelle, P., Sampieri, F., Rochat, H. and Darbon, H. (1997) Eur. J. Biochem. 247, 1118–1126.
[21] Uttenweiler-Joseph, S., Moniatte, M., Lambert, J., Van Dorsselaer, A. and Bulet, P. (1997) Anal. Biochem. 247, 366–375.
[22] Elder, J.H. and Alexander, S. (1982) Proc. Natl. Acad. Sci. USA 79, 4540–4544.
[23] Kubelka, V., Altmann, F., Staudacher, E., Tretter, V., März, L., Hard, K., Kamerling, J.P. and Vliegenthart, J.F.G. (1993) Eur. J. Biochem. 213, 1193–1204.
[24] Parekh, R.B., Dwek, R.A., Rademacher, T.W., Opdenakker, G. and Van Damme, J. (1992) Eur. J. Biochem. 203, 135–141.
[25] Goochee, C.F., Gramer, M.J., Adersen, D.C., Bahr, J.B. and Rasmussen, J.R. (1991) BioTechnology 9, 1347–1355.
[26] Lis, H. and Sharon, N. (1993) Eur. J. Biochem. 218, 1–27.
[27] Kornfeld, R. and Kornfeld, S. (1985) Annu. Rev. Biochem. 54, 631–664.
[28] Van Die, I., Van Tetering, A., Bakker, H., Van Der Eijnden, D.H. and Joziasse, D.H. (1996) Glycobiology 6, 157–164.
[29] Rochat, H., Rochat, C., Sampieri, F., Miranda, F. and Lissitzky, S. (1972) Eur. J. Biochem. 28, 381–388.
[30] Francis, B., Coffield, J.A., Simpson, L.L. and Kaiser, I.I. (1995) Arch. Biochem. Biophys. 318, 481–488.
[31] Tamkun, M.M. and Hessinger, D.A. (1981) Biochim. Biophys. Acta 667, 87–98.
[32] Roepstorff, P. and Fohlman, J. (1984) Biomed. Mass Spectrom. 11, 601.
[33] Bieman, K. (1990) Methods Enzymol. 193, 886–887.