The Structure of a Toxic Octapeptide, containing 4 D-Amino-acids, from the Larvae of a Sawfly, *Lophyrotoma interrupta*

Dudley H. Williams,[a] Sitthivet Santikarn,[a] Peter B. Oelrichs,[b] Francesco De Angelis,[a] John K. MacLeod,[c] and Richard J. Smith[a]

[a] *University Chemical Laboratory, Lensfield Road, Cambridge CB2 1EW, U.K.*
[b] *Department of Primary Industries, Animal Research Institute, Yeerongpilly, Queensland, Australia*
[c] *Research School of Chemistry, Australian National University, Canberra, Australia*

The structure of a toxic octapeptide from the larvae of a sawfly, *Lophyrotoma interrupta*, is determined as PhCO-D-Ala-D-Phe-L-Val-L-Ile-D-Asp-L-Asp-D-Glu-L-Gln.

The poisoning of cattle by ingestion of larvae of an Australian species of sawfly (*Lophyrotoma interrupta*) is a serious problem in several of the grazing areas of Queensland.[1] The first report of the isolation of a toxic factor is due to Leonard.[2] An improved method for isolation of an essentially pure toxin with an $L.D._{100}$ of 2 mg/kg was reported by Oelrichs *et al.*[3]

Amino-acid analysis showed the presence of the following amino-acids (molar ratios given in parentheses): Asp (1.89), Glu(1.93), Ala(1.00), Val(1.16), Ile(0.82), and Phe(0.92) (relative to Ala = 1.00). Electron impact (EI) mass spectrometry established benzoyl as an N-terminal blocking group, and determined[3] the partial sequence PhCO-Ala-Phe-Val-Ile. . . . We now report the determination of the complete structure as PhCO-D-Ala-D-Phe-L-Val-L-Ile-D-Asp-L-Asp-D-Glu-L-Gln.

The sequence of amino-acids in the toxin was determined by the recently developed method of fast atom bombardment (FAB) mass spectrometry.[4] The positive ion FAB mass spectrum of the toxin is reproduced in Figure 1 (spectra were obtained as described previously,[5] except that xenon atoms were used as bombarding particles). The presence of the abundant MH^+ ion at m/z 1040, in conjunction with the amino-acid analysis, is in accord with the presence of four carboxy-groups and one primary amide in the toxin. Four kinds of sequence ions are observed in the spectrum. Those sequence ions which contain the N-terminal portion of the peptide correspond to acylium ions (**1**), or protonated amide ions (**2**).[6] Ions corresponding to (**1**) occur at m/z 105, 176, 323, 422, 535, 650, 765, and 894 (and MH^+ at 1040) and indicate the sequence PhCOAlaPheValIleAspAspGluGln; or the C-terminal Gln could equally be replaced by the isomeric amino-acid carrying the amide on the backbone carbonyl, and the carboxy-function in the sidechain (iso-Gln). Ions corresponding to (**2**) at m/z 667, 782, and 911 (in conjunction with MH^+) confirm the nature of the three amino-acids constituting the C-terminal portion.

$$\underset{(\mathbf{1})}{R\text{–}\overset{+}{C}\text{=}O} \qquad \underset{(\mathbf{2})}{(RCONH_2)H^+}$$

Those sequence ions which contain the C-terminal portion of the peptide correspond to protonated amine fragments (**3**),[6] or the ion-radical fragments (**4**). Ions corresponding to (**3**) at m/z 865, 718, 619, and 506 (Figure 1) indicate the sequence XPheValIleY, where the masses of X and Y are in accord with the proposed sequence. Very low abundance ions, corresponding to (**4**), at m/z 849, 702, 603, and 490 support this sequence.

$$\underset{(\mathbf{3})}{(H_2NR)H^+} \qquad \underset{(\mathbf{4})}{(R\dot{C}HCO\text{–})H^+} \qquad \underset{(\mathbf{5})}{RCONH^-}$$

A number of the structural conclusions derived from the positive ion FAB spectrum could be confirmed from the negative ion FAB spectrum. This shows an extremely abundant $(M - H)^-$ ion at m/z 1038. A series of sequence ions, most generally expressed in terms of the anion (**5**), occur at m/z 338, 437, 550, 665, 780, and 909. These, together with the $(M - H)^-$ ion at m/z 1038, indicate the sequence XValIleAspAspGluGln (or, as before, the C-terminal residue may be iso-Gln).

The ambiguity of Gln or iso-Gln at the C-terminus of the toxin was removed by carrying out a Hofmann degradation on the toxin with [bis(trifluoroacetoxy)iodo]benzene.[7] The resulting peptide was totally hydrolysed and the products subjected to amino-acid analysis. The analysis established

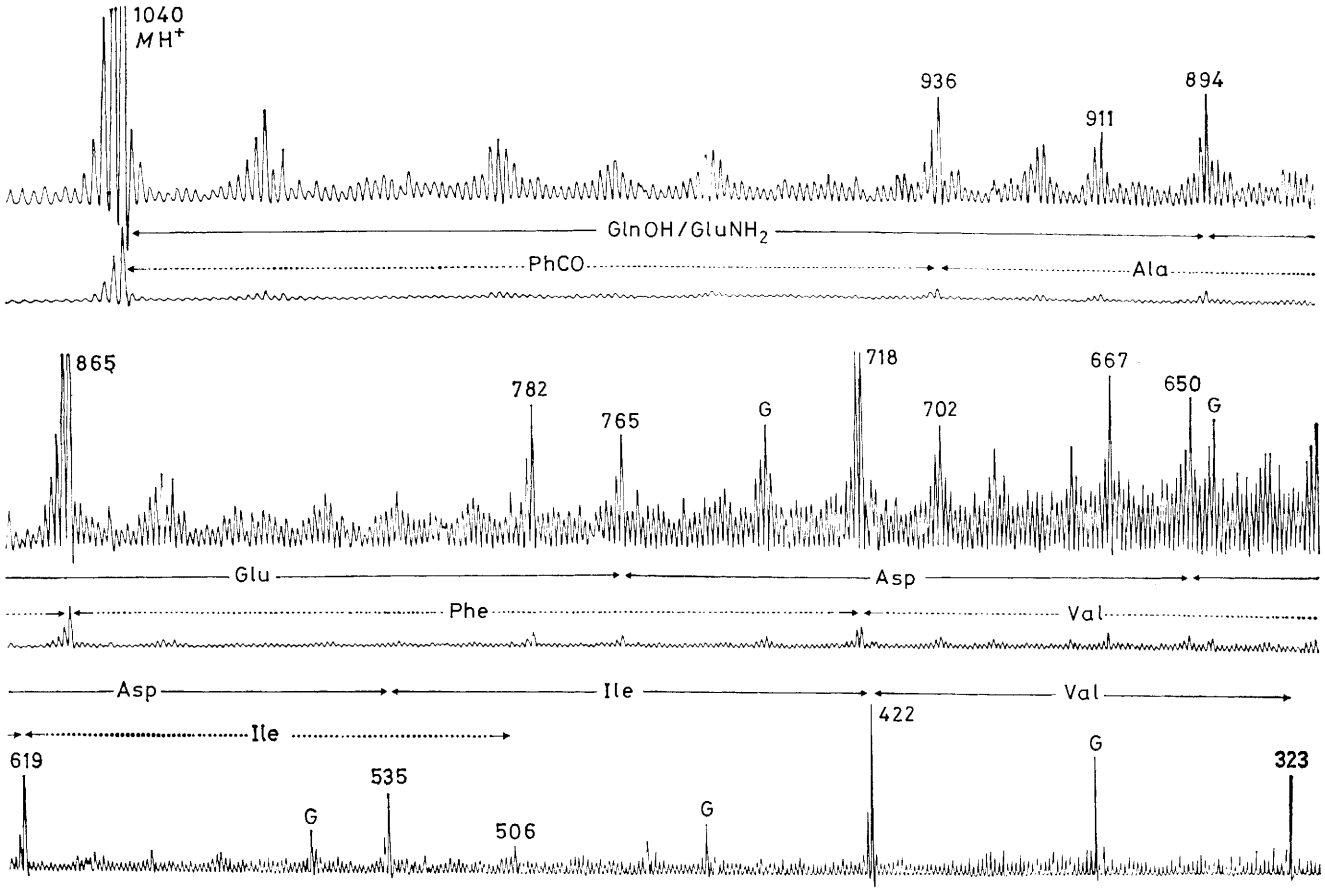

Figure 1. The positive ion FAB mass spectrum (obtained with Xe atoms) of the octapeptide (m/z 320—1050). The peaks marked G represent protonated polymers of glycerol which are 92 a.m.u. apart.

the production of (molar ratios in parentheses, relative to Ala = 1.00) Ala(1.00), Phe(0.96), Val(0.94), Ile(1.01), Asp(1.93), Glu(1.19), and a product with the same retention time as 2,4-diaminobutyric acid (DAB, 0.89). These experiments establish that the C-terminal residue is Gln and not iso-Gln.

The absolute configurations of the amino-acids in the toxin were determined by gas chromatography on a chiral column. A sample of toxin was subjected to total acid hydrolysis, and the resulting amino-acids were converted into their *N*-trifluoroacetyl isopropyl esters. Gas chromatography of these derivatives on a 'Chirasil-Val' column, and a comparison of their retention times with those of authentic samples, established that Ala, Phe, one Asp, and Glu or Gln, have the D-configuration; and that Val, Ile, one Asp, and Glu or Gln have the L-configuration. Peak assignments were confirmed by EI and chemical ionisation (CI) mass spectrometry.

The ambiguity of absolute configurations at Glu[7] and Gln[8] was resolved by carrying out a total acid hydrolysis on the previously mentioned crude product of Hofmann degradation, and analysing the *N*-trifluoroacetyl isopropyl ester derivatives of the resulting amino-acids on the chiral g.c. column. Co-chromatography with authentic derivatives of D and L Glu, and D and L DAB, established the hydrolysate to contain D and L Glu in the approximate ratio 93:7, and D and L DAB in the approximate ratio 5:95. Thus, the C-terminal residue of the toxin is L-Gln, and the penultimate residue Glu[7], has the D-configuration.

The ambiguity of absolute configurations at Asp[5] and Asp[6] was resolved by analysing products of partial acid hydrolysis of the toxin. Cleavage by dilute acid hydrolysis of a peptide residue involving an aspartate residue is known to be a facile process compared with other cleavages.[8] Partial acid hydrolysis of the toxin was effected by 10% aqueous HCl at 100 °C for 1.5 h. Fractions were separated by reversed-phase preparative h.p.l.c., and the fractions analysed by FAB mass spectrometry. One fraction (fraction 28) was observed to give very abundant ions in its positive ion FAB mass spectrum at *m/z* 493 and 247. These ions correspond to MH^+ signals from the peptides PheValIleAsp and IleAsp. Support for the assignment of the peak at *m/z* 493 came from the fragmentation pattern which established loss of C-terminal Asp [loss of 116 to give ion type (**2**), and loss of 133 to give ion type (**1**)]. Additionally, esterification of the peptide with MeOH–HCl gave, by FAB analysis, a diester (MH^+ at *m/z* 521); and acetylation with Ac_2O–H_2O gave a mono-acetyl derivative (MH^+ at *m/z* 535), as required by the proposed sequence. The *m/z* 247 ions also showed the anticipated increments in these experiments.

Total hydrolysis of a sample of the above fraction, and analysis of the *N*-trifluoroacetyl isopropyl ester derivatives of the resulting amino-acids on the chiral g.c. column, established that the D:L Asp ratio was approximately 5:1. Allowing for traces of other peptides in fraction 28, these data indicate that the absolute configuration at Asp[5] is D; and, in the light of data obtained previously, the absolute configuration at Asp[6] must be L.

We have therefore determined that the structure of the toxin, for which the name Iophyrotomin has been suggested previously,[3] is PhCO-D-Ala-D-Phe-L-Val-L-Ile-D-Asp-L-Asp-D-Glu-L-Gln.

The presence of four D-amino-acids in a peptide isolated from an animal is, so far as we are aware, without precedent. However, at this stage we cannot be certain that the peptide, or some or all of its D-amino acid components, is biosynthesised by one or more simpler organisms with which the larvae may interact.

We thank the S.E.R.C. (U.K.) for financial support of this work, Dr. E. Houghton and Mr. P. Teale for g.c.–m.s. results obtained at Race Course Security Services, Newmarket, U.K., and Dr. John Shaw, D.S.I.R., Palmerston North, New Zealand, for some of the 'Chirasil-Val' column results.

Received, 20th August 1982; Com. 1012

References

1 F. H. S. Roberts, *Queensl. Agric. J.*, 1932, **37**, 41; L. L. Callow, *ibid.*, 1955, **81**, 155.
2 G. J. Leonard, *Toxicon*, 1972, **10**, 597.
3 P. B. Oelrichs, P. J. Vallely, J. K. MacLeod, J. Cable, D. E. Kiely, and R. E. Summons, *Lloydia*, 1977, **40**, 209.
4 M. Barber, R. S. Bordoli, R. D. Sedgwick, and A. N. Taylor, *J. Chem. Soc., Chem. Commun.*, 1981, 325.
5 D. H. Williams, C. V. Bradley, G. Bojesen, S. Santikarn, and L. C. E. Taylor, *J. Am. Chem. Soc.*, 1981, **103**, 5700.
6 D. H. Williams, C. V. Bradley, S. Santikarn, and G. Bojesen, *Biochem. J.*, 1982, **201**, 105.
7 A. D. Radakrishna, M. E. Parham, R. M. Riggs, and G. M. Loudon, *J. Org. Chem.*, 1979, **44**, 1746.
8 S. M. Partridge and H. F. Davis, *Nature*, 1950, **165**, 62.