The Structure of Cyanoginosin-LA, a Cyclic Heptapeptide Toxin from the Cyanobacterium *Microcystis aeruginosa*

Dawie P. Botes,* Albert A. Tuinman, Philippus L. Wessels, Cornelis C. Viljoen, and Heléne Kruger
National Chemical Research Laboratory, CSIR, P.O. Box 395, Pretoria, 0001, Republic of South Africa
Dudley H. Williams, Sitthivet Santikarn, Richard J. Smith, and Stephen J. Hammond
University Chemical Laboratory, Lensfield Road, Cambridge, CB2 1EW

The structure of cyanoginosin-LA (previously referred to by us as toxin BE-4) is cyclo(-D-Ala-L-Leu-*erythro*-β-methyl-D-isoAsp-L-Ala-Adda-D-Glu-*N*-methyldehydroAla) (**17a**), where Adda refers to the novel β-amino acid residue of 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (**11**). The connectivity of Adda is deduced from ^{1}H- and ^{13}C-n.m.r. data, but the stereochemistry at carbons 2, 3, 8, and 9 remains unknown. The residue sequence is derived from the mass spectra of cyclic and linear derivatives of the toxin. A new nomenclature for the related toxins of *Microcystis aeruginosa* is discussed.

Since the first reported instance of algal poisoning of cattle in South Australia[1] and the direct link established between *Microcystis aeruginosa* blooms and livestock deaths in South Africa,[2] several reports of similar episodes have been recorded.[3,4] The pathological perturbations arising from acute toxicity to mammals are well documented.[4,5] Because *Microcystis* toxins escape destruction by the normal chlorination, flocculation, and filtration procedures used in water treatment,[6] sporadic blooms of the algae may also represent a considerable hazard to humans. Human illness implicating cyanobacteria has been investigated.[7,8]

Attempts by various laboratories to characterize the toxic principle from *Microcystis* blooms have been made over a period of more than three decades.[2,9–12] Although the peptide-like nature was recognized in 1959,[13] compositional consistency with respect to the amino acid content of toxins isolated from different blooms was only obtained considerably later.[4,14] In previous communications from this laboratory the configurational assignments of the amino acid residues was discussed,[15] and a partial structure was reported for one of the toxins.[16] The *N*-methyldehydroalanine residue (Mdha) was tentatively positioned as an exocyclic branch to the cyclohexapeptide.

The present report corrects that tentative assignment and places Mdha within a monocyclic heptapeptide (**17a**) as well as unequivocally establishing the iso-linkage of the β-methyl-aspartic acid (Masp) residue.

Nomenclature

Previously employed designations such as microcystin,[17] aeruginosin,[18] or simply microcystis toxin with either alphabetical or numerical suffixes to indicate chromatographic elution order,[14–16] seem unsatisfactory since neither the cyanobacterial origin nor the chemical structure are unequivocally defined by these names. We propose the generically derived designation 'cyanoginosin' with a two-letter suffix to indicate the two variant amino acid residues in the otherwise invariant[15] cyclic heptapeptide. The first letter of the suffix indicates the residue *N*-terminal to the invariant β-methyl-iso-Asp; *i.e.* (**17a**) is termed cyanoginosin-LA. If the positions of the variants are unknown, the suffix is placed in parentheses.

Results and Discussion

Amino Acid Residues.—The summed masses of the six previously reported[15] amino acid residues of cyanoginosin-(AL) (previously referred to by us as toxin BE-4[14–16]) is 562 daltons (Table 1). The fast-atom-bombardment mass spectra (f.a.b.m.s.) of the native and borohydride-reduced (Mdha→ *N*-MeAla) toxins indicate molecular masses of 909 and 911 daltons, respectively, leaving a mass of 347 daltons unaccounted for.

The 500 MHz ^{1}H n.m.r. spectrum of cyanoginosin-(AL) in D_2O is reproduced in Figure 1. A striking feature of the spectrum is the olefinic and aromatic resonance patterns between δ 5—8 not accounted for by the amino acid residues listed in Table 1. No resonances occur beneath the HDO signal as no peaks were observed in this region in a spectrum obtained in $(CD_3)_2SO$ solution. Six proton resonances are resolved between δ 3.8—4.6 where signals for C_α protons are expected.[19]

Analysis of the resolution-enchanced 500 MHz ^{1}H spectrum in D_2O, involving numerous homonuclear proton–proton decoupling experiments, leads to the partial structures (**1**)—(**5**), thus confirming five of the six entries in Table 1. Only about half of the observed signals have thus far been accounted for, and two additional partial structures, (**6**) and (**7**), can be deduced from the remainder. Selective irradiation of the aromatic resonances at δ 7.28—7.36 sharpened the aliphatic resonances at δ 2.94 and δ 2.71. The three-proton signal at δ 1.72 is characteristic of a methyl group on a double bond.[20] The olefinic proton which resonates at δ 5.54 (d, *J* 9.7 Hz) must be adjacent to a quaternary carbon which probably also carries the methyl group (δ 1.72). The chemical shifts of the three olefinic protons in (**6**) and (**7**) suggest that they are part of a conjugated system, in line with the reported chemical shifts for the corresponding protons in 1,2,4-trisubstituted buta-1,3-dienes.[21] The partial structures (**6**) and (**7**) can therefore be combined as in (**8**).

Additional proof for structure (**8**) was obtained from the negative nuclear Overhauser effects (n.O.e.s)[22] observed in the 400 MHz spectrum of cyanoginosin-(AL) in $(CD_3)_2SO$ at 303 K. The chemical shifts for the protons in the olefinic portion are indicated in structure (**8**). Irradiation at δ 6.09 gave a −15% n.O.e. to the proton at δ 5.46 and an n.O.e. of the same size was observed in the opposite direction for these protons. The methyl group at δ 1.55 received a −12% n.O.e. when the proton resonating at δ 5.36 was selectively irradiated. These observations establish the *trans–trans* configuration of the diene (**8**).

As yet unassigned resonances in the spectrum are two one-proton doublets at δ 5.91 and δ 5.55 with splittings of 1.6 Hz, and two three-proton singlets at δ 3.28 and δ 3.40, the latter chemical shifts being characteristic of OCH_3 and/or NCH_3 protons. The doublets and one of the singlets are accounted for by the *N*-methyldehydroalanine residue known to be present in

Table 1. Amino acids identified after total hydrolysis of cyanoginosin-(AL)[14]

Amino acids	Residue structure	Residue abbreviation	Residue mass/dalton	Comment
D-Alanine	NH–CH(CO)–CH_3	D-Ala	71	invariant
erythro-β-Methyl-D-aspartic acid	NH–CH–(CO)–CH(CH_3)–CO	Masx	112	invariant
D-Glutamic acid	NH–CH(CO)–CH_2–CH_2–CO	D-Glx	112	invariant
N-Methyldehydroalanine	N(CH_3)–C(CO)=CH_2	Mdha	83	invariant
L-Alanine	NH–CH(CO)–CH_3	L-Ala	71	variant
L-Leucine	NH–CH(CO)–CH_2–CH(CH_3)$_2$	L-Leu	113	variant

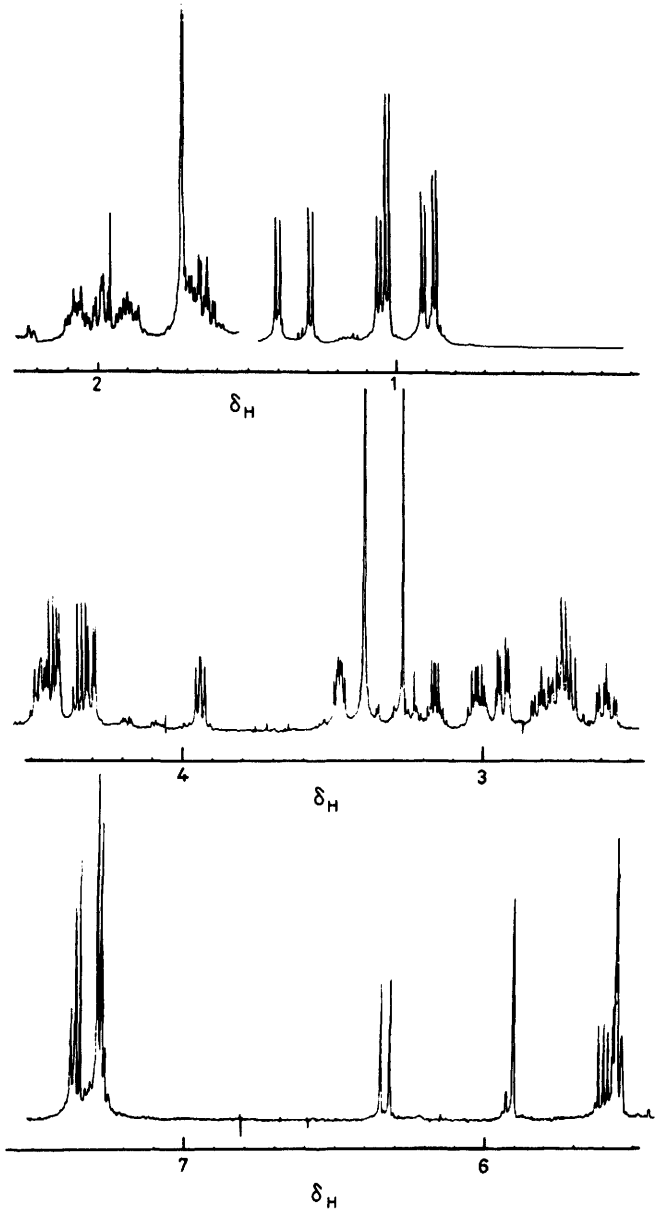
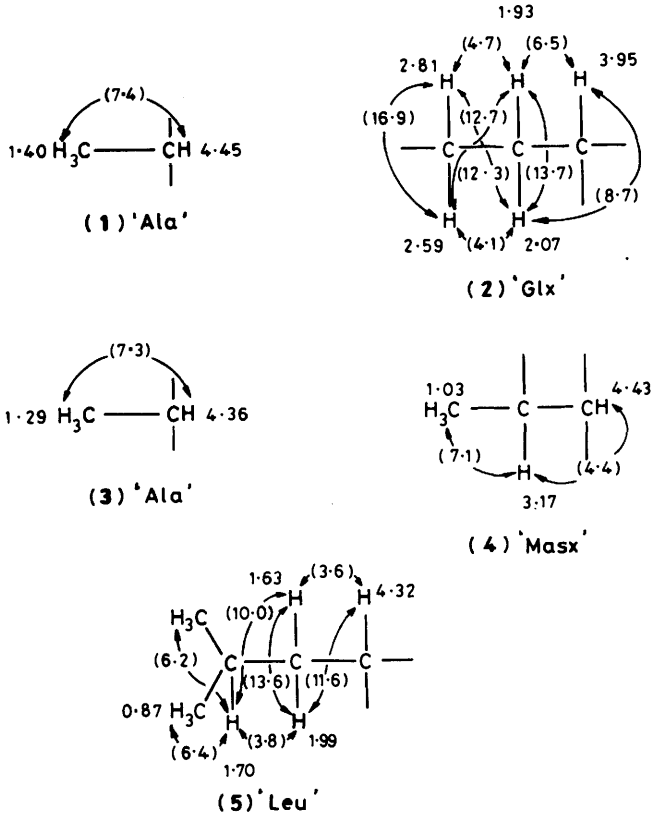


Figure 1. The 500 MHz ^{1}H n.m.r. spectrum of cyanoginosin-(AL) in D_2O

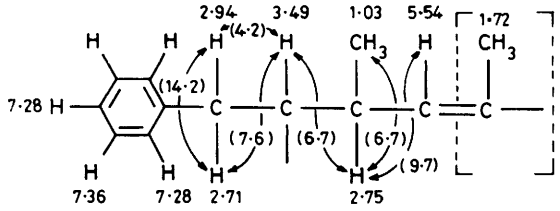


the toxin (*cf.* Table 1). Thus, the only proton resonance not accounted for by the residues listed in Table 1 and the partial structure (**8**) is a three-proton singlet arising from OCH_3 or NCH_3.

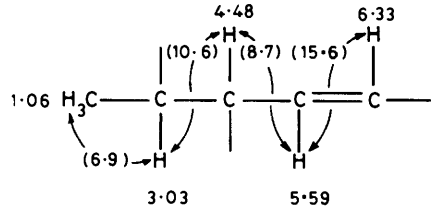
Proton–proton decoupling experiments in $(CD_3)_2SO$ at 303 K show unequivocally that the five C_α protons of the residues of Table 1, and the C-3 proton of (**8**), are coupled to exchangeable protons with resonance positions (δ 7.5—8.7) characteristic of amide protons. The substituent Y of (**8**) is therefore the NH of an amide.

The substituents X and Z of (**8**) may be inferred from the 125 MHz ^{13}C n.m.r. spectrum of cyanoginosin-(AL). The numbers of methyl, methylene, methine, and quaternary carbon resonances observed (Table 2) are in agreement with the structural units identified previously, but not all the carbonyl resonances could be detected under the experimental conditions employed. The methyl carbon at δ_C 58.16 [^1J(CH) 142.1 Hz and $^{>1}J$(CH) 5.0 Hz] is characteristic of an OCH_3 group directly bonded to a methine carbon. This, together with the chemical shift of δ 3.49 for the C-9 proton of (**8**), indicates that the substituent **Z** is methoxy. The six methine carbon resonances observed between δ_C 48—58 are attributed to the five α-carbon atoms of the Table 1 residues (Mdha has no methine α-carbon) and C-3 of (**8**). The C-8 of (**8**) resonates at δ_C 44.98, being adjacent to a double bond and the methoxy substituent. The signals at δ_C 42.26 must then be due to C-2 of (**8**). With a proton chemical shift of δ 3.03, the substituent Z in (**8**) is most likely a carbonyl group.

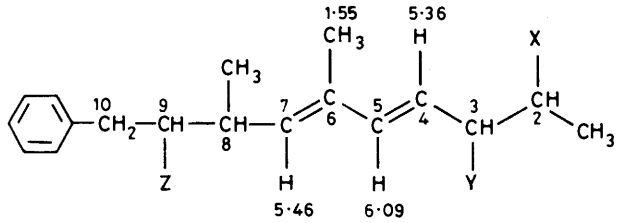
The partial structure (**8**), with X, Y, and Z assigned as



(6)



(7)



(8)

described above, is fully substantiated by the electron-impact mass spectra (e.i.m.s.) of permethylated and perdeuteriopermethylated cyanoginosin-(AL). The methoxy group β to phenyl is attested by abundant ions at *m/z* 91 and 135.0831 (calc. for $C_9H_{11}O$: *m/z* 135.0810) in both spectra. Further abundant ions at *m/z* 326 (329) and 272 (275) fit the expected fragments (**9**) and (**10**) respectively.

The component of cyanoginosin-(AL) not included in Table 1 is therefore the residue of 3-amino-9-methoxy-2,6,8-trimethyl-10-phenyldeca-4,6-dienoic acid (**11**) (Adda). The configurations at asymmetric carbon atoms 2, 3, 8, and 9 of this β-amino acid residue have not as yet been determined.

With the connectivity of this novel residue (mass 313 daltons) known, and with all resonances in the ^{1}H- and ^{13}C-n.m.r. spectra accounted for, it is clear that the remaining 34 daltons required to make up the molecular mass of 909 must comprise two OH groups forming two carboxylic acids. This was verified by the f.a.b.m.s. of borohydride-reduced toxin treated with methanol–hydrogen chloride, *i.e.* an $(M + H)^+$ ion at 940 daltons shows the esterification of two carboxylic acid moieties of the toxin.

The total amino acid composition of cyanoginosin-(AL) therefore comprises the entries in Table 1 plus the β-amino acid residue Adda (**11**), with two of the nine carbonyl moieties present as parts of carboxylic acid groups. To accommodate two not fully substituted entities, the structure must be monocyclic.

Sequence.—Pulse hydrolysis of reduced cyanoginosin-(AL) (6M HCl; 5 min; 100 °C) afforded a mixture of linearised peptides which successfully underwent two cycles of Edman degradation establishing Ala-Leu as the *N*-terminal sequence. The third step of the Edman degradation failed, but the third residue was identified as Masx (**4**) by dansylation,* total hydrolysis, and t.l.c. comparison with authentic Dns-β-methylaspartic acid. On high-performance liquid chromatography (h.p.l.c.) two products, (**12**) and (**13**), were isolated, with molecular weights of 812 and 794 daltons, respectively, by f.a.b.m.s. Glutamic acid is indicated as the *C*-terminal residue in each case by the loss of 130 daltons from the quasi-molecular ion.[23] Identical products were obtained when the pulse hydrolysis was carried out on the deuteriated analogue [containing $N(CH_3)$–$CH(CO)$–CH_2D], showing that this hydrolysis cleaved Mdha from the peptide as well as opening the monocyclic structure. Partial sequence (**12**) summarises these results.

Ala-Leu-Masx-X-X-Glx
(**12**)
(**13**) = (**12**) − H_2O

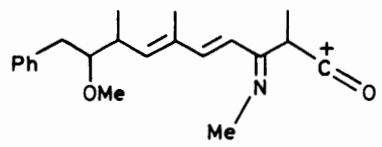
The linear hexapeptide (**12**) was acetylated with a mixture of [$^1H_6$]- and [$^2H_6$]-acetic anhydride (1:1) and then perdeuterio-

Table 2. The 125 MHz ^{13}C n.m.r. data of cyanoginosin-LA in D_2O

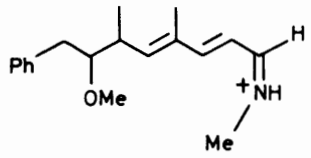
δ[a]	Multiplicity	^1J(CH)/Hz	$^{>1}J$(CH)/Hz	Assignment
178.10	s			carbonyls[b]
177.89	s			
176.37	s			
167.39	s			
144.58	s			
140.47	s			
138.72	d	160.9		Adda and C_α of Mdha
136.95	d	155		
134.37	s			
130.58	d	159		
129.48	d	163		
127.26	d	162.1	8	
125.70	d	157.3		
116.43	t	165.1		C_β of Mdha
87.78	d	151.8		C-9 of Adda
58.16	q, d	142.1	5	OCH_3
57.62	d	144		C_α of D-Ala, L-Ala, Leu, Glx, Masx and C-3 of Adda
56.87	d	*ca.* 140		
56.19	d	138.9		
54.71	d	137.4		
50.59	d	142.1		
48.81	d	136.3		
44.98	d	136.7		C-8 of Adda
42.26	d	131		C-2 of Adda
40.36	t	129		C_β of Leu
38.86	q	140.5		NCH_3
37.87	t	124.0		C-9 of Adda
36.71	d	123.8		C_β of Masx
32.72	t	126		C_γ of Glx
27.70	t	127		C_β of Glx
25.35	d	127.4		C_γ of Leu
23.39	q	126		C_δ of Leu
20.99	q	126		C_δ of Leu
17.38	q	129.2		C_β of D-Ala and L-Ala
17.14	q	130.6		
16.71	q	130		CH_3 of Masx
15.41	q	125.4		CH_3 of Adda
14.99	q	130.0		
13.10	q	126		

[a] Relative to $SiMe_4$ measured from internal dioxane and corrected by using $\delta_C(SiMe_4) = \delta_C(\text{dioxane}) + 67.8$ p.p.m. [b] Not all the carbonyl resonances could be detected under the experimental conditions used.

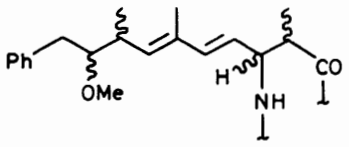
* Dansyl = 5-dimethylaminonaphthalene-1-sulphonyl



(9) *m/z* 326·2112
(calc. 326·2120)

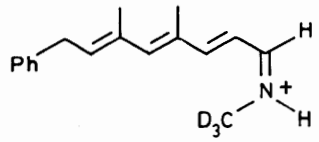


(10) *m/z* 272·2011
(calc. 272·2014)

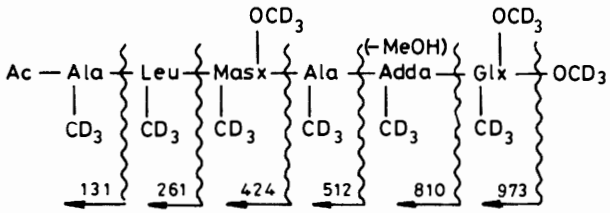


(11) Adda

methylated. The e.i.m.s. (Figure 2) of the product shows a very abundant fragment ion at *m/z* 243 indicating the loss of methanol from Adda during hydrolysis, *i.e.* due to ion (**14**). Furthermore, pairs of acylium sequence ions at *m/z* 131/134, 261/264, 424/427, 512/515, 810/813, and 973/976 establish the sequence (**15**) for the pulse-hydrolysed product. The Mdha which is 'missing' in (**15**) could originally have been attached to the free carbonyl of Masx, or at either of the *C*-terminal Glx carbonyls.



(14) *m/z* 243·1946
(calc. 243·1941)



(15)

Figure 3. H.p.l.c. trace of reduced cyanoginosin-(AL) after pulse hydrolysis with trifluoroacetic acid

Milder pulse hydrolysis of reduced cyanoginosin-(AL) using trifluoroacetic acid (TFA) resulted in several products which were separated by h.p.l.c. (Figure 3). One of these (peak C) was shown to be the ring-opened peptide which has not undergone loss of the Mdha residue or methanol from Adda. The f.a.b.m.s. negative-ion fragmentation pattern[23] (Figure 4) indicates the *C*-terminal sequence -Glu-Mdha-OH, thus associating Mdha and Glx rather than Masx in the cyclic toxin. The major product of hydrolysis (D, Figure 3) was the linearised peptide with loss of methanol from Adda, displaying abundant ions at *m/z* 896, 824, and 681, analogous to those in Figure 4.

Ala-Leu-Masx-Ala-Adda-Glu-Mdha
(**16**)

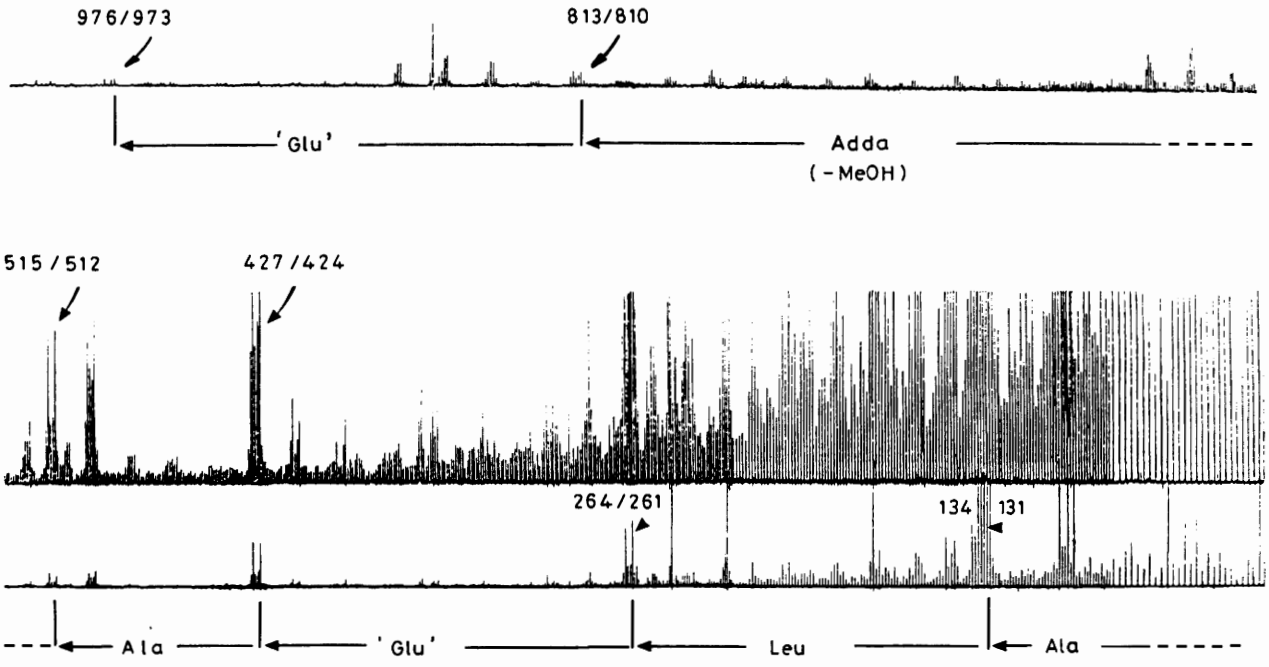
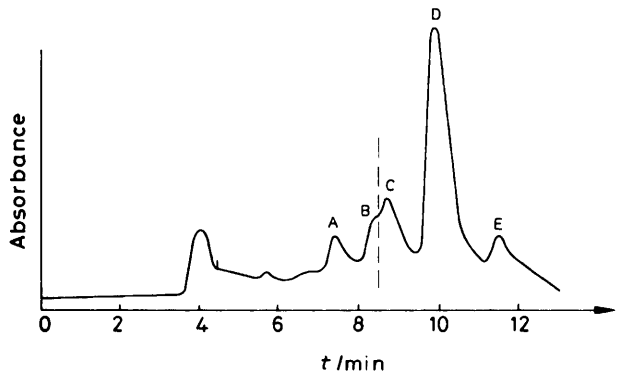


Figure 2. Electron-impact mass spectrum of the ($^1H_6/^2H_6$) acetylated derivative (**15**)



Having thus established the residue sequence for linearised cyanoginosin-(AL) as (**16**), we need to answer three questions. (1) Which of the two Ala residues [at positions 1 and 4 of (**16**)] is the variant (L-configuration) and which is the invariant (D-configuration)? (2) Is the cyclic linkage from the *N*-terminal of Ala to (*a*) the free carbonyl of Masx, (*b*) the free carbonyl of Glx, or (*c*) the carbonyl of Mdha? (3) Are the Masx and Glx residues α- or iso-linked in the cyclic toxin? The first question was resolved by the following experiment. Analysis of the trifluoroacetyl isopropyl esters of the amino acids derived from total hydrolysis of the pulse-hydrolysed product (**13**) by a g.c.-m.s system employing a chiral capillary column showed D-Ala:L-Ala in the ratio 1:1.5. After one Edman degradation cycle on (**13**) the f.a.b.m.s. of the remaining peptide[24] showed that *N*-terminal Ala was cleanly removed. Analysis of the remaining amino acid components as above showed L-Ala as before, but no D-Ala was detected. Thus, D-Ala is the *N*-terminal residue and L-Ala is the fourth residue from the *N*-terminus of the sequence (**16**). The imprecise nomenclature, cyanoginosin-(AL), henceforth becomes cyanoginosin-LA.

Questions (2) and (3) above are addressed *via* ^{1}H n.m.r. spectra recorded at varying pH.[19] Spectra recorded at intervals after gradual addition of small amounts of NaOD to a solution of cyanoginosin-LA in D_2O showed a significant upfield shift (Δδ 0.28) of only one signal: the C_α proton of the glutamic acid residue. However, when DCl was added to a freshly prepared solution of the toxin, the C_α protons of both Masx (Δδ +0.19) and Glx (Δδ +0.40) as well as the C_β protons of Glx (δ 2.07, Δδ +0.27) experienced substantial downfield shifts. The other proton signals are hardly affected (Δδ −0.08 to 0.06). This suggests that the two free carboxylic acid moieties in cyanoginosin-LA are at the α-positions of these residues. It thus establishes the residues containing free carboxy groups (Masx = Masp, Glx = Glu), the iso-linkage of both Masp and D-Glu, and the cyclisation from D-Ala to Mdha, these being the only remaining residues with a free *N*- and *C*-terminus, respectively.

Further conclusive evidence supporting these results is provided by base-catalysed tritium incorporation according to Matsuo and Narita.[25] By this method tritium is generally only introduced into *C*-terminal residues. However, for Asp and Glu, incorporation will also be effected in non-*C*-terminal positions, provided these are iso- rather than α-linked. Table 3 presents the results of tritium incorporation of cyanoginosin-LA (**17a**), its reduced analogue (Mdha→*N*-MeAla) (**17b**), and the latter after ring opening with trifluoroacetic acid, *i.e.* compound (**16**). Strong incorporation into D-Glu and Masp for all three experiments verifies the carboxylic nature and iso-linkage of these residues. The dramatic increase in incorporation into *N*-MeAla after ring opening substantiates the previously assigned (Figure 4) *C*-terminal position of Mdha in (**16**) and thus the cyclising bond from Mdha to D-Ala in the intact toxin.

The complete structures (excluding stereochemistry in Adda) of cyanoginosin-LA and its reduced and permethylated derivatives are thus represented by (**17**)—(**19**).

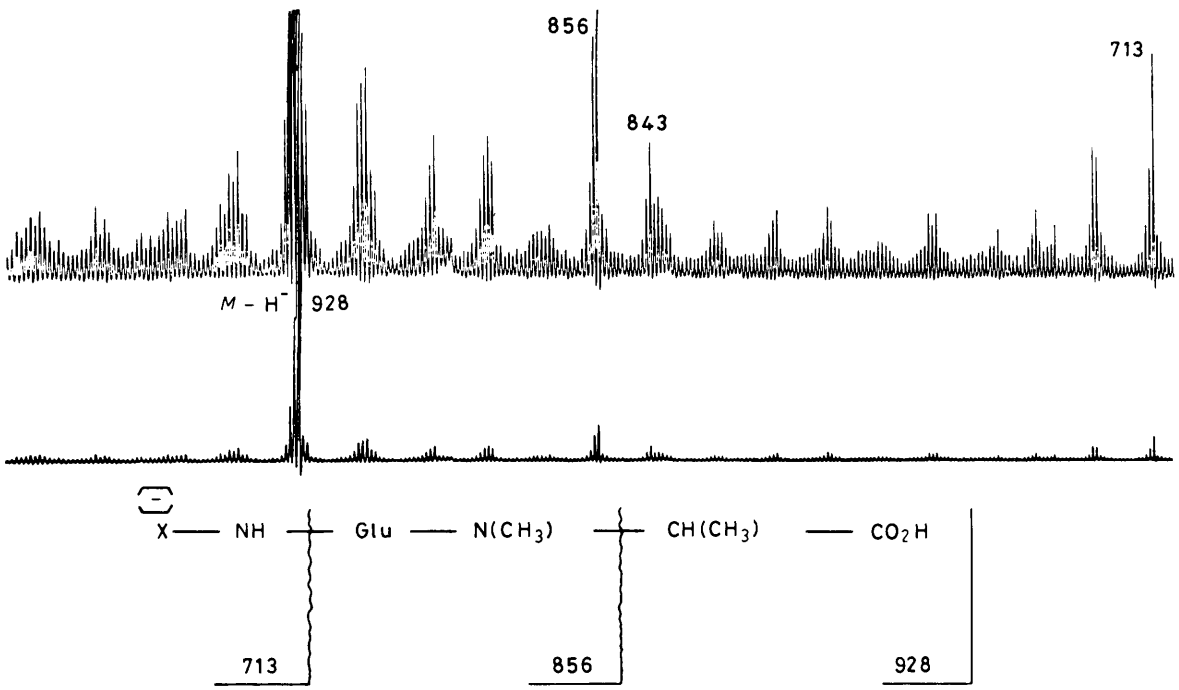
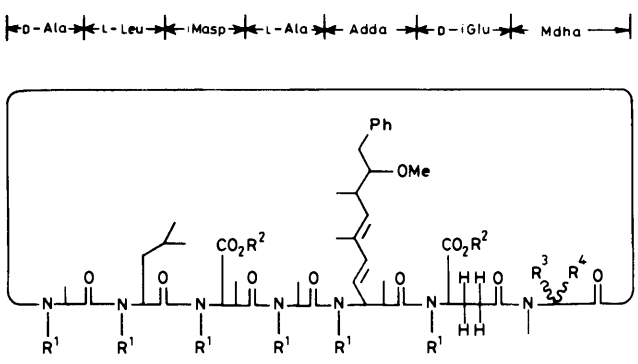


Figure 4. Negative fast-atom-bombardment mass spectrum of a pulse-hydrolysis product (peak C, Figure 3) of reduced cyanoginosin-(AL)

Table 3. Radioactivity counts[a,b] on isolated amino acids after tritium labelling[25] of cyanoginosin-LA derivatives and subsequent total hydrolysis

Derivative	(17a)	(17b)	(16)
D-Glu	4 850 (1.00)	13 127 (1.00)	57 964 (1.00)
Masp	1 848 (0.99)	3 692 (0.73)	15 171 (0.68)
N-MeAla		1 763 (0.11)	56 869 (0.81)
2 × Ala	228 (0.02)	188 (0.01)	305 (0.00)
Leu	366 (0.03)	245 (0.01)	493 (0.00)

[a] Counts min^{-1}. [b] Values in parentheses are relative tritium incorporations after correction for the amino acid sensitivity to the reaction.[25] Model compounds for the correction factors are *C*-terminal Glu, Asp, and Ala for D-Glu, Masp, and *N*-MeAla, respectively.



(17) **a**; $R^1 = R^2 = H$, $R^3R^4 = CH_2$
b; $R^1 = R^2 = R^3 = H$, $R^4 = CH_3$
c; $R^1 = R^3 = H$, $R^2 = R^4 = CH_3$
(18) **a**; $R^1 = R^2 = CH_3$, $R^3R^4 = CH_2$
b; $R^1 = R^2 = R^3 = CH_3$, $R^4 = H$
(19) **a**; $R^1 = R^2 = CD_3$, $R^3R^4 = CH_2$
b; $R^1 = R^2 = CD_3$, $R^3 = H$, $R^4 = CH_3$

Confirmation of Sequence.—The amino acid residue sequence (**17a**) relies on the partial sequence (**16**), which was deduced solely from pulse-hydrolysed derivatives of cyanoginosin-(AL). Transannular bond-formation during acid treatment may result in rearrangement of the residues in the hydrolysed product.[26] It is therefore important to compare results obtained from such hydrolysis products with data from the intact cyclic peptide.

Selected ions from the electron-impact mass spectra of three cyclic derivatives of cyanoginosin-LA were determined by a computer programme[27] to be uniquely assignable to the partial structures given in Table 4. The residue combinations thus determined fit the sequence as previously established (Figure 5)

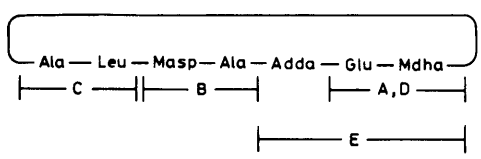


Figure 5. Partial sequences corresponding to fragments A—E of Table 4

and a rearrangement during the hydrolysis seems unlikely. Fragment E of Table 4 is particularly significant since it directly associates the residues Adda, Mdha, and 'Glu' (either D-Glu or the isomeric Masp), providing further evidence that the tentative 'branched' structure previously proposed[16] is incorrect. A plausible fragmentation pathway explaining the loss of NCH_3, H, and $PHCH_2CHOCH_3$ from Adda is proposed in Figure 6.

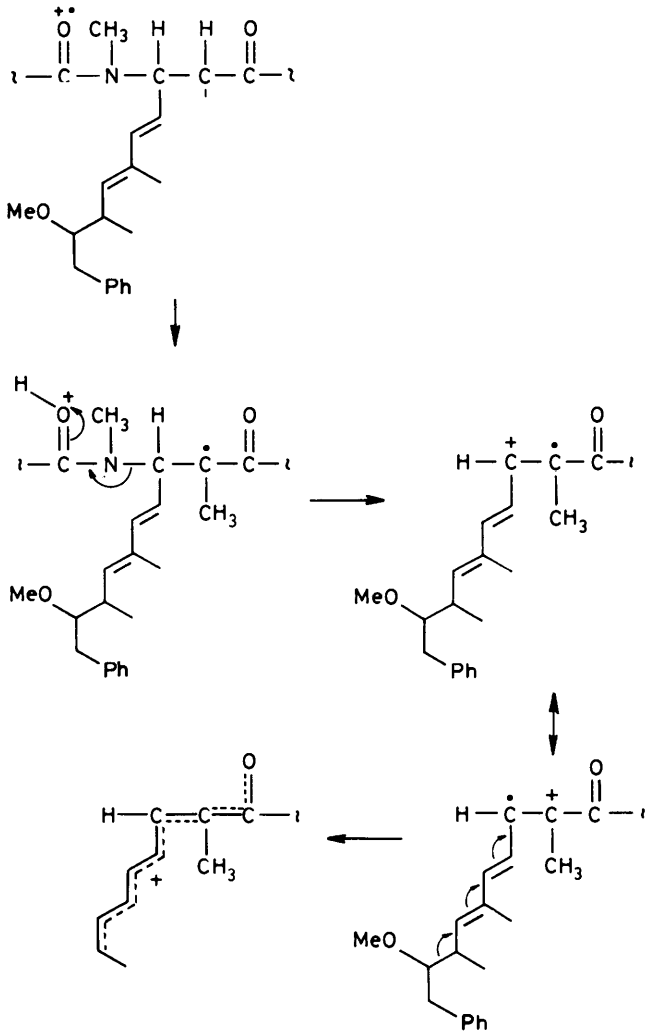


Figure 6. Possible mechanism for the loss of $Ph-CH_2-CH(OMe)$, NCH_3, and H from Adda-containing fragments

Experimental

Isolation of the Toxin.—Cyanoginosin-LA (**17a**) and its borohydride-reduced analogue (**17b**) were isolated as previously described,[14,15] *m/z* 910 and 912 respectively (MH^+) (by f.a.b.m.s.).

N.m.r. Spectra.—Proton n.m.r. spectra were recorded in D_2O or $(CD_3)_2SO$ solutions using Bruker WM500 and WH400 spectrometers. The ^1H chemical shifts in $(CD_3)_2SO$ are relative to $SiMe_4$ using the formula $\delta\ (SiMe_4) = \delta\ [(CD_3)_2SO] + 2.49$ with the $(CD_3)_2SO$ as internal reference. The ^{13}C chemical shifts are relative to $SiMe_4$ using dioxane as internal reference at δ 67.8 p.p.m. The ^{13}C chemical shift values were obtained from a broad-band proton-decoupled 125 MHz ^{13}C n.m.r. spectrum and the multiplicities and carbon–proton coupling constants from DEPT* sequences with and without broad-band proton decoupling during acquisition, respectively.

Mass Spectra.—Mass spectra were recorded on a Kratos Ms-50 spectrometer fitted with a highfield magnet (3 000 dalton at

* DEPT (Distortionless Enhancements by Polarization Techniques), D. M. Doddrell, D. T. Pegg, and M. R. Bendall, *J. Magn. Reson.*, 1982, **48**, 323.

Table 4. Selected masses and derived partial sequences from cyclic derivatives of cyanoginosin-LA

Fragment	Derivative (18a)	Derivative (18b)	Derivative (19a)	Calc.	Interpretation[a]
A	156 (156.0677)	156	159 (159.0864)	159.0849 ($C_7H_7D_3NO_3$)	'Glu'–Mdha, – H, – NCD_3, – $C(CH_2)CO$[b]
B	212 (212.0925)	212	218	212.0923 ($C_{10}H_{14}NO_4$)	(Ala + 'Glu'),[b] – NCH_3, – H
C	241	241 (241.1556)	247	241.1552 ($C_{12}H_{21}N_2O_3$)	(Ala + Leu), + CO, + H
D	241	243 (243.1338)	247	243.1345 ($C_{11}H_{19}N_2O_4$)	(Mdha + 'Glu'),[b] + H
E	402 (402.2158)	404	408	402.2155 ($C_{22}H_{30}N_2O_5$)	(Adda + Mdha + 'Glu'),[b] – NCH_3, – H, – R[c]

[a] These are the only possible assignments using standard *cyclo*-peptide fragmentation patterns.[28–31] [b] 'Glu' signifies D-Glu or the isomeric Masp. [c] – R signifies the loss of $Ph–CH_2–CH–OCH_3$ from Adda.

full accelerating voltage) and a Varian-MAT 212. Electron-impact mass spectra were obtained from *ca.* 100 nmol samples in a quartz crucible at probe temperature intervals of *ca.* 10 °C. Electron energy was held at 70 eV. Samples for fast-atom-bombardment spectra (*ca.* 500 pmol) were evaporated on a copper probe tip from a methanol solution. The glycerol matrix (*ca.* 1 μl) was added to form a thin layer over the entire surface of the probe tip which was then bombarded with 6 keV xenon atoms in the spectrometer source at ambient temperature.

High-pressure Liquid Chromatography (*h.p.l.c.*).—Reverse-phase h.p.l.c. was carried out using a Du Pont 850 instrument fitted with a μ-Bondapak C_{18} column (0.7 × 30 cm; Waters Associates). A linear gradient of acetonitrile (45%→70% in 30 min) in 0.1% aqueous THF at 2 ml min^{-1} was employed at 35 °C. Peaks were detected by u.v. absorption above 205 nm.

Esterification of Reduced Cyanoginosin-LA (**17b**).—Dried hydrogen chloride gas was bubbled through redistilled anhydrous methanol. The chloride content was determined by potentiometric titration with $AgNO_3$, and methanol was added to adjust the final concentration to *ca.* 15 nM. Reduced cyanoginosin-LA (**17b**) (*ca.* 100 μg) was dissolved in this solution (1 ml) and kept for 24 h at room temperature under nitrogen. Evaporation under reduced pressure followed by addition of water (200 μl) and freeze drying gave compound (**17c**), *m/z* 940 (MH^+) (by f.a.b.m.s.).

Permethylation of Cyanoginosin-LA (**17a**).—Sodium dimesylate [$CH_3S(O)CH_2^- Na^+$] was prepared by heating oil-free sodium hydride (250 mg) in dimethyl sulphoxide (5 ml) at 60 °C for 30 min and centrifuging the mixture. The supernatant liquid (40 μl) was added to a solution of cyanoginosin-LA (**17a**) (*ca.* 100 μg) in dimethyl sulphoxide (30 μl), followed by iodomethane (40 μl) one minute later. After 15 min the reaction mixture was diluted with water (1 ml), and the chloroform layer was washed with water (2 × 1 ml). The resultant product, octamethylcyanoginosin-LA (**18a**), was slightly contaminated with a higher (M^+, 1 035) and a lower (M^+, 1 007) homologue: *m/z* 1 021 (M^+, 0.02%), 907 (0.05), 539 (0.6), 487 (1.2), 441 (1.3), 402 (1.3), 326 (1.7), 283 (3), 255 (3), 241 (4), 212 (11), 162 (15), 156 (12), 135 (41), 100 (35), 91 (38), and 58 (100).

Permethylation of Reduced Cyanoginosin-LA (**17b**).—The reaction as described above, starting with (**17b**), afforded compound (**18b**), *m/z* 1 023 (M^+, 0.03%), 909 (0.09), 485 (1.2), 441 (1.5), 404 (5), 326 (2), 283 (2), 255 (5), 243 (4), 241 (4), 212 (12), 162 (14), 156 (14), 135 (30), 100 (56), 91 (47), and 58 (100).

Deuterio-permethylation of Cyanoginosin-LA (**17a**).—The reaction as described above, but using trideuterioiodomethane, afforded compound (**19a**), *m/z* 1.045 (M^+, 0.10%), 928 (0.17), 557 (0.6), 496 (1.7), 453 (1.6), 408 (1.6), 329 (1.9), 289 (4), 261 (3), 247 (4), 218 (18), 162 (18), 159 (14), 135 (37), 103 (44), 91 (42), and 61 (100).

Pulse Hydrolysis of Reduced Cyanoginosin-LA (**17b**).—(*a*) The reduced toxin (**17b**) (*ca.* 500 μg) was dissolved in 6M HCl (*ca.* 2 ml), heated at 100 °C under N_2 for 5 min, and then dried *in vacuo*.

(*b*) Two cycles of Edman degradation were carried out on the crude product from (*a*) to afford PTH-Ala and PTH-Leu respectively.* The third cycle failed to produce any further degradation.

(*c*) Two cycles of Edman degradation were performed on the crude product from (*a*) as above. The residual peptide was dansylated according to Gray's method,[32] and hydrolysed in 6M HCl for 4 h at 110 °C in a sealed tube. Comparative t.l.c. [Merck Kieselgel 60 F254 t.l.c. plates; developer (solvent system IV[33]) EtOAc–MeOH–AcOH (20:1:1); detection by 254 nm fluorescence] with authentic Dns-Ala (R_F 0.60) Dns-Leu (R_F 0.78), Dns-*N*-methyl-Ala (R_F 0.74), Dns-Glu (R_F 0.38) and Dns-Masp (R_F 0.25 and 0.31; two values due to *threo* and *erythro* forms) showed Masp to be the only dansylated amino acid derived from (**17b**).

(*d*) The crude product from (*a*) was subjected to h.p.l.c. under conditions described above, and afforded compound (**12**), *m/z* 813 (MH^+) and 683, as well as compound (**13**), *m/z* 795 (MH^+) and 665 (by f.a.b.m.s.).

Acetylation and Permethylation of (**12**) *to give compound* (**15**).— To a solution of the pulse hydrolysis product (**12**) in water (200 μl) containing a trace of triethylamine was added a mixture of acetic anhydride and perdeuterioacetic anhydride (1:1; 200 μl) in three portions at 5 min intervals while the mixture was continuously agitated. After a further 15 min the mixture was freeze-dried, redissolved in water (300 μl), and freeze-dried again. Permethylation [as described previously for (**17a**), but using perdeuterioiodomethane for only 1 min] afforded compound (**15**). The e.i.m.s. of compound (**15**) is reproduced in Figure 2.

Hydrolysis and Chiral Column G.c.-M.s. of Compound (**13**).—(*a*) A solution of the peptide (**13**) in 6M HCl was heated at

* PTH = phenylthiohydantoin derivative of amino acid.

110 °C for 24 h in a sealed tube. The reaction mixture was freeze-dried, dissolved in *ca.* 5M HCl in dry propan-2-ol (1 ml), and the solution heated in a sealed tube at 110 °C for 20 min. After evaporation under reduced pressure, the mixture was treated with dichloromethane (0.5 ml) containing TFA (150 µl) and reheated in a sealed tube at 110 °C for 10 min; excess of reagents were removed in a gentle stream of N_2 at 0 °C. The derivatised amino acids were separated on a Finnigan 4000 g.c.-m.s. system using a Chirasil-Val capillary column (Applied Science, 25 m × 0.25 mm) with a linear temperature programme from 80 to 180 °C at 4 °C min^{-1}. Detection was by negative-ion chemical ionisation mass spectrometry using NH_3 as reagent gas. The ratio of D-Ala to L-Ala was 1:1.5.

(*b*) The peptide (**13**) was subjected to one cycle of substractive Edman degradation and then to the hydrolysis, derivatisation, and chiral column g.c.-m.s. as described in (*a*) above. No D-Ala was detected in the mixture, but L-Ala was present.

Hydrolysis of Reduced Cyanoginosin-LA (**17b**) *with Trifluoroacetic Acid.*—Reduced toxin (**17b**) (*ca.* 200 µg) was dissolved in TFA (LR grade; *ca.* 1 ml) and the solution was kept at room temperature for 24 h. Freeze-drying of the reaction mixture and h.p.l.c. of the residue afforded compound (**16**) (peak C of Figure 3), *m/z* 928 $[(M - H)^-]$, 856, and 713 (by negative-ion f.a.b.m.s.) and its demethanolated analogue (peak D of Figure 3), *m/z* 896 $[M - H^-]$, 824, and 681 (by negative-ion f.a.b.m.s.).

Tritium Labelling of Cyanoginosin-LA (**17a**) *and its Reduced* (**17b**) *and Ring-opened* (**16**) *Analogues.*—To a solution of the peptide (50—150 µg) in tritiated water (5 µl; 500 µCi) containing pyridine (10µl) at 0 °C was added acetic anhydride (10 µl). After 5 min at 0 °C, the mixture was warmed to 20 °C for 15 min then recooled to 0 °C for the addition of further acetic anhydride (20 µl) and pyridine (20 µl). Tritiated water (5 µl; 500 µCi) was again added after 1 h at 20 °C, and the reaction was allowed to proceed for a further 1 h at this temperature. Repeated (6 ×) addition of 5% aqueous acetic acid (100 µl) and evaporation under reduced pressure were followed by total hydrolysis in 6M HCl at 110 °C for 6 h. The resulting amino acids were separated on a Waters amino acid analyser with a starting buffer of pH 3.0. Individual peaks were collected after detection by the fluorescent detector, mixed with Instagel (Packard; 15 ml), and counted on a Packard Scintillation Counter Model 2660. Results of the scintillation count are given in Table 3.

Acknowledgements

We thank the S.E.R.C. for support for the work carried out in Cambridge, and Dr. E. Houghton and Mr. P. Teale, Race Course Security Services, Newmarket, for g.c.-m.s. data.

References

1 G. Francis, *Nature* (*London*), 1978, **18**, 11.
2 D. G. Steyn, *S. Afr. J. Sci.*, 1945, **41**, 243.
3 J. Gentile, in 'Microbial Toxins,' Academic Press, New York, 1971, vol. 7, p. 27.
4 T. C. Elleman, I. R. Falconer, A. R. B. Jackson, and M. T. Runnegar, *Aust. J. Biol. Sci.*, 1978, **31**, 209.
5 P. R. Gorham, and W. W. Carmichael, *Pure Appl. Chem.*, 1980, **52**, 165.
6 J. R. H. Hoffman, *Water S. Afr.*, 1976, **2**, 58.
7 M. Schwimmer and D. Schwimmer, in 'Algae, Man, and the Environment,' ed. D. F. Jackson, Syracuse Univ. Press, Syracuse, New York, 1968, p. 279.
8 A. T. C. Bourke and R. B. Hawes, *Med. J. Aust.*, 1983, **1**, 491.
9 P. G. L. Louw, *S. Afr. Ind. Chem.*, 1950, **4**, 62.
10 J. R. Murthy and J. B. Capindale, *Can. J. Biochem.*, 1970, **48**, 508.
11 P. Rabin and A. Darbre, *Biochem. Soc. Trans.*, 1975, **3**, 428.
12 U. A. Kirpenko, I. I. Perevozchenko, K. A. Sirenko, and L. F. Lukind, *Dopov. Akad. Nauk Ukr. RSR., Ser. B.*, 1975, p. 359.
13 C. T. Bishop, E. F. L. J. Anet, and P. R. Gorham, *Can. J. Biochem. Physiol.*, 159, **37**, 453.
14 D. P. Botes, H. Kruger, and C. C. Viljoen, *Toxicon*, 1982, **20**, 945.
15 D. P. Botes, C. C. Viljoen, H. Kruger, and P. L. Wessels, *Toxicon*, 1982, **20**, 1037.
16 S. Santikarn, D. H. Williams, R. J. Smith, S. J. Hammond, D. P. Botes, A. A. Tuinman, P. L. Wessels, C. C. Viljoen, and H. Kruger, *J. Chem. Soc., Chem. Commun.*, 1983, 652.
17 H. Konst, P. D. McKercher, P. R. Gorham, A. Robertson, and J. Howell, *Can. J. Comp. Med. Vet. Sci.*, 1965, **29**, 221.
18 R. P. Gregson and R. R. Lohr, *Comp. Biochem. Physiol.*, 1983, **74C**, 413.
19 K. Wüthrich, in 'NMR in Biological Research: Peptides and Proteins,' North Holland Publishing Company, Amsterdam, 1976.
20 E. Pretch, T. Clerc, J. Siebl, and W. Simon, in 'Tabellen Zur Strukturaufkläring Organischer Verbindingen Mit Spektroskopischen Methoden,' Springer-Verlag, Berlin, 1976.
21 W. Brügel, in 'Handbook of NMR Spectral Parameters,' Heyden, London, 1979, vol. 1. p. 143.
22 A. A. Bothner-By, in 'Biological Applications of Magnetic Resonance,' ed. R. G. Schulman, Academic Press, New York, 1979, p. 177.
23 D. H. Williams, C. V. Bradley, S. Santikarn, and G. Bojesen, *Biochem. J.*, 1982, **201**, 105.
24 C. V. Bradley, D. H. Williams, and M. R. Hanley, *Biochem. Biophys. Res. Commun.*, 1982, **104**, 1223.
25 H. Matsuo, and K. Narita, in 'Protein Sequence Determination,' ed. S. B. Needleman, Springer-Verlag, New York, 1975, p. 104.
26 Th. Wieland, G. Lüben, H. Ottenheym, J. Faesel, J. X. de Vries, A. Prox, and J. Schmid, *Angew. Chem., Int. Ed. Engl.*, 1968, **7**, 204.
27 A. A. Tuinman, unpublished results.
28 B. J. Millard, *Tetrahedron Lett.*, 1965, 3041.
29 B. V. Rozyhov, V. M. Burikov, V. V. Shilin, and A. A. Kiryushkin, *Zh. Obshch. Khim.*, 1968, **38**, 2690.
30 F. Compernolle, H. Vanderhaeghe, and G. Janssen, *Org. Mass. Spectrom.*, 1972, **6**, 151.
31 P. S. Steyn, A. A. Tuinman, F. R. van Heerden, P. H. van Rooyen, P. L. Wessels, and C. J. Rabie, *J. Chem. Soc., Chem. Commun.*, 1983, 47.
32 W. R. Gray, in 'Methods of Enzymology,' ed. C. H. Hirs, Academic Press, London, 1967, p. 139.
33 K. Crowshaw, S. J. Jessup, and P. W. Ramwell, *Biochem. J.*, 1967, **103**, 79.

Received 8*th February* 1984; *Paper* 4/223