Peptides from Australian Frogs. Structures of the Caeridins from *Litoria caerulea*

Russell J. Waugh,[a] David J. M. Stone,[a] John H. Bowie,*[,a] John C. Wallace[b] and Michael J. Tyler[c]
[a] *Department of Organic Chemistry, The University of Adelaide, South Australia, 5001, Australia*
[b] *Department of Biochemistry, The University of Adelaide, South Australia, 5001, Australia*
[c] *Department of Zoology, The University of Adelaide, South Australia, 5001, Australia*

The secretions of the parotoid glands of the Australian tree frog *Litoria caerulea* contain twenty-five peptides of which six, the caeridins, have molecular weights in the 1100–1550 dalton range. The paper describes the primary structures of the caeridins, and indicates how the exudate's caeridin composition varies with the geographic location of the species.

We have reported the structures of the seven peptides contained in the secretion of the rostral and parotoid glands of the magnificent green tree frog of Australia, *Litoria splendida*.[1–3] Three types of peptide are present: (i) the hypotensive peptide caerulein, **1**, (ii) the caerin peptides (M > 2000 daltons), of which caerin 1.1, **2**, is the major component [although structurally quite different, the caerins may be placed in the same general category of amphibian peptides as the magainins[4,5] (from *Xenopus laevis* of Africa) and the bombinins[6,7] (from *Bombina* species of Europe): some of these peptides show powerful antimicrobial activity], and (iii) caeridin 1, **3**, a small peptide of M 1139 daltons, which is neither derived from, nor shows structural resemblance to, the caerin peptides.

Pyr-Gln-Asp-Tyr(SO_3H)-Thr-Gly-Trp-Met-Asp-Phe(NH_2) **1**

Gly-Leu-Leu-Ser-Val-Leu-Gly-Ser-Val-Ala-Lys-His-Val-Leu-Pro-His-Val-Val-Pro-Val-Ile-Ala-Glu-His-Leu(NH_2) **2**

Gly-Leu-Leu-Asp-Gly-Leu-Leu-Gly-Thr-Leu-Gly-Leu(NH_2) **3**

While *Litoria splendida* is confined to the Kimberley area of the extreme north of Western Australia and the adjacent Northern Territory, the closely related green tree frog *Litoria caerulea* is widely distributed throughout the northern half of the Australian continent. We have isolated and characterised twenty-five peptides from the rostral and parotoid gland secretions of *Litoria caerulea*, including caerulein **1** (previously isolated from *Litoria caerulea*[8]), six different caeridins, and eighteen caerins. In this paper we describe the primary structures of the six caeridins, and indicate how the exudate's caeridin composition varies with the geographical location of the species.

Results and Discussion

The green tree frog *Litoria caerulea* was named and briefly described by White in 1790.[9] It is a large frog measuring up to 11.5 cm in length, and is found throughout Australia except in the southern states of Victoria and Tasmania. The peptide material investigated in this study is contained in the large parotoid glands which are situated upon the dorsal surface at the rear of the head. We use the benign method of mild electrical stimulation of the skin glands to effect release of the glandular secretions:[10] this procedure may be used on average once a month to produce a regular supply of secretion without any apparent adverse effect on the animals. On average, each 'milking' produces about 50 mg of peptide material.

Table 1 Caeridin content of *Litoria caerulea* as a function of geographical location

Location[a]	Caeridin[c]					
Town (State)[b]	1	2	3	4	5	6
Derby (W.A.)	*					
Kununurra (W.A.)	*	*				
Darwin (N.T.)	*	*				
Melville Isl. (N.T.)	*	*	*			
Groote Eyl. (N.T.)	*	*	*			
Borroloola (N.T.)	*	*	*			
Camooweal (Q.)	*	*	*			
Mt. Isa (Q.)	*	*				
Cairns (Q.)				*	*	
Proserpine (Q.)				*	*	*
Clermont (Q.)				*	*	*
Ban Ban (Q.)				*	*	*

[a] The locations commence in the north of Western Australia, proceed generally in an easterly direction through the Northern Territory and the north of Queensland, then south down the Queensland coastline. [b] W.A., Western Australia; N.T., Northern Territory; Q., Queensland. [c] Relative molecular masses of the caeridins are as follows: 1 (1140), 2 (1408), 3 (1428), 4 (1504), 5 (1355) and 6 (1371).

Live specimens of *Litoria caerulea* were collected from a number of localities ranging from the northwestern coastal region of Western Australia, through the Northern Territory, the northern part of Queensland, and finally along the eastern coastal strip of Queensland. The actual localities are listed in Table 1. Collections made for this project ranged over a 5-year period. Collections were generally made on two separate occasions at any one location. HPLC traces from the glandular secretions of *Litoria caerulea* specimens are complex and show subtle variations depending on the locality from which the collection was made, although frogs collected on different occasions from the same location show virtually identical HPLC traces. We estimate from the HPLC data that each 'milking' produces, on average, 10 mg of caerulein **1**, 35 mg of caerins, and 5 mg of caeridins. The caeridins have quite different structural features from the larger caerins: their structures are considered in this paper.

Caeridins 1–6 have been separated and purified by HPLC techniques (see Experimental section) and characterised: their distribution depends upon geographic location. Frogs found within a radius of 1500 km from Darwin (including Melville Island and Groote Eylandt) contain some or all of the caeridins 1–3, while frogs collected within the eastern coastal strip of Queensland, from Cairns in the north to Brisbane (some 1500 km south), do not contain caeridins 1–3. Instead, some or all of

Table 2 Primary structures of the caeridins from *Litoria caerulea*

Peptide	
Caeridin 1 (1140) Gly-Leu-Leu-Asp-Gly-Leu-Leu-Gly-Thr-Leu-Gly-Leu(NH_2)	**3**
Caeridin 2 (1408) Gly-Leu-Leu-Asp-Val-Val-Gly-Asn-Leu-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)	**4**
Caeridin 3 (1428) Gly-Leu-Phe-Asp-Ala-Ile-Gly-Asn-Leu-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)	**5**
Caeridin 4 (1504) Gly-Leu-Leu-Asp-Val-Val-Gly-Asn-Val-Leu-His-Ser-Leu-Gly-Leu(NH_2)	**6**
Caeridin 5 (1355) Gly-Leu-Leu-Gly-Met-Val-Gly-Ser-Leu-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)	**7**
Caeridin 6 (1371) Gly-Leu-Leu-Gly-Phe-Val-Gly-Ser-Leu-Leu-Gly-Gly-Leu-Gly-Ile(NH_2)	**8**

the caeridins 4–6 are present. There are also clearly definable differences between the caerin peptide content of frogs collected in these two areas.[11] This must mean either that the frogs located in the two defined areas are different species, or that they are of the same species with some evolutionary trend affecting the peptide distribution. Currently, the frogs collected at all sites are regarded as being of the same species: the zoological significance of these results will be reported elsewhere.

The Primary Structures of Caeridins 1–6.—The structures of the caeridins were determined primarily by fast-atom-bombardment (FAB) mass spectrometry and associated techniques which have been described previously.[3] In summary, the relative molecular mass of each peptide was determined by positive-ion FAB mass spectrometry, the manual Edman/FAB MS technique[12] was used to determine as many amino acid residues from the *N*-terminal end of the peptide as possible, and the MS/MS method[13] provided the remaining sequence information. Enzymatic digest experiments provided valuable information; for example, none of the caeridins are degraded by either trypsin or carboxypeptidase Y, indicating that Arg and Lys are absent, and that none of the peptides has a free carboxy group in the *C*-terminal position. Manual *C*-terminal degradation/FAB ms[14] provides sequencing information from the *C*-terminal end of a peptide which has a free carboxy group in that position. This method does not degrade the caeridins at the *C*-terminal position since they have $CONH_2$ groups there, but it does confirm the position(s) of those residues having carboxy groups in their α side-chains, *e.g.* Asp(4) in caeridins 1–4. Combination of all of the above data generally provided the full sequence of each peptide, with the proviso that isomeric Leu and Ile are not differentiated by the mass spectrometric techniques used. This issue was resolved for all of the caeridins by use of an automated sequencer,[15] which also confirmed, in each case, the sequence proposed by the mass spectrometric methods. Caeridin 1 was synthesized[3] by using the *t*-BOC method.[16,*]

Caeridin 1. Caeridin 1 has been reported previously,[3] and has structure **3** (Table 2). Caeridin 1 is the only representative of the caeridin family to have been synthesized and, in marked contrast to the caerins, it shows no antimicrobial activity.

Caeridin 2. Fast-atom-bombardment (FAB) mass spectrometry showed that caeridin 2 has a relative molecular mass of 1408 daltons.† It contains fifteen amino acid residues. Manual Edman/FAB mass spectrometry[12] identified the first eight amino acids from the *N*-terminal end of the peptide as Gly-Leu‡-Leu-Asp-Val-Val-Gly-Asn. The *C*-terminal degradation method[14] did not affect residue 15, but it did form a thiohydantoin-containing peptide with an MH^+ ion at *m/z* 1166, confirming Asp(4).§ The collisional activation mass spectrum (MS/MS) of the MH^+ species (*m/z* 1408) of caeridin 2 is shown for illustrative purposes in Fig. 1. The 'b' fragmentations (those forming RCO^+ ions) are shown in schematic form at the base of the Figure (for a review of peptide fragmentations see ref. 13). Combination of the manual Edman/FAB ms and MS/MS data suggested structure **4** for caeridin 2. This structure was confirmed by automated sequencing.

Caeridin 3. Caeridin 3 has a relative molecular mass of 1428 daltons, contains fifteen amino acid residues, and differs from caeridin 2 at residues 3–6. Manual Edman/FAB MS experiments show the first eight residues to be Gly-Leu-Phe-Asp-Ala-Leu-Gly-Asn. The MS/MS data for the MH^+ ion, *m/z* 1428, are listed in Table 3. Caeridin 3 has structure **5** (Table 2): automated sequencing showed residue 6 to be Ile.

Caeridin 4. Caeridin 4 has a relative molecular mass of 1504 daltons, and fifteen amino acid residues. Manual Edman/FAB MS data give the first eight residues as Gly-Leu-Leu-Asp-Val-Val-Gly-Asn. This, together with the MS/MS data listed in Table 3, suggested structure **6** (Table 2), which was confirmed by automated sequencing.

Caeridin 5. Caeridin 5 has a relative molecular mass of 1355 daltons, and contains fifteen residues. Manual Edman/FAB MS experiments established the first ten amino acid residues in this instance, *viz.* Gly-Leu-Leu-Gly-Met-Val-Gly-Ser-Leu-Leu. MS/MS data (Table 3) suggested the remaining sequence to be Gly(11)-Gly-Leu-Gly-Leu(NH_2), giving structure **7** (Table 2), again confirmed by automated sequencing.

Caeridin 6. Caeridin 6 has a relative molecular mass of 1371 daltons, has fifteen amino acid residues, and differs from caeridin 5 at residues 5 and 15. Manual Edman/FAB MS degradations gave the first six residues as Gly-Leu-Leu-Gly-Phe-Val while MS/MS data (Table 3) suggested the partial sequence Leu(9)-Leu-Gly-Gly-Leu-Gly-Leu(NH_2). This left 144 daltons unaccounted for, probably Gly and Ser. Automated sequencing confirmed this, determined the relative order Gly(7)-Ser(8), and identified Ile(NH_2) as residue 15. Caeridin 6 has structure **8** (Table 2).

Other Structural Features.—Each of the caeridins may be fitted to a well behaved α-helix by using an Edmundson wheel

* We have not confirmed experimentally that the amino acid residues all have the L-configuration. However, most amphibian peptides contain only L-amino acid residues. Exceptions include the opioid peptides the dermorphins[17] and deltorphins[18] which each contain one D-amino acid residue at position 2. Synthetic caeridin 1, **3**, (in which all amino acid residues are L) has the same HPLC retention time as that of its natural counterpart.

† Molecular masses are given to the nearest integer; *m/z* values are given as nominal masses *i.e.* the mass given by summation of the integral masses of the individual amino acid residues.

‡ This method does not differentiate between Leu and Ile. For brevity, Leu is used in this context throughout the paper: differentiation between the isomers was achieved by automated sequencing, and the correct structures are listed in Table 2.

§ The same procedure confirmed the presence of Asp(4) in caeridins 3 and 4.

Table 3 MS/MS data for caeridins 3–6[a]

Caeridin 3 m/z 1428	('b' ions)	m/z 1298, 1241, 1128, 1071, 1014, 901, 788, 674 and 617 [Gly(7)-Asn-Leu-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)]
	('Y + 2' ions)	m/z 1371, 1258, 1111, 996, 925, 812 and 755 [Gly(1)-Leu-Phe-Asp-Ala-Leu-Gly(7)]
Caeridin 4 m/z 1504	('b' ions)	m/z 1374, 1317, 1204, 1117, 980, 867 and 768 [Val(9)-Leu-His-Ser-Leu-Gly-Leu(NH_2)]
	('Y + 2' ions)	m/z 1447, 1334, 1221, 1106, 1007, 908 and 851 [Gly(1)-Leu-Leu-Asp-Val-Val-Gly(7)]
Caeridin 5 m/z 1355	('b' ions)	m/z 1225, 1168, 1055, 998, 941, 828 and 715 [Leu(9)-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)]
	('Y + 2' ions)	m/z 1298, 1185, 1072 and 785 [Gly(1)-Leu-Leu-Gly-(Met-Val)]
Caeridin 6 m/z 1371	('b' ions)	m/z 1241, 1184, 1071, 1014, 957, 844 and 731 [Leu(9)-Leu-Gly-Gly-Leu-Gly-Leu(NH_2)]
	('Y + 2' ions)	m/z 1314, 1201, 1088 and 884 [Gly(1)-Leu-Leu-(Gly-Phe)]

[a] For a recent review of peptide fragmentations see ref. 13.

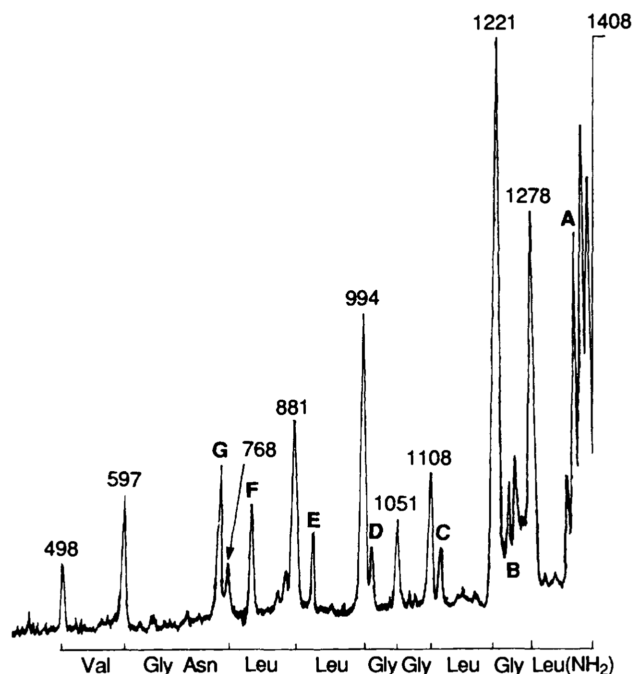

Fig. 1 MS/MS data for caeridin 2 4. Collisional activation (CA) FAB mass spectrum, obtained using a VG ZAB 2HF spectrometer with an Ion Tech FAB gun operating with argon at 6–8 kV using a current of 1 mA. The magnet was set to focus MH^+ (m/z 1408), helium was added to the second collision cell to a measured pressure of 2×10^{-6} mmHg (to produce a main beam reduction of 30%), and an electric sector scan produced the spectrum shown above. The schematics at the bottom of the Figure refer to the numbered peaks, and indicate the 'b' cleavages [*i.e.*, those cleavages which yield RCO^+ ions (*e.g.*, $RCONH_2CH_2CO_2H \longrightarrow RCO^+ + NH_2CH_2CO_2H$)].[13] The peaks designated **A** (m/z 1351), **B** (1238), **C** (1125), **D** (1010), **E** (911), **F** (812) and **G** (755) are formed by 'Y + 2' cleavages and indicate the sequence Gly(1)-Leu-Leu-Asp-Val-Val-Gly(7). ['Y + 2' cleavages are those which form RNH_3^+ ions, *e.g.* $NH_2CH_2CO\overset{+}{N}H_2R \longrightarrow RNH_3^+ + NHCH_2CO$].[13,19]

projection.[20] This is illustrated for caeridin 2, 4, in Fig. 2: the representation indicates clearly delineated hydrophilic and hydrophobic zones. All of these peptides have dominant hydrophobic zones: such systems are suggested[21] to have appreciable hydrophobic moments and, as a consequence, may bind to the surface of biological membranes.

It has been suggested[22] that amphibian peptides should have mammalian analogies. Using the FASTA search procedure[23]

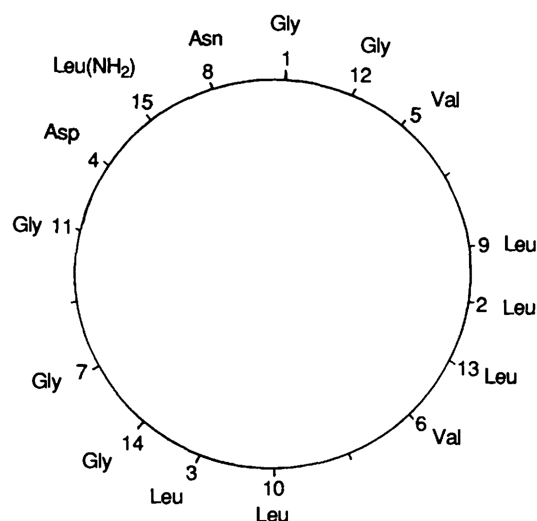

Fig. 2 Edmundson wheel representation for caeridin 2, 4. Residues 12 ⟶ 14 (anticlockwise) are those with 'hydrophilic' α side-chains.

[with GenBank (release #73.0, EMBL (release #31.0), SWISS-PROT (release #22.0) and GenPept (release #72.0)], we have found that caeridin 1, 3, shares some 70% identity with sections of lipid-associated proteins (apolipoprotein B,[24] nicotinic acetylcholine receptor α-5 chain[25] and rabbit skeletal-muscle sarcoplasmic reticulum p53 glycoprotein[26]) although z-scores[27] suggest that the similarities are probably not significant.

Similarly, caeridins 2–4 and 6 (compounds **4**–**6** and **8**) share as high as 76% identity with 13- or 14-residue motifs in a range of unrelated membrane-associated proteins (including a putative transmembrane segment of human PSF-2, a member of the multidrug-resistance family of transporters[28]), but z-scores (between 4.8 and 7) are not indicative of a high level of significance.[27] Finally, caeridin 5, 7, shares a 75% identity with twelve residues of a bombinin peptide from the yellow-bellied toad (*Bombina variegata*).[6] This peptide has antimicrobial activity and is derived by proteolytic cleavage from an approximately 137-residue polypeptide precursor, but again, the z-score is relatively low (5.1).

Experimental

The experimental methods used in this study are identical with those described, in detail, in the previous paper in this series.[3]

Details of the FAB CA MS/MS procedure are outlined in the legend to Fig. 1.

HPLC Separation of the Glandular Secretions.—Preparative separation was effected with a Brownlee preparative column [aquapore octyl, 20 μm spherical (100 × 10 mm)] equilibrated with acetonitrile/aq. trifluoroacetic acid (0.1% TFA [1:10]). Each aqueous injection contained ~2 mg of the peptide mixture: some fifty separations were necessary for the separation and identification of the caeridins. Caeridins 1–3 were eluted by using a non-linear gradient generated by a Waters gradient controller and dual pump system, increasing from 10 to 45% acetonitrile over a period of 20 min at a flow rate of 4 $cm^3 min^{-1}$. The eluent was monitored by UV absorbance at 215 nm by using a Waters 481 variable-wavelength detector. Fractions were collected, concentrated and dried under reduced pressure for analysis. The retention times of caeridins 1–3 are as follows: 1 (15.43), 2 (19.50) and 3 (18.72 min). Caeridins 4–6 were eluted by using a non-linear gradient increasing from 10 to 50% acetonitrile: retention times are 4 (16.59), 5 (14.99) and 6 (15.50 min).

Acknowledgements

We thank the University of Adelaide for the award of a postdoctoral fellowship to D. J. M. S.; the Australian Research Council for grants (to M. J. T.) to allow the collection of specimens; the Conservation Commission of the Northern Territory, the Queensland National Parks and Wildlife Service, and the Western Australian Department of Conservation and Land Management for permission to collect and export frogs; and Dr. Margaret Davies for expert assistance with the collection of frogs.

References

1 M. J. Tyler, J. H. Bowie, D. J. M. Stone and Luminis Pty. Ltd., *Austr. Pat.*, PJ 8549, 1991.
2 D. J. M. Stone, J. H. Bowie, J. C. Wallace and M. J. Tyler, *J. Chem. Soc., Chem. Commun.*, 1992, 1224.
3 D. J. M. Stone, R. J. Waugh, J. H. Bowie, J. C. Wallace and M. J. Tyler, *J. Chem. Soc., Perkin Trans. 1*, in the press.
4 B. W. Gibson, L. Poulter, D. H. Williams and J. E. Maggio, *J. Biol. Chem.*, 1986, **261**, 5341; M. G. Giovannini, L. Poulter, B. W. Gibson and D. H. Williams, *Biochem J.*, 1987, **243**, 113.
5 M. Zasloff, *Proc. Natl. Acad. Sci. USA*, 1987, **84**, 5449; M. Zasloff, B. Martin and H.-C. Chem, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 910.
6 B. W. Gibson, D. Tang, R. Mandrell, M. Kelly and E. R. Spindel, *J. Biol. Chem.*, 1991, **266**, 23103.
7 M. Simmaco, D. Barra, F. Chiarini, L. Noviello, P. Melchiorri, G. Kreil and K. Richter, *Eur. J. Biochem.*, 1991, **199**, 217.
8 A. Anastasi, V. Erspamer and R. Endean, *Arch. Biochem. Biophys.*, 1968, **125**, 57.
9 J. White, *Journal of a Voyage to New South Wales*, Debrett, London, 1790.
10 M. J. Tyler, D. J. M. Stone and J. H. Bowie, *J. Pharm. Methods*, in the press.
11 D. J. M. Stone, R. J. Waugh, J. H. Bowie, J. C. Wallace and M. J. Tyler, unpublished observations.
12 C. V. Bradley, D. H. Williams and M. R. Hanley, *Biochem. Biophys. Res. Commun.*, 1982, **104**, 1223.
13 K. Biemann and S. Martin, *Mass Spectrom. Rev.*, 1987, **6**, 1; K. Biemann, *Biomed. Environ. Mass Spectrom.*, 1988, **16**, 99.
14 P. Schlack and W. Kumpf, *Hoppe-Seyler's Z. Physiol. Chem.*, 1926, **154**, 125; F. E. Dwulet and F. R. N. Gurd, *Int. J. Pept. Protein Res.*, 1979, **13**, 122; L. Meuth, D. E. Harris, F. E. Dwulet, M. L. Crowl-Powers and F. R. N. Gurd, *Biochemistry*, 1982, **21**, 3751.
15 M. W. Hunkapiller, R. M. Hewick, W. J. Dreyer and L. E. Hood, *Methods Enzymol.*, 1983, **91**, 399.
16 R. J. Merrifield, *J. Am. Chem. Soc.*, 1963, **85**, 2149; I. Clark-Lewis, R. Aebersold, H. Zittnener, J. W. Schrader, L. E. Hood and S. B. H. Kent, *Science*, 1986, **231**, 134.
17 P. C. Montecucchi, R. de Castiglione, S. Piana, L. Gozzini and V. Erspamer, *Int. J. Pept. Protein Res.*, 1981, **17**, 275; R. de Castiglione, F. Faoro, G. Persea and S. Piana, *Int. J. Pept. Protein Res.*, 1981, **17**, 263; P. C. Montecucchi, R. de Castiglione and V. Erspamer, *Int. J. Pept. Protein Res.*, 1981, **17**, 316.
18 G. Kreil, D. Barra, M. Simmaco, V. Erspamer, G. F. Erspamer, L. Negri, G. Severini, R. Corsi and P. Melchiorre, *Eur. J. Pharmacol.*, 1989, **162**, 123.
19 P. T. M. Penny, K. Nomoto and R. Orlando, *Rapid Commun. Mass Spectrom.*, 1992, **6** 95.
20 M. Schiffer and A. B. Edmundson, *Biophys. J.*, 1967, **7**, 121.
21 D. Eisenberg, *Annu. Rev. Biochem.*, 1984, **53**, 595; D. Eisenberg, R. M. Weiss and T. C. Terwilliger, *Nature (London)*, 1982, **299**, 371.
22 C. L. Bevins and M. Zasloff, *Annu. Rev. Biochem.*, 1990, **59**, 395.
23 W. R. Pearson and D. J. Lipman, *Proc. Natl. Acad. Sci. USA*, 1988, **85**, 2444.
24 A. Sudarickov and A. Surguchov, *Nucleic Acids Res.*, 1988, **16**, 8187.
25 J. Boulter, A. O'Shea-Greenfield, R. M. Duvoisin, J. G. Connolly, E. Wada, A. Jensen, P. D. Gardner, M. Ballivet, E. S. Deneris, D. McKinnon, S. Heinemann and J. Patrick, *J. Biol. Chem.*, 1990, **265**, 4472.
26 E. Leberer, J. H. M. Charuk, D. M. Clarke, N. M. Green, E. Zubrzycka-Gaarn and D. H. MacLennan, *J. Biol. Chem.*, 1989, **264**, 3484.
27 D. J. Lipman and W. R. Pearson, *Science*, 1985, **227**, 1435.
28 S. Bahram, D. Arnold, M. Bresnahan, J. L. Strominger and T. Spies, *Proc. Natl. Acad. Sci. USA*, 1991, **88**, 10094.

Paper 2/05413B
Received 9*th October* 1992
Accepted 30*th October* 1992