Biochem. J. (1991) 280, 147–150 (Printed in Great Britain)

Structural studies on the [Bu^t-Cys^{18}](19–37)-fragment of human β-calcitonin-gene-related peptide

Jasbir K. SAGOO,* Christopher BOSE,† Nigel R. A. BEELEY† and Saul J. B. TENDLER*‡
*Department of Pharmaceutical Sciences, University of Nottingham, Nottingham NG7 2RD, U.K., and †Celltech Ltd., 216 Bath Road, Slough, Berks. SL1 4EN, U.K.

High-field n.m.r. studies were undertaken upon a peptide fragment of the *C*-terminal region of human β-calcitonin-gene-related peptide (β-hCGRP). Studies on the antigenic [Bu^t-Cys^{18}]β-hCGRP-(19–37)-fragment revealed that several elements of secondary structure were present when the peptide was dissolved in [$^2H_6$]dimethyl sulphoxide. In particular an unspecified turn in the region of Ser^{19}-Gly^{20} and a type I β-turn in the region of Asn^{31}-Val^{32}-Gly^{33} were identified. Through-space connections between the terminal Phe^{37} amide group and the β-protons of Thr^{30} suggest that the peptide may be folded into a loop-type conformation. These structural elements appear to overlap with the epitopes of a number of monoclonal antibodies and provide a molecular basis for understanding the role of the terminal Phe^{37} amide residue in the immune recognition of β-hCGRP.

INTRODUCTION

Calcitonin-gene-related peptide (CGRP) is a 37-amino acid-residue peptide with a disulphide bridge at the *N*-terminus and a *C*-terminal phenylalanine amide residue. Two forms of human CGRP exist that differ by three amino acid residues: α- and β-hCGRP (Table 1), which are produced by the alternative splicing of mRNA from the α- and β-calcitonin gene complexes (Zaidi *et al.*, 1987; MacIntyre *et al.*, 1987; Breimer *et al.*, 1988). Studies have demonstrated that, apart from minor differences, both the α- and β-hCGRP behave similarly (Goodman & Iversen, 1986; Breimer *et al.*, 1988; Beglinger *et al.*, 1988).

hCGRP has extensive distribution in the body and is involved in modulating a variety of physiological processes (Goodman & Iversen, 1986; Cereceda *et al.*, 1987; Zaidi *et al.*, 1987; Beglinger *et al.*, 1988; Yamaguchi *et al.*, 1988; Wada *et al.*, 1988). It is the most potent vasodilator of coronary, cerebral and systemic vasculature known to date (Brain *et al.*, 1985; Zaidi *et al.*, 1987; Breimer *et al.*, 1988). It also has a profound effect on the cardiovascular system, producing an increase in the force and rate of contraction of the heart. hCGRP may also have a role as a neuromodulator and neurotransmitter in the spinal cord and other central neurons.

Structure–activity studies have established that the intact molecule appears to be required for activity *in vitro* (Manning, 1989; Chiba *et al.*, 1989; Zaidi *et al.*, 1990), suggesting that all of the peptide is important for hCGRP activity. High-field n.m.r. studies undertaken on hCGRP in a mixed solvent system of trifluoroethanol/water (1:1, v/v) have identified a well-defined α-helix between residues 8 and 18. The *C*-terminus was, however, shown to be disordered, although a suggestion of a turn was identified in the region of residues 19 and 21 (Breeze *et al.*, 1991). Although the *C*-terminal region appears to have few elements of secondary structure present, peptide fragments of this region have been shown to assume similar structural conformations to intact hCGRP and are important in the immune recognition of the peptide (Andrew *et al.*, 1990). In order to study further the requirements for antibody recognition of peptides, we have investigated the conformation of a fragment of the *C*-terminus region, namely [Bu^t-Cys^{18}]β-hCGRP-(19–37)-fragment or peptide I (Table 1). This peptide, prepared for immunization studies, includes the epitopes for a number of hCGRP-specific monoclonal antibodies (Andrew *et al.*, 1990).

MATERIALS AND METHODS

Peptide I {[Bu^t-Cys^{18}]β-hCGRP-(19–37)-fragment} was synthesized by standard solid-phase synthesis techniques on an Applied Biosystems 430A instrument with the use of fluoren-9-ylmethoxycarbonyl/t-butyl chemistry and Rink resin. The peptide was deprotected with trifluoroacetic acid and purified by gel filtration and preparative reverse-phase h.p.l.c. Amino acid analysis after hydrolysis indicated the correct molar proportions to be present. Samples for n.m.r. analysis contained 5.5 mM-peptide and 0.07 mM-tetramethylsilane in [$^2H_6$]dimethyl sulphoxide. ^{1}H-n.m.r. spectra were recorded at 400 MHz and 500 MHz on Bruker AM400 and AM500 n.m.r. spectrometers respectively. Chemical-shift values (δ) (accurate to 0.01 p.p.m.) were reported relative to the internal standard tetramethylsilane (δ 0.00 p.p.m.). One-dimensional spectra were collected in 16384 data points and routinely zero-filled to 32768 points. The range used for the temperature study was from 293 K to 313 K with increments of 5 K. With each increment the spectrometer was left to equilibrate for 20 min before the next experiment was recorded. The two-dimensional correlation spectroscopy (COSY) experiments were analysed as described by Wüthrich (1986). The

Table 1. Sequences of α-hCGRP, β-hCGRP and peptide I {[Bu^t-Cys^{18}]β-hCGRP-(19–37)-fragment}

The one-letter notation for amino acids is used.

Peptide	Sequence
	5 10 15 20 25 30 35
α-hCGRP	ACDTATCVTHRLAGLLSRSGGVVKNNFVPTNVGSKAF
β-hCGRP	ACNTATCVTHRLAGLLSRSGGMVKSNFVPTNVGSKAF
Peptide I	CSGGMVKSNFVPTNVGSKAF

Abbreviations used: CGRP, calcitonin-gene-related peptide; hCGRP, human CGRP; Bu^t, t-butyl; COSY, correlation spectroscopy; HOHAHA, homonuclear Hartman–Hahn; ROESY, rotating-frame Overhauser enhancement spectroscopy.

‡ To whom correspondence should be addressed.

homonuclear Hartman–Hahn (HOHAHA) experiments were undertaken at 303 K with spin locking times of 56, 135 and 200 ms with 2048 data points recorded for each of the 400 increments, in a phase-sensitive mode with time-proportional phase increment. The rotating-frame Overhauser enhancement spectroscopy (ROESY) experiments were performed at 303 K with a spin locking time of 200 ms; 2048 data points were collected for each of the 400 increments.

Secondary-structure calculations were performed utilizing the University of Wisconsin Genetic Computer Group Package (Devereux *et al.*, 1984) on the S.E.R.C. SEQNET Database (Daresbury Laboratory, Warrington, U.K.) and the Celltech Ltd. Vax computer system. The Chou & Fasman (1978) model was used to predict regions of secondary structure.

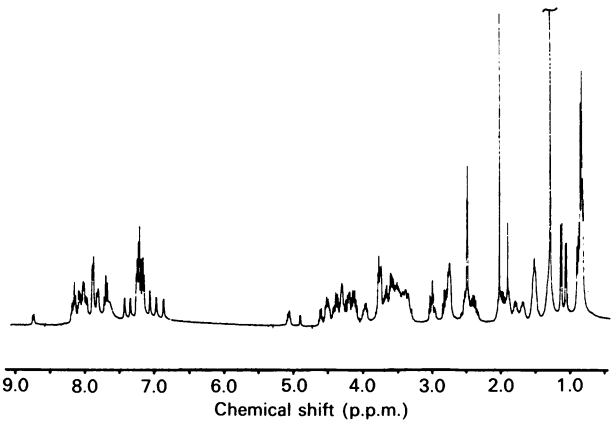

Fig. 1. ^{1}H-n.m.r. spectrum of peptide I

The 400 MHz spectrum of the peptide dissolved in [$^2H_6$]dimethyl sulphoxide recorded at 303 K is shown. The water resonance was suppressed by pre-saturation. The resonance at 2.5 p.p.m. corresponds to the residual proteo-solvent and the resonance at 1.29 p.p.m. corresponds to the t-butyl protecting group on Cys^{18}.

RESULTS AND DISCUSSION

The resonances of peptide I dissolved in dimethyl sulphoxide were initially assigned based upon scalar connections in one-dimensional and two-dimensional COSY and HOHAHA experiments. Although considerable chemical-shift degeneracy was found to be present in the amide region, as shown in Fig. 1, many of the resonances were found at non-random-coil chemical-shift frequencies. The amide-to-aliphatic connectivities of an HOHAHA experiment are displayed in Fig. 2. The sequence-specific assignments of the residues were based mainly upon through-space connections in the ROESY experiment between the α-proton of residue i and the amide proton of the adjacent residue $i+1$. The ROESY experiment was utilized because the tumbling rate of the peptide gave rise to few through-space connections in the nuclear-Overhauser-enhancement spectroscopy experiment (results not shown). The assignments of the peptide are given in Table 2 and the amide region of the ROESY experiment is displayed in Fig. 3. Cys^{18}, Met^{22}, Pro^{29}, Thr^{30} and Ala^{36} were directly assigned as they are the only amino acid residues of their type present in the peptide. Asn^{31} was assigned on the basis of a through-space connection between its amide proton and that of the α-proton of Thr^{30}, allowing us to assign Asn^{26} by a process of elimination. Val^{32} was assigned on the basis of a through-space connection between its amide proton and the α-proton of Asn^{31}. This sequence of assignments was extended by the observation of a through-space connection between the α-proton of Val^{32} and the amide proton of Gly^{33}. Further through-space connections between α-proton to amide proton allowed us to assign Ser^{34}, Lys^{35} and Phe^{37} specifically. The assignment of Phe^{37} was confirmed by the presence of through-space connections between its α- and terminal amide protons. We were therefore able to assign Lys^{24} and Phe^{27} by a process of elimination. Because of their degenerate nature, the specific assignments of the aromatic phenylalanine resonances have not been determined. The assignment of Pro^{29} was confirmed with the observation of a through-space connection between its α-proton and the amide proton of Thr^{30}. Val^{23} was assigned on the basis of a through-space connection between its amide resonance and the α-proton of Met^{22}, allowing the assignment of Val^{28} by elimination. The previous assignment of Lys^{24} was confirmed by the presence of through-space connections with the residue Val^{23}. Ser^{25}, which is degenerate with Ser^{34}, was assigned on the basis of a strong nuclear Overhauser effect between its α-proton resonance and that for the amide resonance of Asn^{26}. Hence by elimination Ser^{19}, with a particularly low-field amide proton resonance, can be assigned. Gly^{20} was assigned on the basis of through-space connections between its amide proton and that of the α-proton of Ser^{19}, leaving Gly^{21}, which was assigned by elimination and confirmed by through space-connections between its α-proton and the amide resonance of Met^{22}.

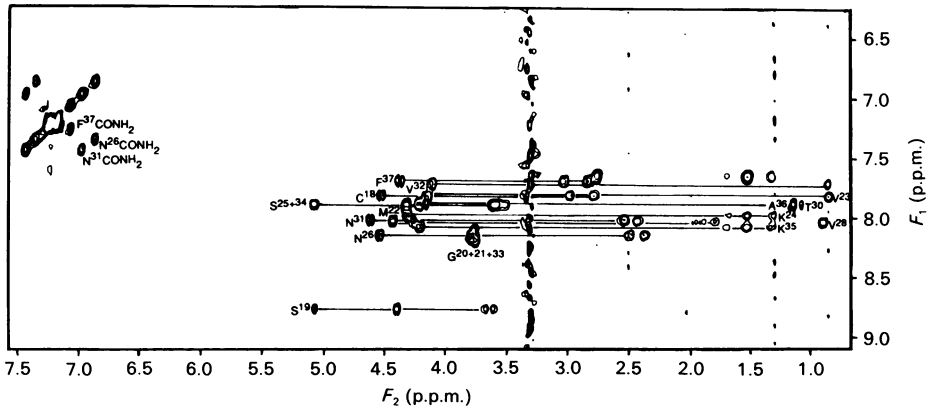

Fig. 2. Amide-to-aliphatic region of the ^{1}H-HOHAHA spectrum of peptide I

The 500 MHz HOHAHA spectrum of the peptide dissolved in [$^2H_6$]dimethyl sulphoxide recorded at 303 K is shown. The capital letters with superscript numbers are the amino acid residues given in the one-letter notation with the positions in the sequence given by the superscript number.

Table 2. Chemical shifts of the assigned amino acid residues for peptide I recorded at 303 K

Abbreviation: N.D., not determined.

Residue	Chemical shift (p.p.m.) NH	α-H	β-H	Others
Cys^{18}	7.83	4.53	2.98, 2.80	
Ser^{19}	8.76	4.42	3.69, 3.62	OH 5.08
Gly^{20}	8.17	3.80, 3.80		
Gly^{21}	8.12	3.76, 3.76		
Met^{22}	8.04	4.43	1.92, 1.80	γ-CH_2 2.42 SCH_3 2.03
Val^{23}	7.83	4.17	1.98	γ-CH_3 0.85
Lys^{24}	7.99	4.32	1.69, 1.53	γ-CH_2 1.33 δ-CH_2 1.53 ε-CH_2 2.73 NH_2 7.65
Ser^{25}	7.90	4.32	3.53, 3.60	OH 5.08
Asn^{26}	8.15	4.53	2.50, 2.27	$CONH_2$ 6.90, 7.40
Phe^{27}	7.64	4.33	2.77, 2.99	Ar N.D.
Val^{28}	8.04	4.26	1.97	γ-CH_3 0.90
Pro^{29}	–	4.51	2.05, 1.92	γ-CH_2 1.80 δ-CH_2 3.57, 3.69
Thr^{30}	7.90	4.22	4.02	γ-CH_3 1.08 OH 4.93
Asn^{31}	8.03	4.62	2.45, 2.55	$CONH_2$ 6.98, 7.45
Val^{32}	7.72	4.12	2.03	γ-CH_3 0.85
Gly^{33}	8.22	3.76, 3.76		
Ser^{34}	7.90	4.32	3.69, 3.53	OH 5.08
Lys^{35}	8.08	4.20	1.69, 1.53	γ-CH_2 1.33 δ-CH_2 1.53 ε-CH_2 2.73 NH_2 7.65
Ala^{36}	7.89	4.17	1.13	
Phe^{37}	7.69	4.38	3.03, 2.83	NH_2 7.07, 7.27 Ar N.D.

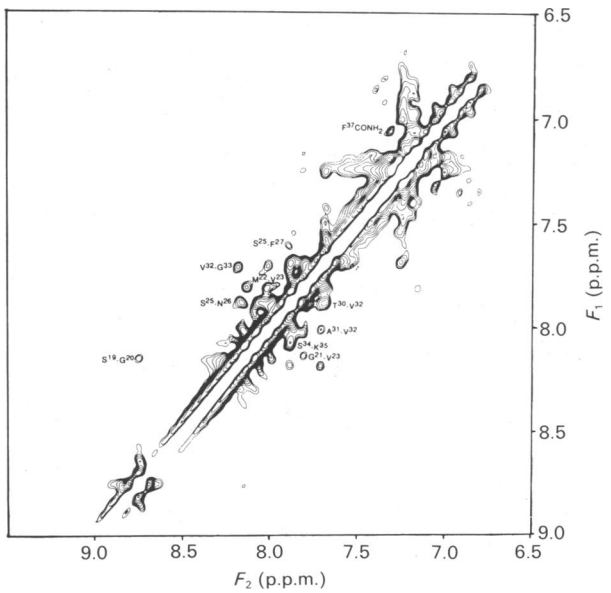

Fig. 3. Low-field portion of ^{1}H-ROESY spectrum of peptide I

The 500 MHz ROESY spectrum of the peptide dissolved in [$^2H_6$]dimethyl sulphoxide recorded at 303 K is shown. The through-space connectivity is given using capital letters with superscript numbers to indicate the amino acid residues in the one-letter notation and the positions in the sequence.

Table 3. Summary of dipolar connections observed in peptide I

The lines below the sequence represent the sequential $C_{(\alpha)}H_{(i)}$–$NH_{(i+1)}$ ($d_{\alpha N+1}$), $NH_{(i)}$–$NH_{(i+1)}$ (d_{NN+1}) and $NH_{(i)}$–$NH_{(i+2)}$ (d_{NN+2}) nuclear Overhauser effects observed. The one-letter notation for amino acids is used.

	Sequence
	20 25 30 35
Peptide I	C S G G M V K S N F V P T N V G S K A F
$d_{\alpha N+1}$	— — — — — — — — — — — — — — —
d_{NN+1}	— — — — — — — — —
d_{NN+2}	— — — — — —

In order to probe the presence of secondary structure in the peptide the temperature-dependence of the amide proton resonances was investigated. Because of general degeneracy in this region of the spectrum the only amide resonance that was studied extensively was Ser^{19}, which has a low-field amide resonance at 8.76 p.p.m., indicating that it may be perturbed to a significant extent. The temperature studies, however, show that it exhibits a temperature coefficient of 0.00624 p.p.m./K, suggesting that the proton is freely accessible to solvent.

It is now known that structurally unrestricted peptides exist in multiple conformations in solution (Dyson *et al.*, 1990). Because of the n.m.r. time scale this technique provides an averaged weighted structure for these molecules, containing a number of structural elements that may not be present simultaneously. Nuclear Overhauser effect $NH_{(i)}$–$NH_{(i+1)}$, $NH_{(i)}$–$NH_{(i+2)}$ and $C_{(\alpha)}H_{(i)}$–$NH_{(i+1)}$ connections present in the ROESY spectra, as shown in part in Fig. 3 and listed in Table 3, provide evidence for these elements of secondary structure in peptide I. A number of $NH_{(i)}$–$NH_{(i+1)}$ connections are found for the peptide, suggesting the presence of turns. Of particular note is a strong through-space connection between the amide protons of Ser^{19} and Gly^{20}; this datum, in conjunction with the low-field nature of the Ser^{19} amide resonance, may indicate the presence of the β-turn that was predicted in the secondary-structure calculations made with the use of the Chou & Fasman (1978) model (Fig. 4). However, because of the absence of further connectivities the nature of this turn cannot be identified. These results are in agreement with those obtained by Breeze *et al.* (1991), who found an unspecified turn in this region of the peptide.

The secondary-structure calculations, as shown graphically in Fig. 4, also predict that a turn may be present in the region of

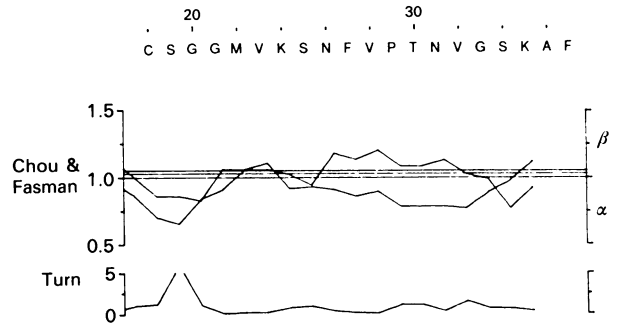

Fig. 4. Graphical representation of Chou & Fasman secondary-structure prediction undertaken on peptide I

Secondary-structure calculations were performed utilizing the University of Wisconsin Genetic Computer Group Package.

residues 32–34 of the peptide. This prediction is supported by the presence of through-space connections between the amide protons of Asn^{31}–Val^{32}–Gly^{33}, suggesting a type I β-turn to be present in this region. The expected through-space connection between the α-proton of Asn^{31} and the amide proton of Gly^{33} is not, however, observed. This may be due to the relatively short sequential amide-to-amide distances in the type I β-turn compared with that of the Asn^{31} α-proton and the Gly^{33} amide proton. In a type I β-turn comprising four residues, the theoretical amide-to-amide $i+1$–$i+2$ and $i+2$–$i+3$ distances are 0.26 and 0.24 nm respectively; these are considerably shorter than the $i+1$–$i+3$ α-proton–amide distance of 0.36 nm (Wüthrich *et al.*, 1984; Wüthrich, 1986).

In addition to the through-space connections noted for adjacent amide proton resonances, a number of through-space connections between $NH_{(i)}$ and $NH_{(i+2)}$ protons were observed for Gly^{21}–Val^{23}, Ser^{25}–Phe^{27} and Thr^{30}–Val^{32}, regions of the peptide where $NH_{(i)}$–$NH_{(i+1)}$ connectivities are found. Such $NH_{(i)}$–$NH_{(i+2)}$ through-space connections are not normally observed in type I or type II β-turns (Wüthrich *et al.*, 1984; Wüthrich, 1986), and the observation of such connectivity is thought to support the presence of some type of nascent helices or modified turns (Dyson *et al.*, 1990). The only long-range through-space connection observed in this study is a connection between a terminal amide proton of Phe^{37} and the β-protons of Thr^{30}. This result suggests that the *C*-terminus of the peptide backbone may fold back on itself; however, if this were the case more long-range nuclear Overhauser effects would be expected.

Although extreme caution must be used when investigating the solution structure of conformationally non-restricted systems by n.m.r. spectroscopy, this study indicates that some elements of secondary structure are to be found in the observed n.m.r. averaged weighted structure of peptide I, in particular the element of secondary structure, possibly a turn, in the region of Ser^{19}–Gly^{20} and a type I β-turn in the region of Asn^{31}–Val^{32}–Gly^{33}. The secondary structure is extended by the finding that the terminal Phe^{37} amide group is close in space to the β-protons of Thr^{30}. Studies probing the antibody recognition of β-hCGRP with monoclonal antibodies have revealed a number of epitopes in the 19–37 region of the peptide (Andrew *et al.*, 1990), including the so-called group 6 binding site from residues 19–28, the group 7 binding site from residues 19–37, the group 8 binding site from residues 19–37 and the group 9 binding site from residues 28–37. Previous n.m.r. studies have identified that β-turns are sites of immune recognition (Williamson *et al.*, 1986; Tendler, 1990); it is tempting to suggest that, as the epitopes each appear to include at least one of the two turn regions, these structural features may be specifically recognized by the monoclonal antibodies. The finding that Phe^{37} is close in space to Thr^{30} may also help to explain the observation by Andrew *et al.* (1990) that the terminal Phe^{37} residue forms a vital part of the epitope of group 8 and 9 antibodies.

J.K.S. acknowledges the financial support of an S.E.R.C. CASE award with Celltech Ltd. The 500 MHz n.m.r. experiments were undertaken at the S.E.R.C. Biological N.M.R. Centre at The University of Leicester. We thank Dr. I. D. Campbell for kindly providing a preprint of his paper.

REFERENCES

Andrew, D. P., Bidgood, T. D., Bose, C., Brown, D., Galfre, G. & Sherwood, M. (1990) J. Immunol. Methods **154**, 87–94
Beglinger, C., Born, W., Hildenbrand, P., Ensinck, J. W., Burkhardt, F., Fischer, J. & Gyr, K. (1988) Gastroenterology **95**, 958–965
Brain, S., Williams, T. J., Tippens, J. R., Morris, H. R. & MacIntyre, I. (1985) Nature (London) **313**, 54–57
Breeze, A. L., Harvey, T. S., Bazzo, R. & Campbell, I. D. (1991) Biochemistry **30**, 575–582
Breimer, H., MacIntyre, I. & Zaidi, M. (1988) Biochem. J. **255**, 377–390
Cereceda, A. F., Gennari, C., Nami, R., Agnusdei, D., Pernow, J., Lundberg, J. M. & Fischer, J. A. (1987) Circ. Res. **60**, 393–397
Chiba, T., Yamaguchi, A., Yamatani, T., Nakamura, A., Morishita, T., Inui, T., Fukase, M., Noda, T. & Fujita, T. (1989) Am. J. Physiol. **256**, E331–E335
Chou, P. Y. & Fasman, G. D. (1978) Adv. Enzymol. Relat. Areas Mol. Biol. **47**, 45–147
Devereux, J., Haeberli, P. & Smithies, O. (1984) Nucleic Acids Res. **12**, 387–395
Dyson, J. H., Satterthwait, A. C., Lerner, R. A. & Wright, P. E. (1990) Biochemistry **29**, 7828–7837
Goodman, E. C. & Iversen, L. L. (1986) Life Sci. **38**, 2169–2178
MacIntyre, I., Alevizaki, M., Bevis, P. J. R. & Zaidi, M. (1987) Clin. Orthop. Relat. Res. **217**, 45–55
Manning, M. C. (1989) Biochem. Biophys. Res. Commun. **160**, 388–392
Tendler, S. J. B. (1990) Biochem. J. **267**, 733–737
Wada, C., Hashimoto, C., Kameya, T., Yamaguchi, K. & Ono, M. (1988) Virchows Arch. B **55**, 217–223
Williamson, M. P., Hall, M. J. & Handa, B. K. (1986) Eur. J. Biochem. **158**, 527–536
Wüthrich, K. (1986) NMR of Proteins and Nucleic Acids, John Wiley and Sons, New York
Wüthrich, K., Billeter, M. & Braun, W. (1984) J. Mol. Biol. **180**, 715–740
Yamaguchi, A., Chiba, T., Yamatani, T., Inui, T., Morishita, T., Nakamura, A., Kadowaki, S., Fukase, M. & Fujita, T. (1988) Endocrinology (Baltimore) **123**, 2591–2596
Zaidi, M., Breimer, L. H. & MacIntyre, I. (1987) Q. J. Exp. Physiol. **72**, 371–408
Zaidi, M., Brain, S. D., Tippens, J. R., Marzo, V. D., Moonga, B. S., Chambers, T. J., Morris, H. R. & MacIntyre, I. (1990) Biochem. J. **269**, 775–780

Received 11 February 1991/21 May 1991; accepted 30 May 1991