The primary structure of ratfish insulin reveals an unusual mode of proinsulin processing

J. Michael Conlon*, Eva Dafgård, Sture Falkmer and Lars Thim+

*Clinical Research Group for Gastrointestinal Endocrinology of the Max-Planck-Gesellschaft, University of Göttingen, Goßlerstraße 10d, D-3400 Göttingen, FRG, Department of Tumor Pathology, Karolinska Institute, Stockholm, Sweden and +Novo Research Institute, Bagsvaerd, Denmark

Received 6 October 1986

The primary structure of insulin from the Holocephalan fish, *Hydrolagus colliei* (the ratfish), has been established by automated Edman degradation as:

```
                5         10        15        20
A chain: G I V E Q C C H N T C S L A N L E G Y C N
                5         10        15        20        25        30
B chain: V P T Q R L C G S H L V D A L Y F V C G E R G F F Y S P K P I* R E L
```

The presence of a COOH-terminal extension to the B-chain is consistent with the occurrence of a single base mutation in the region of the gene encoding one of the dibasic residue processing sites [Arg^{31}(AGA) ⟶ Ile*(AUA)] with the result that the ratfish has utilised an alternative cleavage site within the C-peptide region of proinsulin.

(Ratfish pancreas) *Holocephalan* *Insulin* *C-peptide* *HPLC*

1. INTRODUCTION

In all species studied so far, proinsulin is converted to insulin and the C-peptide of proinsulin by specific proteolytic cleavages at sites of pairs of basic amino acid residues [1]. The enzyme system responsible for effecting these cleavages in eukaryotes has yet to be characterized fully but it has been postulated that cleavage by a trypsin-like enzyme is followed by removal of the basic amino acid residues by an enzyme with carboxypeptidase B-like specificity [2]. The Holocephalan fishes, represented by three extant families: *Hydrolagus* (ratfishes), *Chimaera* (rabbit fishes) and *Callorhynchus* (elephant fishes), were the first class of vertebrate in evolution to develop a pancreatic gland with both exocrine and endocrine parenchyma [3]. The Holocephalan fishes are phylogenetically related to the Elasmobranchian fishes but diverged from the line of evolution leading to contemporary sharks and rays at least 250 million years ago. Although the primary structures of insulins from several Teleostean fishes have been determined (review [4]), structural information regarding Elasmobranchian insulins is confined to that of the ray, *Torpedo marmorata* [5] and an incomplete sequence of insulin from the spiny dogfish, *Squalus ancanthias* [6]. This study demonstrates that the ratfish employes a hitherto undescribed pathway of posttranslation proteolysis in conversion of proinsulin to insulin.

2. MATERIALS AND METHODS

2.1. *Tissue extraction*

Ratfish were collected at Bamfield Marine Station, Vancouver Island, Canada. Pancreata (203 g) from approx. 600 ratfishes were extracted

Published by Elsevier Science Publishers B.V. (Biomedical Division)
00145793/86/$3.50

with 8 vols ethanol/0.7 M HCl (3:1, v/v) using a Waring blender. The homogenate was stirred at 4°C for 5 h, centrifuged (1600 × *g*, 1 h) and ethanol removed from the supernatant under reduced pressure. After further centrifugation (20000 × *g*, 1 h), peptides were isolated from the supernatant using Sep-Pak C18 cartridges (Waters Associates) [7]. Bound material was eluted with acetonitrile/water/trifluoroacetic acid (70:29:1) and the effluent lyophilized.

2.2. *Purification of insulin*

The pancreatic extract, after Sep-Pak concentration, was redissolved in 1% (v/v) trifluoroacetic acid (2 ml) and chromatographed on a column (90 × 1.6 cm) of Sephadex G-50 fine (Pharmacia) equilibrated with 0.1 M ammonium acetate solution, pH 6.8. The column was eluted at 4°C and at a flow rate of 10 ml/h. Fractions (2.1 ml) with K_{av} between 0.27 and 0.44 (K_{av} of porcine insulin = 0.36) were purified further by reverse-phase HPLC. Samples (1 ml) were injected onto a Vydac 218 TP column (0.46 × 2.5 cm) eluted at 30°C and at a flow rate of 1.5 ml/min with a linear gradient (total volume 45 ml) formed from acetonitrile/water/trifluoroacetic acid (21.0:78.9:0.1) and acetonitrile/water/trifluoroacetic acid (35.0:64.9:0.1). UV absorbance was measured at 214 and 280 nm.

Ratfish insulin was purified to homogeneity on a Supelcosil LC-3DP phenyl column (0.46 × 25 cm) equilibrated with 0.1% (v/v) trifluoroacetic acid at 30°C and a flow rate of 1.5 ml/min. The concentration of acetonitrile in the eluting solvent was increased to 17.5% (v/v) over 5 min followed by an increase to 35% (v/v) over 25 min. Ratfish insulin (10 nmol) was reduced and pyridylethylated according to [8] and the derivatized A- and B-chains were separated by reverse-phase HPLC under the conditions shown in fig.2.

2.3. *Structural analysis*

Amino acid compositions of the A- and B-chains were determined using 2 nmol peptide as described [9]. The primary structures of the peptides were determined using approx. 10 nmol A-chain and 6 nmol B-chain by automated Edman degradation using an Applied Biosystems model 470A gas-phase sequencer [10]. The detection limit for PTH-amino acids was 0.5 pmol.

3. RESULTS

3.1. *Purification of insulin*

Extracts of ratfish pancreas contained only trace amounts of insulin-like immunoreactivity when measured in a radioimmunoassay using an antiserum to porcine insulin indicating that ratfish insulin probably adopts an appreciably different conformation to the mammalian peptide. After concentration on Sep-Pak cartridges, the pancreatic extracts were chromatographed on a Sephadex G-50 gel filtration column (fig.1). The fractions indicated by the hatched bar were individually injected onto a C18 reverse-phase HPLC column and a representative elution profile is shown in fig.2. The peak designated by the star was eluted from the column with a retention time comparable to that of porcine insulin and showed a similar ratio of UV absorbance measured at 214 and 280 nm [11]. Ratfish insulin was separated from an uncharacterized impurity by reverse-phase HPLC on a diphenylmethylsilylsilica column (fig.3).

3.2. *Structural analysis*

The amino acid compositions of purified ratfish insulin A- and B-chains are shown in table 1 and the results of automated Edman degradation of the peptides in table 2. The average repetitive yields

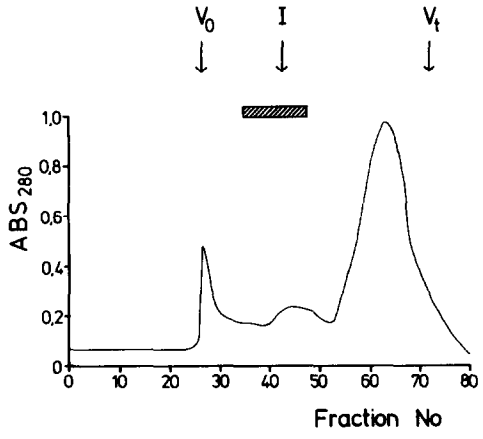

Fig.1. Elution profile on Sephadex G-50 of an extract of ratfish pancreas after partial purification using Sep-Pak C18 cartridges. V_o and V_t refer to the void volume and total volume of the column and *I* denotes the elution volume of human insulin. The fractions indicated by the hatched bar were further purified.

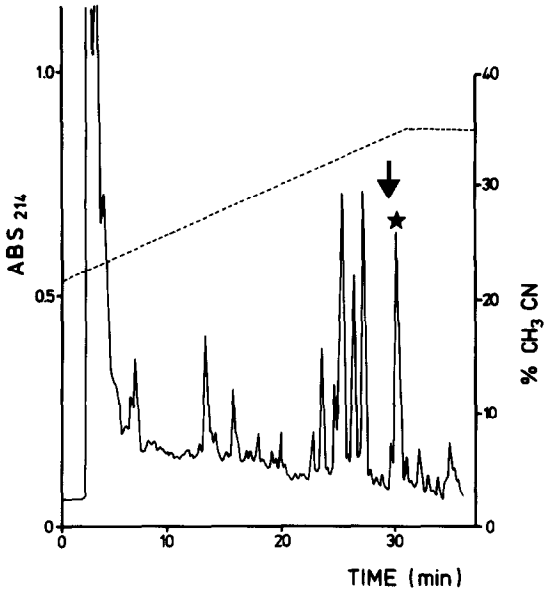

Fig.2. Reverse-phase HPLC on a Supelcosil LC-18-DB column of ratfish insulin after partial purification by gel permeation chromatography. The star denotes the insulin peak and (– – –) the concentration of acetonitrile in the eluting solvent. The arrow denotes the retention time of human insulin.

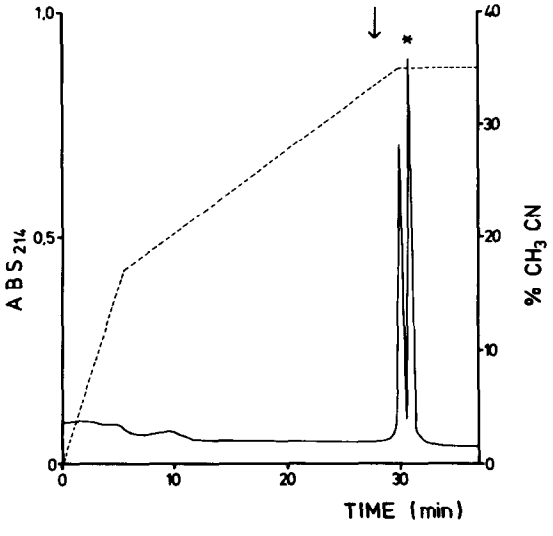

Fig.3. Purification to homogeneity of ratfish insulin (denoted by the asterisk) on a Supelcosil LC-3DP column. The arrow indicates the retention time of human insulin and (– – –) the concentration of acetonitrile in the eluting solvent.

Table 1

Amino acid compositions of A-chain and B-chain of insulin from the ratfish

Residue	Relative amount	
	A-chain	B-chain
Asx	2.90 (3)	1.19 (1)
Thr	0.83 (1)	0.91 (1)
Ser	1.07 (1)	2.08 (2)
Glx	2.91 (3)	4.06 (3)
Pro	– (0)	4.19 (3)
Gly	2.00 (2)	3.00 (3)
Ala	0.91 (1)	0.96 (1)
Val	0.46 (1)	2.73 (3)
Met	– (0)	– (0)
Ile	0.45 (1)	0.75 (1)
Leu	1.79 (2)	5.00 (4)
Tyr	0.89 (1)	1.80 (2)
Phe	– (0)	2.63 (3)
His	0.89 (1)	1.01 (1)
Lys	– (0)	1.23 (1)
Arg	– (0)	3.05 (3)

Data are expressed relative to glycine. Numbers in parentheses represent the values from the sequence determinations

were A-chain (93.1%) and B-chain (90.0%). Unambiguous assignation of 21 amino acid residues of the A-chain and 34 residues of the B-chain was possible. Agreement between the amino acid composition of the A-chain and the proposed sequence was good indicating that the full sequence had been obtained. The low values for the relative amount of Ile and Val are to be expected from the known resistance of the sterically hindered Ile–Val bond to acid hydrolysis. The amino acid composition of the B-chain, however, suggested that a total of 37 amino acids may be present with the possibility of additional Glx, Pro and Leu residues in the molecule. 38 cycles of automated Edman degradation were carried out and no trace of any PTH-amino acid derivative was observed beyond cycle 34. Nevertheless, it is possible that the tripeptide [Glx Pro Leu] may have been washed out of the glass fibre disc of the sequencer at cycle 35 so that further structural studies are required for an unambiguous demonstration that the B-chain of ratfish insulin is not extended beyond Leu 34. The primary struc-

Table 2

Automated Edman degradation of the A-chain and B-chain of insulin from the ratfish

Cycle no.	A-chain		B-chain	
	PTH-amino acid	Yield (pmol)	PTH-amino acid	Yield (pmol)
1	Gly	7423	Val	5182
2	Ile	8577	Pro	3279
3	Val	7688	Thr	2162
4	Glu	3874	Gln	5407
5	Gln	6009	Arg	1513
6	PE-Cys	3633	Leu	4322
7	PE-Cys	3994	PE-Cys	2525
8	His	1385	Gly	3264
9	Asn	4375	Ser	1111
10	Thr	2166	His	1039
11	PE-Cys	2406	Leu	2013
12	Ser	1614	Val	2026
13	Leu	2764	Asp	823
14	Ala	2547	Ala	1481
15	Asn	2696	Leu	1718
16	Leu	2650	Tyr	1282
17	Glu	1435	Phe	1589
18	Gly	2005	Val	1503
19	Tyr	1791	PE-Cys	878
20	PE-Cys	1293	Gly	1227
21	Asn	903	Glu	534
22			Arg	535
23			Gly	698
24			Phe	619
25			Phe	770
26			Tyr	346
27			Ser	89
28			Pro	249
29			Lys	114
30			Pro	162
31			Ile	127
32			Arg	70
33			Glu	53
34			Leu	83

PE-Cys, 4-vinylpyridine derivative of cysteine

ture of ratfish insulin is compared with the insulin from an elasmobranchian fish, *T. marmorata* [5] and with human and guinea pig insulins in table 3.

4. DISCUSSION

The sequence of ratfish insulin shows strong homology with that of the ray, *T. marmorata* [5] in both the A-chain (86%) and corresponding region of the B-chain (70%) (table 3). With the exception of Ala for Phe (A14) and Glu for Pro (B21), the amino acid substitutions are consistent with single base changes in the corresponding sequences of the DNAs. In common with all mammalian insulins yet studied, except that of the coypu [12], the A-chain of ratfish insulin has 21 amino acid residues. In contrast, insulin from the Elasmobranchian fish, *S. acanthias* (Spiny dogfish) has 22 residues [6]. Several substitutions in the proposed receptor-binding region of mammalian insulins that are found in insulins from the torpedo and dogfish, e.g. His for Gln (A5), Pro for Glu (B21), Lys for Arg (B22) and Tyr for Phe (B25) are not found in ratfish insulin. The His residue at B10, important in the formation of zinc-containing hexamers of insulin, that is lost in the insulin of the cyclostome, *Myxine glutinosa* (Atlantic hagfish) [13], is retained in ratfish insulin. The ratfish does, however, share with the hagfish the uncommon substitutions Asn at A15 and Thr at B3. The presence of an Ala residue at A14, a site that is also involved in the hexamerization of insulin, and a Pro residue at B30 are structural features that have not previously been observed in insulins from other species.

As shown in table 3, the tetrapeptide extension to the COOH-terminus of the B-chain of ratfish insulin (Ile Arg Glu Leu) is homologous to the N-terminal region of the C-peptide of proinsulin from the guinea pig (Arg Arg Glu Leu) and human (Arg Arg Glu Ala). The substitution by Ile in ratfish insulin of Arg in mammalian insulin at B31 requires only a single base change in the corresponding region of the DNA (AUA for AGA). It is proposed, therefore, that the increased length of the ratfish insulin B-chain is a consequence of a mutation in ratfish proinsulin at the processing site linking the B-chain to the C-peptide. An alternative processing then occurs within the C-peptide region of proinsulin. This cleavage is probably not effected by the trypsin-like enzyme normally responsible for the proinsulin processing between the B-chain and the C-peptide. Tager et al. [14] have provided evidence for a chymotrypsin-like enzyme participating in normal proinsulin processing in the rat by a selective cleavage of the Gln 22–Thr 23 peptide bond of

Table 3

A comparison of the sequence of ratfish insulin with insulins from the guinea pig, human and an elasmobranchian fish, *Torpedo marmorata*

A-chain

	5					10					15					20			
Ratfish	Gly	Ile	Val	Glu	Gln	Cys	Cys	His	Asn	Thr	Cys	Ser	Leu	Ala	Asn	Leu	Glu	Gly	Tyr	Cys	Asn
Guinea pig	–	–	–	Asp	–	–	–	Thr	Gly	–	–	Thr	Arg	His	Gln	–	Gln	Ser	–	–	–
Human	–	–	–	–	–	–	–	Thr	Ser	Ile	–	–	–	Tyr	Gln	–	–	Asn	–	–	–
Torpedo	–	–	–	–	His	–	–	–	–	–	–	–	–	Phe	Asp	–	–	–	–	–	–

B-chain

	5						10					15					20					25					30				
Ratfish	Val	Pro	Thr	Gln	Arg	Leu	Gly	Ser	His	Leu	Val	Asp	Ala	Leu	Tyr	Phe	Val	Cys	Gly	Glu	Arg	Gly	Phe	Phe	Tyr	Ser	Pro	Lys	Pro	Ile	Arg	Glu	Leu
Guinea pig	Phe	Val	Ser	Arg	His	–	–	–	–	Asn	–	–	Glu	Thr	–	–	Ser	–	–	Gln	Asp	Asp	–	–	–	–	Ile	–	–	Asp	Arg	–	–	–	Glu . .
Human	Phe	Val	Asn	–	His	–	–	–	–	–	–	–	Glu	–	–	–	Leu	–	–	–	–	–	–	–	–	–	Thr	–	–	Thr	Arg	–	–	Ala	Glu . .
Torpedo	Leu	–	Ser	–	His	–	–	–	–	–	–	–	Glu	–	–	–	–	–	–	–	Pro	Lys	–	–	Tyr	–	Leu	–	–	Ala					

The boxed residues are the dibasic processing site followed by the N-terminal part of the C-peptide of proinsulin

the C-peptide. A further chymotryptic cleavage has also been found to occur between Leu 24 and Ala 25 of the C-peptide [15].

Although processing at monobasic Arg residues is common in prohormone processing (review [16]), the enzyme system in the ratfish β-cell seems unable to process at the single Arg B32 residue. The failure of the enzyme system in the human pancreatic β-cell to process at a monobasic residue has been described for patients with hyperproinsulinemia [17]. In this case, a mutation had occurred in the gene segment encoding the Lys-Arg processing site between the C-peptide and the A-chain, resulting in a Lys-X sequence which could not be processed and which led to glucose intolerance in the patients. The deleterious metabolic effects arising from the inability of the organism to process proinsulin to insulin have been overcome in the ratfish by utilising an alternative chymotryptic cleavage site within the C-peptide region of proinsulin. A recent structure-function study [18] using analogues of porcine insulin modified at the COOH-terminus of the B-chain has demonstrated that the Phe residue at B25 is of critical importance in the interaction of insulin with its receptor. This interaction is, however, influenced by steric hindrance involving the COOH-terminal domain of the B-chain. Thus, the presence of the COOH-terminal extension to the B-chain of ratfish insulin is expected to modify, at least to some extent, the bioactivity of the molecule. Confirmation of this hypothesis is underway.

ACKNOWLEDGEMENTS

This work was supported by the Stiftung Volkswagenwerk and the Swedish Medical Research Council. We thank Professor N. Hilschmann, Max-Planck-Institut für Experimentelle Medizin, Göttingen, for providing facilities for amino acid analysis.

REFERENCES

[1] Steiner, D.F., Cunningham, D.D., Spigelman, L. and Aten, B. (1967) Science 157, 697–700.
[2] Steiner, D.F., Kemmler, W., Tager, H.S. and Peterson, J.D. (1974) Fed. Proc. 33, 2105–2115.
[3] Fujita, T. (1962) Z. Zellforsch. 57, 487–494.
[4] Cutfield, J.F., Cutfield, S.M., Carne, A., Emdin, S.O. and Falkmer, S. (1986) Eur. J. Biochem. 158, 117–123.
[5] Conlon, J.M. and Thim, L. (1986) Gen. Comp. Endocrinol. 64, in press.
[6] Bajaj, M., Blundell, T.L., Pitts, J.E., Wood, S.P., Tatnell, M.A., Falkmer, S., Emdin, S.O., Gowan, L.K., Crow, H., Schwabe, C., Wollmer, A. and Strassburger, W. (1983) Eur. J. Biochem. 135, 535–542.
[7] Conlon, J.M., Deacon, C.F., O'Toole, L. and Thim, L. (1986) FEBS Lett. 200, 111–116.
[8] Andrews, P.C. and Dixon, J.E. (1981) J. Biol. Chem. 256, 8267–8270.
[9] Conlon, J.M. and McCarthy, D.M. (1984) Mol. Cell. Endocrinol. 38, 81–86.
[10] Moody, A.J., Thim, L. and Valverde, I. (1984) FEBS Lett. 172, 142–148.

[11] Schmidt, W.E., Mutt, V., Carlquist, M., Kratzin, H., Conlon, J.M. and Creutzfeld, W. (1985) FEBS Lett. 191, 264–268.
[12] Horuk, R., Goodwin, P., O'Connor, K., Neville, R.W.J., Lazarus, N.R. and Stone, D. (1979) Nature 279, 439–440.
[13] Chan, S.J., Emdin, S.O., Kwok, S.C.M., Kramer, J.M., Falkmer, S. and Steiner, D.F. (1981) J. Biol. Chem. 256, 7595–7602.
[14] Tager, H.S., Emdin, S.O., Clark, J.L. and Steiner, D.F. (1973) J. Biol. Chem. 248, 3476–3482.
[15] Markussen, J. and Sundby, F. (1972) Eur. J. Biochem. 25, 153–162.
[16] Schwartz, T.W. (1986) FEBS Lett. 200, 1–10.
[17] Robbins, D.C., Blix, P.M., Rubenstein, A.H., Kanazawa, Y., Kosaka, K. and Tager, H.S. (1981) Nature 291, 679–681.
[18] Nakagawa, S.H. and Tager, H.S. (1986) J. Biol. Chem. 261, 7332–7341.