Isolation and structural characterization of insulin and glucagon from the holocephalan species *Callorhynchus milii* (elephantfish)

Ben C. BERKS, Craig J. MARSHALL, Alan CARNE, Susan M. GALLOWAY and John F. CUTFIELD*

Department of Biochemistry, University of Otago, Dunedin, New Zealand

Both insulin and glucagon from the pancreas of the holocephalan cartilaginous fish *Callorhynchus milii* (elephantfish) have been isolated and purified. Two reverse-phase h.p.l.c. steps enabled recovery of sufficient material for gas-phase sequencing of the intact chains as well as peptide digestion products. The elephantfish insulin sequence shows 14 differences from pig insulin, including two unusual substitutions, Val-A14 and Gln-B30, though none of these is thought likely to influence receptor binding significantly. The insulin B-chain contains 31 residues, one more than mammalian insulins, but markedly less than that of the closely related ratfish with which it otherwise exhibits high sequence similarity. Elephantfish and pig glucagons differ at only four positions, but there are six changes from the ratfish glucagon-36 (normal glucagon contains 29 residues) sequence. It is apparent that different prohormone proteolytic processing mechanisms operate in the two holocephalan species.

INTRODUCTION

The most primitive exocrine pancreas possessing islets of Langerhans is found in the holocephalan cartilaginous fish (the earliest jawed vertebrates) of which relatively few species exist (Conlon *et al.*, 1988). This pancreas is a solid gland, joined to the spleen and connected to the gut by a long duct (Falkmer *et al.*, 1981). Immunohistochemical investigations of two species, the Pacific ratfish (*Hydrolagus colliei*) and the rabbit fish (*Chimaera monstrosa*), have shown both the pancreatic duct and pancreas to contain three types of islet hormone cells, namely those producing insulin, glucagon and somatostatin (Falkmer *et al.*, 1984). A third species, the elephantfish (*Callorhynchus milii*) possesses, additionally, pancreatic polypeptide cells (Falkmer *et al.*, 1984); thus its pancreas is a 'four-hormone organ' similar to that found in the phylogenetically more recent placostomian cartilaginous fish (elasmobranchs such as sharks and rays), as well as in the higher-order vertebrates. This evidence suggests that elephantfish is the most highly developed of the holocephalans, at least as far as the endocrine pancreas is concerned.

The insulins from both Pacific ratfish and rabbit fish have recently been isolated and sequenced (Conlon *et al.*, 1986; Conlon & Thim, 1987). Their primary structures are evidently identical and, unusually, show a considerable B-chain C-terminal extension of at least seven residues compared with insulins from other species. Because this extension shows sequence similarity with the N-terminal region of human insulin C-peptide, save for the crucial substitution of isoleucine-B31 for arginine, it was concluded by Conlon & Thim (1987) that a proinsulin processing site (the dibasic B31–B32) linking B-chain to C-peptide has been lost. An alternative cleavage has evidently taken place further into the C-peptide, although the precise location has not yet been unambiguously determined.

An extended glucagon molecule has also been identified in the ratfish pancreas (Conlon *et al.*, 1987). The 36-residue peptide (normal glucagon contains 29 residues) bears some resemblance to intestinal oxyntomodulin (glucagon-37), and the suggestion was made (Conlon *et al.*, 1987) that this form of proglucagon post-translational processing may be linked with the production of the higher- M_r insulin found in the same species.

The pancreatic hormones from the separate family Callorhynchidae are therefore of particular interest, given the phylogenetic significance of its endocrine pancreas and the unusual pancreatic hormone processing seen in the closely related ratfish and rabbit fish. This investigation is concerned with the isolation and primary structure determination of both insulin and glucagon from elephantfish (*Callorhynchus milii*), a species found in southern oceans.

EXPERIMENTAL

Extraction of pancreatic protein

The method employed to extract pancreatic protein was a modification of that described by Mirsky (1973) and is optimal for insulin isolation. Pancreatic tissue (112 g from 20 elephantfish) was removed, immediately frozen, and stored at -80°C until required. The frozen tissue was homogenized in 4 vol. of ice-cold ethanolic 2% HCl plus 1 vol. of water, containing 0.5 mM-phenylmethanesulphonyl fluoride, and then stirred for 90 min at 4°C . The homogenate was centrifuged at 12000 *g* for 1 h, and the supernatant was retained as the primary extract. The pellet was re-extracted with acid/ethanol overnight and centrifuged to give the secondary extract. Both extracts were processed separately. This involved neutralization with aq. 6 M- NH_3 to pH 8.2, removal of precipitated protein by centrifugation, pH adjustment to 3.5, reduction of volume by rotary evaporator at 28°C ,

* To whom correspondence and reprint requests should be sent.

pH adjustment to 5.3, addition of 2 M-ammonium acetate (0.03 ml/ml of extract), and finally addition of ethanol (6 vol.) and diethyl ether (10 vol.). After the mixture had been left overnight, the lipid-containing supernatant was decanted and the precipitate dried under a stream of air.

Purification

The precipitate was resuspended in 15 ml of 2 M-acetic acid, then centrifuged before loading on to a Sephadex G-50 (fine grade) column (90 cm $\times$ 5 cm) eluted with 1 M-acetic acid at a flow rate of 35 ml/h. Fractions (7 ml each) were pooled in the approx. 3000–7000 M_r range, and the volume was reduced by rotary evaporation before freeze-drying. Fractions were also assayed for insulin activity.

Reverse-phase h.p.l.c. was employed as the main purification step after gel filtration. Two main systems were employed: (a) a Waters μ Bondapak C_{18} RadPak cartridge in conjunction with a linear gradient of 50 mM-ammonium acetate (pH 7.65)/acetonitrile; (b) a Brownlee C_8 Aquapore RP-300 3 cm column with a 0.1 % trifluoroacetic acid/acetonitrile gradient. The latter system was used to further purify peptides isolated from (a). Fractions corresponding to peaks of absorbance at 280 nm and 214 nm were collected manually in silanized borosilicate tubes. Volumes were reduced under a stream of nitrogen and samples that contained ammonium acetate were freeze-dried.

Peptides for sequencing

For insulin it was first necessary to separate the A and B chains, in this case by oxidative sulphytolysis using the method employed by Paynovich & Carpenter (1979), adapted here for use on a small scale. Insulin (about 1 nmol) obtained from h.p.l.c. system (a) was dissolved in 50 μ l of 8 M-urea/0.1 M-Tris/HCl, pH 7.5, then a few grains each of sodium sulphite and sodium tetrathionate (Fluka) were added. After a 30 min incubation at room temperature, the reaction mixture was separated by h.p.l.c. method (a) into the S-sulphonated forms of the A and B chains. The B chain was further cleaved with *Staphylococcus aureus* V8 proteinase (Sigma), essentially as described by Drapeau (1977). Freeze-dried insulin B chain (about 200 pmol) was dissolved in 100 μ l of fresh 50 mM- NH_4HCO_3 , pH 7.8, and V8 proteinase (1 mg/ml in 1 mM-HCl stock) added in a ratio (w/w) of 1:50. The reaction mixture was incubated for 18 h at 37 °C before reverse-phase h.p.l.c. separation of products, this time using a Bakerbond C_8 wide-pore column with an acetonitrile gradient in 0.1 % trifluoroacetic acid.

Glucagon, purified by h.p.l.c. methods (a) and (b), was also fragmented, by exploiting a single methionine cleavage site, with CNBr. Approx. 60 pmol of glucagon was dissolved in aq. 70 % formic acid (AnalaR) and allowed to react with 1 μ l of CNBr (2 g/ml in acetonitrile) at room temperature for 15 h in the dark. The mixture was dried under vacuum, then washed with 20 μ l of water and dried (four cycles) before being dissolved in 30 μ l of 0.05 % trifluoroacetic acid/aq. 50 % acetonitrile and applied to a precycled glass-fibre disc containing Biobrene.

Peptides were submitted to sequence analysis on a Model 470A gas-phase sequencer with an on-line model 120A PTH analyser (Applied Biosystems, Foster City, CA, U.S.A.). They included elephantfish insulin A and B chains, V8 proteinase-generated fragments of insulin B

Fig. 1. Gel-filtration profile of elephantfish pancreatic extract on Sephadex G-50 (fine grade)

The bar indicates insulin-containing fractions as determined by the rat fat-cell bioassay. The column was calibrated with M_r markers: bovine serum albumin (66000), α -chymotrypsinogen (26000), lysozyme (14500) and bovine insulin (6000). Fractions corresponding to the approx. 3000–7000- M_r range were pooled, reduced in volume by rotary evaporation, and freeze-dried.

Fig. 2. Reverse-phase h.p.l.c. separation of elephantfish pancreatic peptides

Material pooled from gel filtration was further fractionated on a Waters C_{18} μ Bondapak column using an acetonitrile gradient in 50 mM-ammonium acetate. Peaks 1, 2 and 3 refer to peptides subsequently identified by sequence analysis. The arrow marks the elution position of bovine insulin.

Fig. 3. Separation of the A and B chains of elephantfish insulin by reverse-phase h.p.l.c.

Insulin from Peak 2 (Fig. 2) was subjected to oxidative sulphytolysis, and the products separated using the same h.p.l.c. system as for Fig. 2.

chain, intact glucagon, CNBr-treated glucagon and several other peptides from h.p.l.c.

Insulin assay

Individual fractions from the gel-filtration step were assayed for insulin activity using the 'in vitro' bioassay of Moody *et al.* (1974), which measures the insulin-stimulated incorporation of D-[3-³H] glucose into lipid in isolated rat fat-cells. The assay was employed simply to establish those fractions containing significant insulin activity.

Model-building

The elephantfish insulin sequence was model-built into the known pig insulin crystal structure with the assistance of Dr. R. E. Hubbard (University of York), using the programs QUANTA (Polygen Corp.) and CHARMM (Brooks *et al.*, 1983) running on a Silicon Graphics Iris 4D/70GT graphics workstation.

RESULTS

Gel-filtration chromatography of the primary elephantfish pancreatic protein extract (Fig. 1) indicated the fractions likely to contain insulin according to *M_r* standards, and this was subsequently confirmed by the rat fat-cell bioassay. Conventional radioimmunoassay using anti-(porcine insulin) antibodies was not used because of the known poor cross-reactivity between

Fig. 4. Sequence information (automated Edman degradation) for elephantfish insulin and glucagon

No yield is recorded for the half-cystine residues in insulin, as these were in the *S*-sulphonated form. The actual sequences are shown in Fig. 5.

mammalian and fish insulins (Yalow & Berson, 1964; Conlon *et al.*, 1986). Further fractionation of the insulin-containing region by *C₁₈* reverse-phase h.p.l.c. (Fig. 2) revealed a number of peptide components, of which peak 2, by its *A₂₈₀*/*A₂₁₄* ratio, and by its similar elution position to that of bovine insulin, was identified as likely to be elephantfish insulin. This was confirmed both by bioassay and by subjecting the sample to oxidative sulphytolysis followed by h.p.l.c. (Fig. 3) to give the predicted separation of the A and B chains. The additional broad peak evident in the chromatogram showed weak absorbance at 214 nm and seems to be associated with the sulphytolysis reagents present in excess. The estimated yield of insulin from the preparation was about 3–4 nmol, rather lower than anticipated particularly in comparison with other cartilaginous fish (Conlon & Thim, 1985; Conlon *et al.*, 1986; Bajaj *et al.*, 1983). This may have been due to the presence of extraneous tissue, to the length of storage time of the frozen tissue (up to 18 months) or to the method of extraction employed. The secondary extract also contained insulin, but only about half of that found in the primary extract.

The A-chain (350 pmol) and B-chain (220 pmol) peptides were subjected to automated Edman degra-

Insulin A chain

		5	10	15	20
Lamprey	=	=	=	R K	= I Y D M = N = =
Hagfish	=	=	=	K R	= I Y = Q N = =
Elephantfish	G	I	V	E	Q C C H N T C S L V N L E G Y C N
Ratfish	=	=	=	=	= A = = = = =
<i>Torpedo</i>	=	=	=	H	= = = = F D = = = =
Dogfish (spiny)	=	=	=	H	= = = = Y D = = = = Q
Salmon (coho)	=	=	=	=	= K P = N I F D = Q N = =
Pig	=	=	=	=	= T S I = = = Y Q = = N = =

Insulin B chain

					-1					5					10					15					20					25					30									
Lamprey	S	A	L	T	G	A	G	G	T	H	=	=	=	=	=	E	=	=	V	=	=	=	=	D	=	=	=	=	T	=	S	K	T											
Hagfish						R	T	=	G	H	=	=	=	K	D	=	=	N	=	=	=	I	A	=	=	V	=	=	=	=	D	=	T	K	M									
Elephantfish						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	Q	I								
Ratfish						=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	=	P	=	R	E	L	E	P	L	?	
Torpedo						L	=	S	=	H	=	=	=	=	=	=	E	=	=	=	=	=	=	=	=	P	K	=	=	Y	=	L	=	=	A									
Dogfish						L	=	S	=	H	=	=	=	=	=	=	E	=	=	=	=	=	=	=	=	P	K	=	=	Y	=	L	=	=	B	Z	V							
Salmon (coho)						A	A	A	=	H	=	=	=	=	=	=	=	=	=	=	=	=	=	=	L	=	=	=	=	K	=	=	=	=	N	=	=							
Pig						F	V	N	=	H	=	=	=	=	=	=	E	=	=	=	=	=	=	=	L	=	=	=	=	=	=	=	=	=	T	=	=	T						

Glucagon

						5						10																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
--	--	--	--	--	--	---	--	--	--	--	--	----	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Fig. 5. Comparisons of insulin and glucagon sequences

The species listed are, in order, representatives of: the cyclostomes or jawless fishes (lamprey, hagfish); the cartilaginous fishes (elephantfish, ratfish, *Torpedo*, dogfish); the teleosts or bony fishes (salmon); mammals (pig).

dation in a gas-phase sequencer. The complete A-chain sequence was determined unambiguously in duplicate runs through to position 21. The B chain was sequenced to position 31, with uncertainties at positions B1–B3 and low yields at the C-terminal end. The six positions (A6, A7, A11, A20, B7 and B19) corresponding to half-cystine residues were assigned by comparison with all known insulin sequences, there being no extraction and hence no signal for the highly charged S-sulphonated form of cystine. The elephantfish B-chain sequence determined thus far contained only one glutamic acid residue, at position B21, as is also seen in ratfish insulin, but the latter also possesses Glu-B33 and Glu-B35 in the extended region. Hence V8 proteinase was selected as the cleavage reagent most likely to resolve the uncertainty at the end of the B chain. About 200 pmol of elephantfish B chain were digested with the enzyme, the products separated by C_8 reverse-phase h.p.l.c., and then sequenced. Two fragments were isolated, their sequences corresponding to B22–B31 and B1–B21 with respect to elution order. These sequences were free of ambiguities and confirmed that the B chain terminated after residue 31. Fig. 4 summarizes the sequence data for elephantfish insulin, the actual sequence being shown in Fig. 5.

Previous work in this laboratory suggested that glucagon is generally eluted close to insulin under the

relatively steep acetonitrile gradient conditions employed in the initial h.p.l.c. fractionation (Fig. 2). Peak 3, by virtue of its proximity to the insulin peak, by its higher A_{280}/A_{214} ratio, and by its lower yield, was isolated and further purified by reverse-phase (C_8) h.p.l.c. About 40 pmol of this peptide was sequenced and identified subsequently as glucagon. There was some doubt as to the identity of the last two residues (28 and 29), because of the low yields; hence the unequivocal determination of the C-terminus remained to be resolved. The presence of a single methionine residue at position 27 afforded the opportunity for a unique CNBr cleavage. After treatment of about 60 pmol of peptide, the resulting mixture was applied to the disc for sequencing. The two peptide sequences (B1–B27, B28–29) were easily identified, with B28 being confirmed as serine, B29 as threonine and no suggestion of any extension beyond B29. Sequencing data for elephantfish glucagon are seen in Fig. 4 and refer to the sequence shown in Fig. 5.

The dominant peak (1) of Fig. 2, which is obviously the main contributor to the M_r -3000–7000 gel-filtration peak (Fig. 1), containing the insulin and glucagon molecules, was also submitted to sequencing. A partial N-terminal sequence (33 residues) was obtained, sufficient to suggest that this peptide was a pancreatic secretory trypsin inhibitor (Kazal-type), but further confirmation

Fig. 6. Comparison of part of the pig insulin crystal structure (left) with that of model-built elephantfish insulin (right)

The region shown is the interface between adjacent dimers in the 2-zinc-insulin hexamer (Baker *et al.*, 1988), the arrow indicating an approximate two-fold axis of symmetry. Some side chains have been omitted for clarity.

is required. Other peaks from h.p.l.c. (Fig. 2) were also partially sequenced, but were not identified.

Despite having 14 sequence differences when compared with pig insulin, the elephantfish insulin sequence was easily model-built into the known crystal structure of pig insulin (Baker *et al.*, 1988). Most of the substitutions are conservative in nature and do not affect the structural integrity of the molecule. A small rotation of the B17 phenylalanine ring was required in order to relieve some close contacts with B14 alanine and with its symmetry-related equivalent residue in the hexameric state (see Fig. 6 and the Discussion).

DISCUSSION

The amino acid sequence of elephantfish insulin reveals strict conservation of the residues known to be involved in maintaining the three-dimensional structure of insulin (Baker *et al.*, 1988). These include twelve invariant residues (the six half-cystine residues, glycine residues at B8 and B23, leucine residues at B6, B11 and B15, and valine-B12) as well as the conservative sites Ile-A2, Leu-A16, Tyr-B16 and Val-B18. Similarly, the group of surface residues believed to be involved in receptor binding and expression of biological activity (the invariant Gly-A1, Tyr-A19 and Asn-A21, together with the commonly found Ile-A2, Val-A3, Gln-A5, Arg-B22, Phe-B24, Phe-B25 and Tyr-B26) are all present, suggesting a 'normal' insulin. Thus it would be expected that elephantfish insulin would possess reasonably high potency with respect to mammalian insulins in '*in vitro*' mammalian assay systems, as has been found for some other species of fish (Cutfield *et al.*, 1986). The amino acid sequence of elephantfish insulin is compared with some other representative insulin sequences in Fig. 5. It is apparent that the insulins from the cartilaginous fish

(holocephalans and elasmobranchs) are characterized by certain structural determinants, namely the sequence -His-Asn-Thr- in the disulphide-enclosed loop A8-10, a non-polar residue at A14, and a glycine at A18. In the absence of activity measurements and three-dimensional structural information on these insulins, the effects of such substitutions cannot yet be fully evaluated.

The features which characterize elephantfish insulin in particular are the sequences at the beginning (B1-B5) and end (B30-B31) of the B-chain, as well as the hitherto-unseen substitution of valine at A14. These residues are not considered important as regards biological activity, but some of them are involved in the assembly of zinc-containing hexamers (co-ordinated through His-B10) along with B17, which is phenylalanine in this case, but normally leucine, as well as the conserved Ala-B14 and Val-B18. Model-building using computer graphics (Fig. 6) suggests that these changes can be accommodated, with only minor adjustments, within the mammalian hexamer structure, though it does not necessarily prove that hexamerization will occur. One can be more certain, however, that the molecule will dimerize, as the surface involved, comprising B12, B16, B20, B24, B26 and B28, is totally conserved. The ability of insulins to dimerize and, frequently, to aggregate further into zinc-containing hexamers is a property widespread amongst vertebrates and is important in the processing and storage of the hormone.

Among the holocephalan group (ratfish, rabbit fish and elephantfish), there are only two amino acid substitutions (at A14 and B30), but the major difference lies in the length of the B-chain. Elephantfish insulin terminates at Ile-B31, whereas in ratfish (and rabbit fish) the chain is extended further, to at least residue 37 (Conlon & Thim, 1987). Those authors have proposed that the usual processing site at B31-B32 has been lost,

owing to a single codon base change (AUA for AGA) at B31. The results of the present study on a closely related species indicate that cleavage can take place after Ile-B31. What is not known, however, is whether the sequence immediately following begins with a normal dibasic cleavage site or just the single arginine residue seen in ratfish. If it were the latter, then elephantfish proinsulin would be unusual amongst insulins in having a monobasic processing site. Knowledge of the preproinsulin gene or cDNA sequence for both elephantfish and ratfish would therefore be of considerable interest. A slightly longer insulin B-chain seems to be a feature of phylogenetically older species, including the jawless fish lamprey and hagfish, the holocephalan elephantfish and the elasmobranch spiny dogfish. In addition, lamprey insulin (Plisetskaya *et al.*, 1988) has an *N*-terminal B-chain extension apparently associated with different prepeptide processing. Ratfish and rabbit fish insulins, however, stand apart with their major C-terminal extensions.

Glucagon is a highly conserved molecule whose crystal structure has been determined, but has no well-defined structure in solution (Blundell *et al.*, 1982), hence it is difficult to equate individual amino acid substitutions with conformational determinants. The amino acid sequence of elephantfish insulin contains several substitutions not seen in mammalian species, but shared by other cartilaginous fish, namely the acidic and basic residues at positions 3 and 20 respectively (Conlon *et al.*, 1987). There are just four changes from that of pig glucagon, three from the elasmobranch *Torpedo*, but interestingly six differences when compared with the first 29 residues of ratfish glucagon-36 (Fig. 5). It has not yet been unequivocally established that the ratfish pancreas also contains significant amounts of normal glucagon-29 (presumably truncated glucagon-36), but in any case this number of sequence differences between these two closely related species is perhaps higher than expected. Thus whereas elephantfish glucagon appears relatively 'normal', ratfish glucagon is both elongated (i.e. processed differently) and less constrained in its sequence requirements. Again, as with the insulins, the relative potencies of both the ratfish and elephantfish glucagons need to be determined before assessing the notion of a 'compensatory' pair of pancreatic hormones required to maintain glucose homeostasis (Seino *et al.*, 1986; Conlon *et al.*, 1987).

Isolation and characterization of individual peptides, together with prohormone cDNA sequence data, should allow the processing pathway of the precursor to be deduced with some measure of confidence. However, the identification of all the peptides released during proteolytic processing is not trivial, and indeed the levels of those which are characterized may depend on the particular purification strategy employed. In the present study a purification scheme optimal for insulin was chosen, and there was the additional restriction of a defined M_r range (about 3000–7000). In addition to insulin and glucagon, there were also several other peptides evident in the h.p.l.c. profile. The predominant polypeptide found was a pancreatic-secretory-trypsin-inhibitor-like molecule, according to partial sequence data. There was, however, no obvious presence of any other glucagon-like peptides, as found in the pancreas of various bony fish (reviewed in Pollock *et al.*, 1988). Moreover, the C-peptide of elephantfish insulin appar-

ently lies outside the M_r range chosen. In order to characterize the full precursor molecules, we have recently made a cDNA library from elephantfish pancreas polyadenylated RNA which awaits screening with oligonucleotide probes based on the amino acid sequences determined in the present study.

The assistance of the University of Otago Portobello Marine Laboratory staff in acquiring the elephantfish is much appreciated. Support from the Otago Medical Research Society (Summer Scholarship to B.C.B.), from the N.Z. University Grants Committee (to S.M.G.) and from the University of Otago Research Committee is acknowledged.

REFERENCES

- 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. & Strassburger, W. (1983) *Eur. J. Biochem.* **135**, 535–542
- Baker, E. N., Blundell, T. L., Cutfield, J. F., Cutfield, S. M., Dodson, E. J., Dodson, G. G., Hodgkin, D. M. C., Hubbard, R. E., Isaacs, N. W., Reynolds, C. D., Sakabe, K., Sakabe, N. & Vijayan, N. M. (1988) *Philos. Trans. R. Soc. London Ser. B* **319**, 369–456
- Blundell, T. L., Pitts, J. E. & Wood, S. P. (1982) *CRC Crit. Rev. Biochem.* **13**, 141–213
- Brooks, B. R., Bruccoleri, R. E., Olafson, B., Stales, D. J., Swaminathan, S. & Karplus, M. (1983) *J. Comput. Chem.* **4**, 187–217
- Conlon, J. M. & Thim, L. (1985) *Gen. Comp. Endocrinol.* **60**, 398–405
- Conlon, J. M. & Thim, L. (1987) in *Protein Purification: Micro-to-Macro* (Burgess, R., ed.), pp. 255–261, A. R. Liss, New York
- Conlon, J. M., Dafgård, E., Falkmer, S. & Thim, L. (1986) *FEBS Lett.* **208**, 445–450
- Conlon, J. M., Dafgård, E., Falkmer, S. & Thim, L. (1987) *Biochem. J.* **245**, 851–856
- Conlon, J. M., Reinecke, M., Thorndyke, M. C. & Falkmer, S. (1988) *Horm. Metab. Res.* **20**, 406–410
- Cutfield, J. F., Cutfield, S. M., Carne, A., Emdin, S. & Falmer, S. (1986) *Eur. J. Biochem.* **158**, 117–123
- Drapeau, G. R. (1977) *Methods Enzymol.* **47**, 189–191
- Falkmer, S., Carraway, R. E., El-Salhy, M., Emdin, S. O., Grimelius, L., Rehfeld, J. F., Reinecke, M. & Schwartz, T. F. W. (1981) *UCLA Forum Med. Sci.* **23**, 21–42
- Falkmer, S., El-Salhy, M. & Titlbach, M. (1984) in *Evolution and Tumour Pathology of the Neuroendocrine System* (Falkmer, S., Hakanson, R. & Sundler, F., eds.), pp. 59–87, Elsevier Science Publishers, Amsterdam
- Mirsky, I. A. (1973) in *Methods in Investigative and Diagnostic Endocrinology* (Berson, S. A. & Yalow, R. S., eds.), vol. 2B, pp. 823–876, North-Holland, Amsterdam
- Moody, A. J., Stan, M. A., Stan, M. & Gliemann, J. (1974) *Horm. Metab. Res.* **6**, 12–16
- Paynovich, R. C. & Carpenter, F. H. (1979) *Int. J. Peptide Protein Res.* **13**, 113–121
- Plisetskaya, E. M., Pollock, H. G., Elliott, W. M., Youson, J. H. & Andrews, P. C. (1988) *Gen. Comp. Endocrinol.* **69**, 46–55
- Pollock, H. G., Kimmel, J. R., Ebner, K. E., Hamilton, J. W., Rouse, J. B., Lance, V. & Rawitch, A. B. (1988) *Gen. Comp. Endocrinol.* **69**, 133–140
- Seino, S., Welsh, M., Bell, G. I., Chan, S. J. & Steiner, D. F. (1986) *FEBS Lett.* **203**, 25–30
- Yalow, R. S. & Berson, S. A. (1964) *N. Engl. J. Med.* **270**, 1171–1178