Eur. J. Biochem. 25 (1972) 153–162

Rat-Proinsulin C-Peptides

Amino-Acid Sequences

Jan MARKUSSEN and Finn SUNDBY

Novo Research Institute, København

(Received September 6/November 5, 1971)

The amino acid sequences of two rat C-peptides corresponding to two rat proinsulins have been elucidated. Both C-peptides have 31 amino acids. There are differences in two positions, namely, in position 8 where rat-I has a proline and rat-II an alanine residue, and in position 17 where rat-I has a glutamic acid, rat-II a glycine residue. The additional acidic residue in rat-I C-peptide is counteracted by the additional basic residue in the insulin moiety of rat-I proinsulin (lysine B 29), which provides the two proinsulins with the same net charge at neutral pH.

The amino acid sequences are:

$$\text{Glu-Val-Glu-Asp-Pro-Gln-Val-}\frac{\text{I: Pro}}{\text{II: Ala}}\text{-Gln-Leu-Glu-Leu-Gly-Gly-Gly-Pro-}\frac{\text{I: Glu}}{\text{II: Gly}}\text{-Ala-Asp-Gly-}$$
$$\text{Leu-Gln-Thr-Leu-Ala-Leu-Glu-Val-Ala-Arg-Gln}$$

The rat C-peptides contain threonine and arginine, neither of which is present in the C-peptides of man, pig and ox.

A comparison of the 5 sequences shows only 10 identical positions out of 31.

In proinsulin, the A- and B-chains of insulin are linked together by "connecting peptide" [1]. The terminals of "connecting peptide" consist of the basic dipeptides, arginyl-arginine and lysyl-arginine, which serve as targets in the proteolytic conversion of proinsulin to insulin and C-peptide [2], the latter is defined as connecting peptide minus the basic dipeptides. Since C-peptide occurs in the pancreas in large amounts as compared to proinsulin, C-peptides have served as convenient materials for the elucidation of the primary structure of bovine [3,4] and human [5,6] proinsulin whereas proinsulin itself was used in the elucidation of the porcine sequence [7].

Methods for preparing ox and pig [8], human [9] and rat C-peptides [10] have been published. The rat C-peptides were found to have 31 amino acids like human C-peptide [10] in contrast to pig and ox C-peptides having 29 and 26 amino acids, respectively.

MATERIALS AND METHODS

Rat C-Peptides

The C-peptide preparations X and Z, previously shown to be rat-II and rat-I C-peptide respectively [10], were each dissolved in 2 ml of water. Amino acid analysis showed the total amounts to be 6 mg of rat-I and 4 mg of rat-II C-peptide.

Enzymes. Trypsin (EC 3.4.4.4); chymotrypsin (EC 3.4.4.5); pepsin (EC 3.4.4.1); papain (EC 3.4.4.10); leucine aminopeptidase (EC 3.4.1.1); carboxypeptidase A (EC 3.4.2.1); carboxypeptidase B (EC 3.4.2.2).

Gel Filtration of Enzymatic Digests

A column (1.5×86 cm) packed with Sephadex G-25 and eluted with 1 mM Na_2HPO_4 was used throughout for gel filtrations and desaltings. The phosphate buffer allowed detection of peptide by measuring the absorbance at 215 nm. In order to keep the column free of contaminations it was operated at 4 °C, and the flow (8 ml/h) was continued between experiments. Fractions of 2 ml were collected.

The distribution coefficient of a peptide, K_d, was calculated from

$$K_d = \frac{V_e - V_0}{V_s - V_0},$$

where V_e is the elution volume of the peptide, V_0 is the void volume, and V_s is the elution volume of salt [11].

The pooled fractions were concentrated prior to enzyme digestion and subsequent gel filtration by rotary evaporation *in vacuo* at 30 °C.

Tryptic Digestion of C-Peptides

The rat C-peptides were digested with trypsin (treated with L-(1-tosylamido-2-phenyl)ethyl chloromethyl ketone, Calbiochem, No 64852) in 5 mM $CaCl_2$, 0.1 M Tris buffer pH 7.8 for 3 h at 37 °C. The ratio of enzyme to substrate was 1:100 (w/w). The digestion mixtures were subjected to gel filtration.

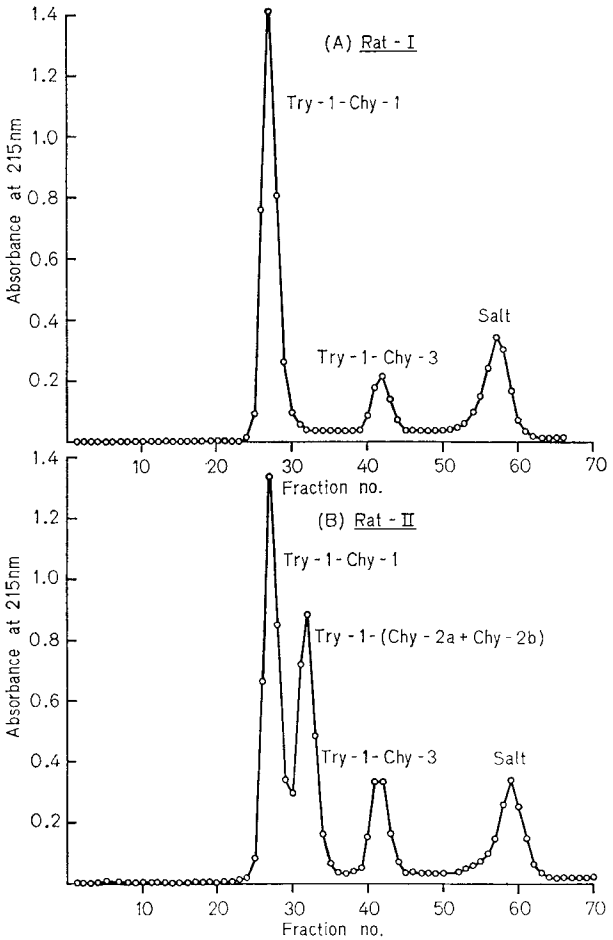

Fig. 1. *Gel filtration on Sephadex G-25 of the mild chymotrypsin digestion of rat-I—Try-1 (A) and rat-II—Try-1 (B).* Both peptides yield Try-1—Chy-3 peptides with K_d values of 0.50. Rat-I yields a single peak in the void volume, Try-1—Chy-1, which in the case of rat-II has been partly split to yield a mixture of two peptides, Try-1—(Chy-2a + Chy-2b)

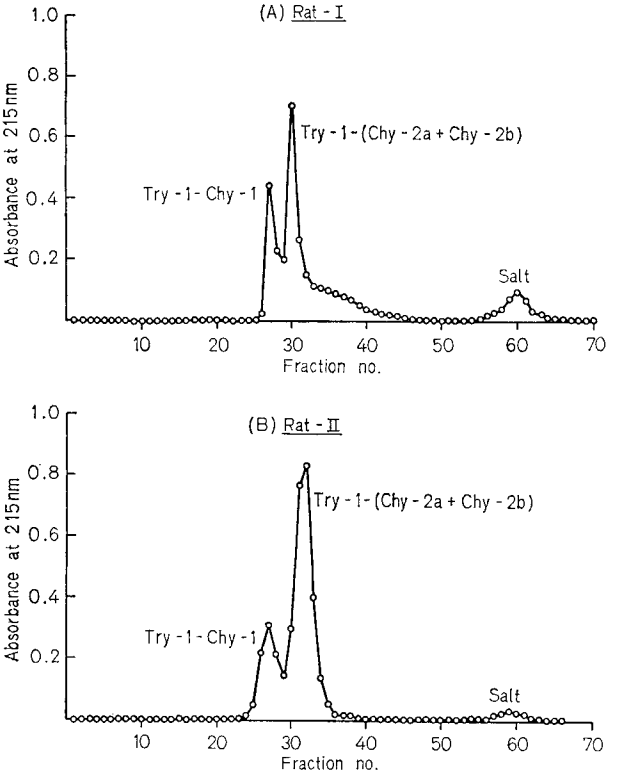

Fig. 2. *Gel filtration of repetitive and heavier chymotrypsin digestion of the Try-1—Chy-1 peptides.* Chy-3 peptide is not seen in any of the digests. In rat-I, some conversion of Try-1—Chy-1 to Try-1—(Chy-2a + Chy-2b) has been accomplished, in rat-II 75% Try-1—Chy-1 has been converted to Try-1—(Chy-2a + Chy-2b)

Both peptides yielded a major peak Try-1 and a minor peak Try-2, the latter including the Tris buffer.

Mild Chymotrypsin Digestion of the Try-1 Peptides

Mild chymotrypsin digestion of the Try-1 peptides, enzyme to substrate ratio 1:100 (w/w), was performed in 0.01 M $CaCl_2$, 0.2 M Tris buffer pH 7.8 at 37 °C for one hour with crystallized α-chymotrypsin obtained from Novo Enzyme Research Division. Gel filtration revealed differences in susceptibility to chymotrypsin between the two peptides (Fig. 1). Rat-I—Try-1 yielded a major peak, rat-I—Try-1—Chy-1, eluting in the same position as rat-I—Try-1, and a minor peak, rat-I—Try-1—Chy-3. Rat-II—Try-1 yielded an additional peak, rat-II—Try-1—(Chy-2a + Chy-2b), eluting between the two other peaks. Rat-II—Try-1—(Chy-2a + Chy-2b) was later resolved into two peptides (rat-II—Try-1—Chy-2a and rat-II—Try-1—Chy-2b) by anion-exchange chromatography on QAE-Sephadex.

Chymotryptic Digestion of the Try-1—Chy-1 Peptides

The rat-I—Try-1—Chy-1 and rat-II—Try-1—Chy-1 peptides were subjected to heavy chymotryptic digestion. After evaporation to dryness, the peptides were taken up in 2 ml 0.1 M Na_2HPO_4 buffer pH 8.0. Rat-I—Try-1—Chy-1 (0.6 mg) was digested for 3 h at 37 °C with an enzyme to substrate ratio of 1:20; rat-II—Try-1—Chy-1 (0.55 mg) was digested for 2 h at 37 °C with an enzyme to substrate ratio of 1:30. Gel filtration of the digests (Fig. 2) showed in both peptides a conversion of most of the Try-1—Chy-1 peptide to the Try-1—(Chy-2a + Chy-2b) mixture, whereas no Try-1—Chy-3 peptides were formed by the second chymotryptic treatment.

Chymotryptic Digestion of C-Peptides

Digestion with chymotrypsin of the C-peptides was carried out as the chymotrypsin digestion of the Try-1—Chy-1 peptides, employing a 0.1 M Na_2HPO_4 buffer pH 8.0 and an enzyme to substrate ratio of 1:30 (w/w). The rat-II C-peptide was converted to a high degree to a (Chy-2a + Chy-2b) mixture and Chy-3, whereas rat-I C-peptide yielded 60% of a Chy-1 peptide, some (Chy-2a + Chy-2b) mixture, 100% Chy-3, and a small peptide, Chy-4, in low yield. Repetitive digestions of Chy-1 were performed in order to produce more Chy-2a + Chy-2b. The results of the experiments with trypsin and chymotrypsin are summarized in Table 1.

Separation of Try-1—(Chy-2a + Chy-2b) and Chy-2a + Chy-2b by Anion-Exchange Chromatography

The mixtures of peptides not resolved by gel filtration (Chy-2a + Chy-2b) were separated on a column (1.4×7 cm) packed with QAE-Sephadex A-25 and equilibrated with 0.1 M NaCl at 4 °C. The peptides dissolved in 1 mM Na_2HPO_4 (gel-filtration buffer) adsorbed completely to the ion-exchanger. Elution was carried out at 4 °C at a rate of 8 ml/h with 0.1 M NaH_2PO_4 buffer adjusted to pH 6.13 with NaOH. Fractions of 4 ml were collected. When the first peak had eluted (Chy-2b, Fig.3), the buffer was changed to 0.1 M NaH_2PO_4 adjusted to pH 2.96 with H_3PO_4.

Pepsin Digestion of the Chy-2b Peptides

60 nmol rat-II—Chy-2b in 0.1 M NaH_2PO_4 buffer was digested at 37 °C with 25 µg pepsin (Sigma, 2×cryst., No P-6875) after adjusting the pH to 2.2. After 3 h the digestion was stopped by adjusting the pH to 8.0. The digest was gel-filtered yielding 2 peptides, Chy-2b—Pep-1 and Chy-2b—Pep-2. Data of the peptic peptides are compiled in Table 2.

Pepsin Digestion Followed by Carboxypeptidase Digestion of Rat-II—Chy-2b

60 nmol rat-II—Chy-2b was digested with pepsin as described above. After 3 h the pH was adjusted to 8.0 and 150 µl 2 M $CaCl_2$ was added in order to precipitate the phosphate. To the suspension was added 400 µg carboxypeptidase A (Schwarz No 0110-14) and digestion allowed to occur for 3 h at 37 °C. The suspension was cleared by adding solid EDTA before gel filtration. A peptide, Chy-2b—(Pep + Carbox)-1, was isolated (Table 2).

Hydrazinolysis of Chy-2b—(Pep + Carbox)-1

The peptide Chy-2b—(Pep + Carbox)-1 was evaporated to dryness *in vacuo* and subjected to hydrazinolysis with 100 µl anhydrous hydrazine in a sealed ampoule for 20 h at 80 °C [12]. The hydrazine was removed by evaporation *in vacuo*, and the products of reaction examined directly on the amino acid analyser.

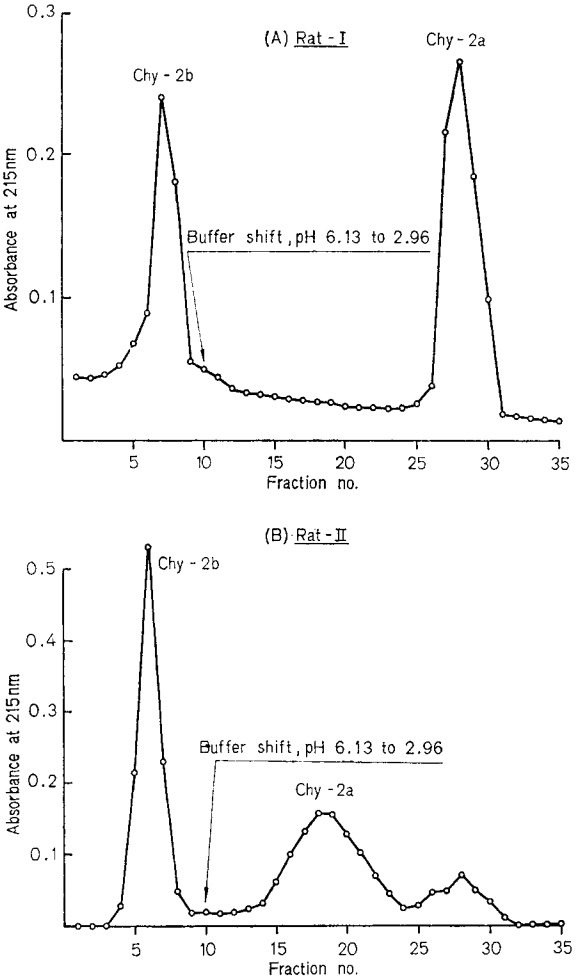

Fig.3. *Separation of the mixtures of Chy-2a + Chy-2b on a column (1.4×7 cm) of QAE-Sephadex A-25 employing two 0.1 M NaH_2PO_4 buffers with pH values adjusted to 6.13 and 2.96, respectively*

Papain Digestion of the Chy-2b Peptides

60 nmol rat-II—Chy-2b in 0.1 M NaH_2PO_4 buffer was adjusted to pH 8.0. 2-Mercaptoethanol was added to 1 mM followed by 8 µg papain (Sigma, 2×cryst., No P-3125). After 3 h at 37 °C, the solution was gel-filtered, yielding two peptides, Chy-2b—Pap-1, and Chy-2b—Pap-2.

There was no cleavage of rat-I—Chy-2b under the same conditions. The data of the papain peptides of rat-II—Chy-2b are compiled in Table 2.

Table 1. *Data on the tryptic and chymotryptic*

C-peptide	Substrate (S)	Amount	Enzyme (E)	Ratio E:S (w/w)	Time of digestion at 37 °C	Peptides isolated by gel filtration	Yield
		mg			h		%
Rat-I	C-peptide	1.5	Trypsin	1:100	3	Try-1	90
						Try-2	75
	Try-1	1.0	Chymotrypsin	1:100	1	Try-1—Chy-1	90
						Try-1—Chy-3	95
	Try-1—Chy-1	0.7	Chymotrypsin	1: 20	3	Try-1—Chy-1	20
						Try-1—Chy-2a, Try-1—Chy-2b	50
	C-peptide	1.5	Chymotrypsin	1: 30	3	Chy-1	60
						Chy-2a, Chy-2b	30
						Chy-3	85
						Chy-4	25
Rat-II	C-peptide	1.0	Trypsin	1:100	3	Try-1	90
						Try-2	65
	Try-1	0.75	Chymotrypsin	1:100	1	Try-1—Chy-1	55
						Try-1—Chy-2a, Try-1—Chy-2b	35
						Try-1—Chy-3	90
	Try-1—Chy-1	0.4	Chymotrypsin	1: 30	2	Try-1—Chy-1	25
						Try-1—Chy-2a, Try-1—Chy-2b	75
	C-peptide	2.0	Chymotrypsin	1: 30	5	Chy-1	<5
						Chy-2a, Chy-2b	95
						Chy-3	100

[a] No corrections made for losses during acid hydrolysis.
[b] Too-low values for leucine probably caused by carboxypeptidase impurities in the chymotrypsin preparation.

Digestion by Leucine Aminopeptidase of Chy-2a Peptides

Leucine aminopeptidase (Sigma, No L-9876) was used to degrade the Chy-2a peptides up to the imide bond of the first proline residue in the Chy-2a peptides. The desalted peptides (150 nmol) were taken up in 0.01 M $MgSO_4$, 0.1 M Tris buffer adjusted to pH 8.5 with HCl and digested with 80 μg leucine aminopeptidase for 20 h at 37 °C. The reaction mixture was gel-filtered, and the isolated peptides, Chy-2a—Lap-1 were used for Edman degradation studies. Data on the peptides are shown in Table 2.

Determinations of N-Terminal Sequence with Leucine Aminopeptidase

Portions of peptides (4—10 nmol) dissolved in 1 ml 0.01 M $MgSO_4$, 0.1 M Tris buffer pH 8.5 were digested with 6—200 μg leucine aminopeptidase for different periods of time. The reactions were stopped by the addition of 150 μl 1 N HCl, and the digests examined directly on the amino acid analyser.

Determinations of C-Terminal Sequence with Carboxypeptidase

Portions of peptides (5—50 nmol) dissolved in 1 ml 0.01 M $CaCl_2$, 0.2 M Tris buffer pH 7.8 were digested at 37 °C with 10—100 μg of carboxypeptidase A for varying periods of time. The reaction was stopped with 200 μl 1 N HCl, and the digests examined directly on the amino acid analyser.

The Try-1—Chy-3 peptides with C-terminal arginine were digested with a mixture of carboxypeptidase A and carboxypeptidase B (Schwarz, No 0110-17). 20 nmol peptide was digested with 3 μg of each of the carboxypeptidases in 1 ml 0.01 M $CaCl_2$ 0.2 M Tris buffer pH 7.8 for 3 h at 37 °C. The digestion was stopped by the addition of 200 μl 1 N HCl, and the mixture diluted to 2.4 ml for analysis on both columns on the amino acid analyser.

Degradation According to Edman

The adopted variety of the Edman procedure was that of Gray [13] using 5 nmol per step. Subtractive amino acid analysis was used to determine

peptides derived from the rat C-peptides

Distribution coefficient, K_d, on Sephadex G-25	Amino acid composition of peptides										
	Arg	Asp	Thr[a]	Glu	Pro	Gly	Ala	Val	Leu	Total	NH_3
	residues/molecule										
0											
1				1						1	
0		2.0	0.8	7.3	2.6	4.1	1.0	2.2	3.8	24	
0.50	0.9			1.1			2.0	1.1	0.9	6	
0											
0.09											
0.13											
0											
0.09		1.0		4.8	1.9			2.1	1.8[b]	12	
0.13		0.9	0.7	2.1	1.0	3.8	1.1		1.5[b]	12	
0.43	0.9			2.1			2.0	1.0	1.0	7	1.0
0.75			0.9	0.9					1.2	3	
0											
1				1						1	
0											
0.12		1.0		4.0	0.8		1.0	2.1	1.0	10	
0.12		1.0	0.9	2.2	0.9	4.8	1.1		2.9	14	
0.50	1.0			1.1			2.1	1.0	0.9	6	
0											
0.12											
0.12											
0											
0.12		1.0		3.8	1.1		1.2	1.9	0.9	10	
0.12		0.9	0.8	2.0	1.1	4.8	1.0		2.9	14	
0.43	1.0			2.0			2.0	1.0	0.9	7	1.0

the sequence. In one case where only 1 nmol was available per step (rat-II—Chy-2b—Pap-1), dansylation was employed to establish the sequence as described by Gros and Labouesse [14].

Electrophoresis of Peptides

When assignments of the amides were impossible by enzymatic release of the amino acids, electrophoresis of the peptides obtained from the Edman degradations of 20 nmol peptide was performed on Whatman paper No 1 in 0.2 M pyridine acetate pH 6.5 for 2 h at 4 °C using a voltage gradient of 30 V/cm.

The paper was dried at 80 °C, dipped in a ninhydrin solution (0.3 g/100 ml acetone), and allowed to develop at room temperature.

Amino-Acid Analysis

The Jeolco Model 5-AH automatic amino acid analyser was employed using the dual-column system [15]. Acid hydrolysis was carried out with 500 µl 6 N HCl in vacuum-sealed ampoules for 20 h at 110 °C.

RESULTS

The amino acid compositions of the prepared peptides are given in Tables 1 and 2. The additional data that establish the amino acid sequences of rat-I and rat-II C-peptides are the following:

Rat-I and Rat-II, Try-2: Gln (Residue 31)

The tryptic fractions, Try-2, were examined directly on the amino acid analyser. A peak was observed in the position of serine. After acid hydrolysis glutamic acid was found, proving the amino acid released to be glutamine. Since only one basic amino acid was present, glutamine must be the C-terminal amino acid and arginine the penultimate, that is No 30, in both C-peptides.

Rat-I and Rat-II, Try-1—Chy-3: Ala-Leu-Glu-Val-Ala-Arg (Residues 25—30)

The presence of arginine in Try-1—Chy-3 places the peptides as the C-terminal part of Try-1.

Table 2. *Data on peptides obtained by the action of pepsin, papain and leucine aminopeptidase on the chymotryptic peptides of C-peptide*

C-peptide	Substrate	Enzyme	Peptides isolated by gel filtration	Distribution coefficient, K_d, on Sephadex G-25	Amino acid composition of peptides								
					Asp	Thr[a]	Glu	Pro	Gly	Ala	Val	Leu	total
					residues/molecule								
Rat-I	Chy-2b	Papain	Chy-2b[b]	0.13	0.9	0.6	2.2	1.2	3.8	1.0		1.4	12
	Chy-2a	Leucine-aminopeptidase	Chy-2a—Lap-1	0.13	1.1		3.1	2.3			1.0	1.6[c]	9 (8)
	Chy-2b	Pepsin	Chy-2b—Pep-1	0.16	1.0		1.2	1.3	4.8	1.0		1.8	11
			Chy-2b—Pep-2	0.75		1.0	1.1					0.9	3
	Chy-2b	Pepsin + Carboxypeptidase	Chy-2b—(Pep + Carbox)-1	0.22	1.0		1.1	1.0	4.9	1.0		1.1	10
Rat-II	Chy-2b	Papain	Chy-2b—Pap-1	0.23	1.0	0.8	1.4	1.1	3.8	0.9		1.9	11
			Chy-2b—Pap-2	0.47			1.1		1.1			0.8	3
	Chy-2a	Leucine aminopeptidase	Chy-2a—Lap-1	0.34	1.1		2.0	1.0		1.1	0.8	0.2[c]	7 (6)

[a] No corrections made for losses during hydrolysis.
[b] No degradation of rat-I—Chy-2b with papain.
[c] Due to carboxypeptidase activity in the leucine aminopeptidase, C-terminal leucine has been released, rendering the values for leucine too small (1.6 for 2.0 and 0.2 for 1.0).

Edman Degradation, Rat-I

First step: Glu 1.0; **Ala 1.0**; Val 1.0; Leu 1.0
Second step: Glu 1.0; Ala 0.9; Val 1.0; **Leu 0.2**
Third step: **Glu 0.5**; Ala 1.0; Val 1.0; Leu 0.1
Fourth step: Glu 0.5; Ala 1.0; **Val 0.8**; Leu 0.1

Rat-II

First step: Arg 0.9; Glu 1.2; **Ala 1.1**; Val 0.9; Leu 0.9
Second step: Arg 0.8; Glu 1.1; Ala 1.0; Val 1.1; **Leu 0.3**
Third step: Arg 1.1; **Glu 0.5**; Ala 1.0; Val 0.9; Leu 0.0
Fourth step: Arg 1.1; Glu 0.4; Ala 0.9; **Val 0.4**; Leu 0.0

Carboxypeptidase A + Carboxypeptidase B

Amino acids released, 3 h:
Rat-I: Arg 1.0; Ala 0.7; Val 0.6
Rat-II: Arg 1.0; Ala 0.9; Val 0.9

The mixture of carboxypeptidases failed to release amino acids beyond the third residue from the C-terminal. The next residue is glutamic acid rather than glutamine, because the ammonia analysis of the Chy-3 heptapeptides (25—31), Table 1, gives only one mole of ammonia which arises from glutamine 31, and consequently No. 27 must be glutamic acid.

The sequences of the two hexapeptides are identical and fully established by the four Edman cycles in combination with the knowledge of arginine being C-terminal. The results of the carboxypeptidase digestions confirm the sequences.

Rat-I—Chy-1 (Residues 1—24)

This peptide proved more resistant to chymotrypsin than the corresponding rat-II peptide. It was used for terminal studies only, allowing positioning of the constituent chymotryptic peptides, Chy-2a and Chy-2b.

Carboxypeptidase A. Amino acids released after 3 h at 37 °C with 9 μg enzyme to 7 nmol peptide were leucine (1.2), threonine (0.8) and a peak in the position of serine (0.6). Gel filtration of the mixture of digestion followed by acid hydrolysis of the free amino acids in the salt peak gave glutamic acid, showing the third amino acid released to be glutamine. Hence, the sequence Leu-Gln-Thr-Leu (residues 21—24) is likely to occur in the C-terminal part of Chy-1, taking into account the specificity of chymotrypsin, which places the first leucine in position 24.

Leucine Aminopeptidase. Amino acids released after 6 h at 37 °C with 6 μg enzyme to 7 nmol peptide were glutamate (1.6) and valine (1.0). This establishes the N-terminal sequence in Chy-1 (and C-peptide) to be (Glu-Val)Glu.

Rat-I—Chy-4: Gln-Thr-Leu (Residues 22—24)

This peptide obtained from the chymotryptic digestion of rat-I C-peptide was found to be identical to the peptic peptide rat-II—Chy-2b—Pep-2 (see later).

Rat-II—Chy-2a:
Glu-Val-Glu-Asp-Pro-Gln-Val-Ala-Gln-Leu
(Residues 1—10)

Edman Degradation

First step: Asp 1.1; **Glu 3.3**; Pro 1.1; Ala 1.0; Val 1.8; Leu 1.0
Second step: Asp 1.2; Glu 3.2; Pro 1.0; Ala 1.0; **Val 1.0**; Leu 0.8
Third step: Asp 1.0; **Glu 2.8**; Pro 1.0; Ala 1.0; Val 1.1; Leu 1.0
Fourth step: **Asp 0.5**; Glu 2.7; Pro 1.1; Ala 1.0; Val 1.0; Leu 0.9

Carboxypeptidase. Amino acids released at 37 °C with 10 μg enzyme to 7 nmol Chy-2a were:
1 h, Gln 0.5; Ala 0.20; Val 0.16; Leu 1.0.
2 h, Gln 0.8; Ala 0.56; Val 0.50; Leu 1.0.
7 h, Gln 1.0; Ala 1.0; Val 0.95; Leu 1.0.

The sequence Val-Ala-Gln-Leu (residues 7—10) is suggested from the carboxypeptidase experiments.

The sequence Glu-Val-Glu-Asp is established by the Edman degradation. Since the N-terminal amino acids found in Chy-1 were glutamic acid and valine, Chy-2a is the N-terminal of the chymotryptic peptides.

From Chy-2a the peptide Chy-2a—Lap-1 was prepared (Table 2). The first three residues are removed quantitatively. The lack of leucine can be explained by impurities of carboxypeptidase in the leucine aminopeptidase preparation.

Edman Degradation of Chy-2a—Lap-1

First step: **Asp 0.2**; Glu 2.1; Pro 1.0; Ala 1.0; Val 0.9.
Second step: Glu 2.1; **Pro 0.3**; Ala 0.9; Val 0.8.
Third step: **Glu 1.6**; Ala 1.0; Val 0.9.
Fourth step: Glu 1.4; Ala 1.0; **Val 0.3.**
Fifth step: Glu 1.4; **Ala 0.5.**

The two overlapping Edman degradations established the sequence of residues 1—8 and confirmed the carboxypeptidase experiments.

The free amino acids in the salt pool from the gel filtration of the digest of Chy-2a with leucine aminopeptidase were analysed: Glu 2.0; Val 1.0; Leu 0.3. Hence, glutamic acids can be assigned to positions 1 and 3.

Electrophoresis of Chy-2a—Lap-1
and Its Edman Derivatives

Chy-2a—Lap-1 and the products of the three first Edman degradations hereof were subjected to electrophoresis in order to determine the position of the amide group around the proline residue. A drop in anodic mobility was seen between Chy-2a—Lap-1 and the first Edman derivative, but the first, second and third derivatives were in line, meaning that residue No 4 is aspartic acid and No 6 glutamine.

Rat-II—Chy-2b: Glu-Leu-Gly-Gly-Gly-
Pro-Gly-Ala-Asp-Gly-Leu-Gln-Thr-Leu
(Residues 11—24)

Edman Degradation

First step: Asp 1.0; Thr 1.1; **Glu 1.4**; Pro 1.0; Gly 4.9; Ala 1.2; Leu 2.7.
Second step: Asp 1.0; Thr 0.9; Glu 1.3; Pro 1.1; Gly 4.8; Ala 1.2; **Leu 1.9.**
Third step: Asp 1.0; Thr 0.9; Glu 1.3; Pro 1.1; **Gly 4.1**; Ala 1.2; Leu 1.9.
Fourth step: Asp 1.0; Thr 0.8; Glu 1.3; Pro 1.1; **Gly 3.2**; Ala 1.1; Leu 1.8.
Fifth step: Asp 1.0; Thr 0.8; Glu 1.3; Pro 1.1; **Gly 2.6**; Ala 1.0; Leu 1.7.
Sixth step: Asp 1.0; Thr 0.8; Glu 1.3; **Pro 0.4**; Gly 2.3; Ala 1.0; Leu 1.7.

The above establish the first six residues of the peptide.

Leucine Aminopeptidase. Amino acids released at 37 °C from 10 nmol Chy-2b with 20 μg leucine aminopeptidase were:
3 h, Glu 0.5; Leu 0.4; Gly 0.3.
6 h, Glu 0.9; Leu 0.6; Gly 0.8.
24 h, Glu 1.0; Leu 1.3; Gly 2.6; Ala 0.6; Asp 0.3; Thr 0.1; Gln 0.1; Gly-Pro dipeptide present.

Digestion of 5 nmol Chy-2b for 24 h at 37 °C with 200 and 100 μg leucine aminopeptidase released the following amino acids:
200 μg: Glu 1.0; Leu 1.8; Gly 3.5; Ala 0.9; Asp 0.4; Thr 0.4; Gln 0.4; Gly-Pro dipeptide present.
100 μg: Glu 1.0; Leu 1.3; Gly 3.1; Ala 0.8; Asp 0.3; Thr 0.15; Gln 0.15; Gly-Pro dipeptide present.

The N-terminal is glutamic acid and not glutamine. A peak on the amino acid analyser is observed just after leucine, which is identified as the dipeptide Gly-Pro [3,5]. Leucine aminopeptidase is capable of by-passing a proline residue when glycine is positioned before proline. Hence, the sequence Gly-Pro (residues 15—16) is confirmed. From the experiments with leucine aminopeptidase the sequence Glu-Leu-Gly-Gly-Gly-Pro-Gly-Ala-(Asp-Gly) (residues 11—20) is suggested, the first six residues agreeing with the Edman degradations.

Carboxypeptidase. Digestion of 50 nmol Chy-2b with 10 μg carboxypeptidase at 37 °C for 6 h released leucine (1.6), threonine (0.9) and glutamine (0.9). Half of the digest was gel-filtered, and the salt peak subjected to acid hydrolysis. Amino acid analysis showed glutamic acid, confirming that glutamine was released. The amino acids released correspond

to the C-terminal of Chy-1, meaning Chy-2b is the C-terminal peptide of Chy-1.

Rat-II—Chy-2b—Pep-1: Glu-Leu-Gly-Gly-Gly-Pro-Gly-Ala-Asp-Gly-Leu (Residues 11—21)

Carboxypeptidase. Heavy treatment with carboxypeptidase (100 μg to 12 nmol peptide for 20 h at 37 °C) released only the C-terminal leucine.

Carboxypeptidase plus Hydrazinolysis. The residual peptide from carboxypeptidase treatment of Chy-2b—Pep-1 released upon hydrazinolysis glycine (residue 20).

Rat-I—Chy-4 and Rat-II—Chy-2b—Pep-2: Gln-Thr-Leu (Residues 22—24)

Upon standing in the gel-filtration buffer 1 mM Na_2HPO_4), the peptide cyclized into the pyrrolidone form, indicating that glutamine is N-terminal in agreement with the results from carboxypeptidase digest of rat-I—Chy-1 and rat-II—Chy-2b. Dansylation of the peptide before the cyclization was complete, revealed glutamic acid. Carboxypeptidase released leucine in agreement with the results from Chy-1 and Chy-2b.

Rat-II—Chy-2b—Pap-1: Gly-Gly-Pro-Gly-Ala-Asp-Gly-Leu-Gln-Thr-Leu (Residues 14—24)

Six cycles of Edman degradation using 1 nmol peptide per cycle followed by dansylation, revealed the sequence Gly-Gly-Pro-Ala-Asp-Gly (residues 14 to 20).

Rat-II—Chy-2b—Pap-2: Glu-Leu-Gly (Residues 11—13)

Dansylation of 3 nmol peptide revealed glutamic acid. Carboxypeptidase digestion with 10 μg enzyme to 9 nmol peptide did not result in release of any amino acids, indicating that glycine is the C-terminal amino acid. The tripeptide was observed on the amino acid analyser in the position of isoleucine.

Rat-I—Chy-2a: Glu-Val-Glu-Asp-Pro-Gln-Val-Pro-Gln-Leu-Glu-Leu (Residues 1—12)

Rat-I—Chy-2a was obtained in a purity of 75%, the impurity being the peptide 1—21. Thus the Leu-Gln bond (residues 21—22) apparently has a susceptibility comparable to the Leu-Gly bond (residues 12—13). No cleavage was found at the Leu-Glu bond (residues 10—11). Corrections for the impurity were introduced by means of the single residue alanine in the peptide 1—21 (residue 18).

Edman Degradation

First step: Asp 1.0; **Glu 3.8**; Pro 1.9; Val 1.9; Leu 1.8.
Second step: Asp 1.0; Glu 3.8; Pro 1.9; **Val 1.0**; Leu 1.8.
Third step: Asp 1.0; **Glu 3.1**; Pro 1.9; Val 1.0; Leu 1.8.
Fourth step: **Asp 0.7**; Glu 3.1; Pro 1.9; Val 0.9; Leu 1.8.
Fifth step: Asp 0.6; Glu 3.3; **Pro 1.4**; Val 1.0; Leu 1.8.
Sixth step: Asp 0.6; **Glu 2.9**; Pro 1.3; Val 1.0; Leu 1.8.
Seventh step: Asp 0.6; Glu 2.9; Pro 1.3; **Val 0.8**; Leu 1.8.
Eighth step: Asp 0.6; Glu 2.9; **Pro 1.0**; Val 0.8; Leu 1.8.

This established the N-terminal sequence of rat-I C-peptide to be Glu-Val-Glu-Asp-Pro-Gln-Val-Pro, differing from rat-II in position 8, where rat-II exhibits alanine. Positions 9, 10 and 11 were assigned by analogy to rat-II only.

Carboxypeptidase. Digestion of 5 nmol Chy-2a with 10 μg carboxypeptidase for 4 h at 37 °C released only leucine (1.0).

Rat-I—Chy-2a—Lap-1: Asp-Pro-Gln-Val-Pro-Gln-Leu-Glu-Leu (Residues 4—12)

Heavy digestion with leucine aminopeptidase followed by gel filtration resulted in a preparation, rat-I—Chy-2a—Lap-1 (residues 4—12). The low value for leucine (Table 2) is most likely a result of some carboxypeptidase activity in the leucine aminopeptidase used. Electrophoresis of Chy-2a—Lap-1 and its first three Edman derivatives established the position of the amide group around the proline residue to be **Asp**-Pro-**Gln** (residues 4—6) as in rat-II C-peptide.

Rat-I—Chy-2b: Gly-Gly-Gly-Pro-Glu-Ala-Asp-Gly-Leu-Gln-Thr-Leu (Residues 13—24)

Leucine Aminopeptidase. Four samples of 4 nmol peptide were digested with 10 μg leucine aminopeptidase each at 37 °C for 1, 2, 4 and 20 h. A fifth sample was digested 20 h with 20 μg.

1 h, Gly 0.8.
2 h, Gly 1.2.
4 h, Gly 1.4; Glu 0.15; (Gly-Pro) dipeptide just visible.
20 h, Gly 2.3; Glu 0.75; Ala 0.75; Asp 0.3; Leu 0.2; Gln 0.1; Thr 0.1; (Gly-Pro) dipeptide present.
20 h, 20 μg, Gly 2.5; Glu 0.9; Ala 0.9; Asp 0.5; Leu 0.5; Gln 0.4; Thr 0.2; (Gly-Pro) dipeptide present.

The sequence Gly-Gly-Gly-Pro-(Glu-Ala)-(Asp-Gly) can be deduced, knowing that Gly-Pro is released as a dipeptide.

Edman Degradation

First step: Asp 1.0; Thr 0.5; Glu 2.4; Pro 1.2; **Gly 3.1**; Ala 1.0; Leu 1.5.
Second step: Asp 1.0; Thr 0.5; Glu 2.3; Pro 1.2; **Gly 2.1**; Ala 1.0; Leu 1.4.
Third step: Asp 0.9; Thr 0.6; Glu 2.3; Pro 1.1; **Gly 1.4**; Ala 1.0; Leu 1.4.

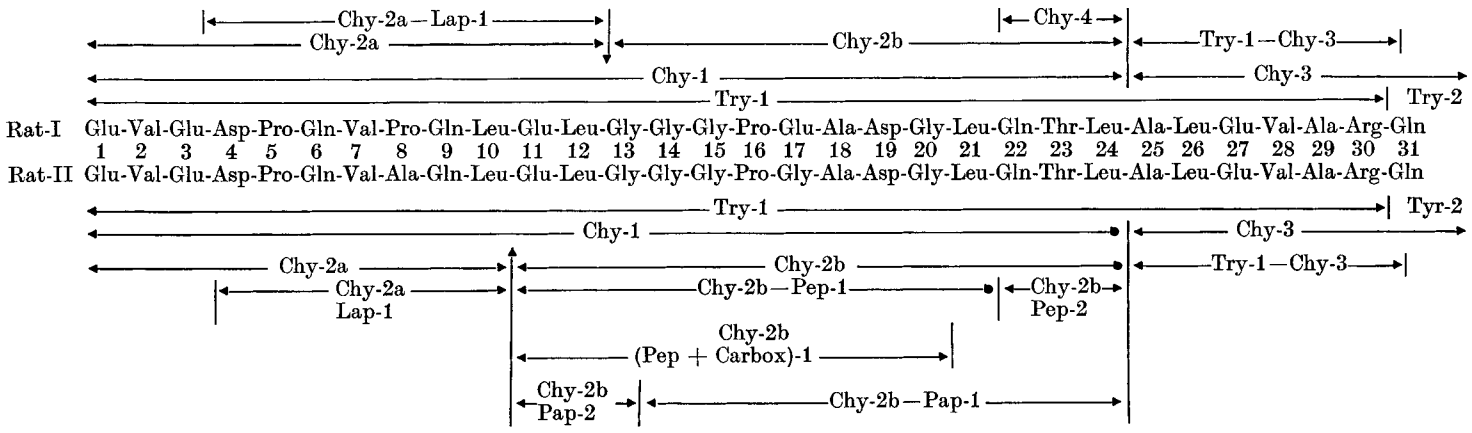

Fig. 4. *Alignment of peptides isolated from rat C-peptides by the action of trypsin, chymotrypsin, pepsin, papain, leucine aminopeptidase, and carboxypeptidase.* The difference in cleavage with chymotrypsin is indicated with vertical arrows

Position	1		3			6	8			11	12		14	15	17											26	27				31
Man	Glu	Ala	Glu	Asp	Leu	Gln	Val	Gly	Gln	Val	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Ser	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Gly	Ser	Leu	Gln
Pig	Glu	Ala	Glu	Asn	Pro	Gln	Ala	Gly	Ala	Val	Glu	Leu	Gly	Gly	Gly	Leu	Gly	Gly	—	—	Leu	Gln	Ala	Leu	Ala	Leu	Glu	Gly	Pro	Pro	Gln
Ox	Glu	Val	Glu	Gly	Pro	Gln	Val	Gly	Ala	Leu	Glu	Leu	Ala	Gly	Gly	Pro	Gly	Ala	Gly	Gly	—	—	—	—	—	Leu	Glu	Gly	Pro	Pro	Gln
Rat-I	Glu	Val	Glu	Asp	Pro	Gln	Val	**Pro**	Gln	Leu	Glu	Leu	Gly	Gly	Gly	Pro	**Glu**	Ala	Asp	Gly	Leu	Gln	Thr	Leu	Ala	Leu	Glu	Val	Ala	Arg	Gln
Rat-II	Glu	Val	Glu	Asp	Pro	Gln	Val	**Ala**	Gln	Leu	Glu	Leu	Gly	Gly	Gly	Pro	**Gly**	Ala	Asp	Gly	Leu	Gln	Thr	Leu	Ala	Leu	Glu	Val	Ala	Arg	Gln

Fig. 5. *Comparison of the sequences of human, pig, ox and rat C-peptides, showing the 10 positions where no variation has been observed.* (Positions 1, 3, 6, 11, 12, 14, 15, 26, 27, and 31). The differences between the two rat C-peptides are shown (positions 8 and 17) in bold-face type

Fourth step: Asp 0.9; Thr 0.6; Glu 2.2; **Pro 0.5**; Gly 1.3; Ala 1.0; Leu 1.4.
Fifth step: Asp 0.9; Thr 0.6; **Glu 1.8**; Pro 0.3; Gly 1.3; Ala 1.0; Leu 1.5.

The sequence found, Gly-Gly-Gly-Pro-Glu (residues 13—17), combined with the leucine aminopeptidase experiments confirms that glutamic acid is in position 17. Rat-II had glycine in this position.

DISCUSSION

The peptides prepared and sequenced are aligned in Fig. 4. The differences between the two C-peptides are seen in position 8 (Pro→Ala) and in position 17 (Glu→Gly). The additional negative charge in rat-I C-peptide (C-17 glutamic acid) counteracts the additional positive charge in rat-I insulin (B-29 lysine), giving the two proinsulins the same net charge at neutral pH values [10]. Both differences can be explained by single base mutation in the corresponding codons. The sequences fit with the reported amino acid compositions of the C-peptides [10].

The differences in susceptibility and point of cleavage with chymotrypsin are interesting. Rat-II C-peptide is cleaved more easily than rat-I C-peptide, and the bonds being cleaved are in rat-II: leucine-glutamic acid (residues 10—11), and in rat-I: leucine-glycine (residues 12—13). The most likely explanation is that proline (residue 8) distorts the rat-I C-peptide chain near the leucine (residue 10), impairing enzyme-substrate complex formation decreasing the cleavage of the leucine-glutamic acid bond (residues 10—11). Instead, the point of cleavage is moved two residues to the leucine-glycine bond (residues 12—13), where proline is not part of the complex formation site. Human C-peptide with valine in position 10 is cleaved as rat-I between leucine-glycine (residues 12—13) [5]. The most sensitive point of cleavage to chymotrypsin, leucine-alanine (residues 24—25) is common for human, pig, rat-I and rat-II C-peptides.

A comparison of the five known C-peptides is seen in Fig. 5. The number of identical amino acid positions are 10 out of 31 possible. The amino acids in the rat C-peptides are, with one exception, related to those in the three other species by single base mutations, the exception being proline (residue 8) in rat-I. In position 17, rat-II C-peptide shows identity with the three other species, whereas rat-I is related with a single base mutation. Combined, these results indicate a possibility of rat-II C-peptide being closer to the common ancestor C-peptide than rat-I C-peptide. C-peptides might prove useful in phylogenetic studies because of the occurrence of a high degree of variation between species.

The six-side-chain carboxyl groups and the one arginine residue found in rat-II C-peptide are consistent with the same electrophoretic mobility exhibited by human C-peptide (5 side-chain carboxyl groups and no basic residues), and the faster mobility of rat-I C-peptide fits with the 7 side-chain carboxyl groups and the single arginine residue [10]. The separation of the Chy-2a and Chy-2b peptides on the anion-exchanger fits well with the charge distribution between these peptides. Rat-II has 2 and 3 side-chain carboxyl groups in Chy-2b and Chy-2a, respectively; rat-I has 2 and 4 side-chain carboxyl groups in the analog peptides.

The mechanism of C-peptide for the folding of proinsulin [16] has not been clarified by the addition of two more sequences. Of the ten invariable amino acids, four residues are found to be glutamic acid (residues 1, 3, 11 and 27) and possibly important for the formation of ionic bonds to the basic residues in the B-chain and to the basic residues at the terminals of connecting peptide.

The amino acid compositions of the rat proinsulins [10] showed one lysine and three arginine residues more than the calculated sum of the residues from insulins and C-peptides. Most likely, the rat C-peptides are connected to the insulin B-chain through the dipeptide Arg-Arg, and to the insulin A-chain through the dipeptide Lys-Arg, as has been shown to be the case for the proinsulins examined so far.

The authors are indebted to Mrs Lene Drube and Mrs Edith Thomsen for their excellent technical skills.

REFERENCES

1. Steiner, D. F., Hallund, O., Rubenstein, A. H., Cho, S., and Bayliss, C., *Diabetes*, 17 (1968) 725.
2. Kemmler, W., Peterson, J., Borg, J., Nehrlich, S., and Steiner, D. F., *Diabetes*, 20 (1971) 332.
3. Salokangas, A., Smyth, D. G., Markussen, J., and Sundby, F., *Eur. J. Biochem.* 20 (1971) 183.
4. Steiner, D. F., Cho, S., Oyer, P. E., Terris, S., Peterson, J. D., and Rubenstein, A. H., *J. Biol. Chem.* 246 (1971) 1365.
5. Ko, A. S. C., Smyth, D. G., Markussen, J., and Sundby, F., *Eur. J. Biochem.* 20 (1971) 190.
6. Oyer, P. E., Cho, S., Peterson, J. D., and Steiner, D. F., *J. Biol. Chem.* 246 (1971) 1375.
7. Chance, R. E., Ellis, R. M., and Bromer, W. W., *Science (Washington)*, 161 (1968) 165.
8. Sundby, F., and Markussen, J., *Horm. Metab. Res.* 2 (1970) 17.
9. Markussen, J., Sundby, F., Smyth, D. G., and Ko, A. S. C., *Horm. Metab. Res.* 3 (1971) 229.
10. Sundby, F., and Markussen, J., *Eur. J. Biochem.* 25 (1972) 147.
11. Flodin, P., *Dextran Gels and Their Applications in Gel Filtration*, Pharmacia, Uppsala 1962, p. 47.
12. Fraenkel-Conrat, H., and Tsung, C. M., *Methods Enzymol.* 11 (1967) 151.
13. Gray, W. R., *Methods Enzymol.* 11 (1967) 469.
14. Gros, C., and Labouesse, B., *Eur. J. Biochem.* 7 (1969) 463.
15. Moore, S., Spackman, D. H., and Stein, W. H., *Anal. Chem.* 30 (1958) 1195.
16. Markussen, J., *Int. J. Protein Res.* 3 (1971) 201.

J. Markussen and F. Sundby
Novo Research Institute
Nordre Fasanvej 215, DK-2200 København N, Denmark