PRIMARY STRUCTURE OF THE A CHAIN OF PORCINE RELAXIN.

Christian Schwabe, J. Ken McDonald,
Department of Biochemistry
Medical University of South Carolina
Charleston, South Carolina 29401
and
Bernard G. Steinetz
Ciba-Geigy Corporation
Ardsley, New York

Received March 19,1976

SUMMARY: Relaxin, a peptide hormone previously purified from ovaries of pregnant hogs, has been reduced, alkylated, and separated into its component A and B chains. The A chain contains 22 residues and the B chain 30 residues. Evidence is presented here that the sequence of the A chain of relaxin is Arg-Met-Thr-Leu- Ser-Glu-Lys-Cys-Cys-Glu-Val-Gly-Cys-Ile-Arg-Lys-Asp-Ile-Ala-Arg-Leu-Cys (2500 daltons). While no homology to the insulin A chain exists the relative distribution of the cysteine residues is identical if the N-terminal glycine of the insulin A chain is aligned with the threonine residue of the A chain of relaxin.

Relaxin is an ovarian peptide hormone responsible for the dilation of the birth canal in humans and most mammals (1, 2). The hormonal activity has been recognized since 1926 (3) but only recently has it been possible to isolate sufficient material for chemical analysis (4). An efficient method for the separation of the A and B chain of relaxin, recently devised in our laboratory, has opened the way for sequence analysis.

MATERIALS AND METHODS: Relaxin was purified from an acetone-HCl extract (supplied by the National Institute of Arthritis, Metabolism, and Digestive Diseases) by ion exchange chromatography on DEAE cellulose and preparative exclusion chromatography on Sephadex G-50 and G-75 Superfine in the presence of 6 M guanidine·HCl. A second sample purified from fresh hog ovaries by a method described by Sherwood and O'Byrne (4) proved to be identical by all chemical criteria. The purity of the preparations was ascertained by SDS acrylamide gel electrophoresis and analytical exclusion chromatography in 6 M guanidine·HCl. Quantitative amino acid analyses, performed after 24, 48, and 72 hours of hydrolysis on a Durrum D-500 amino acid analyzer, were evaluated by the method of Alt et al. (5) for best fitting factors (F). A Hewlett-Packard 3820-A programmable desk top calculator was used to obtain F values according to the expression

$$F = \Sigma \frac{\Delta_i}{R_i} \Big/ \Sigma \frac{0.5}{R_i}$$

where $\Delta_i = |(R_i - R_{ig})|$, the difference between integer R_i and the next integer R_{ig}. A low F value signifies a high correlation of an assumed molecular weight with a set of integral values.

The A and B chains of relaxin were separated after reduction and alkylation of the intact hormone in 6 M guanidine·HCl with dithiothreitol (DTT) and ^{3}H iodoacetic acid respectively. During these experiments it was discovered that the A and B chain could be separated by utilizing their solubility difference in dilute acidic acid at pH 4.0. After reduction and alkylation in 6 M guanidine·HCl the reaction mixture was diluted with water and acidified to pH 4.0. The insoluble B chain could then be collected by low speed centrifugation whereas all of the A chain remained in the supernatant. Amino acid residues occurring only in either the A or B chain were utilized to ascertain the completeness of separation. Each chain, contaminated to less than 1% by the companion relaxin chain, was then subjected to analytical chromatography on a column of Sephadex G-50 Superfine in order to determine their molecular weight. Amino acid analyses were performed on a Durrum Model D-500 amino acid analyzer. Samples to be analyzed were hydrolyzed in 6 N HCl for 24 hours at 120°. The HCl had been purified by several hours of refluxing over sodium dichromate followed by distillation. The hydrolysis vials were evacuated three times followed by nitrogen flushes and finally sealed under vacuum. Hydrolyses were performed in a thermostated block heater (Pierce Chem. Co.). Phenylthiohydantoin derivatives of amino acids were hydrolyzed in 6 N HCl (containing 0.5 µl mercaptoethanol per ml) at 150° for 24 hours. Sequence analyses of the purified A chain were performed on a Beckman Model 890-C automatic protein sequencer. All solvents (sequence grade) were purchased from Beckman, Palo Alto, California.

Carboxypeptidase C and aminopeptidase M were purchased from Boehringer, Ingelheim, Germany, while dipeptidyl aminopeptidase (DAP I) was purified in our laboratory according to McDonald et al. (6). Enzymatic degradations were carried out in Reactivials (Pierce Chem. Co.) in volumes between 20 and 50 microliters. The enzyme:substrate weight ratios for aminopeptidase M, carboxypeptidase C, and DAP I digests were 0.2, 1, and 0.1 respectively. Digests were analyzed at varying times by spotting aliquots corresponding to 2 nmoles onto TLC plates (microgranular cellulose, Brinkmann Chem. Co.). Detection of amino acids on TLC plates was achieved by a polychromatic copper-ninhydrin spray (7). Aliquots (10 nmoles) were placed onto the amino acid analyzer column for direct identification after the volatile buffers (0.1 M N-ethylmorpholine pH 8.5 for AMP or 0.1 M ammonium acetate pH 4.7 for carboxypeptidase) had been evaporated. Tyrosylated relaxin was prepared by the reaction of tyrosine N-carboxyanhydride with relaxin. Total substitution amounted to 10 residues of tyrosine per molecule of relaxin of which an average of 7 residues was linked through the primary amine of the N-terminal residues and 3 residues were linked through the ε-amine groups of lysine. A tryptic digest of the A chain of relaxin (3 mg) was prepared in 0.5 M ammonium bicarbonate with an initial substrate:enzyme ratio of 100:1, followed by a second addition of trypsin after 5 hours to increase the ratio to 50:1. The digestion was allowed to proceed at room temperature for a total of 18 hours.

Separation of peptides was achieved on a 2 x 100 cm column of Bio-Gel P-2 (200-400 mesh) using 0.05 M ammonium bicarbonate as eluant. Further purification was achieved by one-dimensional electrophoresis on a high voltage electrophoresis apparatus (Gilson Medical Electronics) using pyridine acetate buffer at pH 3.6. Scission of the A chain by cyanogen bromide (CNBr) was performed according to Stark (8) in 0.1 M HCl with a 10-fold excess of CNBr over the A chain by weight. The reaction was allowed to proceed for 24 hours at room temperature whereafter the reaction mixture was lyophilized. The CNBr-treated A chain of relaxin was then placed on a column containing Sephadex G-50 Superfine and eluted with guanidinium chloride in order to separate the two major fragments expected. The major fragment was subsequently desalted on a Sephadex G-15 column equilibrated with 0.05 M ammonium bicarbonate. This CNBr fragment as well as the unmodified chain were subjected to automatic sequence analysis.

RESULTS: The purified relaxin yielded a single band on SDS acrylamide gel electrophoresis corresponding to an average molecular weight of 3,000 for the

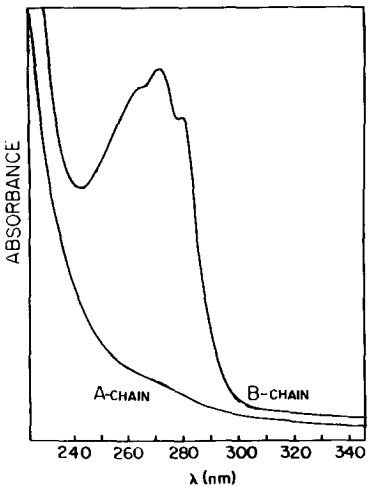

Fig. 1 Ultraviolet spectra of the A and B chains of relaxin. The A chain contains no aromatic residues while the B chain contains two tryptophan residues and one phenylalanine residue. The A chain appears to be free of B chain contamination. Measurements were performed on a Cary 15 spectrophotometer. The concentration of each chain was 0.5 mg/ml.

A and B chain mixture. Native relaxin eluted in a single sharp peak with an elution time similar to insulin from Sephadex G-50 and G-75 Superfine while the reduced and alkylated A and B chains eluted at a position corresponding to 2,500 and 3,200 daltons respectively. The effectiveness of the A and B chain separation by solubility differences is documented by the ultraviolet spectrum of both chains (Fig. 1). In addition, the amino acid analyses showed no B-chain-phenylalanine in the A chain and only a trace of A-chain-methionine in the B chain (4). The relaxin A chain (2.5 mg or 1000 nmoles) was dissolved in 200 µl of 35% acetic acid and transferred to the sequencer cup. The sample was taken through a drying cycle whereafter sequencing commenced (fast peptide program, Beckman #102974, DMAA, benzene, butyl chloride). A double-coupling procedure was used for the first cycle but the remaining 19 cycles were run without modification. The yield of the first residue was only about 50% on a dry-weight basis and declined fairly fast. After identification of the first few residues it became clear that the CNBr fractionation had been only about 70% complete. Double residues were obtained at every step in a 70:30 percent ratio indicating that two identical chains

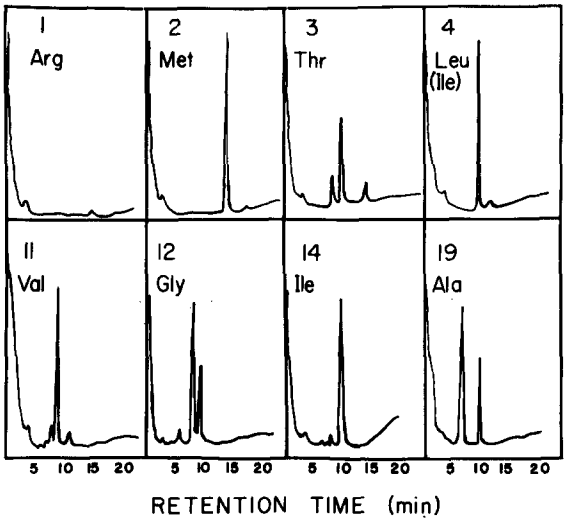

Fig. 2 Identification by gas chromatography of selected PTH residues derived from the A chain of relaxin. The experiments were performed on a Hewlett Packard gas chromatograph (Model 5710 A) fitted with a 4 ft. glass column (2 mm I.D.) and containing Sp-400 (Beckman). The program included the following steps: (1) isothermal, 175°, 4 min; (2) temperature rise 8°/min; (3) isothermal, 300°, 4 min. The residues Leu and Ile were not separated by this method and Thr gave rise to double peaks.

were sequenced two residues out of register. This was confirmed by a separate experiment with the unmodified A chain of relaxin.

The first residue of the unmodified A chain of relaxin identified in the water phase was arginine as demonstrated on a TLC plate developed in n-butanol: 50% formic acid (70:30), and identified with phenanthrenequinone, a fluorogenic reagent specific for arginine (9). Back hydrolysis and amino acid analysis clearly showed arginine to be present in the water phase of residue one. The ethyl acetate phase contained no PTH derivative (Fig. 2). The next residues were methionine, threonine, and leucine followed by serine, glutamic acid, and lysine. The analysis did not give rise to any difficulties until positions 8 and 9, where two cysteines were located. Since cysteine had been converted to ^{3}H CM-cysteine residues the appearance of radioactivity confirmed the positions. In Table I and Figure 2 the sequence of the remaining residues is shown.

The N-terminal arginine was furthermore confirmed by the inability of

TABLE I

YIELD OF PTH OR SILYLATED PTH DERIVATIVES FROM THE AUTOMATIC SEQUENCE ANALYSIS OF THE A CHAIN OF RELAXIN.

per cent yield (based on first residue = 100)

CNBr Fragment

Step	Arg	Met	Thr	Leu	Ser	Glu	Lys	Cys	Cys	Glu	Val	Gly	Cys	Ile	Arg	Lys	Asp	Ile	Ala
1			100	tr															
2				92															
3				(7)	4														
4					90														
5					(5)	61													
6						(5)	10+												
7						(5)		7+											
8						(2)	(8)	(4)	29										
9								(3)	(9)	30									
10									(13)	(8)	41								
11									(5)	(4)	(12)	7+							
12											(5)	(2)	36						
13													(11)	++					
14													(3)		5				
15													(2)			tr	2		
16																	tr	30	
17																		(12)	23

Unaltered A chain

1	++				
2		100			
3		(7)	92		
4			(21)	45	
5				(17)	16

\+ Supported by radioactive data (^{3}H CM-Cys).
++ Determined by TLC with the aid of phenanthrenequinone.
tr Trace.
Values in parenthesis are due to washover or contamination.

The relative yield for steps 1-17 for the CNBr fragment and 1-5 for the unaltered A chain are depicted. Note that step 1 for the CNBr fragment cleaves residue # 3 in the total sequence. The data presented were verified by at least one additional method, i.e., TLC for Thr, Ser, Glu, CM-Cys, Arg, and back-hydrolysis for Arg, Met, Ile, Leu, Ala, Gly, Val, and Asp.

DAP I to attack the A chain directly. The specificity of DAP I is such that arginine at the N-terminal position will prohibit removal of a N-terminal dipeptide whereas arginine in the penultimate position makes for a very favorable substrate. The arginine position was confirmed using tyrosylated relaxin which could be expected to have odd and even numbers of tyrosyl residues at the N

terminus. From this modified molecule DAP I readily produced Tyr-Tyr peptides and the peptide Tyr-Arg. In addition Met-Thr, Leu-Ser, and Glu-Lys were demonstrated after elution from the TLC plate, hydrolysis, and quantitative amino acid analysis. Overlapping peptides Thr-Leu and Ser-Glu were obtained from the action of DAP I on the CNBr fragment. Evidence obtained from aminopeptidase M digests of the relaxin A chain was inconclusive by itself but was consistent with the results of the automatic sequencer.

In contrast, carboxypeptidase C gave very conclusive results concerning the sequence of residues at the C terminus of the A chain. Time-course aliquots of the digests used for amino acid analysis and thin-layer chromatography clearly showed, first, the appearance of carboxymethyl cysteine (a characteristic greenish spot with polychromatic ninhydrin spray) followed by leucine, arginine, alanine, and isoleucine. Traces of aspartic acid and lysine could also be observed. Thus, clear evidence for the sequence: CM-Cys, Leu, Arg, and Ala (beginning at the C terminus) was obtained in addition to suggestive evidence for Ile, Asp, and Lys. This overlaps nicely with the analysis obtained from the sequencer indicating that the residues, progressing from number 15, are Arg, Lys, Asp, Ile, Ala, Arg, Leu, and Cys.

Further confirmation of the structure was obtained from the tryptic digest of the A chain. The separation of the digest on Bio-Gel P-2 (200-400 mesh in 0.5 M ammonium bicarbonate) is shown in Figure 3. The radioactivity indicates that three distinct CM-cysteine-containing fractions exist. Fraction 1 emerges at the exclusion limit of the column and is probably undigested A chain whereas fraction II contains a peptide including residue 8 to residue 15 or 16. The second small peak of radioactivity contains leucyl-CM cysteine. The final fraction showing 230 nm absorbance only contains free lysine and arginine. The peptide distribution on the column, shown in Figure 3, confirms the original sequence analysis. The amount of free lysine and arginine is as small as expected since the trypsin affinity for N-terminal and C-terminal basic residues is not very great. The N-terminal arginine will

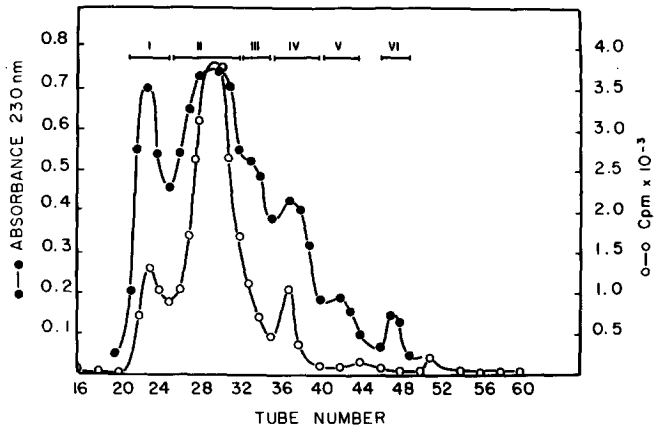

Fig. 3 Chromatography of a tryptic digest of 3 mg of the A chain of relaxin on a Bio-Gel P-2 column (200-400 mesh) in 0.05 M ammonium bicarbonate. The fractions indicated in the graph contained the following peptides:

I Large fragment, not identified.

II Residues Cys_8 to Arg_{15}, and Cys_8 to Lys_{16} (about 10%).

III Residues Arg_1 to Lys_7 and Asp_{17} to Arg_{20}.

IV Leu_{21} to Cys_{22}.

V Not identified.

VI Free Arg and Lys.

therefore not be hydrolyzed extensively and the Arg-Lys bond, although giving rise to some free lysine, is no longer vulnerable once either the Arg-Lys or the Lys-Asp bond has been hydrolyzed. Consequently we should expect a peptide containing Lys-Asp-Ile-Ala-Arg and a peptide Cys-Cys-Glu-Val-Gly-Cys-Ile-Arg-Lys. These assumptions are borne out by our experimental findings.

DISCUSSION: Scarcity of material prohibited several desirable repetitions of these experiments but the data presented are self consistent and support the proposed structure. The positions of various amino acids throughout the disulfide bond region will be once more examined when the structure of the B chain is known and when peptides around the disulfide bonds are isolated to determine the position of the crosslinks. Figure 4 compares the primary structure of the relaxin A chain with the A chain of insulin. It is inter-

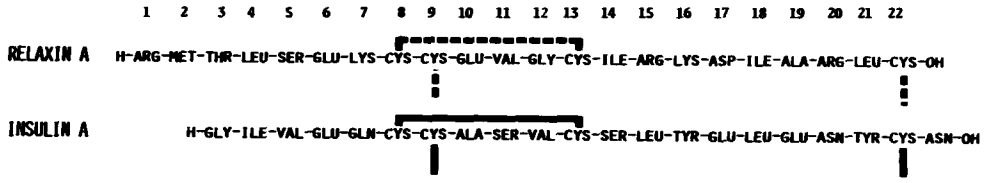

Fig. 4 The sequence of the A chain of relaxin and insulin A are compared. The dotted lines connecting the cysteine residues in the relaxin chain only show potential similarity. The actual crosslinks have not yet been determined.

esting to note that the same relative distribution of the cysteinyl residues occurs in the two chains even though homology is lacking in all other positions. In the absence of information about the B chain sequence and the actual position of its cysteinyl residues, it is tempting to speculate concerning the possibility of a "crosslink analog" to insulin (Fig. 4). We have not been able to crosslink the two chains with superimidate analogous to the successful crosslinking of the insulin A and B chains, but this might be due to the lack of a lysine residue in the relaxin B chain equivalent to Lys 29 of the insulin B chain rather than to a difference in conformation. Some difference in conformation is suggested however by the fact that DAP I was able to attack the N terminus of the A chain of unreduced relaxin, whereas insulin is not susceptible to attack by DAP I until its chains are separated (6).

ACKNOWLEDGMENT

The expert technical assistance of Sandra Halsey and Fred Bishop is gratefully acknowledged.

REFERENCES

1. Hisaw, F. L. and Zarrow, M. X. (1950) Vitamins and Hormones 8: 151-178.
2. Steinetz, B. G., Beach, V. L., Tripp, L. V., and DeFalco, R. J. (1964) Acta Endocrinol. 47: 371-384.
3. Hisaw, F. L. (1926) Proc. Soc. Exptl. Biol. Med. 23: 661-664.
4. Sherwood, O. D. and O'Byrne, E. M. (1974) Arch. Biochem. Biophys. 160: 185-196.
5. Alt, J., Heymann, E. and Krisch, K. (1975) Eur. J. Biochem. 53: 357-369. (also E. Heymann, personal communication).

6. McDonald, J. K., Callahan, P. X. and Ellis, S. (1972) in Methods in Enzymology (C. H. W. Hirs and S. N. Timasheff, eds.) Academic Press, New York, 25, pp. 272-281.
7. Moffat, E. D. and Lytle, R. I. (1959) Anal. Chem 31: 926-928.
8. Stark, G. R. (1970) in Advances in Protein Chemistry (C. B. Anfinsen, Jr., J. T. Edsall, and F. M. Richards, eds.) Academic Press, New York, 24, pp. 261-308.
9. Yamada, S. and Itano, H. A. (1966) Biochim. Biophys. Acta 130: 538-540.