PRIMARY STRUCTURE OF THE B-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 February 14,1977

SUMMARY: The sequence of the B-chain of relaxin, an ovarian peptide hormone isolated from ovaries of pregnant sows, has been shown to have the following primary structure: PCA-Ser-Thr-Asn-Asp-Phe-Ile-Lys-Ala-Cys-Gly-Arg-Glu-Leu-Val-Arg-Leu-Trp-Val-Glu-Ile-Cys-Gly-Val-Trp-Ser (2820 daltons). The heterogeneity of relaxin observed during purification procedures is likely to be due to variations in the C-terminal region of the B-chain, in particular the substitution of Gln for Glu_{20}, and the possible addition of arginine or serylarginine at the C terminus. The B-chain exhibited a distribution of sulfhydryl residues relative to one another that is identical to that found in the B-chain of insulin. A similar analogy has already been demonstrated for the A-chains of relaxin and insulin.

The primary structure of the relaxin A-chain has been reported recently (1) and was found to resemble insulin A with respect to size and cysteine residue distribution. Proof for a PCA* residue at the N terminus of the B-chain of relaxin has also been reported (2). Evidence presented here indicates that the B-chain of relaxin similarly exhibits the same cysteine residue distribution as the analogous chain in insulin; again lacking homology in other positions. It appears that relaxin and insulin can be considered "disulfide homologs".

MATERIALS AND METHODS: Porcine relaxin A- and B-chains were purified and tested for purity by methods previously described (1). Trypsin (three times crystallized) was purchased from Worthington Biochemical Corporation and treated with TPCK as described by Carpenter (3) in order to remove chymotryptic activity. Stock solutions of TPCK-treated trypsin (1 mg/ml) were prepared in 1 mM HCl-10 mM $CaCl_2$. To perform a tryptic digest the carboxymethylated B-chain (5 mg) was dissolved in 200 µl 0.2 M N-ethylmorpholine·HCl (pH 8.5) to which 50 µl of the trypsin solution was added. A thin-layer plate (microcrystalline cellulose, Brinkmann) was spotted with 1 µl of the digest at zero time, 30 min and 60 min to estimate progress of the reaction. The reaction

*Abbreviations: PCA = pyroglutamic acid, TPCK = L-(1-tosylamido-2-phenyl) ethyl chloromethyl ketone; PTH = phenylthiohydantoin; BNPS-skatole = [2-(2-nitrophenylsulfenyl)-3-methyl-3'-bromoindolenine]; HPLC = high-pressure liquid chromatography; 4-SPITC = 4-sulfo phenyl isothiocyanate.

ISSN 0006-291X

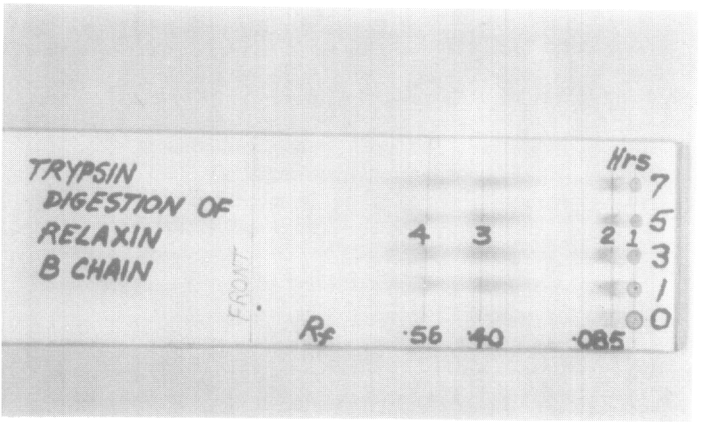

Fig. 1 Thin-layer chromatogram showing separation of four tryptic peptides derived from the B-chain of relaxin. In this experiment 5 mg of relaxin B was dissolved in 500 µl of a 1:1 mixture of N-ethylmorpholine 0.2 M (pH 8.5) and 0.005 M NH_4HCO_3. Trypsin (50 µg) was added in a 50 µl volume of 10^{-3} M HCl-10^{-2}M in $CaCl_2$. The digestion mixture (1 µl) was spotted at time intervals given on the right hand side of the figure. The plate was developed in n-butanol, 50% formic acid (7:3). Peptides designated 1, 2, 3, and 4 are: (1) C-terminal peptide (same as HPLC #3 in Table 1), (2) Ala-Cys-Gly-Arg, (3) N-terminal peptide (HPLC #2 in Table 1), (4) Glu-Leu-Val-Arg.

was allowed to proceed for 3 hr at 40° and was then fractionated by high-pressure liquid chromatography using a discontinuous gradient of acetonitrile (10%, 20%, 40%, 50%) in 0.05 M NH_4HCO_3. Peptides were collected in a fraction collector and lyophilized. Amino acid analyses were performed as reported earlier (1). The method of removing the pyroglutamyl residue at the N terminus of the relaxin B-chain has already been described in detail (2). Enzymatic treatments of the various tryptic peptides with carboxypeptidase C (Henley & Co., New York), bovine carboxypeptidase A (DFP-treated, Sigma Chemical Co.), lysosomal carboxypeptidase (prepared according to McDonald and Ellis (4)), and aminopeptidase M (Henley & Co.) were performed as previously described (1).

PTH-amino acids were identified by either gas chromatography or high-pressure liquid chromatography. Cleavage of the C-terminal tryptic peptide at the tryptophan residues was performed with BNPS-skatole (5) in 70% acetic acid (10-fold molar excess over tryptophan) for 14 hr at 30°. Succinylated B-chain was prepared by addition of succinic anhydride (50-fold molar excess) to the B-chain of relaxin immediately following reduction and alkylation. The reaction mixture was continuously stirred and the pH maintained at 8.0 by addition of 1 M NaOH.

RESULTS AND DISCUSSION: The purified trypsin-digested B-chain of relaxin yielded four distinct fractions separable by thin-layer chromatography (Fig. 1). In order to obtain a sufficient quantity of the tryptic peptides for automatic sequence analysis, a large preparative run on several 20 x 20 cm TLC plates

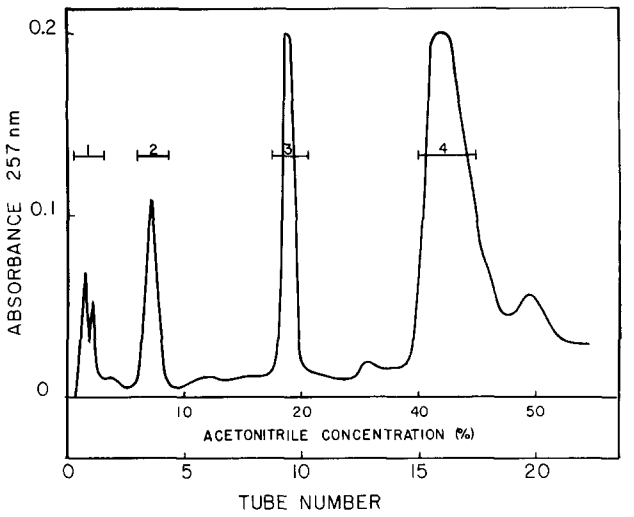

Fig. 2 High-pressure liquid chromatography of a tryptic digest of relaxin B on a Bondpack C_{18} column using acetonitrile in NH_4HCO_3 mixtures as indicated. With longer periods of digestion peak #4 diminished and the earlier peaks increased in size. The HPLC peak numbers are shown in the figure.

was performed. The bulk of the digest was fractionated on a high-pressure liquid chromatograph (Waters Associates, Milford, Mass.) (Fig. 2). The amino acid analyses of the peptides are depicted in Table 1. Sequence analysis of the N-terminal (PCA) peptide (HPLC #2, TLC #3) was accomplished with the aid of carboxypeptidase C and the automatic sequencer (Beckman 890-C). In order to increase hydrophilicity the PCA peptide was reacted with 4-SPITC prior to removal of the PCA group with PCA-peptidase. Table 2 shows the result of the sequence analysis of the des-PCA peptide (as well as the two arginine peptides and the C-terminal fragment). It is apparent that the 4-SPITC allowed excellent recovery at each step of degradation. The enzymatic digest with carboxypeptidase C confirmed the terminal residues -Asp-Phe-Ile-Lys.

The amino acid analyses clearly showed two peptides of the composition Ala-Cys(CM)-Gly-Arg (TLC #2, HPLC #1) and Glu-Leu-Val-Arg (TLC #4, HPLC #1) with excellent integral values. Both peptides remained in the sequencer cup (DMAA program #102974, Beckman) to the last residue arginine. The arginine

Table 1

THE TRYPTIC PEPTIDES OF THE B-CHAIN OF RELAXIN

		Isolation Method				
Amino Acid	B-Chain	TLC3	TLC2	HPLC2	HPLC3	HPLC4*
	Residues					
CMCys	2		1		1	∿1
Asx	2			2		
Thr	1			1		
Ser	2			1	1	2
Glx	3	1		1	1	2.4
Gly	2		1		1	2.5
Ala	1		1			0.3
Val	3	1			2	2.2
Ile	2			1	1	1
Leu	2	1			1	1.3
Phe	1			1		
Lys	1			1		
Arg	2	1	1			0.7
Trp	2				2	
Total	26	4	4	8	10	

*Mixture of TLC2 and TLC3 still attached to HPLC3. Hydrolysis was incomplete due to solubility problems and the position on the HPLC column can only be explained by strong hydrophobic peptide-peptide interaction.

Variation from integral values for the remaining peptides was in the order of ± 0.1 residue.

was confirmed by examination of the waterphase using TLC and fluorogenic phenanthrenequinone reagent (6) for detection.

The C-terminal peptide containing the two tryptophan residues proved most difficult to work with. Extreme hydrophobicity and low solubility in

Table 2

YIELD OF PTH AMINO ACIDS FROM THE AUTOMATIC SEQUENCE ANALYSIS OF THE TRYPTIC PEPTIDES OF THE B-CHAIN OF PORCINE RELAXIN

per cent yield based on total amount of peptide used

Des-PCA Peptide

Step	Ser-Thr-Asn-Asp-Phe-Ile-Lys
1	30
2	60
3	50
4	45
5	50
6	50 *

TLC Peptide #2

Step	Ala-Cys-Gly-Arg
1	80
2	10*
3	30 *

TLC Peptide #4

Step	Glu-Leu-Val-Arg
1	25
2	70
3	70 *

C-Terminal Peptide (HPLC #3)

Step	Leu-Trp-Val-Glu-Ile-Cys-Gly-Val-Trp-Ser
1	70
2	50
3	40
4	40
5	30
6	*
7	5
8	10 * *

Table 2 Approximal yields of PTH amino acids were obtained by gas chromatography. Positions designated by an asterisk were determined by other methods in addition or exclusively. In case of liquid chromatography interpretation was qualitative and based only on the largest peak observed. Other detection methods were radioactivity for ^{3}H CM cysteine and phenanthrenequinone for arginine. Trp_{25} and Ser_{26} were determined by enzymatic degradation.

solution at pH 7.0 or higher (insoluble below pH 7.0) led to high losses through aggregation, adhesion to glass, and washes from the sequencer cup. Consequently only the first six residues (#17 through #22) off the sequencer were clearly observed by gas chromatography, residue #23 (Gly) and #24 (Val) only marginally,

and #25 (Trp) and #26 (Ser) could not be detected. High-pressure liquid chromatography of the PTH-amino acids from the sequencer however allowed identification of residue #23 (Gly), #24 (Val) and #25 (Trp). This left only one position for serine, i.e., the C terminus. The C-terminal residue in the B-chain was confirmed with carboxypeptidase A and lysosomal carboxypeptidase digestion of whole relaxin to be serine, i.e., -Gly-Val-Trp-Ser. Whole relaxin was used because solubility problems led to very slow reaction rates when instead digestion of the B-chain or the C-terminal peptide was attempted. Carboxypeptidase A did not attack the A-chain while the lysosomal carboxypeptidase produced some Leu (A_{21}), Arg (A_{20}), Ala (A_{19}), and Ile (A_{18}), all of which did not interfere with the detection of Ser (B_{26}), Trp (B_{25}), Val (B_{24}), and Gly (B_{23}). Final confirmation of Ser (B_{26}) was obtained through BNPS-skatole treatment. The reaction mixture of 3 mg relaxin (0.6 µmoles) and 6 µmoles of BNBS-skatole was extracted with ethylacetate to remove excess reagent and then lyophilized. Chromatography on a 1.5 x 60 cm column packed with Sephadex G-15 superfine in 0.05 M NH_4HCO_3 yielded one large and one small particle fraction. The low molecular weight fraction was lyophilized and redissolved in 500 µl of water. An aliquot of 20 µl was placed directly on the amino acid analyzer (Durrum D-500) and a second aliquot of 20 µl was hydrolyzed in 6 N HCl at 115° for 6 hr prior to analysis. Serine values did not change after hydrolysis indicating that Ser (B_{26}) was the C-terminal residue.

Trypsin digestion of the succinylated B-chain provided fragments that allowed proper placement of the two arginine-containing tetrapeptides in the center of the chain. Both TLC and HPLC fractionation of this tryptic digest of the succinyl-relaxin B-chain (Lys_8 modified) clearly indicated that Ala-Cys-Gly-Arg was positioned next to the N-terminal PCA peptide. Ninhydrin treatment revealed only the origin (C-terminal peptide) and spot #4 (Glu-Leu-Val-Arg) while spot #3 in Fig. 3 appeared only with either peptide spray (7) or phenanthrenequinone spray for arginine. Peptide #2 (Ala-Cys-Gly-Arg) could no longer be detected. Peptide #3 also possessed as much radioactivity as did the origin

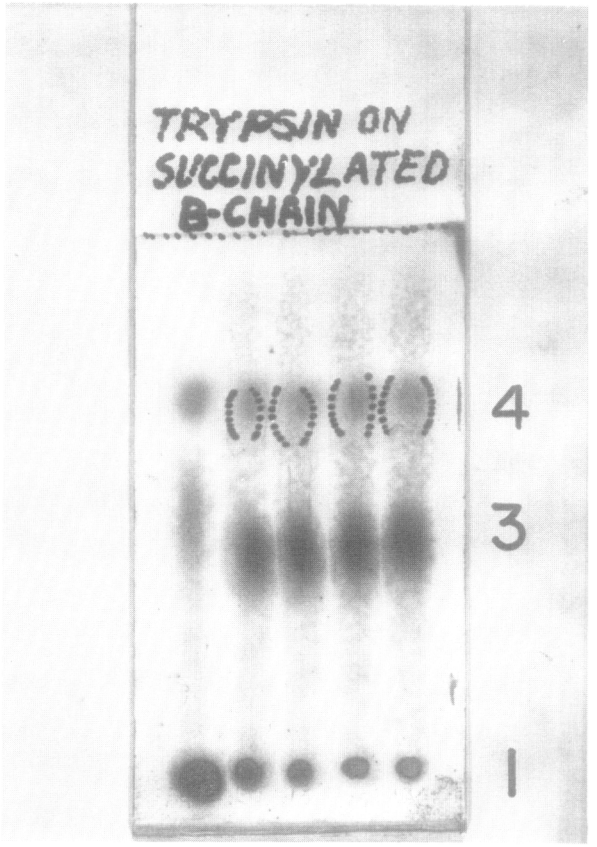

Fig. 3 Thin-layer chromatogram showing separation of three tryptic peptides derived from the B-chain of relaxin following succinylation of the Lys_8 residue. Peptide # 2 (see Fig. 1) has disappeared and peptide #3, no longer ninhydrin-positive, contains radioactivity and becomes fluorescent when treated with the arginine-detecting spray. This observation suggests that peptide #3 in this experiment is a fragment containing the PCA peptide linked to Ala-Cys-Gly-Arg. The plate depicted has been treated with a peptide spray reagent which does not require a free amino group. The spots from left to right are zero digestion, time, 5, 15, 30, and 60 min respectively.

peptide. Finally the amino acid analysis of a new HPLC peptide (10% acetonitrile) yielded Glu-Ser-Thr-Asp_2-Phe-Ile-Ala-Cys-Gly-Arg in precise integral values leaving no doubt that the sequence of the B-chain of porcine relaxin is as given in Fig. 4.

The B-chain of porcine relaxin is shown together with the B-chain of porcine insulin in order to illustrate that the distribution of the cysteine residues relative to each other is identical as previously observed for the

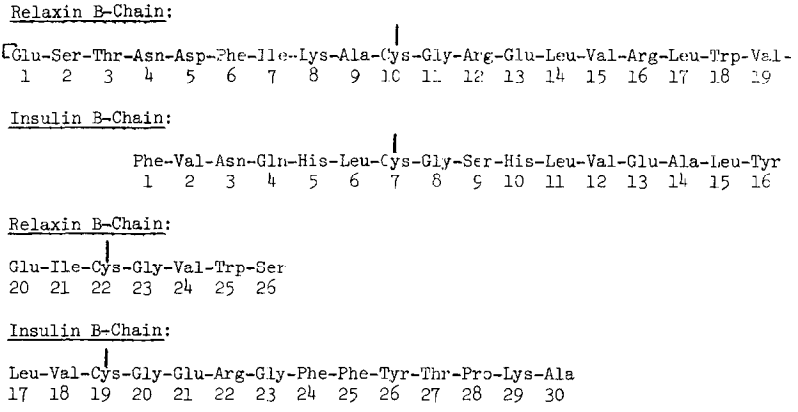

Fig. 4 Comparison of the relative distribution of cysteinyl residues in the B-chains of relaxin and insulin. It is notable that the amino acids following cysteine are glycine in both cases and those preceding cysteine are aliphatic.

A-chains of both hormones. The crosslink pattern still requires further investigation while preliminary evidence suggests that Cys (A_{22}) and Cys (B_{22}) are linked (see lysosomal carboxypeptidase experiment), leaving only a limited number of possibilities for the second crosslink.

ACKNOWLEDGMENT

We thank Dr. R. Crouch for the use of her liquid chromatograph and Dr. Robert E. Fellows for providing the pyrrolidonecarboxylyl peptidase.
The expert technical assistance of Mr. Frederick E. Bishop and Mr. Joseph W. Northington is gratefully acknowledged.
This research was supported in part by NICHD Grant HD-10540 and by the South Carolina state funds for research.

REFERENCES

1. Schwabe, C., McDonald, J. K. and Steinetz, B. G. (1976) Biochem. Biophys. Res. Comm. 70: 397-405.
2. Schwabe, C. and McDonald, J. K. (1977) Biochem. Biophys. Res. Comm. (in press).
3. Carpenter, F. H. (1967) in Methods in Enzymology (C. H. W. Hirs, ed.) Academic Press, New York, 11, 237.
4. McDonald, J. K. and Ellis, S. (1975) Life Sci. 17: 1267-1276.
5. Fontana, A. (1972) in Methods in Enzymology (C. H. W. Hirs and S. N. Timasheff, eds.) Academic Press, New York, 25, 419-432.
6. Yamada, S. and Itano, H. A. (1966) Biochim. Biophys. Acta 130: 538-540.
7. Nitecki, D. E. and Goodman, J. W. (1966) Biochemistry 5: 665-673.