Journal of Chromatography, 508 (1990) 81–96
Elsevier Science Publishers B.V., Amsterdam — Printed in The Netherlands

CHROM. 22 322

Characterization of chemically synthesized human relaxin by high-performance liquid chromatography

ELEANOR CANOVA-DAVIS*, IDA P. BALDONADO and GLEN M. TESHIMA
Analytical Chemistry Department, Genentech, Inc., 460 Point San Bruno Boulevard, South San Francisco, CA 94080 (U.S.A.)
(First received June 13th, 1989; revised manuscript received January 2nd, 1990)

SUMMARY

Highly purified human relaxin, produced by combining chemically synthesized A- and B-chains, was analyzed by reversed-phase high-performance liquid chromatography, ion-exchange chromatography and tryptic mapping in order to ascertain purity of the material, presence of uncleaved protecting groups, correctness of disulfide linkages and presence of deamidated or oxidized variants. It was shown by a variety of analytical methods that the product was of high purity; a minimum purity level as judged by the most discriminating assay was greater than 98%. Components of the relaxin preparation removed during the purification were identified to be variants containing deletions arising from incomplete coupling reactions in the solid phase peptide synthesis and/or oxidized methionine residues.

INTRODUCTION

Recent elucidation of the primary structure of relaxin has revealed its homology to insulin essentially by its strikingly similar disulfide bond structure. The relaxin and insulin molecules are each composed of two non-identical peptide chains linked by two disulfide bridges with an additional intrachain disulfide bridge in the smaller A-chain. The amino acid sequences of relaxin are known for a number of species including porcine[1], rat[2], sand tiger shark[3], spiny dogfish shark[4], human[5,6], skate[7], minke whale[8] and Bryde's whale[8]. From protein sequencing data on the purified ovarian hormones and nucleotide sequence analysis data of cDNA clones[9,10] it appears that the relaxins are expressed as single chain peptide precursors with the overall structure: signal peptide–B-chain–C-peptide–A-chain. Since the sites of *in vivo* processing of human preprorelaxin are not identified, they were postulated by analogy to the processing of porcine and rat preprorelaxins[6]. Hence, the human A-chain was chemically synthesized by solid phase methods as a 24-amino acid polypeptide and the B-chain as 33 amino acids in length.

The chemical synthesis of small polypeptide hormones has been extensively described and shown to be invaluable for confirmation of proposed sequences and for

0021-9673/90/$03.50 © 1990 Elsevier Science Publishers B.V.

structure–function studies[11]. In particular, peptide synthesis has been shown to be a very useful technique for studies of the chemical and biological properties of human relaxin since it has not been possible to obtain the natural hormone from human tissue[12]. However, the use of chemical synthesis as opposed to recombinant-DNA techniques for the manufacture of small polypeptides (50 residues or less) for pharmaceutical application is an area of lively debate[13,14]. Examples to date include calcitonin (32 amino acids), pentagastrin and tetracosa adrenocorticotropin peptide[13]. In addition to considerations such as speed, cost, complexity of the technology and yield, the requirement of high purity for human testing is an important aspect, as a product of chemical synthesis has a variety of potential side reactions that are not generated by a biosynthetic process; *e.g.*, racemization, deletions and residual blocking groups.

With the advent of biotechnology and the capability to produce recombinant proteins in good yield with a high degree of purity, has come the requirement for sensitive and varied analytical methods for product characterization[15]. Recommendations issued by the Federal Drug Administration for protein pharmaceuticals have included a specification for the characterization of any variant present at the 1% level. Therefore, this study was an attempt to characterize a chemically synthesized protein in some detail, *i.e.*, using various analytical methods in addition to the routine amino acid and liquid chromatographic analyses generally employed[12] to investigate the relative merits of the two technologies. This is the first instance in which a chemically synthesized protein of the size of relaxin (57 amino acids) has been characterized to conform with the Federal Drug Administration's requirements. Emphasis was placed upon reversed-phase high-performance liquid chromatography (HPLC) analysis as it is a widely accepted high-resolution procedure[16–19]. In addition, the recent technique of fast atom bombardment-mass spectrometry (FAB-MS)[20–21] was utilized to more thoroughly examine the protein for any chemical modifications.

In recombinant DNA-derived protein pharmaceuticals, the potential variants have typically been deamidated, oxidized, proteolytically degraded and aggregated species. In addition to the possibility of these forms existing in a chemically synthesized protein, amino acid deletions, truncations, racemization, incomplete removal of blocking groups and re-acetylation or alkylation can also occur. These side reactions inherent in a chemical synthesis become more significant as the length of the synthesized peptide increases. Since the chemical synthesis method chosen involved the use of *tert.*-butyloxycarbonyl amino acids, the probability of racemization occurring to any extent is minimized[22]. The following investigation is an in-depth study of the chemical characteristics of such a prepared human relaxin molecule.

EXPERIMENTAL

Chemical synthesis of human relaxin

The A- and B-chains and a B-chain analogue of human relaxin were synthesized in the Peptide Group at Genentech by the solid-phase synthesis method of Merrifield[23] using *tert.*-butyloxycarbonyl amino acid derivatives and a phenylacetamidomethyl resin. The side chain protecting groups were as follows: cysteine, 4-methylbenzyl; aspartic and glutamic acid, cyclohexyl; lysine, chlorocarbobenzoxy; asparagine and glutamine, xanthyl; arginine, tosyl; serine and threonine, benzyl and tyrosine,

2,6-dichlorobenzyl. The crude chains were purified, combined to form relaxin and the resultant protein was isolated as previously described[24]. Briefly, the individual chains were purified by reversed-phase HPLC using a Vydac C_4 column with a 0.1% trifluoroacetic acid (TFA)–acetonitrile linear gradient. The isolated chains were subsequently combined in a reaction mixture containing 100 m*M* glycine (pH 10.5), 1 m*M* EDTA, 2.5 m*M* dithiothreitol (DTT), 3% *n*-propanol, 3% acetonitrile and 1 *M* urea. The solution was stirred open to the air for 28 h at ambient temperature. This material was then purified by reversed-phase HPLC using a SynChropak RP-C_4 column with a 0.05% TFA–acetonitrile linear gradient and finally by ion-exchange chromatography in the Process Development Group at Genentech. The purified material was shown to be biologically active in the mouse pubic symphysis ligament *in vivo* bioassay[25] and the rat uterine contractility *in vitro* bioassay[26].

Amino acid and amino terminal sequencing analyses

The acid hydrolysates of 6–8 μg of relaxin preparations were analyzed using a Beckman Model 6300 analyzer with the standard three-buffer elution system and post-column ninhydrin detection. To analyze for cysteine, samples were treated with performic acid according to Hirs[27] prior to acid hydrolysis. Thioglycolic acid (7%) was added to samples before acid hydrolysis in order to quantitate tryptophan.

The A-chain amino terminal of the relaxin molecule was deblocked by treatment with pyroglutamate aminopeptidase (Boehringer Mannheim) as described by Podell and Abraham[28]. The treated protein was chromatographed by reversed-phase HPLC on a Waters gradient liquid chromatography system (connected to a Nelson Analytical 6000 data system) that consisted of two 510 pumps, a 720 controller, a WISP injector and 440 absorbance detector with an extended wavelength module for dual-wavelength detection (214 and 280 nm) fitted with a Vydac C_4 column (250 mm × 4.6 mm I.D., 5 μm, 300 Å). An acetonitrile elution was begun with 15% acetonitrile containing 0.1% TFA for 10 min followed by a 40 min linear gradient to 35% acetonitrile. The recovered peptide-containing fractions were pooled and subjected to Edman degradation using an Applied Biosystems 447A protein sequencer, with online phenylthiohydantoin detection (Applied Biosystems 120A analyzer).

Extinction coefficient

The ultraviolet spectrum of relaxin in 10 m*M* sodium citrate, pH 5.0, and isotonic saline was obtained. The extinction coefficient was determined according to Beaven and Holiday[29] as modified by Wetlaufer[30] for a 0.1% solution of the protein and found to be 2.04.

Size-exclusion chromatography

Gel filtration of relaxin (50 μl) was conducted using a TSK G2000 SWXL column (300 mm × 7.5 mm I.D.) equilibrated with 10 m*M* sodium citrate, pH 5.0, in 0.25 *M* sodium chloride. The flow-rate was 0.5 ml/min and the protein was monitored at 214 nm.

Hydrophobic interaction chromatography (HIC)

These separations were done using a Hewlett-Packard 1090M liquid chromatograph. This system consisted of an autoinjector with a 25-μl sample loop and a

photodiode array detector. Ultraviolet absorbance was monitored at 214 nm. The TSK-phenyl-5 PW column was purchased from HP Genenchem. This column was 75 mm × 7.5 mm I.D. and was packed with a resin of 10-μm particle diameter with an average pore diameter of 1000 Å. Buffer A consisted of 0.1 *M* potassium phosphate, pH 7.0, containing 2 *M* ammonium sulfate. Buffer B was 0.1 *M* potassium phosphate, pH 7.0. The column was equilibrated with 50% buffer A and 50% buffer B and a decreasing linear salt gradient was generated to 100% B in 50 min. The flow-rate was maintained at 1 ml/min and the temperature was ambient. Relaxin concentrations were adjusted to 1 mg/ml with buffer B; 50-μl aliquots were injected.

Tryptic mapping

Human relaxin at a concentration of 0.5 mg/ml was digested with L-1-tosylamido-2-phenylethyl chloromethyl ketone (TPCK)-treated trypsin (Cooper Biomedical) at 37°C for 4 h in 10 m*M* Trisacetate, pH 7.3, 0.2 m*M* calcium chloride. Two aliquots (10 μl each) of enzyme–substrate (1:100, w/w) of trypsin were added: the first at zero time, and the second after 2 h of digestion. Samples were frozen at −20° if not analyzed immediately. The peptide peaks collected manually from the Waters HPLC system were characterized by amino acid analysis after acid hydrolysis and by FAB-MS. For FAB-MS the samples (0.5–1.0 nmol) were lyophilized to remove volatile salts. They were then solubilized in a minimum volume of 0.1% acetic acid and analyzed by direct addition to glycerol on the sample probe of a two-sector JEOL HXIIOHF mass spectrometer with a DA 5000 data system. Fast atom bombardment was performed with 6 keV xenon atoms. Data were acquired over a mass range of 100–4300 daltons.

RESULTS AND DISCUSSION

Amino acid analyses

Typical analyses of relaxin and its composite A- and B-chains are shown in Table I. Residues per mol were normalized to the sum of alanine and leucine. These results demonstrate excellent correlation with the theoretical values[6]. As can be seen in this table, the cysteic acid determination for human relaxin is 87% of the expected theoretical value. Humulin (recombinant human insulin) was analyzed at the same time, as a control for these particular cystine linkages. The cysteic acid value obtained was 80% of the expected value. Amino acid analysis can be used for the partial assessment of the primary structure of a polypeptide. With small peptides, such as relaxin, it was a useful technique for the determination of purity[31]. At the very least, this method served as a confirmation of the presence of the correct amino acids in the proper ratios and suggested that the purified product contained minimal levels of deletion peptides.

Amino terminal sequencing

Synthetic human relaxin was analyzed by Edman degradation to determine the amino acid sequence at the N-termini of the two chains present in the molecule. Since the A-chain contains an N-terminal pyroglutamic acid residue, sequence data could be obtained only after treatment of the molecule with pyroglutamate aminopeptidase. The expected sequences for the two chains were found (Table II). No other N-termi-

TABLE I

AMINO ACID COMPOSITION OF HUMAN RELAXIN AND ITS COMPOSITE A- AND B-CHAINS

Amino Acid	*Composition (residues/mol)*[b]		
	Relaxin	*A-chain*	*B-chain*
CyA[a]	5.18 (6)	3.82 (4)	1.92 (2)
Asx	1.99 (2)	0.98 (1)	0.93 (1)
Thr[c]	1.97 (2)	1.00 (1)	0.98 (1)
Ser[c]	5.83 (6)	1.97 (2)	3.90 (4)
Glx	5.02 (5)	0.98 (1)	3.97 (4)
Pro[a]	0 (0)	0 (0)	0 (0)
Gly	3.00 (3)	1.02 (1)	2.01 (2)
Ala	4.96 (5)	2.95 (3)	2.02 (2)
Val[d]	2.93 (3)	0.98 (1)	1.94 (2)
Met	1.86 (2)	0 (0)	1.82 (2)
Ile[d]	2.74 (3)	0 (0)	2.63 (3)
Leu	6.04 (6)	3.05 (3)	2.98 (3)
Tyr	0.98 (1)	0.97 (1)	0 (0)
Phe	1.05 (1)	1.04 (1)	0 (0)
His	0.99 (1)	1.00 (1)	0 (0)
Lys	3.97 (4)	2.00 (2)	1.93 (2)
Trp[e]	2.01 (2)	– (0)	2.03 (2)
Arg	4.98 (5)	2.01 (2)	2.90 (3)

[a] Performic acid oxidized sample.
[b] Theoretical values are in parentheses.
[c] Extrapolation to zero time of hydrolysis.
[d] After a 72-h hydrolysis.
[e] Determined in the presence of thioglycolic acid.

nal sequences were observed, again indicating the absence of any significant deletion peptides arising from the chemical synthesis as present in the purified material. It was possible to detect a secondary sequence in a B-chain preparation estimated to contain 5% des Asp^1-B-chain by reversed-phase HPLC[32].

Size-exclusion chromatography

When the chemically synthesized relaxin preparation was chromatographed at pH 5.0 at a concentration of 0.4 mg/ml or less it behaved as a monomer. The molecular weight of the monomer was determined by FAB-MS to be 6448.9 as compared to a calculated value of 6449.6. However, when the chromatography was conducted at a concentration of 2 mg/ml, it began eluting from the column as a dimer. This is consistent with the data[33] obtained from sedimentation equilibrium experiments which demonstrated that human relaxin self associates in a reversible manner which is concentration and pH dependent. Analysis with a monomer–dimer association model[34] results in an association constant of 38.2 l/g at pH 5.0.

Reversed-phase HPLC

Relaxin samples were analyzed by reversed-phase HPLC using a shallow gradient of acetonitrile, which resolved relaxin and its component A- and B-chains from

TABLE II

AMINO-TERMINAL SEQUENCE ANALYSIS OF HUMAN RELAXIN AFTER TREATMENT WITH PYROGLUTAMATE AMINOPEPTIDASE

Cycle number	*A-chain*		*B-chain*	
	Sequence	*Yield (pmol)*	*Sequence*	*Yield (pmol)*
1	Leu	120	Asp	64
2	Tyr	67	Ser[a]	–
3	Ser[a]	–	Trp	49
4	Ala	110	Met	76
5	Leu	110	Glu	57
6	Ala	110	Glu	84
7	Asn	47	Val	75
8	Lys	61	Ile	79
9	Cys[b]	–	Lys	62
10	Cys[b]	–	Leu	86
11	His	2.4	Cys[b]	–
12	Val	32	Gly	56
13	Gly	28	Arg	15
14	Cys[b]	–	Glu	36
15	Thr	4.4	Leu	68
16	Lys	23	Val	50
17	Arg	11	Arg	11
18	Ser[a]	–	Ala	66
19	Leu	36	Gln	33
20	Ala	42	Ile	49
21	Arg[c]	–	Ala	89
22	Phe	30	Ile	49
23	Cys[c]	–	Cys[c]	–
24			Gly	32
25			Met	28
26			Ser[c]	–
27			Thr	4.0
28			Trp[c]	–
29			Ser[c]	–
30			Lys[c]	–
31			Arg	3.7
32			Ser[c]	–
33			Leu[c]	–

[a] Identified as the dehydroalanine adduct of dithiothreitol with a concomitant peak in the Ser position.

[b] Identified as the dehydroalanine adduct of dithiothreitol without a concomitant peak in the Ser position.

[c] Not confirmed.

each other. Fig. 1 shows a typical profile of a relaxin preparation obtained before and after reduction with DTT. Relaxin preparations that had been purified by reversed-phase HPLC and cation-exchange chromatography contained greater than 98% of the total integrated area in the main peak when analyzed by reversed-phase HPLC. Recovery from this column was quantitative as judged from the linear response of the total integrated area of relaxin samples from 0–20 μg. The later elution of B-chain as

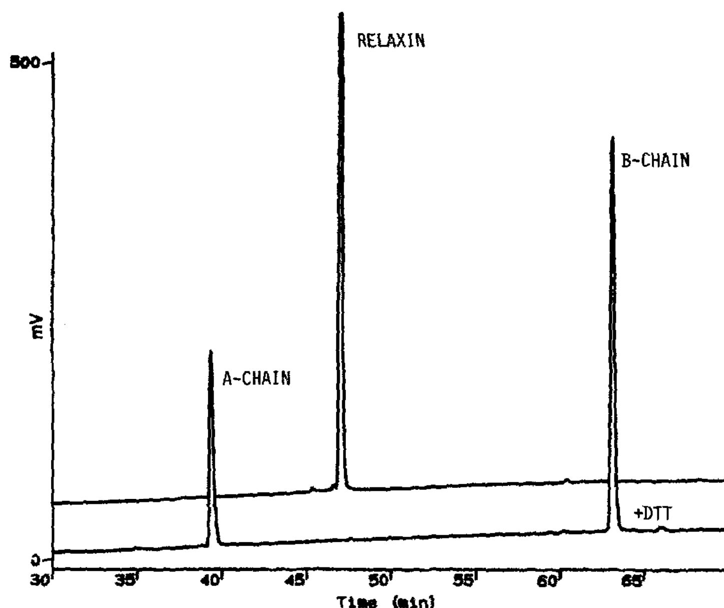

Fig. 1. Reversed-phase HPLC chromatograms of human relaxin before and after reduction with DTT. The chromatography was performed on the Waters HPLC system using a Vydac C_4 column equilibrated with 18% acetonitrile containing 0.1% TFA. Elution was conducted isocratically for 20 min followed by a linear gradient to 50% acetonitrile over 40 min and ending with a 10-min hold at 50% acetonitrile. The flow-rate was 1 ml/min, the protein was monitored at 214 nm, the analog output to the integrator was scaled at 0.5 volts per absorbance unit and the column temperature was maintained at 35°C. An amount of 10 μg of intact relaxin in 5% acetic acid was injected in 50 μl. In addition, relaxin samples were routinely reduced to their component A- and B-chains with 50 m*M* DTT, 4 *M* guanidine in 50 m*M* Tris · HCl, pH 8.5, at ambient temperature. A sample (50 μl) was injected directly onto the column.

compared to the intact hormone suggested greater hydrophobicity for the B-chain. Calculations of the hydrophilicity of the A-chain, the B-chain and relaxin itself yielded values of 1.18, 0.74 and 0.95, respectively[35] consistent with this observation.

Hydrophobic interaction chromatography

Since Gooding *et al.*[36] claimed that substantial changes in retention times and selectivity were seen under HIC conditions and little effect seen with reversed-phase HPLC in analyzing ovalbumin and bovine serum albumin, the HIC of relaxin was investigated. The profiles obtained for relaxin preparations showed less heterogeneity than those that had been observed after reversed-phase chromatography (Fig. 1). Hence, in this case, reversed-phase HPLC was a more discriminating technique than HIC. This is consistent with the recent work of Hoeger *et al.*[37] showing that reversed-phase HPLC is quite efficient in resolving closely related impurities from synthetic polypeptides in the 1500–4200-dalton molecular mass range.

Cation-exchange chromatography

Cation-exchange chromatography profiles of relaxin were generated in 0.05 *M* potassium phosphate at pH 6.0, 6.5 and 7.0. At lower pH values, the protein bound

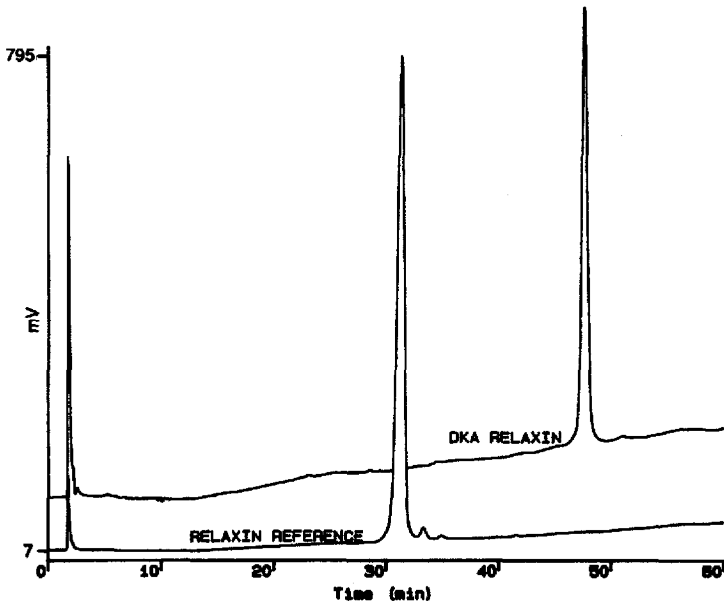

Fig. 2. Comparison of the human DKA relaxin analogue (Lys^{B-4} Ala^{B-25} relaxin) and relaxin by cation-exchange chromatography. The chromatography was performed on a polyaspartic acid (Poly CAT A) column (200 mm × 4.6 mm I.D., 5 μm, 300 Å) purchased from the Nest Group. The buffer used was 0.05 *M* potassium phosphate at pH 7.0 in 25% acetonitrile. A linear gradient was generated with this buffer containing 0.5 *M* sodium chloride at a rate of 5m*M* sodium chloride/min. The flow-rate was maintained at 1 ml/min, the protein was monitored at 220 nm (0.5 volts per absorbance unit), and the temperature was kept at ambient. Relaxin concentrations were adjusted to 0.2 mg/ml with equilibration buffer; 50-μl aliquots were injected.

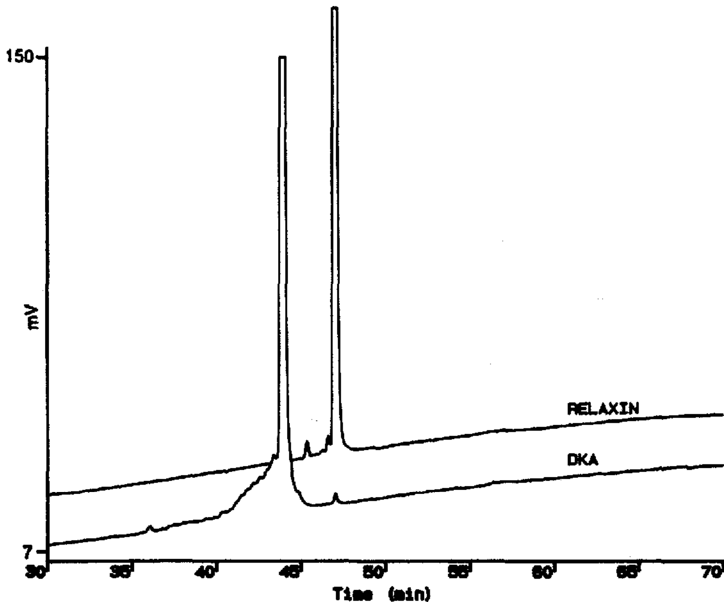

Fig. 3. Comparison of the human DKA relaxin analogue (Lys^{B-4} Ala^{B-25} relaxin) and relaxin by reversed-phase chromatography. The chromatography was performed as described in legend to Fig. 1.

more strongly to the column as expected by virtue of the greater protonation of relaxin at more acidic pH values (data not shown). In all cases, the main peak contained more than 99% of the material by peak area calculations.

An analogue of relaxin (DKA relaxin) was chemically synthesized whereby the methionine residues in the B-chain were replaced by lysine and alanine at positions B-4 and B-25, respectively. As can be seen in Fig. 2 the addition of a basic residue such as lysine results in the elution of the analogue at later retention times. When this analogue is analyzed by reversed-phase HPLC it elutes at an earlier retention time than relaxin itself consistent with its more hydrophilic nature (Fig. 3).

As might be expected, the cation-exchange chromatography was capable of detecting and resolving the charged variant better than reversed-phase chromatography. Cation-exchange chromatography was chosen as opposed to anion-exchange since relaxin has an isoelectric point above 7 (p*I* 10.5) as calculated according to Shire[38], in particular, polyaspartic acid cation-exchange columns were used since they have been reported to be capable of resolving proteins with as little difference as one amino acid[39].

Since this method separates proteins on the basis of their ionic character[40], it is a complementary technique to reversed-phase HPLC which resolves peptides primarily according to their hydrophobicity. Hydrophobic interaction and cation-exchange chromatography were investigated in an attempt to find an analytical procedure that would detect any variants that elude a reversed-phase HPLC step. This was particularly important in view of the preparative reversed-phase HPLC step used to purify the synthetic relaxin.

Tryptic mapping

Although synthetic relaxin was shown to be 98% pure by these different chromatographic techniques it was still possible that deletion peptides could not be detected due to any three dimensional structure of the hormone. Therefore, the tryptic map was a key technique in this study as enzymatic digestion generated nine small peptides that could be readily analyzed by reversed-phase HPLC. In addition, the trypsin digestion could be carried out on the unreduced protein so that information could also be gained concerning the disulfide linkages.

Digestion of the A-chain of human relaxin with trypsin can theoretically result in the release of five fragments (T_1, A1–9; T_2, A10–17; T_3, A18; T_4, A19–22; T_5, A23–24); that of the B-chain in the release of six fragments (T_6, B1–9; T_7, B10–13; T_8, B14–17; T_9, B18–30; T_{10}, B31; T_{11}, B32–33). By analogy with the disulfide pairing established for porcine relaxin[41] peptide T_2 would be expected to be covalently bonded to T_7 and peptide T_5 to T_9, respectively, in human relaxin. A typical tryptic map (digestion at pH 7.3) of relaxin is shown in Fig. 4. The peptide assignments were made after analysis of the peaks by acid hydrolysis for amino acid composition (data not shown) and confirmed by FAB-MS. The observed masses were within $\pm$ 0.7 a.m.u. of the theoretical values. Non-tryptic-like cleavages were seen to occur in the T_9 peptide. These peptides were identified as: T_{9a}, B18–25; T_{9b}, B18–28; T_{9c}, B26–30 as determined by their amino acid composition. The expected atomic mass units were observed for the T_5–T_{9a} peptide upon FAB-MS analysis. However, that expected for the T_{9c} peptide was low by 18 a.m.u., suggesting that a dehydration reaction had occurred. One possibility would be that a serine residue may have been converted to

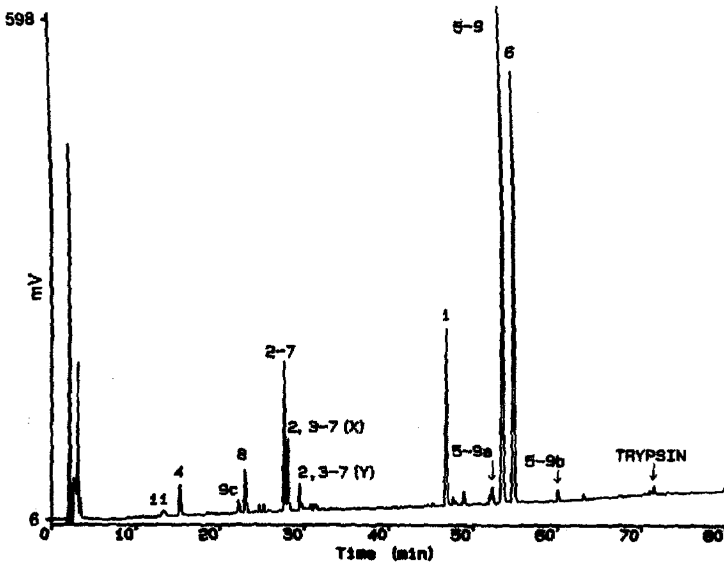

Fig. 4. Tryptic map of human relaxin following enzymatic digestion at pH 7.3. The digestion procedure was conducted as described in the Experimental section. The resulting peptide mixture (200 μl; 0.49 mg/ml) was separated by reversed-phase HPLC and monitored at 214 nm (0.5 volts per absorbance unit). The column was packed with Nucleosil C_{18} resin (150 mm × 4.6 mm I.D., 5μm, 300 Å). Elution was effected with a linear gradient from 3 to 30% acetonitrile containing 0.08% TFA over 54 min then to 70% acetonitrile in 40 min. The flow-rate was 1 ml/min and the column temperature was controlled at 35°C. The tryptic peptides are numbered sequentially from the amino termini of the A- and B-chains as described in the Results and Discussion section.

dehydroalanine. Since the recovery of serine by amino acid analysis is at least 83% in all cases and the FAB-MS spectrum of the intact B-chain shows no evidence for a dehydroalanine-containing peptide (data not shown), it is possible that the T_{9c} peptide underwent dehydration during the FAB-MS analysis. The carboxy terminal amino acid of T_{9b} was presumed to be tryptophan since the peptide T_5–T_{9b} had an appreciable absorbance at 280 nm. This deduction was confirmed by FAB-MS.

No evidence for improperly formed disulfide bonds was found when the tryptic digest was conducted at pH 7.3. However, if digestion was performed at the usual pH of 8.2 additional peptides were observed (Fig. 5). The addition of iodoacetic acid to the digestion reaction at pH 8.2 produced a profile similar to the one obtained when the digestion was conducted at pH 7.3 (Fig. 4). This suggested that disulfide exchanges were responsible for the extra peaks. These peaks were collected and identified by both acid hydrolysis for amino acid composition (data not shown) and by FAB-MS (Table III). It is apparent that a disulfide exchange is occuring during the tryptic digestion at pH 8.2. Not only are new peaks (identified as mixed disulfides) being observed, but the expected T_2–T_7 and T_5–T_9 peaks are decreasing. Dimers of T_2 and T_7 were also identified.

The cystine-containing peaks were isolated and reduced as described in the legend to Fig. 1 before rechromatography by reversed-phase HPLC. In every instance the two cysteine-containing peptides of the individual chains were observed (data not shown). A comparison of the retention times observed for each cysteine-containing

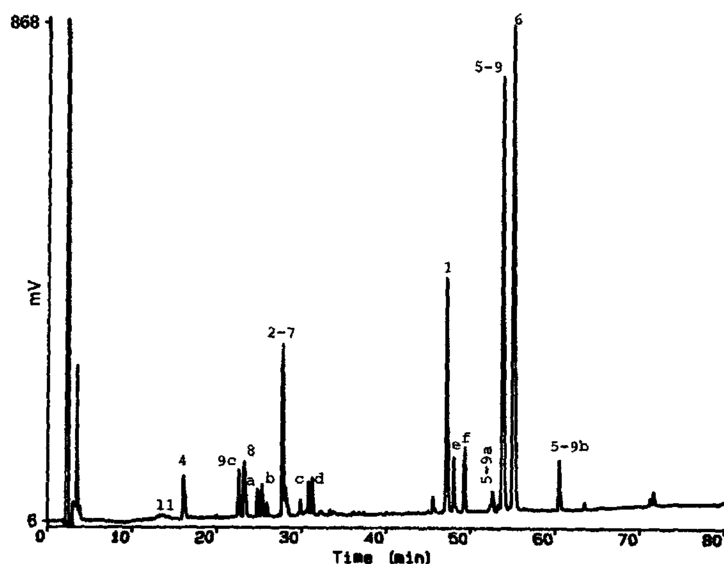

Fig. 5. Tryptic map of human relaxin following enzymatic digestion at pH 8.2. The procedure was conducted as described in the Experimental section and in Fig. 4 legend.

peptide led to the assignment of the component peptides. The identification of the T_2, T_5, T_7 and T_9 peptide peaks was confirmed by amino acid analysis. These cystine-containing peaks were also oxidized by performic acid before amino acid analysis to quantitate the cysteine levels. The correct composition was obtained for each of the peptides indicating that the six cysteine residues are actually present in the intact molecule.

These analyses not only confirmed the correct peptide sequence but also demonstrated the absence of significant coelutions. Thus, synthetic relaxin could be manufactured free of significant amounts of deletion peptides at the threshold of detection

TABLE III

MASS SPECTRAL ANALYSIS OF ADDITIONAL TRYPTIC PEPTIDES OF RELAXIN GENERATED AT pH 8.2.

Peak	*Theoretical mass*	*Observed*	*Assignment*
a	447.6	893.5	T_7
b	1293.8[a]	1293.6	T_2–T_7
		1295.6	
d	848.2[a]	848.2	T_2
		1695.8	
e	2242.1[a]	2243.2	T_2–T_9
f	1841.5	1842.1	T_7–T_9

[a] Molecular ions expected if the Cys^{A-10}–Cys^{A-15} intrachain disulfide bond is intact.

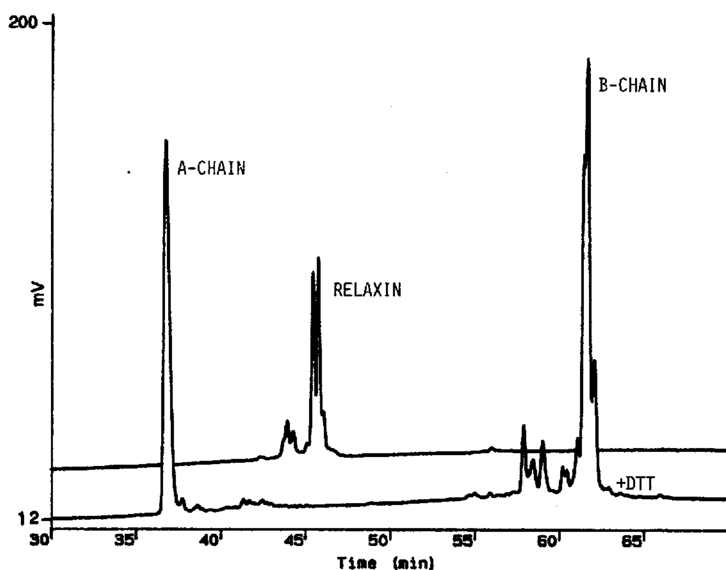

Fig. 6. Reversed-phase HPLC chromatograms of human relaxin side fraction before and after reduction with DTT. The chromatography was performed on a Vydac C_4 column as described in legend to Fig. 1.

in the tryptic map which is approximately 2–5%[42]. In addition, the following disulfide assignments were confirmed: T_2–T_7 and T_5–T_9. This unequivocally established Cys^{A-24}–Cys^{B-23}. The positions of the disulfides in T_2–T_7 will require further analysis to confirm their homology to those of insulin[43].

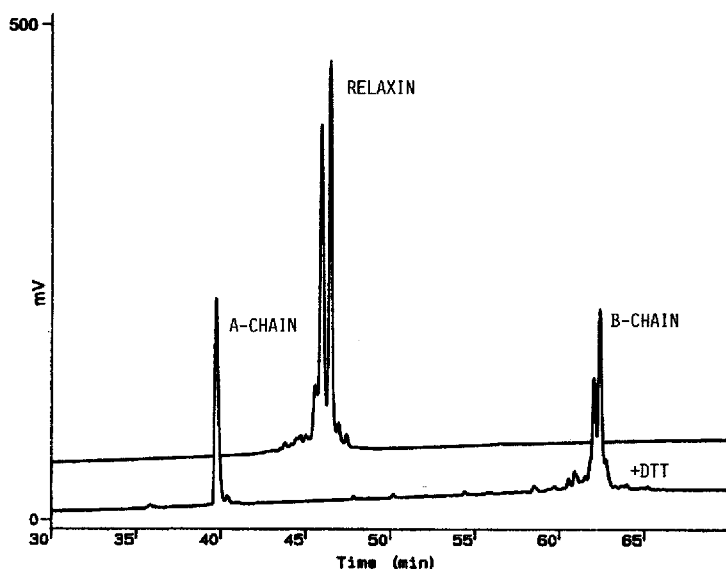

Fig. 7. Reversed-phase HPLC chromatograms of human relaxin preparation after incubation at 40°C before and after reduction with DTT. The chromatography was performed on a Vydac C_4 column as described in legend to Fig. 1.

Since deamidation of asparagine side chains is commonly seen in proteins[44], that at position A-8 was of interest. The T_1 peptide was chemically synthesized with an aspartic acid at position A-8. Its elution position in the tryptic map is as expected, immediately following the T_1 Asn-containing peptide. No detectible amount of the deamidated peptide is visible in the tryptic map of a purified relaxin preparation.

Evaluation of HPLC Techniques

In order to assess the ability of these HPLC methods to discriminate between closely related relaxin molecules, two samples known to be heterogeneous were obtained. One sample was a by-product of the purification procedure. The first step in the purification of relaxin after combination of the A- and B-chains is a reversed-phase HPLC isolation. The leading edge of the relaxin main peak is routinely removed in an effort to produce a homogeneous product. The second sample was generated by incubating a purified relaxin sample at 40°C in 10 m*M* Tris, pH 7.5, and isotonic saline for eight months. The analyses described above were performed on these preparations. Both of these samples contained material which eluted earlier than purified relaxin upon reversed-phase HPLC (Figs. 6 and 7). The profiles obtained from analysis of reduced material suggested that the differences were due to variants of the B-chain.

Although both HIC and cation-exchange chromatography were not effective in demonstrating any heterogeneity in the purified relaxin preparation, these methods did yield profiles of the side fraction and intentionally degraded preparations that consisted of multiple peaks (data not shown). These profiles were distinct from those obtained by reversed-phase HPLC (Figs. 6 and 7) and from each other substantiating the unique discriminating powers of these different chromatographic conditions.

The side fraction from the relaxin purification was also analyzed by tryptic mapping in the manner as described above for the purified fraction. A methionine sulfoxide was found at position B-4 in T_6 and at B-25 in T_9. In addition, four T_6 peptides were identified; namely, des Ser^2 Met^4 SO T_6, des Asp^1 T_6, des Ser^2 T_6 and des (Asp^1 Ser^2) T_6.

The intentionally degraded sample was analyzed similarly. The reversed-phase HPLC profile suggested that the methionine residues may have been oxidized (Fig. 7).

Analytical methods including amino acid analysis, reversed-phase HPLC and FAB-MS were utilized during the synthesis of the individual chains to monitor any synthesis-related problems.

The synthesis of the A-chain was reported[12] to be relatively straightforward compared to that of the B-chain. Not only was the B-chain difficult to synthesize[12,14], but it was also difficult to purify mainly due to its solubility characteristics which are a reflection of its high p*I* (11.5) calculated according to Shire[38] and extreme hydrophobicity[12]. Fortunately, the procedure of combining the A- and B-chains resulted in a product that lacked many of the impurities present in the B-chain component. The modifications of the B-chain shown to be present in the crude preparation such as *tert.*-butyl Trp and benzyl Cys[14] were not detected in the purified relaxin; neither were the B-chain deletions at Thr^{27}, Cys^{23} and Gln^{19} (ref. 14). The des Asp^{1} truncated peptide did combine with the A-chain to form relaxin, but it could be removed during the reversed-phase HPLC purification step. A deletion at Ser^{2}, which was not previously reported, was also found in the combined relaxin, but was removed along with the des Asp^{1} peptide during reversed-phase HPLC used to purify crude relaxin.

Chain combination followed by an optimized recovery process led to a clean profile upon reversed-phase HPLC. In order to demonstrate that this homogeneity is in fact real, and not due to lack of resolution of the technique, a number of tests were applied. The assays typically used for protein characterization, such as amino acid composition and amino terminal sequencing reinforced the conclusion of high purity as suggested by the reversed-phase chromatogram. In addition, degraded and chemically modified relaxin samples were identified in part by their altered retention times upon reversed-phase chromatography. Other chromatographic methods such as HIC and cation-exchange were not as effective as reversed-phase HPLC in detecting any latent heterogeneity.

Since no deamidation at Asn^{8} in the A-chain was detected by tryptic mapping, it is not surprising that the purity of relaxin preparations were seen to be higher when determined by ion-exchange as opposed to reversed-phase chromatography. The lack of evidence for deamidation of relaxin preparations is consistent with the absence of Asn–Gly or Asn–Ser sequences[45] in its primary structure (Table II). An ion-exchange chromatography step was included in the purification procedure to guard against such a possibility. Tryptic mapping, which was instrumental in characterizing the oxidized and truncated forms, also suggested that the final product was quite pure. Hence, it has been possible by the judicious choice of selected analytical techniques to monitor the purity of a chemically synthesized relaxin molecule and to use this knowledge to devise a purification scheme leading to a relaxin preparation that is more than 98% pure.

ACKNOWLEDGEMENTS

The authors wish to thank Dr. John Burnier for the synthesis of the relaxin A- and B-chains; Dr. Ernst Rinderknecht for the production and purification of relaxin from its individual chains; Reed Harris for the amino acid analyses; Karen Williams for the N-terminal sequence analysis; and Dr. John Stults for the FAB-MS analyses. Special thanks are due to Dr. William Hancock for helpful discussions in the preparation of the manuscript.

REFERENCES

1 R. James, H. Niall, S. Kwok and G. Bryant-Greenwood, *Nature (London)*, 267 (1977) 544.
2 M. J. John, B. W. Borjesson, J. R. Walsh and H. D. Niall, in G. D. Bryant-Greenwood, H. D. Niall and F. C. Greenwood (Editors), *Relaxin*, Elsevier/North-Holland, New York, 1981, p. 37.
3 L. K. Gowan, J. W. Reinig, C. Schwabe, S. Bedarkar and T. L. Blundell, *FEBS Lett.*, 129 (1981) 80.
4 E. E. Büllesbach, L. K. Gowan, C. Schwabe, B. G. Steinetz, E. O'Byrne and I. P. Callard, *Eur. J. Biochem.*, 161 (1986) 335.
5 P. Hudson, J. Haley, M. John, M. Cronk, R. Crawford, J. Haralambidis, G. Tregear, J. Shine and H. Niall, *Nature (London)*, 301 (1983) 628.
6 P. Hudson, M. John, R. Crawford, J. Haralambidis, D. Scanlon, J. Gorman, G. Tregear, J. Shine and H. Niall, *EMBO*, 3 (1984) 2333.
7 E. E. Büllesbach, C. Schwabe and I. P. Callard, *Biochem. Biophys, Res. Commun.*, 143 (1987) 273.
8 C. Schwabe, E. E. Büllesbach, H. Heyn and M. Yoshioka, *J. Biol. Chem.*, 264 (1989) 940.
9 J. Haley, P. Hudson, D. Scanlon, M. John, M. Cronk, J. Shine, G. Tregear and H. Niall, *DNA*, 1 (1982) 155.
10 P. Hudson, J. Haley, M. Cronk, J. Shine and H. Niall, *Nature (London)*, 291 (1981) 127.
11 G. Losse, H. Stange, W. Naumann and B. Schwenzer, *Z. Chem.*, 20 (1980) 429.
12 G. W. Tregear, Y-C. Du, K-Z. Wang, C. Southwell, P. Jones, M. John, J. Gorman, B. Kemp and H. D. Niall, in M. Bigazzi, F. C. Greenwood and F. Gasparri (Editors) *Biology of Relaxin and its Role in the Human*, Elsevier, Amsterdam, 1983, p. 42.
13 D. T. Liu, N. Goldman and F. Gates, III, in M. Phillips, S. P. Shoemaker, R. D. Middlekauff and R. M. Ottenbrite (Editors), *The Impact of Chemistry on Biotechnology: Multidisciplinary Discussions* (*ACS Symposium Series*, No. 362), American Chemical Society, Washington, DC, 1988, p. 162.
14 E. Canova-Davis, R. C. Chloupek, I. P. Baldonado, J. E. Battersby, M. W. Spellman, L. J. Basa, B. O'Connor, R. Pearlman, C. Quan, J. A. Chakel, J. T. Stults and W. S. Hancock, *Am. Biotechnol. Lab.*, 6 (1988) 8.
15 W. S. Hancock, E. Canova-Davis, R. C. Chloupek, S.-L. Wu, I. P. Baldonado, J. E. Battersby, M. W. Spellman, L. J. Basa and J. A. Chakel, *Banbury Report 29: Therapeutic Peptides and Proteins: Assessing the New Technologies*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1988, p. 95.
16 M. T. W. Hearn and W. S. Hancock, *Trends Biochem. Sci.*, 4 (1979) N58.
17 P. E. Petrides, R. T. Jones and P. Bohlen, *Anal. Biochem.*, 105 (1980) 383.
18 J. E. Rivier, *J. Liq. Chromatogr.*, 1 (1978) 343.
19 J. B. Shelton, J. R. Shelton and W. A. Schroeder, *Hemoglobin*, 3 (1979) 353.
20 M. Barber, R. S. Bordoli, R. D. Sedgwick and A. N. Tyler, *J. Chem. Soc., Chem. Commun.*, (1981) 325.
21 K. Biemann, *Anal. Chem.*, 58 (1986) 1288A.
22 J. P. Greenstein and M. Winitz, *Chemistry of the Amino Acids*, Robert E. Krieger Publ., Malabar, FL, 1984, p. 763.
23 R. B. Merrifield, *Adv. Enzymol.*, 32 (1969) 221.
24 J. P. Burnier and P. D. Johnston, *Eur. Pat., 0251615A2*, Bulletin 88/01, January 7, 1988.
25 B. G. Steinetz, V. L. Beach, R. L. Kroc, N. R. Stasilli, R. E. Nussbaum, P. J. Nemith and R. K. Dun, *Endrocinology*, 67 (1960) 102.
26 J. St.-Louis, *Can. J. Physiol. Pharmacol.*, 59 (1981) 507.
27 Hirs, C. H. W., *Methods Enzymol.*, 11 (1967) 59.
28 D. N. Podell, and G. N. Abraham, *Biochem. Biophys. Res. Commun.*, 81 (1978) 176.
29 G. H. Beaven, and E. R. Holiday, *Adv. Protein Chem.*, 7 (1952) 319.
30 D. B. Wetlaufer, *Adv. Protein Chem.*, 17 (1962) 303.
31 R. E. Chance, E. P. Kroeff, and J. A. Hoffmann, in J. L. Gueriguian (Editor), *Insulins, Growth Hormone and Recombinant DNA Technology*, Raven Press, New York, 1981, p. 71.
32 E. Canova-Davis, I. P. Baldonado and G. M. Teshima, unpublished results.
33 S. J. Shire, D. L. Foster and E. Rinderknecht, *Biophys. J.*, 53 (1988) 72a.
34 D. C. Teller, *Methods Enzymol.*, 27 (1972) 346.
35 D. Guo, C. T. Mant, A. K. Taneja, J. M. R. Parker and R. S. Hodges, *J. Chromatogr.*, 359 (1986) 499.
36 D. L. Gooding, M. N. Schmuck and K. M. Gooding, *J. Chromatogr.*, 296 (1984) 107.
37 C. Hoeger, R. Galyean, J. Boublik, R. McClintock and J. Rivier, *BioChromatography*, 2 (1987) 134.
38 S. J. Shire, *Biochemistry*, 22 (1983) 2664.
39 A. J. Alpert, *J. Chromatogr.*, 266 (1983) 23.

40 E. A. Peterson and H. A. Sober, *Methods Enzymol.*, 5 (1962) 3.
41 C. Schwabe and J. K. McDonald, *Science (Washington, DC)*, 197 (1977) 914.
42 W. S. Hancock, E. Canova-Davis, J. Battersby and R. Chloupek, in J. L. Gueriguian, V. Fattorusso and D. Poggiolini (Editors), *Biotechnologically Derived Medical Agents*, Raven Press, New York, 1988, p. 29.
43 A. P. Ryle, F. Sanger, L. F. Smith and R. Kitai, *Biochem. J.*, 60 (1955) 541.
44 A. B. Robinson, and C. J. Rudd, in B. L. Horecker and E. R. Stadtman (Editors), *Current Topics in Cellular Regulation,* Acedemic press, New York, 1974, p. 247.
45 M. Bodanszky and J. Z. Kwei, *Int. J. Pept. Protein Res.*, 12 (1978) 69.