The Primary Structure and the Disulfide Links of the Bovine Relaxin-like Factor (RLF)†

Erika E. Büllesbach* and Christian Schwabe

Department of Biochemistry and Molecular Biology, Medical University of South Carolina, 173 Ashley Avenue, PO Box 250509, Charleston, South Carolina 29425

Received August 28, 2001; Revised Manuscript Received November 7, 2001

ABSTRACT: The relaxin-like factor (RLF), produced by the Leydig cells, is an essential link in the chain of events leading to the proper positioning of the testes during fetal development. The primary structure of RLF, as reported in the literature, is based solely upon cDNA sequences with chain lengths determined according to deduced processing sites and with relaxin-like cross-links. Biochemical characterization of bovine testicular RLF shows clearly that the endogenous hormone does consist of a 26 residue A chain and two forms of B chain, one containing 40 residues, the other 45. In addition, both B chains are 9 residues longer at the C terminus than the cDNA-deduced chain, and about 20% of the B chains have an additional 5 residue extension at the N terminus. Sequence analysis in combination with mass spectrometry and tryptic peptide mapping showed unambiguously that RLF is larger than previously assumed and that it has the relaxin-type disulfide bond distribution that makes it a bona fide member of the relaxin family of hormones.

The relaxin-like factor (RLF)[1] mediates testicular descent in humans. It is a major developmental hormone that is produced predominantly by the Leydig cells (*1*) and, to a lesser extent, in the theca cells of ovaries (*2*, *3*). In the rat, RLF appears in the prenatal bloodstream in males, diminishes at birth, and rises again after puberty when functional Leydig cells are present (*4*). The biological activity of prenatal RLF has been elucidated in knockout experiments, which show that $RLF^{-/-}$ males retained testes in the body cavity (*5*, *6*). RLF interacts with specific receptors on the gubernaculum, triggering outgrowth of the gubernaculum which mediates abdominal migration of the testes (*4*, *7*).

Low levels of RLF mRNA were detected in the mouse ovary (*2*), and RLF receptors have been found in uterine tissue (*8*, *9*). The role of RLF in the female is unclear, with one report showing no effect (*5*) and another suggesting impaired fertility in the $RLF^{-/-}$ mice due to prolonged estrous cycles (*6*). While in human and rodent females RLF levels are low throughout life (*4*, *10*), in the ovaries of ruminants they are exceptionally high (*11*, *12*).

The structure of RLF deduced from the cDNA comprises a leader sequence, followed by a B domain, and a C domain which links the C terminus of the B chain and the N terminus of the A chain (*1*). This arrangement, however, does not imply that the message is processed to form a two-chain molecule like relaxin. It could remain unprocessed as in insulin-like growth factors (IGF), and could possess cross-links different from the relaxin configuration. Chemical structure analysis, the method of choice for disulfide bond determination, was used for insulin (*13*), relaxin (*14*, *15*), and IGF (*16*), and, in light of the importance of RLF for human development, it seems appropriate that this be done at least once for this new hormone.

In the present paper we report the isolation and characterization of bovine testicular RLF. In addition, the natural RLF is compared with a synthetic hormone derived from its predicted structure in terms of chemical, physicochemical, and receptor-binding parameters.

EXPERIMENTAL PROCEDURES

Materials. Bovine testes were purchased from Rockland (Gilbertsville, PA) and stored at −20 °C.

Peptide synthesis was performed by Fmoc chemistry on a solid support (*17*). The synthesis of bovine RLF and a Tyr-(A1) analogue was based on the protein structure deduced from the cDNA (*11*) which consisted of a 26 residue A chain and a 31 residue B chain. The use of differential protection for the cysteine side chains allowed the directed formation of the relaxin-like disulfide bonding arrangement. The synthesis was performed as described for human RLF (*18*).

Synthesis of Model Peptides. Different cysteine protecting groups were chosen to synthesize the three possible disulfide isomers of the tryptic fragment A9−17/B6−14. The two cysteines forming the intrachain disulfide loop of the A chain peptide were protected by the acid-stable acetamidomethyl group whereas the cysteine linking to the B chain was protected via the trifluoroacetic acid-labile trityl group. After trifluoroacetic acid treatment, the free sulfhydryl group was reacted with 2,2′-dipyridine disulfide in 1 M acetic acid, yielding the activated 2-pyridinesulfenyl form. The peptide was purified by gel filtration on Sephadex G10 in 1 M acetic

† This work was supported by National Institutes of Health Grant 1-R01-HD40406-01 and by U.S. Department of Agriculture Grant SCR-1999-03614.

* Correspondence should be addressed to this author. Phone: (843) 792-9926; Fax: (843) 792-4322; Email: bullesee@musc.edu.

[1] Abbreviations: Fmoc, 9-fluorenylmethyloxycarbonyl; HPLC, high-performance liquid chromatography; PBS, phosphate-buffered saline; RLF, relaxin-like factor.

10.1021/bi0117302 CCC: $22.00
Published on Web 12/11/2001

acid and lyophilized, and the interchain disulfide bond was established by reaction with the free sulfhydryl group of the B chain peptide under conditions described in the literature (*19*). The peptide was purified, the acetamidomethyl groups removed, and the intrachain disulfide bond formed by treatment with iodine in 80% acetic acid as described (*20*).

Monoclonal antibodies were produced by the antibody facility at the Medical University of South Carolina, using synthetic bovine RLF as antigen. Monoclonal antibody 2-8F (IgG1, κ) was used throughout this study.

^{125}I-3-Mono-iodotyrosyl(A1)-bovine RLF tracer was prepared from Tyr(A1) bovine RLF. The hormone (10 μL, c = 1 μg/μL in water) was placed into the tip of a 500 μL Eppendorf vial to which the following additions were made: 10 μL of 50 mM citrate buffer, pH 5.6, 0.9 mCi of Na^{125}I, 0.8 μL of hydrogen peroxide (1 mM in citrate buffer, pH 5.6), and 0.8 μL of lactoperoxidase (EC 1.11.1.7) (74 μg/100 μL of citrate buffer, pH 5.6). The reaction was performed for 1 min at room temperature; thereafter, the reaction was quenched with 30 μL of 33% acetic acid. The tracer was separated by reversed-phase HPLC using an Aquapore 300 (C_8, 2.1 mm × 30 mm) column. The mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 80% acetonitrile (solvent B). Separation was achieved by applying a 60 min linear gradient from 23 to 34% B at a flow rate of 100 μL/min. Peptides were detected by UV absorbance at 230 nm. Fractions were manually collected into 100 μL of 1% bovine serum albumin in water and the tracer stored at −20 °C.

Isolation of bRLF from Testis. One testicle (563 g) was thawed on ice for 6 h, and the vas deferens and scrotum were discarded. The testis was cut into 1 cm slices, ground in a meat grinder, dropped directly into chilled 1.6 M HCl (289 mL), and further homogenized with a Polytron homogenizer. Chilled acetone (850 mL) was added and the tissue extracted for 5 h at 4 °C. The extract was centrifuged at 4000 rpm in a Beckman J-6B centrifuge for 10 min at 4 °C, and the supernatant (1100 mL) was collected and brought up to 95% acetone (9.2 L). After 20 h at 4 °C, the protein had precipitated and the supernatant was siphoned off; the pellet was suspended and collected by centrifugation at 5000 rpm in a refrigerated centrifuge (Sorvall RC-5B, equipped with a GSA rotor). The pellet was dissolved in 1 M acetic acid, the pH adjusted to 5.5 with ammonia, the precipitate removed by centrifugation, and the supernatant dialyzed against water (3 × 10 L) and lyophilized (yield 1.37 g).

The lyophilized protein (100 mg/run) was dissolved in 1 M acetic acid (5 mL) and loaded onto a Sephadex G50sf column (2.5 × 50 cm), equilibrated with 1 M acetic acid. Separation was performed at a flow rate of 20 mL/h, fractions of 10 mL were collected, and 1 μL of each fraction was analyzed by RIA. The RIA-positive fractions were pooled, lyophilized (overall yield: 182 mg), and further purified by preparative HPLC. About 20 mg of the protein was loaded onto a Synchropak RP-P column (10 mm × 250 mm) equilibrated with 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 83% acetonitrile (solvent B) at 25% B in A. Separation was achieved by applying a 30 min linear gradient from 25 to 45% B at a flow rate of 3 mL/min, and 1 min fractions were collected. RIA-positive fractions were pooled and lyophilized (overall yield of nine identical runs: 9.9 mg). Thereafter the protein was dissolved in 50 mM sodium acetate buffer at pH 5.6 and separated on a CM-cellulose column (1.5 cm × 10 cm), equilibrated in the same buffer. After all UV-absorbing proteins were washed away, the RLF-enriched fraction was eluted with 0.2 M NaCl in sodium acetate buffer, pH 5.6. RIA-positive fractions were pooled, lyophilized and desalted, and rechromatographed by HPLC on a Synchropak RP-P column (10 mm × 250 mm) under the conditions described above. The main RIA-positive fraction was further purified by HPLC on a Bakerbond wide-pore C18 column (4.1 mm × 250 mm). Peptides were eluted with a 30 min linear gradient from 20 to 60% B at a flow rate of 1 mL/min. Peptides were detected by UV absorbance at 220 nm, and fractions were manually collected (yield: 0.33 mg).

Reduction, Carboxymethylation, and Chain Separation. Approximately 50 μg of bovine RLF was reduced in 50 μL of 25 mM dithiothreitol, containing 3 M guanidinium chloride and 0.1 M Tris/HCl at pH 8.6 for 60 min at 37 °C, and reacted with 50 μL of freshly prepared 100 mM iodoacetic acid in 100 mM *N*-methylmorpholine for 15 min at room temperature. The reaction mixture was acidified with 10 μL of glacial acetic acid and loaded onto a Bakerbond column for HPLC (conditions as described above).

Tryptic Digest and Peptide Analysis. In parallel experiments, approximately 10 μg of native and synthetic bovine RLF was digested with trypsin (E:S ratio of 1:25) in 20 μL of 50 mM ammonium hydrogen carbonate for 30 min at 37 °C. The digest was quenched with 50 μL of 0.1% trifluoroacetic acid and loaded onto Aquapore 300 for microseparation, applying a 40 min linear gradient from 0 to 40% B. Peptides were detected by UV absorbance at 220 nm. The peptides were collected manually and concentrated in a vacuum, and 20% of each fraction was hydrolyzed and subjected to amino acid analysis.

Amino Acid Composition. Peptides were hydrolyzed in vapor phase 6 M HCl containing 0.1% phenol for 1 h at 150 °C, modified with phenyl isothiocyanate, and separated by HPLC (Pico·Tag system, Waters).

Sequence analysis was performed on a Procise protein sequencer (Applied Biosystems) connected to an inline PTH analyzer. Analysis of the intact bovine RLF revealed the primary sequence of the A chain and about 20% of an additional sequence which was derived from an N-terminally extended B chain. The majority of the B chain appeared N-terminally blocked. Cysteine positions were identified as holes in the sequence. The primary structure of the A chain, including the positions of the cysteines, was confirmed by sequencing the reduced and carboxymethylated A chain.

The primary structure of the B chain was determined by direct sequencing of the reduced and S-carboxymethylated extended B chain and by sequencing fragments of the tryptic digest of the intact molecule.

Protein Concentration. UV spectroscopy and/or amino acid analysis was used to determine the concentrations of synthetic and highly purified natural RLF. The protein concentration of partially purified natural RLF was determined by weight.

Mass spectrometry was performed at the MUSC mass spectrometry facility. A Perceptive Voyager DE was used for matrix-assisted laser desorption ionization mass spectrometry, and Finnigan LCQ was used for electrospray ionization mass spectrometry.

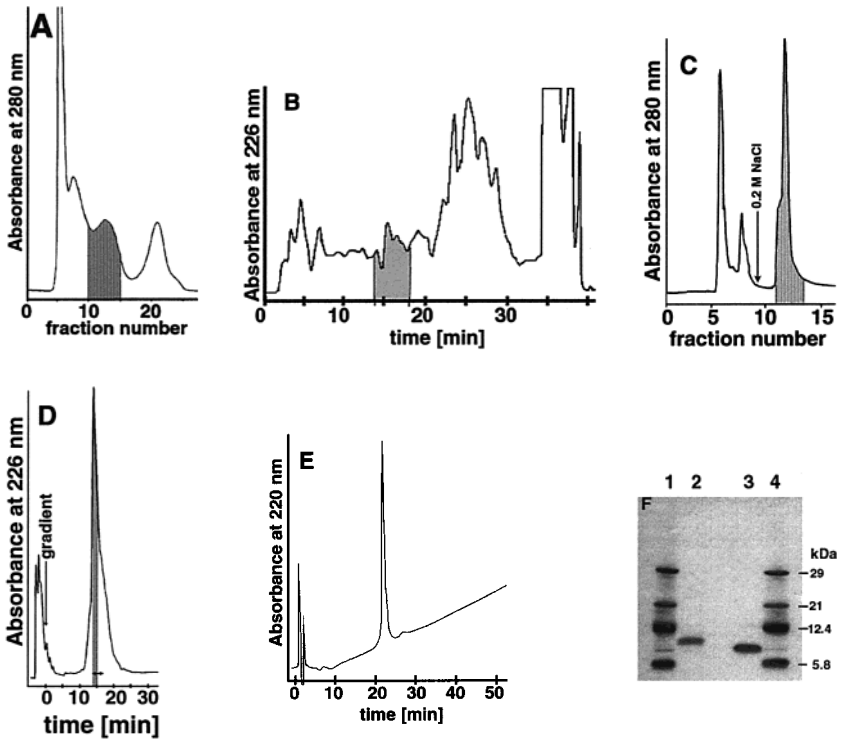

FIGURE 1: Panel A: Size separation of 95% acetone precipitate on Sephadex G50sf in 1 M acetic acid. Panel B: Reversed-phase HPLC purification of size-separated bovine RLF on Synchropak RP-P (C_{18}, 10 × 250 mm) in 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 83% acetonitrile (solvent B). Peptide was eluted with a linear gradient from 25 to 45% B over 30 min. Panel C: Separation of the HPLC-enriched bovine RLF by ion exchange chromatography on CM-cellulose in 50 mM sodium acetate buffer at pH 5.5 using a stepwise NaCl gradient. Panel D: Rechromatography on Synchropak RP-P under conditions described in panel B. The double-headed arrow indicates the RIA-active fractions. A single fraction represented by the shaded area was used for further investigation. Panel E: Analytical HPLC of 2 μg of purified bovine RLF on Aquapore 300 (C8, 2.1 × 30 mm). The mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 80% acetonitrile (solvent B). The peptide was eluted with a linear gradient from 20 to 50% B over 60 min at a flow rate of 100 μL/min. Panel F: Separation of bovine RLF on 16% polyacrylamide gels in SDS containing Tris/glycine buffer. Lane 1, standard; lane 2, native bovine RLF; lane 3, synthetic bovine RLF; and lane 4, standards. The apparent molecular mass of synthetic bovine RLF is about 7800 Da (calculated 6200 Da), and that of native bovine RLF is about 9200 Da (calculated 7200 Da).

Polyacrylamide Gel Electrophoresis (PAGE). For sodium dodecyl sulfate−PAGE, precast Tris−glycine gels (16% polyacrylamide) were purchased from Invitrogen. About 2 μg of native and synthetic bovine RLF was compared with a standard comprising bovine insulin (5800 Da), cytochrome *c* (12 400 Da), soybean trypsin inhibitor (21 000 Da), and carbonic anhydrase (29 000 Da). The protein was denatured in sample buffer at 95 °C for 5 min. Electrophoresis was performed in an Xcell Mini-Gel apparatus (Novex) for 90 min at 125 V. Gels were stained using the colloidal blue staining kit (Novex).

Radioimmunoassays. These assays were performed in pH 7.5 PBS buffer containing 1% BSA. Bovine RLF was dissolved in buffer and serial-diluted. To 100 μL of each concentration were added 100 μL of ^{125}I-bRLF (20 000 cpm) and 100 μL of 2-8F monoclonal antibody in the appropriate dilution, mixed, and incubated for 4 h at room temperature. The assay was terminated with rabbit anti-mouse IgG antibody that was covalently bound to CNBr-activated cellulose and the suspension shaken for 1 h at room temperature. Thereafter 3 mL of 50 mM phosphate buffer, pH 7.5, containing 0.1% Tween 20 was added and the solid collected by low-speed centrifugation for counting.

Receptor-Binding Assays. CD-1 virgin female mice were primed with 5 μg of estrogen cypionate in 100 μL of sesame oil and 5 days later killed in an atmosphere of CO_2. Uteri were collected and crude membranes prepared for receptor-binding assay as described before (*9*). Bovine RLF (synthetic or native) at concentrations from 0 to 200 ng/mL, bovine RLF tracer (100 000 cpm, 275 pM), and 40 μL of membrane preparation were mixed. The total volume of the assay was 100 μL. Each assay was performed in duplicate, and three assays were averaged.

RESULTS

Bovine RLF was isolated from testicles by extraction into acidic 50% acetone and precipitated with 95% acetone. It was purified to homogeneity by size separation on Sephadex G50sf (Figure 1A), followed by preparative reversed-phase HPLC (Figure 1B), ion exchange chromatography on CM-cellulose at pH 5.6 (Figure 1C), and rechromatography by reversed-phase HPLC (Figure 1D) (Table 1), yielding 0.33 mg of bovine RLF. Homogeneity was confirmed by analytical HPLC (Figure 1E), and SDS−polyacrylamide gel electrophoresis (Figure 1F). Electrospray ionization mass spectrometry revealed molecular masses of 7175.4 and 7560.0 Da.

Reduction and S-carboxymethylation yielded three peaks (Figure 2A); the more hydrophilic peptide was the A chain, and the two more hydrophobic peptides were B chains in a ratio of approximately 1:4, corresponding to one with a five amino acid residue N-terminal extension and one with a shorter chain. The A chain was characterized by amino acid

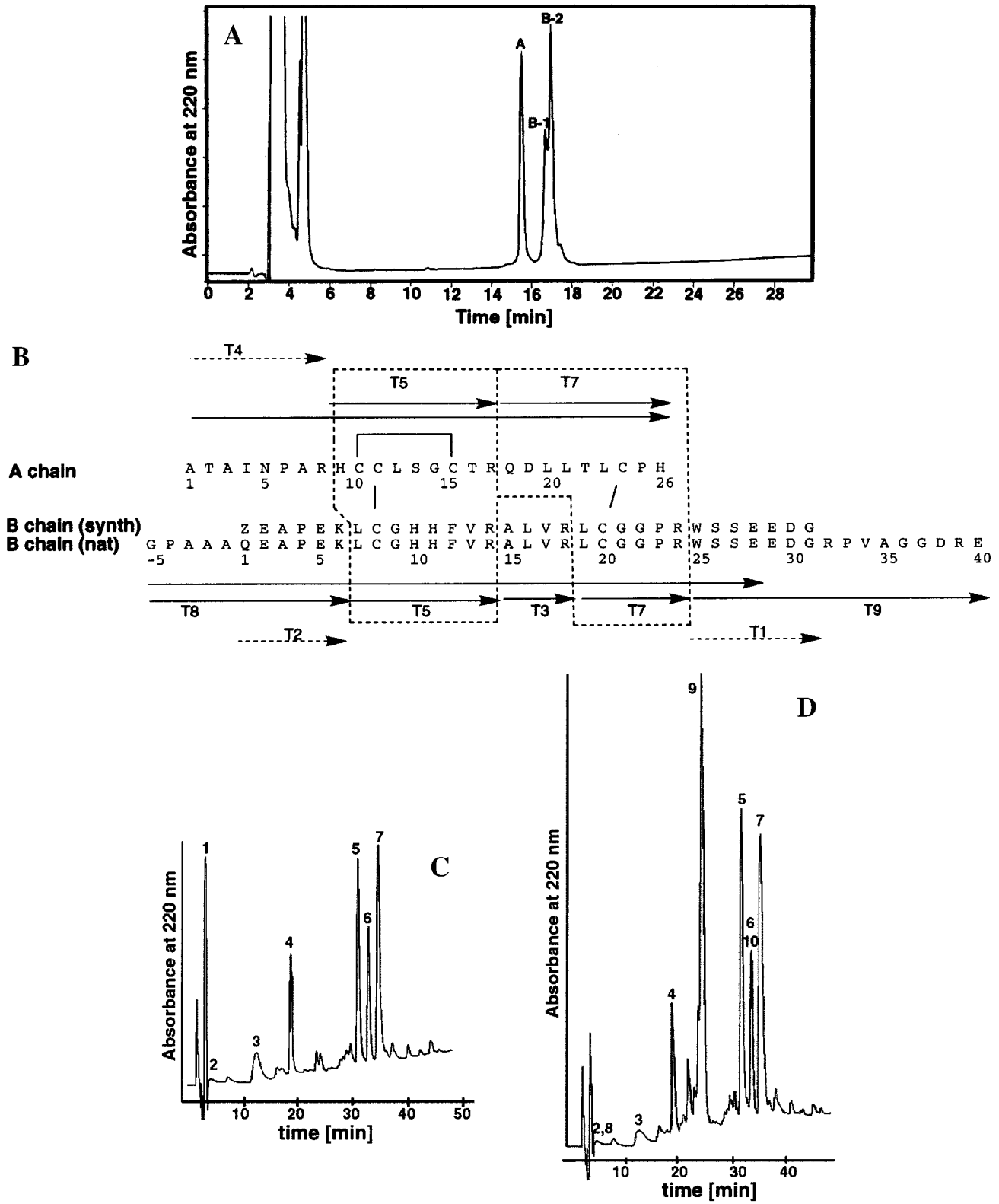

FIGURE 2: Panel A: HPLC separation of the reduced and carboxymethylated bovine RLF chains. Chromatography was performed on Bakerbond wide-pore 4.1 × 250 mm. The mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 83% acetonitrile (solvent B). The peptide was eluted with a linear gradient from 20 to 60% B over 30 min at a flow rate of 1 mL/min. Panel B: Primary structure of isolated bovine RLF as compared to synthetic bovine RLF, which consists of identical A chains but different B chains. The three disulfide bonds in synthetic bovine RLF were specifically synthesized. T1 to T9 are tryptic fragments, isolated and characterized either by protein sequencing (solid lines) or by amino acid composition (dashed lines). The dashed boxes frame the sequences of disulfide-linked peptides. The numbers correlate to the tryptic peptides isolated by reversed-phase HPLC shown in panels C and D. Tryptic digests of synthetic bovine RLF (panel C) and native bovine RLF (panel D). Tryptic fragments were separated on Aquapore 300 (C_8; 2.1 × 30 mm). The mobile phase consisted of 0.1% trifluoroacetic acid in water (solvent A) and 0.1% trifluoroacetic acid in 80% acetonitrile (solvent B). The peptide was eluted with a linear gradient from 0 to 40% B over 40 min at a flow rate of 100 μL/min. Peptides: 1 = B25−31; 2 = B1−6; 3 = B15−18; 4 = A1−8; 5 = A9−17/B7−14; 6 = A9−17/B1−B14 (by amino acid analysis); 7 = A18−26/B19−24; 8 = B(−5−6); 9 = B25−40; 10 = A9−17/B(−5-B14) (sequence analysis).

Table 1: Enrichment of Testicular Bovine RLF during Purification[a]

	protein (mg)	recovery (%)	ED_{50} (ng/mL)	absolute yield (mg)
bRLF (synthetic)	–	–	3.1	
acetone powder	880	100	600	5.7
Sephadex G50	173	26.8	350	1.53
Synchropak RP-P	12	12.6	52	0.72
CM-cellulose and HPLC	0.33	5.8	3.1	0.33

[a] Bovine RLF content was determined by RIA using monoclonal antibody 2-8F and synthetic bovine RLF as standard.

analysis, by mass spectrometry, and by protein sequencing. The A chain consisted of 26 residues containing 4 *S*-carboxymethylcysteines and had a molecular mass of 3028.0 Da (calcd 3027.4 Da).

The sequence of unreduced bovine RLF showed, besides the A chain sequence, a B chain sequence starting with Gly-Pro, which comprised about 20% of the A chain quantity. This indicated that the main fraction of the B chain is N-terminally blocked. Amino acid analyses of the reduced and alkylated B chains showed two *S*-carboxymethylcysteines. Mass spectrometry showed two B chain variants of different molecular mass; the smaller and more hydrophobic main fraction had a molecular mass of 4501.5 Da whereas the larger chain was 4885.5 Da. This is commensurate with a B chain that is nine residues longer than the B domain assigned from the cDNA sequence. In addition, the more hydrophilic (larger) fraction, which comprises approximately 20% of the main component, showed an extension at the N terminus of five amino acid residues, starting with Gly-Pro-Ala-Ala-Ala. This observation is in agreement with the cDNA sequence but puts a processing site into an unexpected region of the pre-peptide. The sequence matches the molecular mass and the Edman degradation data (Figure 2B).

Tryptic digestion of intact bovine RLF yields two hydrophilic fragments of the N terminus of the B chain with very low UV absorbance (Figure 2C,D). The N-terminally blocked hexapeptide (T2) was clearly identified by amino acid composition, and the more hydrophobic component (T8), upon sequencing, showed the five residue extension. The C-terminal tryptic peptide of the B chains gave rise to a strong UV positive peak (T9), which is indicative of the presence of the only tryptophan in the molecule. Trypsin cleaved neither the Arg−Pro bond (B32/33) nor the Arg−Glu bond (B39/40) at the C terminus of the B chain. The derived structure (Figure 2B) matches the calculated mass of the complete protein sequence as observed during mass spectrometry which showed a main component of 7175.4 Da (calcd 7175.1 Da) and a minor component of 7560.0 Da (calcd 7559.5 Da).

The tryptic fragments were also used to identify the disulfide-bonding pattern. For comparison, we digested bovine RLF synthesized according to the published cDNA sequence which consisted of a 26 residue A and a short 31 residue B chain (Figure 2B), and the 3 disulfide bonds as they are in relaxin (*8, 18, 21*). The elution profiles of the tryptic fragments of synthetic and natural bovine RLF showed near-identity (Figure 2C,D). Retention times, amino acid composition, and sequence analysis of the two cysteine-containing peptides (T5 and T7) were identical between the synthetic and isolated molecules, confirming the relaxin-like orientation of the chains. This unambiguously identified the

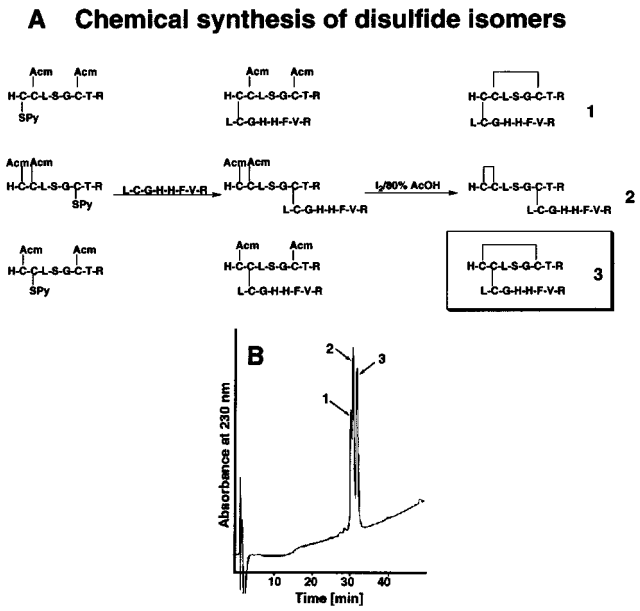

FIGURE 3: Panel A: Synthesis of the three disulfide isomers of the tryptic fragment A9−17/B7−14. Acm = acetamidomethyl, SPy = *S*-pyridinesulfenyl. Panel B: HPLC separation of the three disulfide isomers of the tryptic fragment A9−17/B7−14 on Aquapore 300 (C_8; 2.1 × 30 mm). HPLC conditions are described for Figure 2C,D. Disulfide bonding: 1 = A11/A15 and A10/B8; 2 = A10/A11 and A15/B8; 3 = A10/A15 and A11/B8.

disulfide bond A24/B20 but left a degree of uncertainty for the positions of the two disulfide bonds within fragment T5 (Figure 2B−D). To clearly identify the natural disulfide bonds, we synthesized the three different disulfide isomers of the tryptic fragment T5 (Figure 3A). HPLC analysis (Figure 3B) indicated that the three isomers have different chromatographic properties and only the disulfide isomer-3 with disulfide bonds A10/A15 and A11/B8 coeluted with the tryptic fragment (T5), confirming the relaxin-like disulfide bonding in endogenous bovine RLF.

A side-by-side comparison of synthetic and isolated bovine RLF revealed that both have identical affinities to the monoclonal antibody 2-8F (Figure 4A). In competitive binding experiments with the mouse uterine receptor, synthetic bovine RLF showed an ED_{50} of 0.45 nM whereas the isolated bovine RLF showed an ED_{50} of 1.4 nM (Figure 4B).

The RLF content of bovine testes in terms of immunereactive material varied significantly between specimens. One of four testes worked up in this study contained high RLF levels (5.7 mg total) while the other three contained low RLF levels (59, 22, and 63 μg, respectively). This may indicate a different state of maturity of the animal from which the gonads were recovered.

DISCUSSION

This paper concerns the first isolation and biochemical characterization of the relaxin-like factor from a natural source. While our work establishes RLF as a relaxin-like hormone, there are differences between the endogenous bovine hormone and the cDNA-deduced structure as concerns pro-hormone conversion sites. The endogenous hormone is larger than that deduced from the cDNA, and this alters measurably the affinity of RLF for its receptor. The shorter RLF which we synthesized according to the cDNA shows greater avidity in the receptor-binding assay than the

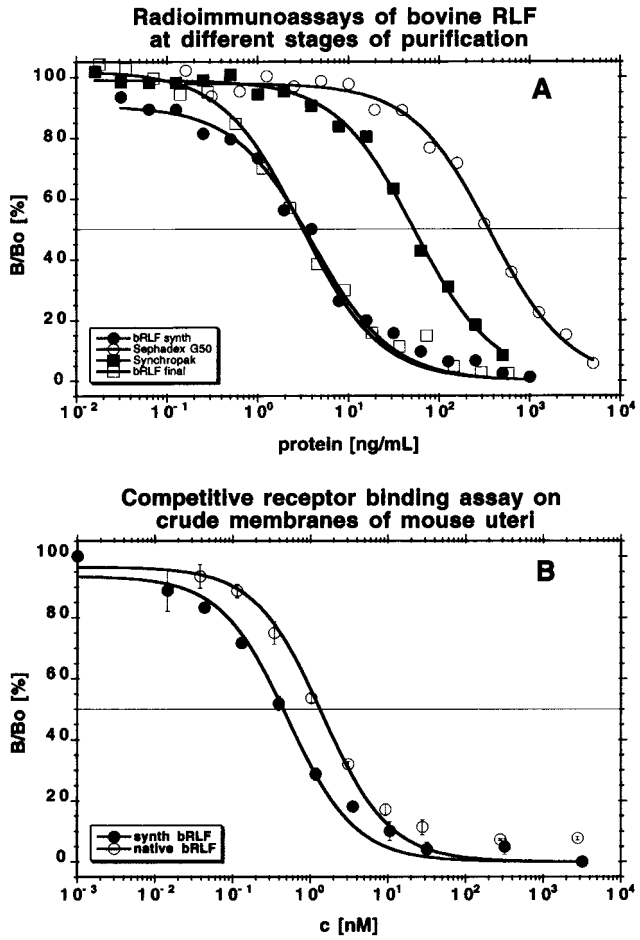

FIGURE 4: Panel A: Radioimmunoassay of bovine RLF, using synthetic ^{125}I-Tyr(A1)bovine RLF for tracer in combination with monoclonal antibody 2-8F. Dose−response curves of native bovine RLF at different stages during the isolation are compared with synthetic bovine RLF, which served as antigen. Panel B: Receptor-binding assay of native (open circles) and synthetic (closed circles) bovine RLF. Assays were performed on crude membrane preparations of uteri of estrogen-primed mice and synthetic ^{125}I-Tyr(A1)-bovine RLF for tracer. Data points were collected in duplicate. The presented data are the mean (±SEM) of three independent experiments.

larger natural version. During the course of our structure/function work, we identified Trp(B25) as a crucial residue for RLF receptor interaction (*8*, *18*, *21*). In the endogenous (larger) hormone, the binding-site tryptophan may actually be sterically hindered by the nine amino acid residue extension. Thus, our shortened synthetic version represents an improvement upon the natural hormone in terms of receptor binding. Clearly, one must be aware of the possibility that this extension may have physiological significance.

Pre-pro-RLF processing requires at least three specific enzymes (Figure 5). The first cleavage is due to a signal peptidase, which produces two N termini, one due to cleavage between leucine and glycine and the other between alanine and glutamine. The resulting pro-RLF is converted to RLF by a furin-like endoprotease with specificity for a pair of basic amino acid residues at the C-peptide−A chain junction. An unknown convertase ignores an arginine in position P2 and cleaves between Glu(B40) and Leu to remove the connecting peptide from the C terminus of the B chain. In fact, all pro-RLF cDNAs (*1*, *2*, *11*, *12*, *22*−*29*), except for canine pro-RLF (*30*), show a Glu-Leu-Leu-sequence in positions P1, P1′, and P2′, which suggests that this enzyme is guided by a -Glu-Leu-Leu- sequence to affect C-peptide removal. This is probably also true for human and marmoset RLF even though these molecules have a pair of basic amino acids eight residues earlier in the C-terminal region of the B chain (EARRP) (*22*, *24*). The Arg−Pro bond in that recognition site is likely resistant to hydrolysis by trypsin-like enzymes.

Unusual pro-hormone−hormone conversions have been observed before. One example is the conversion of porcine pro-relaxin where an unknown enzyme cleaves the B chain/C-peptide junction between Leu(B32) and Ser(B33). Subsequent carboxypeptidase action accounts for the observed series of shorter B chains (*31*, *32*). RLF, in contrast, seems to be processed to a defined product ending in Glu(B40) in bovidae.

The appearance of two different forms of the B chain is due to an N-terminal extension and may be caused by alternative cleavage by the signal peptidase, or due to subsequent removal of the pentapeptide GPAAA by a different enzyme. Intermediates are not detected in RLF preparation so that aminopeptidase action is not likely. Besides the overall structure of signal peptides, the critical residues for enzyme recognition are thought to contain small and neutral residues in positions −3 and −1 (*33*). This

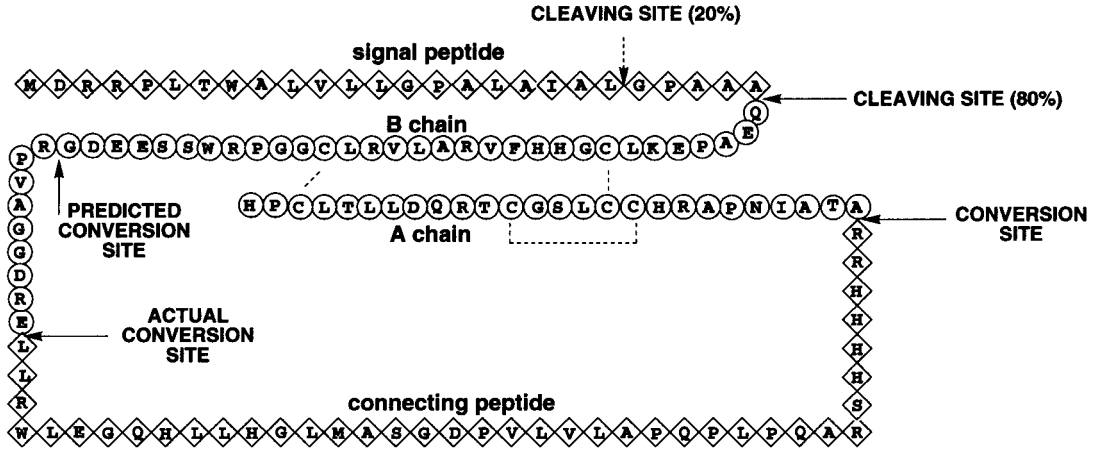

FIGURE 5: Amino acid sequence of bovine pre-pro-RLF as deduced from the cDNA sequence is shown with hypothetical and experimentally determined processing sites.

criterion is matched in bovine RLF in the sequence AAAQE (−3 to 2) which is hydrolyzed between Ala and Gln by a type 1 signal peptidase. The five residue longer B chain is produced by cleavage between Leu and Gly within the sequence IAL-GP where the −3 and −1 residues are hydrophobic and bulky, which is a low-preference cleaving site. Incomplete, partial, or alternative cleavage of signal peptides in secreted proteins is usually affiliated with deletion mutations in the N-terminal region of the mature protein (*34*, *35*), or with mutations within the signal peptide as observed in different disease states (*36*−*38*). In bovine RLF, there is no indication of a mutation in the mature hormone. Consequently, alternative cleavage could be caused by a mutation in the signal or connecting peptide, or could in this case be due to two different enzyme reactions, the first cleaving off the leader peptide at position −5 and the second removing the now N-terminal pentapeptide.

It is clear that one cannot rely upon cDNA structures and conjectures about conversion site when a protein needs to be characterized. At least one member of each group of proteins has to be isolated and characterized to determine its biologically active form(s). A number of such small two-chain proteins that have been discovered by gene technology are awaiting systematic investigation. While we have synthesized the INSL-4 gene product (*39*) guessing, as it were, at the primary structure, disulfide bonds, and chain lengths, the molecule does not take on an orderly structure in aqueous solvents (*40*). On the horizon are INSL-5 and -6 gene products (*41*−*44*), which we have synthesized assuming a relaxin-like structure, and although they show better physical−chemical properties than INSL-4, much work needs to be done before one can be sure that those structures resemble the biologically active proteins. The RLF data presented here demonstrate that uncertainties in our predictive powers should be a real concern.

According to our investigations of the disulfide arrangement, the chain identity and orientation, and the previously published CD spectrum (*8*), the relaxin-like factor is indeed relaxin-like.

ACKNOWLEDGMENT

We thank Robert Bracey and George Fullbright for technical assistance.

REFERENCES

1. Adham, I. M., Burkhardt, E., Benahmed, M., and Engel, W. (1993) *J. Biol. Chem. 268*, 26668−26672.
2. Zimmermann, S., Schottler, P., Engel, W., and Adham, I. M. (1997) *Mol. Reprod. Dev. 47*, 30−38.
3. Balvers, M., Spiess, A. N., Domagalski, R., Hunt, N., Kilic, E., Mukhopadhyay, A. K., Hanks, E., Charlton, H. M., and Ivell, R. (1998) *Endocrinology 139*, 2960−2970.
4. Boockfor, F. R., Fullbright, G., Büllesbach, E. E., and Schwabe, C. (2001) *Reproduction* (in press).
5. Zimmermann, S., Steding, G., Emmen, J. M. A., Brinkmann, A. O., Nayernia, K., Holstein, A. F., Engel, W., and Adham, I. M. (1999) *Mol. Endocrinol. 13*, 681−691.
6. Nef, S., and Parada, L. F. (1999) *Nat. Genet. 22*, 295−299.
7. Emmen, J. M. A., McLuskey, A., Adham, I. M., Engel, W., Grootegoed, J. A., and Brinkmann, A. O. (2000) *Endocrinology 141*, 4720−4727.
8. Büllesbach, E. E., and Schwabe, C. (1995) *J. Biol. Chem. 270*, 16011−16015.
9. Büllesbach, E. E., and Schwabe, C. (1999) *J. Biol. Chem. 274*, 22354−22358.
10. Büllesbach, E. E., Rhodes, R., Rembiesa, B., and Schwabe, C. (1999) *Endocrine 10*, 167−169.
11. Bathgate, R., Balvers, M., Hunt, N., and Ivell, R. (1996) *Biol. Reprod. 55*, 1452−1457.
12. Roche, P. J., Butkus, A., Wintour, E. M., and Tregear, G. (1996) *Mol. Cell. Endocrinol. 121*, 171−177.
13. Ryle, A. P., Sanger, F., Smith, L. F., and Kitai, R. (1955) *Biochem. J. 60*, 541−556.
14. Schwabe, C., and McDonald, J. K. (1977) *Science 197*, 914−915.
15. Stults, J. T., Bourell, J. H., Canova-Davis, E., Ling, V. T., Laramee, G. R., Winslow, J. W., Griffin, P. R., Rinderknecht, E., and Vandlen, R. L. (1990) *Biomed. Environ. Mass Spectrom. 19*, 655−664.
16. Raschdorf, F., Dahinden, R., Maerki, W., Richter, W. J., and Merryweather, J. P. (1988) *Biomed. Environ. Mass Spectrom. 16*, 3−8.
17. Atherton, E., and Sheppard, R. C. (1989) *Solid-phase peptide synthesis, a practical approach. Practical Approach Series* (Rickwood, D., and Hames, B. D., Eds.) IRL Press, Oxford.
18. Büllesbach, E. E., and Schwabe, C. (1999) *Biochemistry 38*, 3073−3078.
19. Castell, J. V., and Tun-Kyi, A. (1979) *Helv. Chim. Acta 62*, 2507−2510.
20. Kamber, B., Hartmann, A., Eisler, K., Riniker, B., Rink, H., Sieber, P., and Rittel, W. (1980) *Helv. Chim. Acta 63*, 899−915.
21. Büllesbach, E. E., and Schwabe, C. (1991) *J. Biol. Chem. 266*, 10754−10761.
22. Burkhardt, E., Adham, I. M., Hobohm, U., Murphy, D., Sander, C., and Engel, W. (1994) *Hum. Genet. 94*, 91−94.
23. Pusch, W., Balvers, M., and Ivell, R. (1996) *Endocrinology 137*, 3009−3013.
24. Zarreh-Hoshyari-Khah, M. R., Einspanier, A., and Ivell, R. (1999) *Biol. Reprod. 60*, 445−453.
25. Hombach-Klonisch, S., Tetens, F., Kauffold, J., Steger, K., Fischer, B., and Klonisch, T. (1999) *Mol. Reprod. Dev. 53*, 135−141.
26. Spiess, A. N., Balvert, M., Tena-Sempere, M., Huhtaniemi, I., Parry, L., and Ivell, R. (1999) *Mol. Reprod. Dev. 54*, 319−325.
27. Hombach-Klonisch, S., Abd-Elnaeim, M., Skidmore, J. A., Leiser, R., Fischer, B., and Klonisch, T. (2000) *Biol. Reprod. 62*, 839−846.
28. Klonisch, T., Froelich, C., Tetens, F., Fischer, B., and Hombach-Klonisch, S. (2001) *Mol. Biol. Evol. 18*, 393−403.
29. Hombach-Klonisch, S., Kauffold, J., Rautenberg, T., Steger, K., Tetens, F., Fischer, B., and Klonisch, T. (2000) *Mol. Cell. Endocrinol. 159*, 147−158.
30. Klonisch, T., Kauffold, J., Steger, K., Bergmann, M., Leiser, R., Fischer, B., and Hombach-Klonisch, S. (2001) *Biol. Reprod. 64*, 442−450.
31. Niall, H. D., John, M., James, R., Kwok, S., Mercado, R., Bryant-Greenwood, G., Bradshaw, R. A., Gast, M., and Boime, I. (1980) in *Insulin: Chemistry, structure and function of insulin and related hormones* (Brandenburg, D., and Wollmer, A., Eds.) pp 719−726, Walter de Gruyter, Berlin.
32. Kemp, B. E., and Niall, H. D. (1984) *Vitam. Horm. (San Diego) 41*, 79−115.
33. Nielsen, H., Engelbrecht, J., Brunak, S., and von Heijne, G. (1997) *Protein Eng. 10*, 1−6.
34. Wiren, K. M., Potts, J. J., Jr., and Kronenberg, H. M. (1988) *J. Biol. Chem. 263*, 19771−19777.
35. Wiren, K. M., Ivashkiv, L., Freeman, M. W., Potts, J. J., Jr., and Kronenberg, H. M. (1989) *Mol. Endocrinol. 3*, 240−250.
36. Racchi, M., Watzke, H. H., High, K. A., and Lively, M. O. (1993) *J. Biol. Chem. 268*, 5735−5740.
37. Karaplis, A. C., Lim, S. K., Baba, H., Arnold, A., and Kronenberg, H. M. (1995) *J. Biol Chem. 270*, 1629−1635.
38. Siggaard, C., Rittig, S., Corydon, T. J., Andreasen, P. H., Jensen, T. G., Andresen, B. S., Robertson, G. L., Gregersen,

N., Bolund, L., and Pedersen, E. B. (1999) *J. Clin. Endocrinol. 84*, 2933−2941.
39. Chassin, D., Laurent, A., Janneau, J. L., Berger, R., and Bellet, D. (1995) *Genomics 29*, 465−470.
40. Büllesbach, E. E., and Schwabe, C. (2001) *J. Pept. Res. 57*, 77−83.
41. Conklin, D., Lofton-Day, C. E., Haldeman, B. A., Ching, A., Whitmore, T. E., Lok, S., and Jaspers, S. (1999) *Genomics 60*, 50−56.
42. Hsu, S. Y. (1999) *Mol. Endocrinol. 13*, 2163−2174.
43. Kasik, J., Muglia, L., Stephan, D. A., and Menon, R. K. (2000) *Endocrinology 141*, 458−461.
44. Lok, S., Johnston, D. S., Conklin, D., Lofton-Day, C. E., Adams, R. L., Jelmberg, A. C., Whitmore, T. E., Schrader, S., Griswold, M. D., and Jasper, S. R. (2000) *Biol. Reprod. 62*, 1593−1599.

BI0117302