Amino Acids (2010) 39:1343–1352
DOI 10.1007/s00726-010-0586-3

ORIGINAL ARTICLE

Design and recombinant expression of insulin-like peptide 5 precursors and the preparation of mature human INSL5

Xiao Luo · Ross A. D. Bathgate · Wei-Jie Zhang · Ya-Li Liu · Xiao-Xia Shao · John D. Wade · Zhan-Yun Guo

Received: 4 January 2010 / Accepted: 26 March 2010 / Published online: 11 April 2010

Abstract Insulin-like peptide 5 (INSL5) is a recently identified insulin superfamily member. Although it binds to and activates the G-protein coupled receptor, RXFP4, its precise biological function remains unknown. To help determine its function, significant quantities of INSL5 are required. In the present work, three single-chain INSL5 precursors were designed, two of which were successfully expressed in *E. coli* cells. The expressed precursors were solubilized from inclusion bodies, purified almost to homogeneity by immobilized metal-ion affinity chromatography, and then refolded in vitro. One precursor could be converted to two-chain human INSL5 bearing an extended N-terminus of the A-chain (designated long-INSL5) by sequential Lys-C endoproteinase and carboxypeptidase B treatment. The 6 residue A-chain N-terminal extension of long-INSL5 was subsequently removed by *Aeromonas* aminopeptidase to yield native INSL5 that was designated short-INSL5. Circular dichroism spectroscopic analysis and peptide mapping showed that the recombinant INSL5s adopted an insulin-like conformation and possessed the expected characteristic insulin-like disulfide linkages. Activity assay showed that both long- and short-INSL5 had full RXFP4 receptor activity compared with chemically synthesized human INSL5. This suggested that extension of the N-terminus of the A-chain of long-INSL5 did not adversely impact upon the binding to or activation of the RXFP4 receptor. However, the single-chain INSL5 precursor was inactive which indicated that a free C-terminus of the B-chain is critical for the activity of INSL5. Our present work thus provides an efficient approach for preparation of INSL5 and its analogs through recombinant expression in *E. coli* cells.

Keywords Activity · INSL5 · Recombinant expression · Single-chain precursor · Refolding

X. Luo · W.-J. Zhang · X.-X. Shao · Z.-Y. Guo (✉)
Institute of Protein Research, Tongji University,
1239 Siping Road, Shanghai 200092, China
e-mail: zhan-yun.guo@tongji.edu.cn

R. A. D. Bathgate · J. D. Wade (✉)
Florey Neuroscience Institutes, Howard Florey Institute,
The University of Melbourne, Melbourne, VIC 3010, Australia
e-mail: john.wade@florey.edu.au

R. A. D. Bathgate
Department of Biochemistry and Molecular Biology,
The University of Melbourne, Melbourne, VIC 3010, Australia

Y.-L. Liu
Central Laboratory, Shanghai East Hospital, Tongji University
School of Medicine, Shanghai 200120, China

J. D. Wade
School of Chemistry, The University of Melbourne,
Melbourne, VIC 3010, Australia

Abbreviations

CD	Circular dichroism
DTT	Dithiothreitol
GSSG	Oxidized glutathione
INSL5	Insulin-like peptide 5
IPTG	Isopropyl β-D-thiogalactoside
TCEP	Tris(2-carboxyethyl)phosphine
TFA	Trifluoroacetic acid
UV	Ultra violet

Introduction

The insulin superfamily is a group of peptide hormones that share an identical cysteine arrangement pattern, but

which have diversified biological functions (Shabanpoor et al. 2009). One member of the family, insulin-like peptide 5 (INSL5), was identified from the Expressed Sequence Tag database in 1999 (Conklin et al. 1999; Hsu 1999), and also named relaxin/insulin-like factor 2 (RIF2) (Hsu 1999). Its mRNA is primarily found in the rectum, colon, and uterus (Conklin et al. 1999; Hsu 1999). INSL5 activates a G-protein coupled receptor, RXFP4 (formerly named GPCR142), (Liu et al. 2005b) that is also activated by relaxin-3 (Liu et al. 2003). A number of INSL5/relaxin-3 hybrid peptides have been designed and prepared by recombinant DNA techniques in order to discriminate the binding properties of the relaxin family peptide receptors (Haugaard-Jönsson et al. 2008; Kuei et al. 2007; Liu et al. 2005a). The transmembrane domains 2, 3, and 5 of RXFP4 are critical determinants of the functional receptor activation by INSL5 (Zhu et al. 2008). So far, the precise biological function of INSL5 remains unknown. However, one group has suggested that homozygotes of INSL5-knockout mice show no conspicuous phenotypes (Shirneshan 2005), whereas published patent applications (WO 2005/124361 A2; US 20080269118) suggest that the phenotype of the RXFP4-knockout mice points to a role for INSL5 in both fat and glucose metabolism. A recent NMR solution structure determination shows that the chemically synthesized INSL5 (Hossain et al. 2008a) adopts an insulin/relaxin-like fold (Haugaard-Jönsson et al. 2009).

To help determine its biological function, significant quantities of INSL5 peptide is needed. In our present work, we designed three INSL5 precursors and successfully expressed two of them in *E. coli* cells. After purification and in vitro refolding, one precursor was enzymatically converted to two-chain human INSL5s with differing lengths of the N-terminus of the A-chain. The recombinant two-chain INSL5s had full RXFP4 activation activity compared with the chemically synthesized INSL5. Our present work provides an efficient approach for the preparation of INSL5 and its analogs through recombinant expression in *E. coli* cells.

Materials and methods

Gene construction, expression, and purification of the single-chain INSL5 precursors

Three single-chain INSL5 precursors were designed in our present work as shown in Fig. 1a. To construct their genes, four chemically synthesized oligonucleotide primers (Fig. 1b) were annealed, elongated by T4-DNA polymerase, cleaved by restriction enzymes *Nde*I and *Eco*RI, and

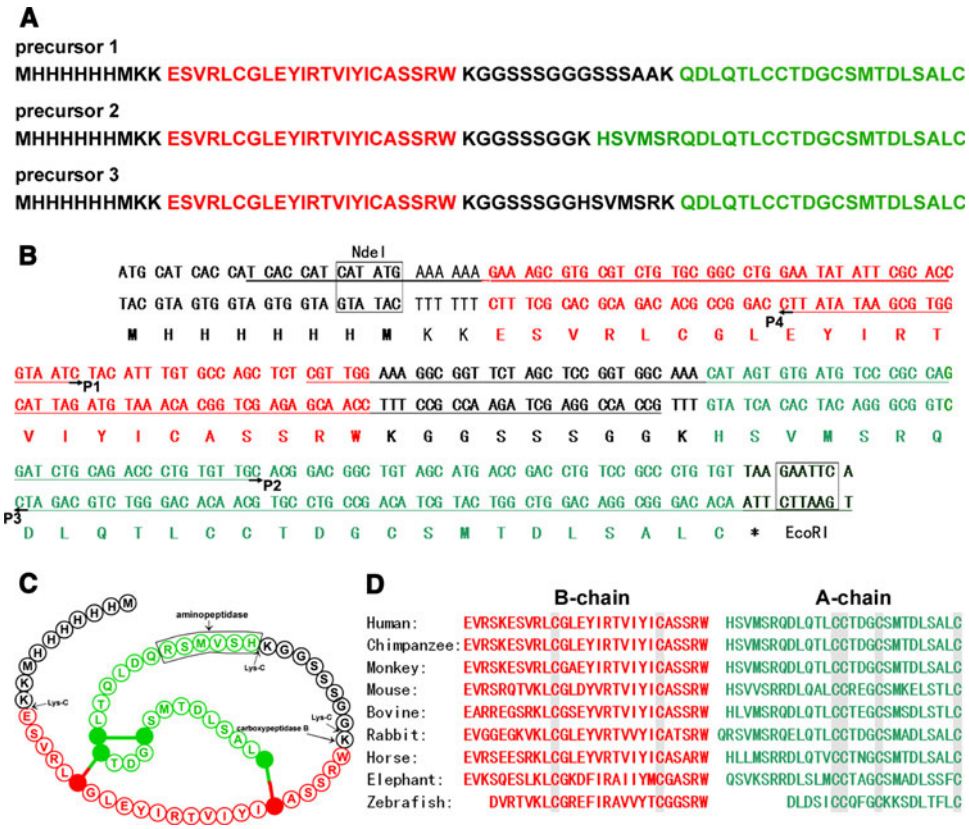

Fig. 1 **a** The amino acid sequence of three INSL5 precursors designed in present work. **b** The amino acid sequence and nucleotide sequence of INSL5 precursor 2. The four nucleotide primers (P1, P2, P3, and P4) used to construct the gene of INSL5 precursor 2 are *underlined* and *labeled*. **c** A cartoon show of the amino acid sequence of the INSL5 precursor 2. The cysteines are shown in *filled circles*. The disulfide bonds are shown as *sticks*. The Lys-C endoproteinase cleavage sites, the additional lysine residue removed by carboxypeptidase B, and the N-terminal extension removed by *Aeromonas* aminopeptidase are indicated. **d** The amino acid sequence of INSL5s from different species

subsequently ligated into a pET vector that was pretreated with same restriction enzymes. The encoding DNA fragment of INSL5 precursor was confirmed by DNA sequencing.

The expression construct (pET/INSL5) was transformed into *E. coli* strain BL21(DE3) star. The transformed *E. coli* cells were cultured in liquid LB medium (with 100 μg/ml ampicillin) to OD600 nm = 1.0 at 37°C with vigorous shaking (250 rpm). Then IPTG stock solution was added to the final concentration of 1.0 mM, and the cells were continuously cultured at 37°C for 4 h with gentle shaking (120 rpm). The *E. coli* cells were harvested by centrifugation (5,000*g*, 5 min), re-suspended in lysate buffer (20 mM phosphate, pH 7.5, 0.5 M NaCl), and lysed by sonication. After centrifugation (10,000*g*, 15 min), the pellet was solubilized by the solubilization buffer (lysate buffer plus 8 M urea). After centrifugation (10,000*g*, 15 min), the supernatant was loaded onto a Ni^{2+} column that was pre-equilibrated with the solubilization buffer. The INSL5 precursor was then eluted from the column by stepwise increase of imidazole concentration in the solubilization buffer. The eluted fraction of INSL5 precursor was manually collected, and directly used for in vitro refolding.

In vitro refolding of the single-chain INSL5 precursors

Dithiothreitol (DTT) stock solution was added to the above eluted INSL5 precursor fraction (peptide concentration was 0.5–1.0 mg/ml) to a final concentration of 20 mM. The reduction reaction was carried out at room temperature for 1 h. Then, the reduced sample was 10-fold diluted into the pre-warmed refolding solution (0.5 M L-arginine–HCl, 1.0 mM EDTA, and 8.0 mM GSSG) with different pH values. The refolding reaction was carried out at different temperatures (0, 8, or 16°C) for 1–2 h. The optimal refolding pH was 10.0 and the optimal temperature was 0°C. Subsequently, the refolding mixture was acidified to pH 2.0 by TFA, centrifuged (12,000*g*, 10 min), and analyzed by C18 RP-HPLC. The peptide was eluted from the C18 reverse-phase column by an acetonitrile gradient composed of solvent A and solvent B. Solvent A was 0.1% aqueous TFA, and solvent B was acetonitrile containing 0.1% TFA. The elution gradient was listed as follows: 0 min, 10% solvent B; 3 min, 10% solvent B; 53 min, 60% solvent B; 55 min, 100% solvent B; 56 min, 100% solvent B; 60 min, 10% solvent B. The flow rate for analytical column (Zorbax 300SB-C18, 4.6 mm × 250 mm) was 0.5 ml/min, and that for semi-preparative column (Zorbax 300SB-C18, 9.4 × 250 mm) was 1.0 ml/min. The elution gradient was also used for other experiments of this work. The eluted peptide was detected by UV absorbance both at 280 nm and at 230 nm. The eluted fractions were manually collected and lyophilized. The molecular mass of INSL5 precursor was measured by electro-spray mass spectrometry.

Enzymatic conversion of INSL5 precursor into two-chain human INSL5s

The refolded INSL5 precursor was dissolved in 50 mM Tris–Cl, 1.0 mM EDTA, and 4.0 M urea (pH 8.3) solution at a final concentration of 1–2 mg/ml. Then endoproteinase Lys-C (from Sigma–Aldrich, St. Louis, USA) was added (one unit of enzyme per ~5 mg of INSL5 precursor), and the digestion was carried out at 30°C overnight. After digestion, an aliquot (~2 μg) was removed and analyzed by C18 RP-HPLC. The eluted peaks were manually collected and their molecular masses were measured by electro-spray mass spectrometry. Thereafter, carboxypeptidase B (from Sigma-Aldrich, St. Louis, USA) was added (enzyme: peptide mass ratio 1:30) to remove the additional lysine residue at the C-terminal of B-chain. The reaction was carried out at 30°C for 1 h. The two-chain INSL5 with an extended N-terminus of the A-chain (designated as long-INSL5) was purified by C18 RP-HPLC and lyophilized. The long-INSL5 was then dissolved in 100 mM Tris–HCl (pH 8.5) buffer and treated by *Aeromonas* aminopeptidase (from Sigma-Aldrich, St. Louis, USA) (5 units of enzyme per ~1 mg of long-INSL5). The digestion was carried out at 37°C overnight. The resultant two-chain native INSL5 with a shorter N-terminus of the A-chain (designated as short-INSL5) was purified by C18 RP-HPLC and lyophilized. The molecular masses of double-chain INSL5s were measured by electrospray mass spectrometry.

Peptide mapping of long-INSL5

Long-INSL5 was dissolved in the digestion buffer (100 mM Tris–HCl, 10 mM $CaCl_2$, pH 7.8), then chymotrypsin (sequencing grade from Sigma-Aldrich, St. Louis, USA) stock solution was added (the mass ratio of enzyme to peptide was 1:50). The digestion was carried out at 37°C. At different reaction times, aliquot (~4 μg) was removed and analyzed by C18 RP-HPLC. The peptide fragments were eluted from the C18 RP-HPLC column by an acetonitrile gradient, the eluted peaks were manually collected, lyophilized, and analyzed by mass spectrometry. The lyophilized peak III and IV were also dissolved in a small volume of solvent A (0.1% aqueous TFA), respectively. Then, TCEP stock solution was added to the final concentration of 50 mM in order to reduce the disulfide bonds. After reduction, the mixture was analyzed by C18 RP-HPLC. The peptide fragments were eluted by an acetonitrile gradient. The eluted peaks were manually collected, lyophilized, and analyzed by electrospray mass spectrometry.

Circular dichroism spectroscopic analyses

The two-chain INSL5s were dissolved in 20 mM phosphate buffer (pH 7.4), while single-chain INSL5s were dissolved in 1.0 mM aqueous HCl due to their low solubility in neutral buffer. Their concentrations were determined by UV absorbance at 280 nm using the extinction coefficient of $\varepsilon_{280\text{nm}} = 8{,}480\ \text{M}^{-1}\ \text{cm}^{-1}$ that was calculated from the number of tryptophan and tyrosine residues in INSL5s. Their final concentrations were adjusted to 20 μM for circular dichroism (CD) spectroscopic measurement that was performed on a Jasco-715 CD spectrometer at room temperature. The spectra were scanned from 250 to 190 nm with a cell of 0.1-cm path length. The software J-700 for Windows Secondary Structural Estimation (Version 1.10.00) was used for secondary structural content evaluation from CD spectra.

cAMP activity assay

The cAMP activity assay using CHO cells transiently transfected with human RXFP4 and a pCRE-reporter gene construct was performed as previously described (Liu et al. 2005a). The data were analyzed using Graphpad Prism 4 and are the mean ± SEM of at least three independent assays performed in triplicate.

Results

Gene construction, expression, and purification of the recombinant INSL5 precursors

To prepare human INSL5 through recombinant expression, we first designed a single-chain INSL5 precursor (precursor 1) (Fig. 1a) in which the B- and A-chains of INSL5 were connected by a long linker peptide (15 residues) that is rich in small hydrophilic glycines and serines. Unfortunately, the precursor could not be efficiently expressed in *E. coli* cells. We therefore constructed another precursor (precursor 2) that contained a longer N-terminus of the A-chain (Fig. 1a) and a shorter linker sequence (9 residues). This precursor expressed well in *E. coli* cells. To obtain INSL5 with a short, native N-terminus of the A-chain, we constructed precursor 3 (Fig. 1a) in which a lysine residue was inserted after the sixth residue of the A-chain. This precursor could also be expressed well in *E. coli* cells. In these precursors, a 6× His tag was fused to the N-terminus of the B-chain in order to facilitate purification. These single-chain INSL5 precursors were expected to be readily converted to two-chain human INSL5s by endoproteinase Lys-C and carboxypeptidase B treatment after refolding. The genes of these INSL5 precursors were constructed from four chemically synthesized oligonucleotide primers, respectively (Fig. 1b), and subsequently ligated into a pET expression vector that carries a 6× His tag. The *E. coli* biased codons were used in order to improve the expression level of these precursors. For clarity, the amino acid sequence and expected disulfide linkages of precursor 2 is shown in Fig. 1c. For interest, the amino acid sequences of INSL5s from different species are shown in Fig. 1d. The A- and B-chains of INSL5s show high sequence homology.

The INSL5 precursors were expressed in *E. coli* strain BL21(DE3) star under IPTG induction, respectively. The expression level of precursor 1 was consistently low under various culture conditions (data not shown). However, both precursors 2 and 3 had good expression levels in *E. coli* cells. As analyzed by tricine SDS–PAGE (Fig. 2a), a 9 kDa-band (indicated by a star) was significantly increased after IPTG induction. This band was expected to be the INSL5 precursor. After the *E. coli* cells were lysed by sonication, the precursor was shown to be present in the pellet (Fig. 2b). After solubilization by 8 M urea solution, the INSL5 precursor 2 was subjected to immobilized metal-ion affinity chromatography (Ni^{2+} column) as shown in Fig. 2c. As analyzed by tricine SDS–PAGE (Fig. 2d), the precursor 2 was eluted from the Ni^{2+} column by 250 mM imidazole and was almost homogeneous. The purification data of the precursor 3 were similar to those of precursor 2.

In vitro refolding of the single-chain INSL5 precursors

The in vitro refolding of INSL5 precursors was difficult because they were prone to aggregation during this step. Various refolding conditions were examined, and it was found that arginine buffer at a high pH and low temperature was essential for efficient refolding of both INSL5 precursors 2 and 3. The above INSL5 precursor fraction that was eluted from the Ni^{2+} column was first analyzed by RP-HPLC. As shown in Fig. 3a (the trace of before refolding), the fraction was a mixture of disulfide isomers of the single-chain INSL5. The sample was then treated with DTT in order to fully reduce the disulfide bonds. Unfortunately, when the fully reduced precursor was subjected to subsequent RP-HPLC analysis, almost no peptide could be eluted from the C18 column (data not shown). The reduced sample was then diluted 10-fold into the refolding buffer and, after refolding, a peptide product (indicated by a star) could be eluted from the column as shown in Fig. 3a (the trace of after refolding). Mass spectrometry analysis showed its measured molecular mass (7,700.0) corresponded well to the expected value (7,700.8) of the refolded INSL5 precursor 2. It was thus assumed that this peak was the refolded INSL5 precursor and it was used for subsequent studies. The eluted peak was manually

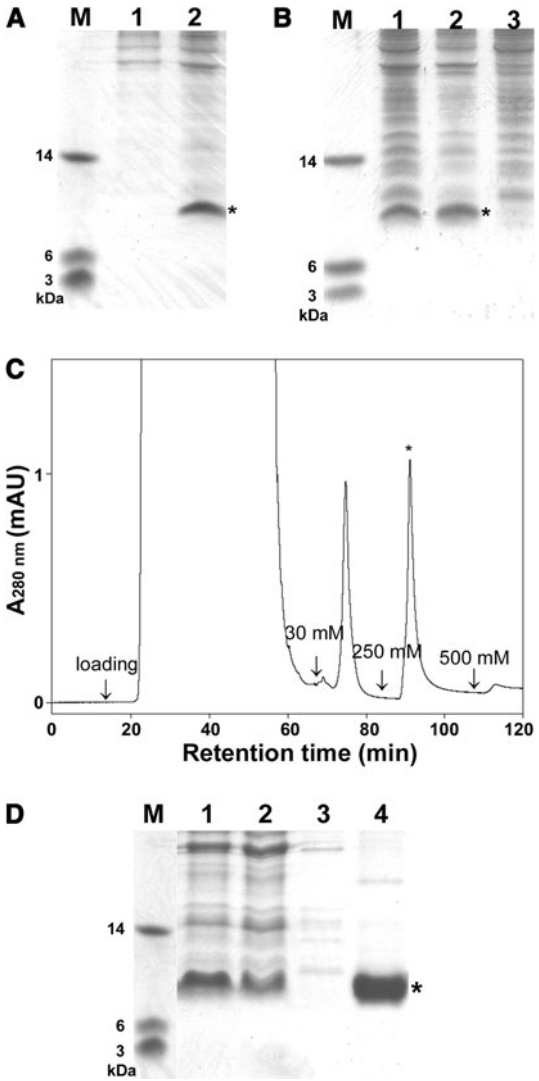

Fig. 2 Expression and purification of INSL5 precursor 2. **a** Analysis of the expression of INSL5 precursor 2 by tricine SDS–PAGE. A 30 µl of culture broth before and after IPTG induction was centrifuged, and the pellet was re-suspended in 15 µl of water, and then mixed with 5 µl of loading buffer that containing 100 mM DTT. After boiling, the sample was loaded onto a 16.5% tricine SDS-gel. After electrophoresis, the gel was stained by Coomassie Brilliant Blue R250. *Lane 1* before induction; *lane 2* after induction. **b** Tricine SDS–PAGE analysis after sonication. The *E. coli* cells expressing INSL5 precursor 2 were lysed by sonication. The total cell lysate (*lane 1*), the pellet (*lane 2*), and the supernatant (*lane 3*) were loaded onto a 16.5% tricine SDS-gel, respectively. **c** Purification of INSL5 precursor 2 by immobilized metal-ion affinity chromatography. The pellet of the cell lysate was solubilized by the solubilization buffer (20 mM phosphate, pH 7.5, 0.5 M NaCl, 8.0 M urea). After centrifugation (10,000*g*, 10 min), the supernatant was loaded onto a Ni^{2+} column (1 cm × 4 cm) and eluted by step-wise increase of imidazole concentration in the solubilization buffer. The eluted peak of INSL5 precursor 2 was indicated by a star. **d** Tricine SDS–PAGE analysis after immobilized metal-ion affinity chromatography. *Lane 1* before loading; *lane 2* flow-through; *lane 3* eluted by 30 mM imidazole; *lane 4* eluted by 250 mM imidazole. The band of INSL5 precursor 2 was indicated by a *star*

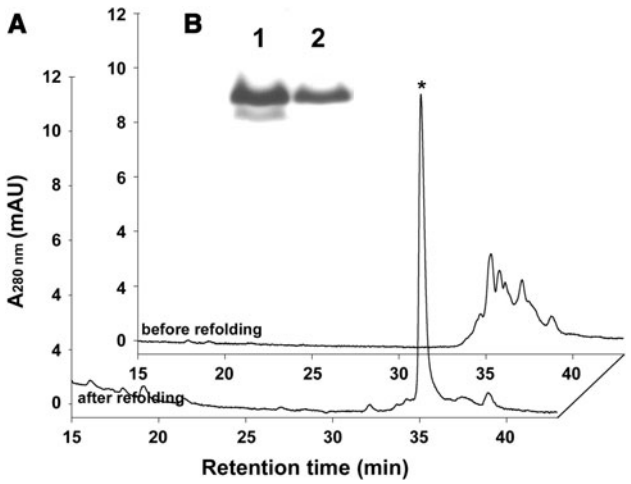

Fig. 3 In vitro refolding of INSL5 precursor 2. **a** Analysis of the INSL5 precursor 2 before and after refolding by RP-HPLC. For the 'before refolding' trace, 8 µl (~4 µg) of the INSL5 precursor fraction eluted from Ni^{2+} column was directed loaded onto the column. For the 'after refolding' trace, 8 µl (~4 µg) of the INSL5 precursor fraction was first mixed with 2 µl of 100 mM DTT solution and incubated at room temperature for 1 h. Subsequently, 90 µl of ice-cold refolding solution (0.5 M L-arginine-HCl, 1.0 mM EDTA, pH 10.0, and 8.0 mM GSSG) was added to initiate refolding. The refolding was carried out on ice for 1 h, then the folding mixture was acidified by TFA, centrifuged, and loaded onto a C18 RP-HPLC column. The eluted folded INSL5 precursor was manually collected, lyophilized, and used for refolding yield analysis in **b**. **b** Analysis of the refolding yield of INSL5 precursor 2 by tricine SDS–PAGE. *Lane 1* 8 µl (~4 µg) of the INSL5 precursor fraction eluted from Ni^{2+} column was mixed with 8 µl of loading buffer that contains 100 mM DTT, then 8 µl of the mixture was loaded on the gel. *Lane 2* the lyophilized refolded INSL5 from **a** was dissolved in 8 µl of water, mixed with 8 µl of loading buffer that contains 100 mM DTT, and then 8 µl of the mixture was loaded onto the gel. The gel was stained by Coomassie Brilliant Blue R250. The band density was analyzed by software Scion Image after scanning

collected, lyophilized, and analyzed by tricine SDS–PAGE as shown in Fig. 3b. As analyzed from the band density before and after refolding, the refolding yield of INSL5 precursor 2 was approximately 60% under optimized refolding condition (0.5 M L-arginine–HCl, pH 10.0, 1.0 mM EDTA, 8.0 mM GSSG, 2.0 mM DTT, at 0°C for 1–2 h). The refolding behavior of INSL5 precursor 3 was similar to that of precursor 2.

Enzymatic conversion of INSL5 precursor into two-chain INSL5s

To prepare two-chain human INSL5, the refolded precursor 2 was first treated with endoproteinase Lys-C that can cleave the peptide bond at the C-terminal side of lysine residues. The digestion mixture was analyzed by RP-HPLC as shown in Fig. 4a. The measured molecular mass of the major peak was 5,759.0, consistent with the expected

molecular mass (5,759.7) of the intermediate that carries an additional lysine at the C-terminal of B-chain. Subsequently, carboxypeptidase B was added to remove this extra lysine residue. The digestion mixture was analyzed by reverse-phase HPLC as shown in Fig. 4b. The measured molecular mass of the major peak was 5,632.0, consistent with the expected value (5,631.5) of two-chain INSL5 that has an extended N-terminus of A-chain (designated as long-INSL5).

The refolded INSL5 precursor 3 was also subjected to endoproteinase Lys-C digestion. Unfortunately, the enzyme could not cleave at the junction site between A-chain and the linker peptide. This failure was attributed to steric hindrance. For this reason, *Aeromonas* aminopeptidase was assessed for its ability to remove the 6 extended N-terminal residues of A-chain in long-INSL5 (Fig. 4c). The major peak had a measured molecular mass of 4,934.0, consistent with the expected value (4,933.7) of the two-chain INSL5 with a short, native N-terminus of A-chain (designated as short-INSL5). As *Aeromonas* aminopeptidase cannot remove glutamate residue, it did not digest the B-chain, the first N-terminal residue of which is a glutamate.

Peptide mapping of long-INSL5

To determine the disposition of the disulfide linkages, long-INSL5 was subjected to peptide mapping. V8 endoproteinase that can cleave the peptide bond at the carboxyl side of both Glu and Asp was evaluated. Unfortunately, the enzyme could not cleave at the carboxyl side of A22D probably due to steric hindrance. Chymotrypsin that can cleave the peptide bond at the carboxyl side of large hydrophobic resides, such as Trp, Tyr, Phe, Met, Leu, Ile, etc., was then tried. As shown in Fig. 5a, chymotrypsin cleavage resulted in 4 major peaks (labeled as I-IV) on RP-HPLC. The peak I had a measured molecular mass of 472.2, corresponding to the N-terminal fragment of A-chain (Table 1). The peak II had two major components. One had a measured molecular mass of 959.4, corresponding to an octapeptide of A-chain, while the other had a measured value of 602.4, corresponding to a pentapeptide of the B-chain N-terminal (Table 1).

Mass spectrometry analysis (Table 1) showed that peak III contained two major components: one had a measured molecular mass of 1,212.4, and other had a measured value of 763.4. The former was deduced to be the fragment containing disulfide A27-B18; the latter was deduced to be a middle fragment of B-chain (Table 1). After treatment with TCEP (to break the disulfide bond), peak III was converted to three major peaks (labeled as peak III-1, III-2, and III-3) as shown in Fig. 5b. The peak III-1 had a measured molecular mass of 392.2, corresponding to the

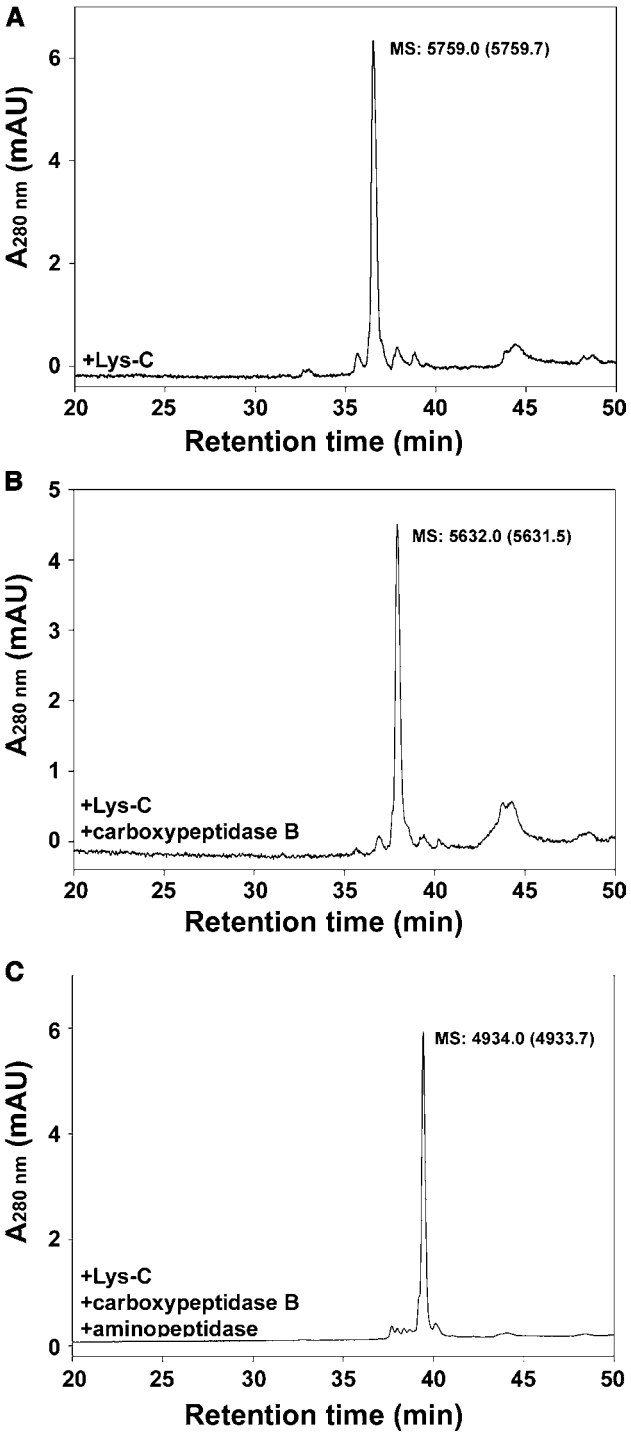

Fig. 4 Enzymatic conversion of INSL5 precursor 2 to two-chain long-INSL5 and short-INSL5. **a** C18 RP-HPLC of endoproteinase Lys-C digested INSL5 precursor 2. **b** C18 RP-HPLC of INSL5 precursor 2 sequentially digested by endoproteinase Lys-C and carboxypeptidase B. **c** C18 RP-HPLC of long-INSL5 treated by *Aeromonas* aminopeptidase. A ~2 μg of digestion mixture was loaded onto a C18 reverse-phase column, and eluted by an acetonitrile gradient. The major peak was manually collected, lyophilized, and its molecular mass (MS) was measured by electro-spray mass spectrometry as shown in **a**, **b**, and **c**, the theoretical values were shown in parenthesis

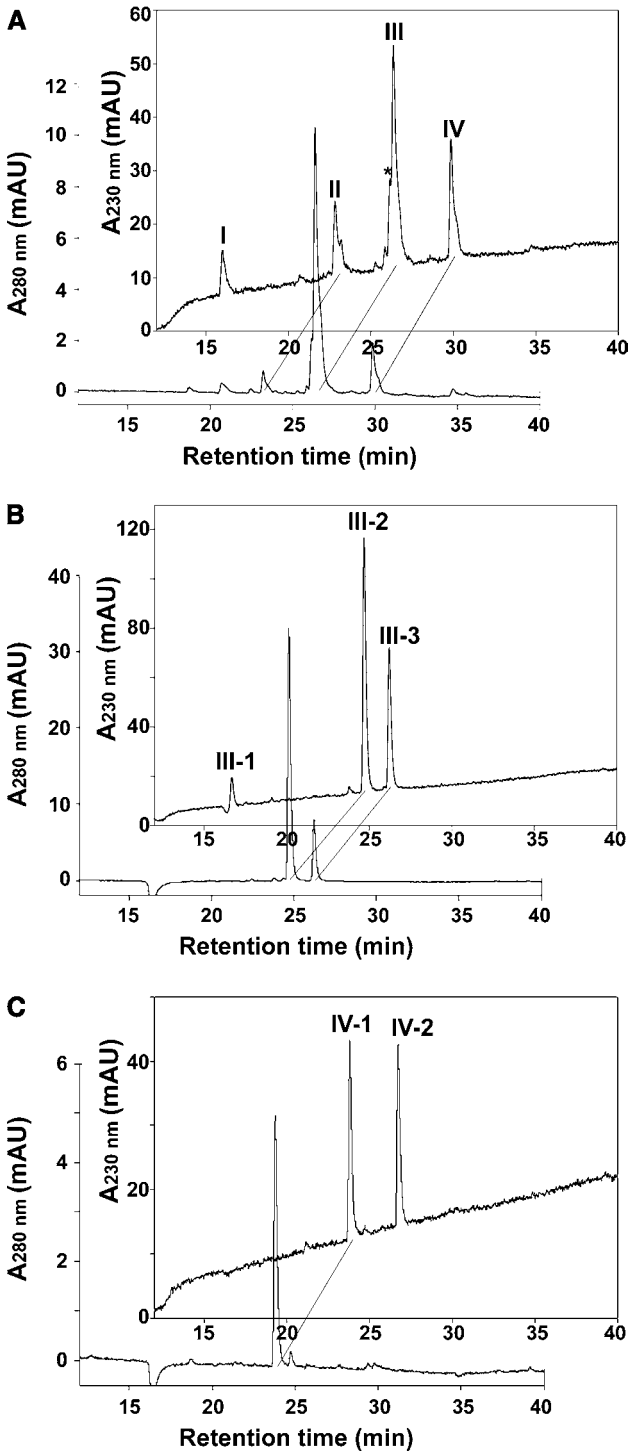

Fig. 5 Peptide mapping of long-INSL5. **a** C18 RP-HPLC of long-INSL5 digested by chymotrypsin at 37°C for 1 h. **b** C18 RP-HPLC of the TCEP-treated peak III in **a**. **c** C18 RP-HPLC of the TCEP-treated peak IV in **a**. The INSL5 fragments were eluted from C18 reverse-phase column by an acetonitrile gradient as listed in "Materials". The eluted peaks were monitored by UV absorbance at both 230 and 280 nm

C-terminal fragment of A-chain (Table 1). This peak had no absorbance at 280 nm due to lack of Trp and Tyr residues. The peak III-2 had a measured molecular mass of 821.3, corresponding to the fragment of C-terminal of B-chain (Table 1). The peak III-2 had strong absorbance at 280 nm due to a Trp residue (Fig. 5b). The peak III-3, the position of which on RP-HPLC was identical to the shoulder (indicated by a star) of peak III, had a measured molecular mass of 763.4, corresponding to the fragment of central B-chain (Table 1). This fragment had moderate absorbance at 280 nm due to presence of a Tyr residue.

Mass spectrometry analysis showed that peak IV (Fig. 5a) had a measured molecular mass of 1,728.2 and was deduced to be the fragment containing two disulfide bonds (Table 1). After treatment with TCEP, the peak IV was converted to two peaks (labeled as peak IV-1 and IV-2) as shown in Fig. 5c. Based on their molecular mass, the peak IV-1 was deduced to be a fragment of B-chain, while the peak IV-2 was deduced to be a fragment of A-chain (Table 1). The peak IV-1 contained a Tyr residue, so it had absorbance at 280 nm. In contrast, peak IV-2 had no absorbance at 280 nm due to lack of Trp and Tyr.

In summary, peptide mapping showed long-INSL5 contained disulfide A27-B18, a critical disulfide bond that tethers the C-terminal of A-chain and the C-terminal of B-chain. The results also showed that long-INSL5 contained another inter-chain disulfide bond, probably A14-B6, as well as an intra A-chain disulfide bond, probably A13-A17.

Circular dichroism spectroscopic study

The secondary structure of the single-chain and two-chain INSL5s was analyzed by circular dichroism (CD) spectroscopy. As shown in Fig. 6a, both long-INSL5 and short-INSL5 had similar CD spectra to human relaxin-3. Their estimated α-helix contents were 42% (long-INSL5) and 44% (short-INSL5), respectively, which is similar to the estimated α-helix content of relaxin-3 (35%). The single-chain INSL5 precursors had a low solubility in neutral phosphate buffer, so their CD spectra were recorded in acidic solvent (1.0 mM HCl) as shown in Fig. 6b. Precursors 2 and 3 had almost identical CD spectra, and their estimated α-helix content was about 20%.

Functional cAMP assay

The recombinant INSL5s were tested for their ability to inhibit forskolin-induced cAMP activity in RXFP4 transfected cells. As shown in Fig. 7, both long-INSL5 (pEC50 = 8.86 ± 0.06; $n = 3$) and short-INSL5 (pEC50 = 9.05 ± 0.02;

Table 1 Molecular masses of the chymotrypsin digested peptide fragments of long-INSL5

Peak no. in Fig. 5	Measured value	Theoretical value	Corresponding fragment
I	472.2	472.6	HSVM
II	959.4	960.0	SRQDLQTL
	602.4	602.7	ESVRL
III	1,212.4	1,212.4	SALC–ICASSRW
	763.4	763.9	IRTVIY
IV	1,728.2	1,728.0	CCTDGCSMTDL–CGLEY
III-1	392.2	392.5	SALC
III-2	821.3	821.9	ICASSRW
III-3	763.4	763.9	IRTVIY
IV-1	583.4	583.7	CGLEY
IV-2	1,148.1	1,148.3	CCTDGCSMTDL

$n = 3$) had full activity compared with the chemically synthesized INSL5 (pEC50 = 8.61 ± 0.07; $n = 5$), suggesting the recombinant two-chain INSL5s were folded correctly. However, the single-chain INSL5 precursor showed no activity, suggesting a free C-terminal of B-chain is critical for receptor-binding and activation of INSL5.

Discussion

In the present work, we established an efficient approach for the preparation of INSL5 peptide through recombinant expression in *E. coli* cells. The recombinant two-chain INSL5 had full activity compared with chemically synthesized INSL5 confirming the efficiency of this recombinant approach. This, in turn, will provide a mean for the preparation of numerous INSL5 analogs through site-directed mutagenesis and subsequent recombinant expression.

The in vitro refolding of the single-chain of INSL5 was difficult due to the propensity of the peptide to aggregate. A similar observation was made during the chemical synthesis of this peptide (Hossain et al. 2008a). In Tris buffer, almost no folded INSL5 precursor could be obtained. However, in arginine buffer that facilitates protein folding as previously reported (Lange and Rudolph 2009; Nakakido et al. 2009; Tsumoto et al. 2004), folded peptide could be obtained. After optimization of the refolding conditions, the yield of correctly folded INSL5 precursor reached about 60%. After purification and in vitro refolding, 1–2 mg of refolded and purified INSL5 precursor typically could be obtained from 1 l of the culture broth. If the *E. coli* cells are cultured in richer medium, where higher cell densities can be obtained, it is expected that the expression level of INSL5 precursor will be further improved.

The solubility of the folded INSL5 precursor was low in neutral or slightly basic solutions. During endoproteinase Lys-C digestion, urea was added to increase its solubility. However, significant aggregation occured after overnight digestion at 30°C at high peptide concentration (greater than 2 mg/ml). The aggregate could not be re-solubilized by changing the pH of the solution. Aggregation could cause ~80% peptide loss. When mixed with thioflavin T, a reagent that specifically binds to protein fibrils, no fluorescence increase was observed suggesting that the aggregation of INSL5 was not caused by peptide fibrillation. At lower peptide concentration (about 1 mg/ml), the aggregation could be prevented. Under these conditions, ~0.5 mg of two-chain INSL5 could be typically obtained from 1.0 mg of precursor.

Both long- and short-INSL5 were equipotent to chemically assembled INSL5 in activating the RXFP4 receptor. This showed that extension of the N-terminus of the A-chain had no effect on either the binding or activation of the INSL5 receptor. However, the single-chain INSL5 precursor was inactive which suggested that a free C-terminus of the B-chain is crucial for the activity of INSL5. These results are consistent with the activity of relaxin-3 on the RXFP3 receptor whereby the length of the N-terminus of the A-chain is not critical for activity (Hossain et al. 2008b; Kuei et al. 2007) and that addition of residues at the C-terminus after the Trp result in a dramatic loss of activity (Kuei et al. 2007). The Trp residue at the C-terminal of the B-chain is therefore very likely to be critical for INSL5 activity on RXFP4 and is invariant in

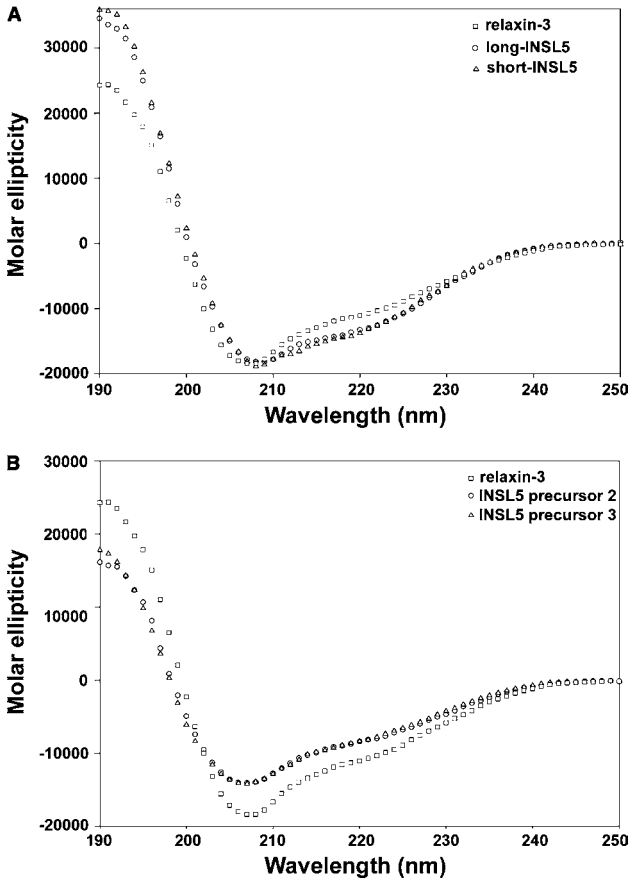

Fig. 6 CD spectroscopic study. **a** Far-UV spectra of two-chain INSL5s. Samples were dissolved in 20 mM phosphate buffer (pH 7.4), respectively. **b** Far-UV spectra of single-chain INSL5 precursors 2 and 3. Samples were dissolved in 1.0 mM aqueous HCl, respectively

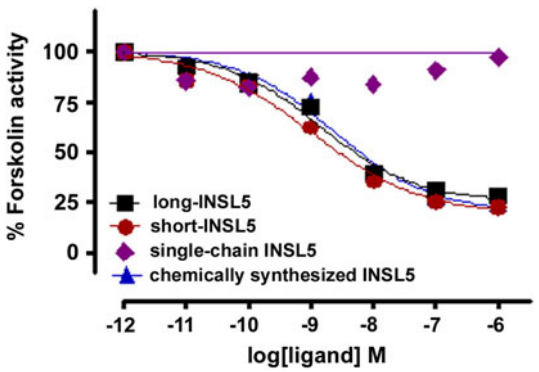

Fig. 7 Inhibition of forskolin (1 μM) induced cAMP activity in RXFP4 transfected cells. The activity of single-chain and two-chain INSL5s was compared to synthetic INSL5. The values are expressed as mean ± SEM ($n = 3$) of three assays performed in triplicate

INSL5s from different species (Fig. 1d). This work suggests that INSL5 may interact with RXFP4 in a similar manner to relaxin-3 interacting with RXFP3. In future work, we will prepare a series of INSL5 analogs through site-directed mutagenesis and recombinant expression in order to reveal the critical residues for INSL5 activity on RXFP4.

Acknowledgments This work was supported by the National Basic Research Program of China (973 Program, no. 2010CB912604) and by the National Natural Science Foundation of China (30970609, 30700124). The studies at the HFI (Melbourne) were supported in part by NHMRC project grants (# 508995 and 509048) to JDW and RADB. We thank Linda Chan (FNI) for the amino acid analyses and Sharon Layfield (FNI) for the bioassays.

References

Conklin D, Lofton-Day CE, Haldeman BA, Ching A, Whitmore TE, Lok S, Jaspers S (1999) Identification of INSL5, a new member of the insulin superfamily. Genomics 60:50–56

Haugaard-Jönsson LM, Hossain MA, Daly NL, Bathgate RA, Wade JD, Craik DJ, Rosengren KJ (2008) Structure of the R3/I5 chimeric relaxin peptide, a selective GPCR135 and GPCR142 agonist. J Biol Chem 283:23811–23828

Haugaard-Jönsson LM, Hossain MA, Daly NL, Craik DJ, Wade JD, Rosengren KJ (2009) Structure of human insulin-like peptide 5 and characterization of conserved hydrogen bonds and electrostatic interactions within the relaxin framework. Biochem J 419:619–627

Hossain MA, Bathgate RA, Kong CK, Shabanpoor F, Zhang S, Haugaard-Jönsson LM, Rosengren KJ, Tregear GW, Wade JD (2008a) Synthesis, conformation, and activity of human insulin-like peptide 5 (INSL5). ChemBioChem 9:1816–1822

Hossain MA, Rosengren KJ, Haugaard-Jonsson LM, Zhang S, Layfield S, Ferraro T, Daly NL, Tregear GW, Wade JD, Bathgate RA (2008b) The A-chain of human relaxin family peptides has distinct roles in the binding and activation of the different relaxin family peptide receptors. J Biol Chem 283:17287–17297

Hsu SY (1999) Cloning of two novel mammalian paralogs of relaxin/insulin family proteins and their expression in testis and kidney. Mol Endocrinol 13:2163–2174

Kuei C, Sutton S, Bonaventure P, Pudiak C, Shelton J, Zhu J, Nepomuceno D, Wu J, Chen J, Kamme F, Seierstad M, Hack MD, Bathgate RA, Hossain MA, Wade JD, Atack J, Lovenberg TW, Liu C (2007) R3(BDelta23 27)R/I5 chimeric peptide, a selective antagonist for GPCR135 and GPCR142 over relaxin receptor LGR7: in vitro and in vivo characterization. J Biol Chem 282:25425–25435

Lange C, Rudolph R (2009) Suppression of protein aggregation by L-arginine. Curr Pharm Biotechnol 10:408–414

Liu C, Chen J, Sutton S, Roland B, Kuei C, Farmer N, Sillard R, Lovenberg TW (2003) Identification of relaxin-3/INSL7 as a ligand for GPCR142. J Biol Chem 278:50765–50770

Liu C, Chen J, Kuei C, Sutton S, Nepomuceno D, Bonaventure P, Lovenberg TW (2005a) Relaxin-3/insulin-like peptide 5 chimeric peptide, a selective ligand for G protein-coupled receptor (GPCR)135 and GPCR142 over leucine-rich repeat-containing G protein-coupled receptor 7. Mol Pharmacol 67:231–240

Liu C, Kuei C, Sutton S, Chen J, Bonaventure P, Wu J, Nepomuceno D, Kamme F, Tran DT, Zhu J, Wilkinson T, Bathgate R, Eriste E, Sillard R, Lovenberg TW (2005b) INSL5 is a high affinity specific agonist for GPCR142 (GPR100). J Biol Chem 280:292–300

Nakakido M, Kudou M, Arakawa T, Tsumoto K (2009) To be excluded or to bind, that is the question: arginine effects on proteins. Curr Pharm Biotechnol 10:415–420

Shabanpoor F, Separovic F, Wade JD (2009) The human insulin superfamily of polypeptide hormones. Vitam Horm 80:1–31

Shirneshan K (2005) Expression, functional and regulation analysis of selected genes from the insulin family. Ph. D. dissertation, University of Gottingen (Germany), http://webdoc.sub.gwdg.de/diss/2006/shirneshan/

Tsumoto K, Umetsu M, Kumagai I, Ejima D, Philo JS, Arakawa T (2004) Role of arginine in protein refolding, solubilization, and purification. Biotechnol Prog 20:1301–1308

Zhu J, Kuei C, Sutton S, Kamme F, Yu J, Bonaventure P, Atack J, Lovenberg TW, Liu C (2008) Identification of the domains in RXFP4 (GPCR142) responsible for the high affinity binding and agonistic activity of INSL5 at RXFP4 compared to RXFP3 (GPCR135). Eur J Pharmacol 590:43–52