PRELIMINARY STRUCTURE-FUNCTION RELATIONSHIP STUDIES ON INSULIN-LIKE PEPTIDE 5 (INSL5)

Alessia Belgi[a,b], Ross A. D. Bathgate[a,b], Geoffrey W. Tregear[a,b], John D. Wade[a,c] and Mohammed Akhter Hossain[a,c*]

[a]Florey Neuroscience Institutes, [b]Department of Biochemistry and Molecular Biology, [c]School of Chemistry, The University of Melbourne, Victoria 3010, Australia.

*Corresponding author:

Dr Mohammed A. Hossain

Florey Neuroscience Institutes

University of Melbourne

Victoria 3010, Australia

Ph: +61 3 83447285 / Fax: +61 3 93481707

E-mail: akhter.hossain@florey.edu.au

Keywords**:** insulin-like peptide 5, INSL5, relaxin, solid phase peptide synthesis, RXFP4, RXFP3

ABSTRACT

Insulin-like peptide 5 (INSL5) is a two-chain, three-disulfide bonded member of insulin/relaxin superfamily of peptides that includes insulin, insulin-like growth factor I and II (IGFI and IGFII), insulin-like peptide 3, 4, 5 and 6 (INSL3, 4, 5 and 6), relaxin-1 (H1 relaxin), -2 (H2 relaxin) and -3 (H3 relaxin). Although it is expressed in relatively high levels in the gut, its biological function remains unclear. However, recent reports suggest a significant orexigenic action and a role in the regulation of insulin secretion and β-cell homeostasis, which implies that both agonists and antagonists of the peptide may have significant therapeutic applications. Modern solid phase synthesis techniques together with regioselective disulfide bond formation were employed for a preliminary structure-function relationship study of mouse INSL5. Two point mutated analogues, mouse INSL5 A-B(R24A, W25A) and mouse INSL5 A-B(K6A, R14A, Y18A) were chemically prepared, where the residues in the B-chain that may be involved in receptor activation and affinity binding, were respectively mutated. Synthetic mouse INSL5 A-B(R24A, W25A) analogue was inactive on RXFP4, the native receptor for INSL5, suggesting Arg^{B24} and Trp^{B25} are probably directly involved in INSL5 receptor activation. Mouse INSL5 A-B(K6A, R14A, Y18A) analogue had both decreased affinity and potency on RXFP4 (pIC_{50} 7.7±0.2, pEC_{50} 7.87±0.18) which indicated that one or both of these residues are critical for the binding to the receptor.

INTRODUCTION

Insulin-like peptide 5 (INSL5) is a peptide hormone of the insulin superfamily that comprises insulin, insulin-like growth factors I and II (IGFI-II), relaxins 1-3 and insulin-like peptides 3-6 (Shabanpoor et al., 2009). It was first identified in 1999 from the screening of expressed sequence tags databases (ETS) (Conklin et al., 1999) and it contains all the common/characteristic features of the peptides of the insulin superfamily: it is constituted by two chains, A- and B-chain, it has the conserved cysteine motif whereby two interchain disulfide bonds take place between the A- and the B-chain and one is intramolecular within the A-chain (Conklin et al., 1999).

In 2005, INSL5 was identified as the specific agonist for the G-protein coupled receptor GPCR142 (Liu et al., 2005) now referred to as relaxin family peptide receptor 4 (RXFP4) according to recent nomenclature (Bathgate et al., 2006). INSL5 and RXFP4 are found to be expressed in the same tissues such as colon, prostate, kidney, placenta and other peripheral tissues (Conklin et al., 1999, Liu et al., 2005, Liu and Lovenberg, 2008) with the highest level in the colon (Liu et al., 2005). Though physiological role of INSL5 is not yet clear, this expression pattern of both INSL5 and its cognate RXFP4 receptor lead to the proposition that INSL5 plays an active role in the stimulation of appetite and in colon motility (Takeda Cambridge Ltd, 2009). The results from most recent study suggest that INSL5 is involved in the regulation of insulin secretion and β-cell homeostasis (Burnicka-Turek et al., 2012). INSL5 and its agonists have therefore the potential to become therapeutics for the management of eating disorders and/or diabetes. Conversely, an INSL5 antagonist would possess the potential to treat obesity either alone or in conjunction with other known feeding regulators.

To better understand the physiological role of INSL5, prior knowledge of the interaction between the peptide and its cognate receptor at a molecular level is necessary. This can be achieved via the design and chemical synthesis of point-mutated analogues to identify the residues within the peptide responsible for the binding to the receptor and its subsequent activation of downstream intracellular pathways. In theory, such studies are technically easier than studies using mutagenesis of the much larger, more complicated, membrane-bound RXFP4 receptor to determine the molecular interactions that take place between the peptide and receptor.

INSL5 is the most recent member of the insulin/relaxin superfamily for which its receptor has been identified and limited information is currently available regarding its mechanism of interaction with RXFP4 receptor (Belgi et al., 2011, Luo et al., 2010). Given the complexity of INSL5 ligand, sequential mutation of the residues from one terminus to the other of each A- and B-chain is achievable but also arduous given the enormous synthesis work that would be required. Such studies may be somewhat simplified by the use of assumptions based on prior studies of a related ligand/receptor pair that has been well-characterized and that is thought to interact on a similar fashion. In this respect, the relaxin-3/RXFP3 system has been very well characterised by us and our collaborators at Johnson & Johnson Pharmaceuticals (USA). Relaxin-3 and INSL5 are closely related as they belong to the same branch in the phylogenetic tree of the relaxin family (Wilkinson et al., 2005) and their cognate receptors, RXFP3 and RXFP4 respectively, belong to the same class of G-protein coupled receptors (Bathgate et al., 2006). INSL5 is therefore assumed to interact with RXFP4 receptor in a similar fashion to relaxin-3 binding to and activation of RXFP3 receptor (Haugaard-Jonsson et al., 2009).

The residues within relaxin-3 involved in the binding and activation of RXFP3 have largely been identified and they are all within relaxin-3 B-chain (Kuei et al., 2007). Moreover, all but one of these same residues are also used by relaxin-3 to bind and activate RXFP4 receptor (Figure 1) (Kuei et al., 2007). The corresponding residues within mouse INSL5 B-chain were identified and two preliminary analogues, analogue 1 and analogue 2, have been designed and synthesized. We wanted to assess if for INSL5 too these residues are important for RXFP4 binding and the activation. The preliminary results indicate that INSL5 uses a mechanism similar to relaxin-3 with the C-terminal residues, Arg^{B24} and Trp^{B25}, important for activity and residues within the B-chain exposed in the α-helical part involved in the binding interaction with the receptor.

MATERIALS AND METHODS

Materials

9-Fluroenylmethoxycarbonyl (Fmoc) protected L-α-amino acids and 2-(1H-benzotriazole-1-yl)-1,1,3,3-tetramethylaminium hexafluorophosphate (HBTU) were purchased from GL Biochem (Shanghai, China). N,N-dimethylformamide (DMF), piperidine and trifluoroacetic acid (TFA) were obtained from Auspep (Melbourne, Australia). Fmoc-PAL-PEG-PS and Fmoc-L-Ala-PEG-PS resins with substitution of *ca.* 0.20 mmol/g were purchased from Applied Biosystems Inc, (Melbourne, Australia), Fmoc-Cys(Acm)-S PHB TentaGel was purchased from Rapp Polymere (Tübingen, Germany). Acetonitrile, dichloromethane, diethyl ether, and methanol, were from Merck (Melbourne, Australia), and 2-pyridyl disulfide (DPDS) was purchased from Fluka (Buchs, Switzerland). Na_2HPO_4 and $MgSO_4$ were purchased from BDH Chemicals (Sydney, Australia). KH_2PO_4 and $MnCl_2$ were purchased from AJAX

Chemicals (Sydney, Australia). $CaCl_2$ was purchased from Mallinckrodt (Melbourne, Australia). Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12), L-glutamine, penicillin/streptomycin and foetal bovine serum were all purchased from Invitrogen (Melbourne, Australia). Chlorophenol red β–D galactopyranoside was purchased from Roche (Mannheim, Germany). Diethylene triamine pentaacetic acid (DTPA)-tetra (tBu ester) was purchased from Macrocyclics (Dallas, Texas, US). All other reagents were from Sigma-Aldrich (Sydney, Australia).

Methods

Preparation of resin-bound base labile handle-linked solubilizing tagged peptide:

Fmoc-PAL-PEG-PS (0.2 mmol) was N^{α}-deprotected and then carried through five cycles of acylation with Fmoc-Lys(Boc)-OH and N^{α}-deprotection as described above. The base labile linker, 4-hydroxymethylbenzoic acid (4 equivalents), was attached via a 1 hour acylation using 4 equivalents of HBTU and 6 equivalents of diisopropylethylamine (DIEA). The C-terminal residue was attached to the resin-bound linker via a double 1 hour acylation with the Fmoc-amino acid symmetrical anhydride (5 equivalents) in the presence of 0.1 equivalents of catalyst, 4-dimethylamino pyridine (DMAP).

Peptide synthesis: The peptides were solid phase synthesized employing Fmoc-chemistry on one of the following instruments: CEM Liberty^{TM} microwave peptide synthesizer (Ai Scientific, Queensland, Australia), or Protein Technologies Tribute batch-wise peptide synthesizer (Tucson, AZ, USA). The side chain-protecting groups of trifunctional amino acids were TFA-labile, except for *tert*-butyl (tBu)-protected and acetamidomethyl (Acm)-protected cysteine Cys in the A-chain and Cys(Acm) in the B-

chain. The peptides were synthesized on a either 0.1 or 0.2 mmol scale using instrument default protocols with either a 4- or 5-fold molar excess of Fmoc-protected amino acids (0.4 or 0.5 mmol for a 0.1 mmol scale; 0.8 or 1.0 mmol for a 0.2 mmol scale) that were activated by using 4- or 5-fold excess of HBTU in the presence of DIEA (10 equivalents). N^{α}-Fmoc protecting groups were removed by treating the resin-attached peptide with piperidine (20% v/v) in DMF. Using the microwave synthesizer, the coupling and deprotection were carried out at 75 °C using 25 watt microwave power for 5 min and 60 watt microwave power for 3 min respectively. For batch synthesis (on the Tribute) the coupling and deprotection were carried out for 30 min and 10 min respectively.

Following the cleavage and purification, as described below, the individual A- and B-chains were combined to give the target two-chain peptides with the cystine framework following the regioselective disulfide bond formation strategy previously described (Hossain et al., 2008). Briefly, the crude [Cys7,12 (S-thiol), Cys8 (tBu), Cys21 (Acm)] A-chain was oxidised with 2-pyridyl disulfide (DPDS) (solution 4mM in methanol) to form the intramolecular disulfide bond Cys7-Cys12. Following isolation via RP-HPLC, the oxidised A-chain was treated with DPDS in trifluoromethanesulfonic acid/TFA (1/4 v/v) to convert Cys8 (tBu) into the more reactive pyridinyl group (Cys8(Pyr)). The A-chain was then combined with the B-chain and the intermolecular disulfide bond CysA8-CysB8 was formed. The A-B peptide [Cys21A(Acm)/Cys20B(Acm)] obtained was then treated with iodine (20mM solution in acetic acid) to simultaneously cleave the Acm protecting groups and form the last disulfide bond CysA20-CysB20.

Peptide-resin cleavage: On completion of solid phase synthesis, the peptides were cleaved from the solid support by treatment with trifluoroacetic acid (TFA) containing

anisole: triisopropylsilane (TIPS): 3, 6-dioxa-1, 8-octanedithiol (DODT): H_2O (93% : 2.5% : 2% : 1.5% : 1%, 20 ml) for 2 hours. The cleaved products were precipitated in ice-cold diethyl ether and centrifuged at 3000 rpm for 3 min. The pellets were washed by re-suspending them in ice-cold diethyl ether and centrifuging them again. This washing process was repeated at least three times.

Peptide purification: All the HPLC purifications and analytical reaction monitoring, were performed using a Vydac C18 column, either analytical or preparative (4.6 x 250 mm or 22 x 250 mm respectively), in a gradient mode with eluant A: 0.1% aq. TFA and eluant B: 0.1% TFA in acetonitrile.

General procedure for the cleavage of the solubilising tag from target peptide: To the purified synthetic solubilising tag-linked peptides was added 1 mL of ice cold 0.1M NaOH. The whole was stirred for 2-5 minutes after which the pH was made acidic by the addition of 0.1% aq. TFA. The solution was then applied directly onto a RP-HPLC column for purification as described above.

Peptide characterization

Matrix-assisted laser desorption ionization time-of-flight mass spectrometry (MALDI-TOF MS) was carried out on a Bruker Ultraflex II instrument (Bruker Daltonics, Bremen, Germany) and used to characterize the peptides at each intermediate step using sinapinic acid (3,5-dimethoxy-4-hydroxycinnamic acid) as matrix. The matrices were made up in 70% acetonitrile containing 0.1% TFA.

The peptide content was determined using vapor-phase acid hydrolysis in 6 M HCl containing 2% phenol at 110^{o} C for 24 hours. The individual amino acids were converted to stable, fluorescent derivatives using Waters AccQ.Tag kit (Waters, Sydney, Australia). The derivatized amino acids were separated using a Shim-Pak XR-

ODS (3 × 75 mm, 2.2 μm) column on Shimadzu RP-HPLC system (Shimadzu, Victoria, Australia).

Functional assays

Binding assays: Chinese hamster ovary CHO-K1 cells transfected with human RXFP4 were plated out at 50000 cells per well per 200 μL in a 96 well ViewPlate with clear bottom and white walls precoated with poly-L-lysine. Competition binding experiments were performed with 5 nM of europium-labelled Eu(A)-mouse INSL5 acid (Belgi et al., 2011) in the presence of increasing amounts of peptides following the protocol described previously (Shabanpoor et al., 2008). GraphPad PRISM 5 (GraphPad Inc., San Diego, USA) was used to analyse the receptor-binding assays data that are expressed as mean ± SEM.

Agonist assays: The peptides were tested for their ability to inhibit cAMP activity in CHO-K1 cells co-transfected with human RXFP4 and a pCRE (cAMP Response Element) β-galactosidase reporter plasmid. On day 1 the cells were plated out in a Corning cell bind 96 well plate at 8000 per well per 100 μL. A transfection with a poly-CRE (cAMP Response Element) β-galactosidase reporter plasmid was carried out on day 2 and on day 3 the cells were stimulated with 1 μM of forskolin and incubated with increasing concentrations of each peptide as previously described (Liu et al., 2001a, Liu et al., 2001b). In brief, cell culture medium was removed from the plate and the cells were stimulated by the addition of cell culture medium additioned with 1 μM of forskolin and increasing concentrations of the peptides being tested (0.1 nM to 1 μM at 10 fold increments). Measurements were made in triplicate wells (100 μL per well) and the plate was incubated for 6 hours at 37 °C in a humified atmosphere of 5% CO_2. After

the incubation, the plate was removed from the incubator, the medium aspirated and the plate was frozen at – 80 °C overnight to assist in cell lysis. The cAMP concentration is assessed with a β-galactosidase colorimetric assay measuring absorbance at 570 nm on Benchmark Plus microplate spectrophotometer (BIO RAD) as previously described (Liu et al., 2001a, Liu et al., 2001b, Scott et al., 2006). Briefly, the cells were lysed with the addition of 0.1x assay buffer A (100 mM Na_2HPO_4, pH 8, 2 mM $MgSO_4$, 0.1 mM $MnCl_2$; 25 μL per well) and incubated for 10 minutes at room temperature. Cells were further incubated for 10 minutes at room temperature with 1x assay buffer B (assay buffer A containing 0.5% of TritonX-100 and 40 mM β–mercaptoethanol; 100 μL per well). The colour is developed with the enzyme substrate solution chlorophenolred β–D galactopyranoside (CPRG) (1mg/mL; 25 μL per well) and quantitated monitoring absorbance at 570 nm on a plate reader. Each concentration point was measured in triplicate and each experiment was performed independently for a total of three times. GraphPad PRISM 5 (GrapdPad Inc., San Diego, USA) was used to analyse the cAMP inhibition assays data that are expressed as mean ± SEM.

RESULTS AND DISCUSSION

Two point-mutated analogues have been designed and synthesized in the study. In analogue **1**, the pair Arg^{B24}-Trp^{B25} expected to be responsible for activation was mutated with Ala to generate mouse INSL5 B(R24A, W25A)-chain and in analogue **2**, Lys^{B6}, Arg^{B14} and Tyr^{B18}, within the B chain, expected to be responsible for receptor affinity were also mutated to Ala to give mouse INSL5 B(K6A, R14A, Y18A)-chain. Both B-chain analogues were then combined with the A-chain to give the two chain-disulfide linked peptides. The residue Val^{B13} in mouse INSL5 B-chain, which corresponds to Ile^{B15} in relaxin-3 B-chain, was not mutated to Ala as this would only change the size of the side chain whereas for Lys^{B6}, Arg^{B14} and Tyr^{B18} the mutation substantially affects both the side chain charge and the size. On the basis of the results of the bioassays for this analogue, especially the competition binding, the role of these residues can be determined and if necessary further analogues involving a mutation of Val^{B13} can be designed to investigate the importance (or contribution) of this residue. More generally, poly-mutations were inserted into both the analogues with the purpose to obtain a loss of function in the activation for mouse INSL5 A-B(R24A, W25A) (analogue 1) and in the affinity for mouse INSL5 A-B(K6A, R14A, Y18A) (analogue 2) respectively. If such loss of function is observed after performing the functional assays, the design and synthesis of analogues bearing a single mutation will be required to assess whether individual residues are responsible for the effects observed or two or more of them are acting synergistically.

Given the beneficial effects of the penta-lysine solubilising tag on both the solid phase assembly and the RP-HPLC purification for mouse INSL5 acid as described previously (Belgi et al., 2011) and firstly applied, as a proof-of-concept, to the case of

human insulin glargine A-chain (Hossain et al., 2009), this approach was applied to the B-chain of analogue **1** as well. As for mouse INSL5 B-chain-solubilising tail, the synthesis of analogue **1** B-chain with the penta-lysine tag at the C-terminus led to a crude chromatographic profile with the major peak identified by MALDI-TOF MS as being the target B-chain (Figure 2a). This supported the hypothesis of the tag as a spacer that prevents or minimizes the aggregation of the growing chain on the solid support during the synthesis. The following purification by preparative HPLC was straightforward and the partially purified B-chain (Figure 2b) was obtained in a good amount (about 26 mg) confirming the positive effect of the polycationic tag in increasing the hydrophilicity of the peptide. The partially purified analogue **1** B-chain with the solubilising tail was then combined with the A-chain. The chain combination and disulfide bond formation were successful and yielded 1.2 mg of the two-chain, three disulfide-linked peptide with the B-chain C-terminally bearing the penta-lysine tail in good yield, 12.9% calculated on the B-chain, and purity as shown in Figure 2c. However, upon removal of the tail with a basic treatment (0.1N NaOH) as specified in the General Methods, the final two-chain analogue was obtained in poor overall yield (0.6% calculated on the B-chain) leading to the major loss of the peptide and formation of side products as showed in Figure 2d. It appears that, unexpectedly, tail cleavage seems to be another "high risk" reaction in addition to iodine oxidation which is considered to be the crucial step due to the harshness of the reagents used. After the positive application of the NaOH treatment to, first, glargine and then mouse INSL5, this was the first case of a poor outcome leading to the high loss of the final target peptide. The three major peaks eluting at 19.96, 21.63 and 23.04 minutes respectively were analysed by MALDI-TOF MS. The peak at 19.96 minutes was identified as the target analogue **1** whereas the other two peaks had lower molecular weight indicating

they were generated by fragmentation of the target peptide. However, the nature of the fragments could not be identified. Only the central fraction of the peak at 19.96 minutes was used to ensure the purity of the analogue, which was confirmed by the mass spectrum (Figure 2e). It was freeze dried to give 9.0 μg of effective peptide following amino acid analysis. The 9.0 μg of final peptide were then used for the *in vitro* characterization. Unfortunately, due to such a small amount of effective peptide, full functional testing with three independent assay replicates could not be performed. The competition binding assay, in fact, was performed only once and although it demonstrated no binding in the concentration range tested (0.01nM-1μM), it would need to be repeated at least twice more to be certain of the results. If this result is reproducible, then this would be in contrast to the H3/RXFP3 system where mutations of C-teminal Trp and Arg in the B chain demonstrated full binding with no activity. Nevertheless the cAMP agonism assay of analogue **1** on CHO-K1/RXFP4 cells, which was of most interest for this analogue, was successfully achieved in three independent assays. The analogue, up to a concentration of 1μM, showed no ability to activate the receptor which was as expected (Figure 3a). This was the first evidence that INSL5 is activating RXFP4 receptor with the B-chain via a major contribution from C-terminal Arg and Trp like relaxin-3 does on RXFP3 and RXFP4 which correlates with the mechanism of action proposed for relaxin-3 where the peptide uses the residues exposed in the α-helical region of the B-chain to bind to the receptor, after which the C-terminus of the peptide is in the optimal position for receptor activation (Kuei et al., 2007). Following the earlier finding that INSL5 requires a C-termini acid for full activity on RXFP4 (Belgi et al., 2011), the B-chain not only needs to be in the acid form but the terminal two residues need to be Arg and Trp as their Ala mutation suppressed the

activity of the peptide completely. However, further studies are required to investigate the individual role of the two residues in the activation, i.e. whether only Trp is important or Arg is also contributing to the overall activity.

At the time of the synthesis of analogue **1**, a re-assessment of the unsuccessful tail cleavage was carried out. We decided to change the order of the reactions in the protocol such that the cleavage of the tail was carried out before the iodine oxidation step to form the last disulfide bond. This new order was applied to the synthesis of an analogue of human glargine A(0)R glargine, however, no significant difference in final yield, relative to the B-chain, was observed compared with cleaving the tail as last step before obtaining the final peptide (Kwak et al., 2010). The base treatment was shortened to 2 minutes as a maximum instead of the 2-5 minutes as indicated above in the Methods and used for human glargine and mouse INSL5 acid (Belgi et al., 2011, Hossain et al., 2009) and the reaction mixture (or peptide solution) after the ice cold base addition was kept at below 0°C. The treatment with NaOH at a lower concentration, 0.01M, was also investigated but after 20 minutes treatment it was found to be not as effective as 0.1M NaOH solution (data not shown). To summarize, a change in the order of the reaction, with the tail cleavage before the formation of the third disulfide bond instead as the last step, is possible and does not generally affect the overall yields which remain similar (Kwak et al., 2010); the optimal conditions found were a shortened treatment time of 2 minutes, instead of a maximum of 5 minutes, and for the reaction mix to be kept on ice upon the addition of ice cold NaOH 0.1M (Kwak et al., 2010).

For the synthesis of the modified B-chain of analogue **2** having a C-terminus acid the preloaded resin Fmoc-Trp(Boc)-PAC-PEG-PS was used. Such synthesis was successful giving the target peptide as the major peak, however its purification by

preparative RP-HPLC was very challenging due to the very poor solubility of the peptide. The Ala mutation, in fact, substituted positively charged residues with a non-charged residue giving a B-chain analogue with an overall charge of 0 compared to an overall charge of +2 for the native sequence of mouse INSL5 B-chain acid. The recovery of partially pure B-chain analogue from the HPLC column was very poor so the synthesis needed to be repeated three times in total to ensure a sufficient amount of B-chain for the combination with the A-chain and formation of the final two-chain peptide. The solubilising tail approach remains undoubtedly a very valuable tool for the synthesis and especially for the post synthesis handling of hydrophobic peptides but its cleavage can be detrimental to some sensitive sequences such as in the case of analogue **1**.

Analogue **2** was obtained in sufficient amount to allow a complete *in vitro* characterisation via the competition binding and the cAMP agonism assays. The peptide was able to inhibit forskolin-stimulated cAMP and hence to activate RXFP4 as a full agonist but with less potency compared to native mouse INSL5 acid with the curve shifted towards the right and a pEC_{50} value of 7.87 ± 0.18 versus pEC_{50} 9.29 ± 0.09 for mouse INSL5 acid (Figure 3a, Table 1).

The relative affinity of analogue **2** competing for binding to RXFP4 with the developed Eu-(A) mouse INSL5 acid (Belgi et al., 2011) is shown in Figure 3b. The analogue had a measured affinity, expressed as pIC_{50} ± SEM, of 7.72 ± 0.2 and showed partial binding as if it were able to compete for one binding site but not for a second one. As shown in Figure 3b, the curve associated with the analogue did not reach zero showing its inability to fully compete with the labelled peptide. This result highlights simultaneously the importance of not only Lys^{B6}, Arg^{B14} and Tyr^{B18} but also of Val^{B13} which was not mutated. It is, in fact, possible that the affinity of the analogue for the

receptor is due to the presence of Val^{B13} and therefore due to hydrophobic interactions between its side chain and the receptor. The decreased potency in the cAMP agonism assay correlates with the decreased binding affinity suggesting that Lys^{B6}, Arg^{B14} and Tyr^{B18} contribute to both binding and activation even though their contribution in the activation is marginal (or moderate) compared to Arg^{B24} and Trp^{B25} whose mutation completely abolishes the activation. The diminished potency can also be a consequence of the non efficient interaction of the peptide with the receptor and therefore the peptide is failing to present Arg^{B24} and Trp^{B25} in the optimal position for activation which, once again, is in line with the mechanism of action suggested for relaxin-3 (Kuei et al., 2007).

To gain a better understanding of INSL5 structure-function relationship, the next candidate that needs to be investigated is Val^{B13} which correspond to Ile^{15} in relaxin-3 B-chain and which in this preliminary screening was not mutated to Ala. It was in fact thought that the Val to Ala mutation only affects the size of the side chain whereas for Lys^{B6}, Arg^{B14} and Tyr^{B18} the mutation was substantially changing the charge/polarity and size of the side chains. Therefore it is possible that Val^{B13} is contributing to INSL5 binding to the receptor with hydrophobic interaction similar to Ile^{15} in relaxin-3. Ile^{15} in relaxin-3 was mutated to either Gly, Ala or Val and it was observed an increased affinity associated with the analogue with increased side chain length confirming the presence of an hydrophobic interaction between the side chain of the residue and an hydrophobic pocket within the receptor (Kuei et al., 2007). To further investigate this hydrophobic interaction, Val^{B13} should be substituted with a charged amino acid such as Lys having a positive charge or the negatively charged Asp (Schwartz et al., 1987) and the effect of these substitutions on the binding tested.

CONCLUDING REMARKS

The chemical syntheses of analogues **1** and **2**, especially the different approaches applied to the B-chain, highlighted the challenges and difficulties associated with the assembly of this insulin superfamily member. Not only the synthesis of the peptide chains has complications but, and foremost, their purification via HPLC is very difficult due to the extreme hydrophobic properties of INSL5 and its analogues. These same purification issues were encountered in the synthesis of human INSL5 via recombinant techniques (Luo et al., 2010).

This is the first structure-function relationship study of INSL5. The preliminary findings showed that like H3 relaxin, INSL5 peptide uses Arg^{B24} and Trp^{B25} residues to activate the receptors and Lys^{B6}, Arg^{B14} and Tyr^{B18} to bind to the receptor. However, it is not conclusive as binding assays for analogue **1** are incomplete. The fact that INSL5 can not activate RXFP3 although it has some affinity, while H3 relaxin has high affinity for RXFP4, suggests that the binding and activation mechanism of INSL5/RXFP4 should be slightly different to that of H3/RXFP3. Further structure-activity studies on INSL5-RXFP4 are required to confirm the binding and activation mechanisms.

ACKNOWLEDGEMENTS

We thank Sharon Layfield for undertaking the cell culture, Tania Ferraro for assisting with the binding assays and Feng Lin for the amino acid analysis. This work was funded in part by a National Health and Medical Research Council of Australia Project grants #1023078 and #1023321 to JDW and MAH and by the Victorian Government's Operational Infrastructure Support Program.

REFERENCES

Bathgate, R. A., Ivell, R., Sanborn, B. M., Sherwood, O. D. & Summers, R. J. 2006. International Union of Pharmacology LVII: recommendations for the nomenclature of receptors for relaxin family peptides. *Pharmacological reviews*, 58**,** 7-31.

Belgi, A., Hossain, M. A., Shabanpoor, F., Chan, L., Zhang, S., Bathgate, R. a. D., Tregear, G. W. & Wade, J. D. 2011. Structure and function relationship of murine insulin-like peptide 5 (INSL5): Free C-terminus is essential for RXFP4 receptor binding and activation. *Biochemistry*, 50**,** 8352-8361.

Burnicka-Turek, O., Mohamed, B. A., Shirneshan, K., Thanasupawat, T., Hombach-Klonisch, S., Klonisch, T. & Adham, I. M. 2012. INSL5-Deficient Mice Display an Alteration in Glucose Homeostasis and an Impaired Fertility. *Endocrinology*.

Conklin, D., Lofton-Day, C. E., Haldeman, B. A., Ching, A., Whitmore, T. E., Lok, S. & Jaspers, S. 1999. Identification of INSL5, a new member of the insulin superfamily. *Genomics*, 60**,** 50-56.

Haugaard-Jönsson, L., Hossain, M., Daly, N., Craik, D., Wade, J. & Rosengren, K. J. 2009. Structure of human insulin-like peptide 5 and characterization of conserved hydrogen bonds and electrostatic interactions within the relaxin framework. *Biochemical Journal* 419**,** 619-627.

Hossain, M. A., Bathgate, R. a. D., Kong, C. K., Shabanpoor, F., Zhang, S., Haugaard-Jönsson, L. M., Rosengren, K. J., Tregear, G. W. & Wade, J. D. 2008. Synthesis, conformation, and activity of human insulin-like peptide 5 (INSL5). *ChemBioChem*, 9**,** 1816-1822.

Hossain, M. A., Belgi, A., Lin, F., Zhang, S., Shabanpoor, F., Chan, L., Belyea, C., Truong, H. T., Blair, A. R., Andrikopoulos, S., Tregear, G. W. & Wade, J. D. 2009. Use of a temporary "solubilizing" peptide tag for the Fmoc solid-phase synthesis of human insulin glargine via use of regioselective disulfide bond formation. *Bioconjugate Chemistry,* 20**,** 1390-1396.

Kuei, C., Sutton, S., Bonaventure, P., Pudiak, C., Shelton, J., Zhu, J., Nepomuceno, D., Wu, J., Chen, J. & Kamme, F. 2007. R3(BΔ23-27)R/I5 chimeric peptide, a selective antagonist for GPCR135 and GPCR142 over relaxin receptor LGR7: in vitro and in vivo characterization. *Journal of Biological Chemistry,* 282**,** 25425-25435.

Kwak, S. Y., Forbes, B. E., Lee, Y. S., Belgi, A., Wade, J. D. & Hossain, M. A. 2010. Solid phase synthesis of an analogue of insulin, A0:R glargine, that exhibits decreased mitogenic activity. *International Journal of Peptide Research and Therapeutics,* 16**,** 153-158.

Liu, C., Kuei, C., Sutton, S., Chen, J., Bonaventure, P., Wu, J., Nepomuceno, D., Kamme, F., Tran, D. T., Zhu, J., Wilkinson, T., Bathgate, R. A., Eriste, E., Sillard, R. & Lovenberg, T. W. 2005. INSL5 is a high affinity specific agonist for GPCR142 (GPR100). *Journal of Biological Chemistry,* 280**,** 292-300.

Liu, C. & Lovenberg, T. 2008. Relaxin-3, INSL5, and their receptors. *Results and Problems in Cell Differentiation,* 46**,** 213-237.

Liu, C., Ma, X.-J., Jiang, X., Wilson, S. J., Hofstra, C. L., Blevitt, J., Pyati, J., Li, X., Chai, W., Carruthers, N. & Lovenberg, T. W. 2001a. Cloning and pharmacological characterization of a fourth histamine receptor (H4) expressed in bone marrow. *Molecular Pharmacology,* 59**,** 420-426.

Liu, C., Wilson, S. J., Kuei, C. & Lovenberg, T. W. 2001b. Comparison of human, mouse, rat, and guinea pig histamine H4 receptors reveals substantial pharmacological species variation. *Journal of Pharmacology and Experimental Therapeutics,* 299**,** 121-130.

Luo, X., Bathgate, R. a. D., Zhang, W. J., Liu, Y. L., Shao, X. X., Wade, J. D. & Guo, Z. Y. 2010. Design and recombinant expression of insulin-like peptide 5 precursors and the preparation of mature human INSL5. *Amino Acids,* 39**,** 1343-1352.

Schwartz, G. P., Burke, G. T. & Katsoyannis, P. G. 1987. A superactive insulin: [B10-Aspartic acid]insulin(human). *Proceedings of the National Academy of Sciences of the United States of America,* 84**,** 6408-6411.

Scott, D. J., Layfield, S., Yan, Y., Sudo, S., Hsueh, A. J. W., Tregear, G. W. & Bathgate, R. a. D. 2006. Characterization of novel splice variants of LGR7 and LGR8 reveals that receptor signaling is mediated by their unique low density lipoprotein class A modules. *Journal of Biological Chemistry,* 281**,** 34942-34954.

Shabanpoor, F., Hughes, R. A., Bathgate, R. a. D., Zhang, S., Scanlon, D. B., Lin, F., Hossain, M. A., Separovic, F. & Wade, J. D. 2008. Solid-phase synthesis of Europium-labeled human INSL3 as a novel probe for the study of ligand-receptor interactions. *Bioconjugate Chemistry,* 19**,** 1456-1463.

Shabanpoor, F., Separovic, F. & Wade, J. D. 2009. The human insulin superfamily of polypeptide hormones. *In:* LITWACK, G. (ed.) *Vitamins and Hormones.*

Takedacambridgelimited. 2009. *Compound for controlling appetite*. PCT/GB2008/003023.

Wilkinson, T. N., Speed, T. P., Tregear, G. W. & Bathgate, R. a. D. 2005. Coevolution of the relaxin-like peptides and their receptors. *Annals of the New York Academy of Sciences,* 1041**,** 534-539.

Table 1 Binding affinity (pIC_{50}), activation (pEC_{50}) and Efficacy maximum (E_{max}) on RXFP4 receptor. ND=not determined; NA=not active in the concentration range tested; (*) partial binding, see Figure 3b. Data are expressed as mean ± SEM of n=3 independent experiments. Emax (%): efficacy maximum of the peptide in the cAMP assay relative to forskolin as 100%.

Peptide	Sequence	pIC_{50} n=3	pEC_{50} n=3	E_{max} (%)
mouse INSL5	RDLQALCCREGCSMKELSTLC SRQTVKLCGLDYVRTVIYICASSRW	8.47 ± 0.03	9.29 ± 0.09	82.09 ± 2.40
1	RDLQALCCREGCSMKELSTLC SRQTVKLCGLDYVRTVIYICASSAA	ND	NA	NA
2	RDLQALCCREGCSMKELSTLC SRQTVALCGLDYVATVIAICASSRW	7.72 ± 0.2 (*)	7.87 ± 0.18	77.00 ± 0.73

Figure legends:

Figure 1 Primary structures of synthetic mouse INSL5 analogues. From top to bottom: (top) relaxin-3 sequence where the residues within the B-chain that were found to be important for binding and activation are highlighted in grey for RXFP3 and in black for RXFP4. Plain arrows indicate residues important for the binding and bold arrows the one involved in activation (Kuei et al., 2007). (Middle) mouse INSL5 sequence with arrows indicating the corresponding residues within the B-chain that in relaxin-3 were found important for interaction with RXFP4; and (bottom) sequences of the two point mutated analogues designed: mouse INSL5 A-B(R24A, W25A) (analogue **1**) and mouse INSL5 A-B(K6A, R14A, Y18A) (analogue **2**).

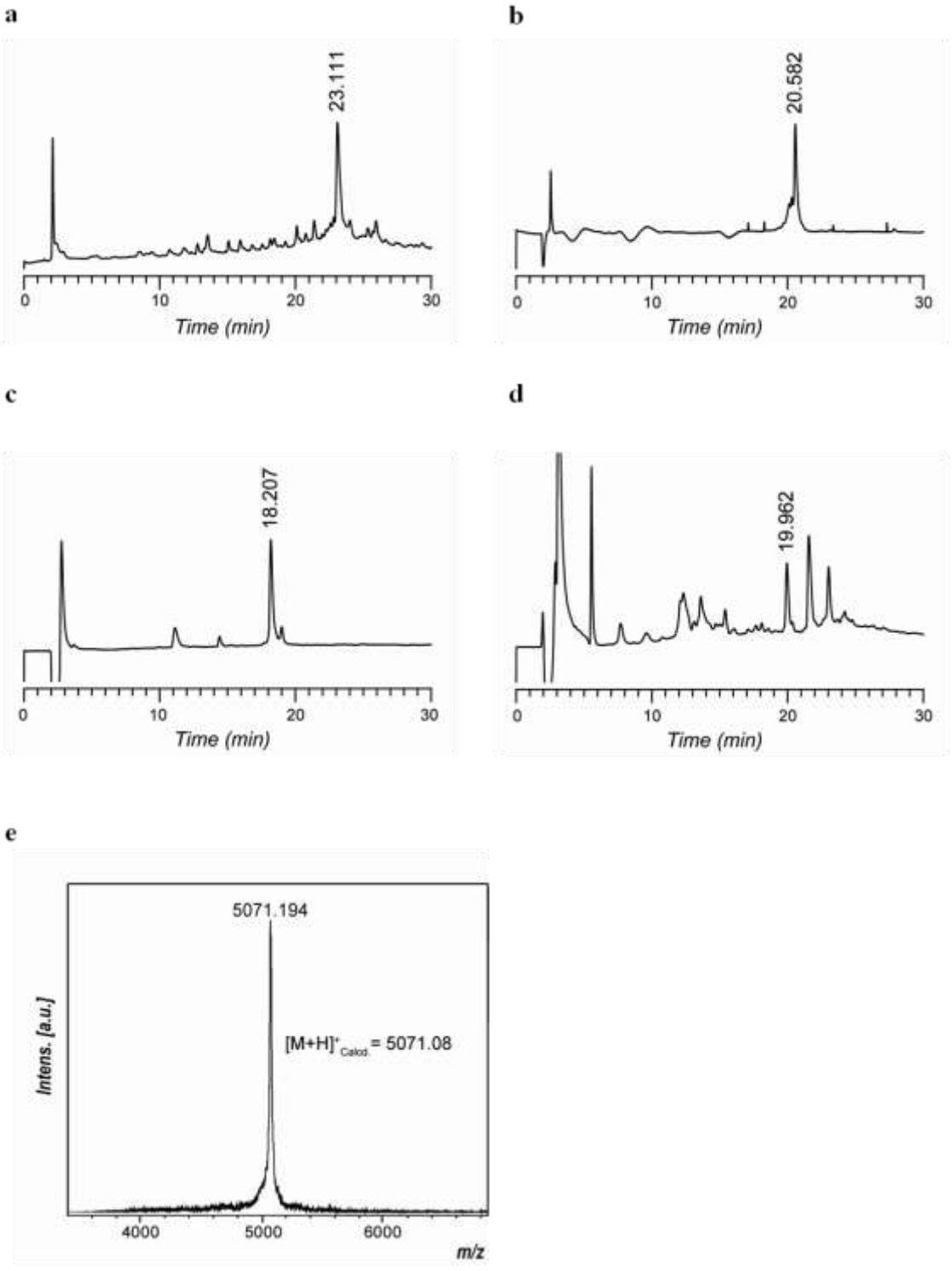
Figure 2 RP-HPLC of synthetic analogue **1** B-chain with C-terminal penta-Lys tail (a) crude and (b) partially purified. Conditions: Vydac C18 column (4.6 x 250 mm, pore size 300Å, particle size 5µ); Buffer A, 0.1% aq. TFA; Buffer B, 0.1% TFA in CH_3CN. Gradient Buffer B 15-45% over 30 minutes for (a) and gradient Buffer B 20-50% over 30 minutes for (b). Analogue **1**: RP-HPLC of (c) synthetic analogue **1**-ST partially purified where the B-chain still bears the C-terminal penta-Lys tail and (d) after the basic treatment to cleave the tail. Conditions: Vydac C18 column (4.6 x 250 mm, pore size 300Å, particle size 5µ); Buffer A, 0.1% aq. TFA; Buffer B, 0.1% TFA in CH_3CN; gradient Buffer B 20-40% over 30 minutes for both (c) and (d). MALDI-TOF MS of analogue **1** after the tail cleavage (peak 19.96 minutes) (e). Analogue **2**: (f) RP-HPLC; conditions: Vydac C18 column (4.6 x 250 mm, pore size 300Å, particle size 5µ); Buffer A, 0.1% aq. TFA; Buffer B, 0.1% TFA in CH_3CN; gradient Buffer B 20-55% over 35 minutes; (g) MALDI-TOF MS.

Figure 3 In vitro RXFP4 receptor bioassays: (a) cAMP inhibition assay of analogue **1**, analogue **2** and mouse INSL5 acid; (b) competition binding curve of analogue **2** and mouse INSL5 acid competing with 5 nM of Eu-(A) mouse INSL5 acid. The data are the result of n= 3 independent experiments and are expressed as mean ± SEM.

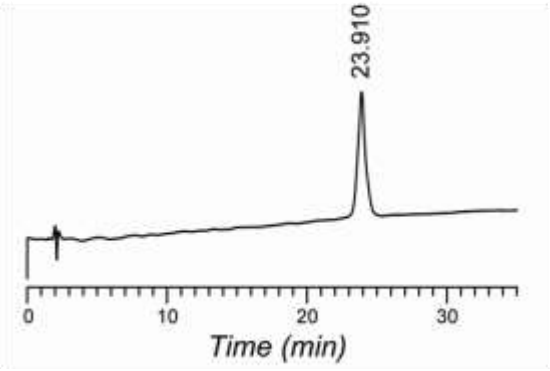
Figure 1


RELAXIN-3
D V L A G L S S S C C K W G C S K S E I S S L C
R A A P Y G V R L C G R E F I R A V I F T C G G S R W
RXFP3
RXFP4
MOUSE INSL5
R D L Q A L C C R E G C S M K E L S T L C
S R Q T V K L C G L D Y V R T V I Y I C A S S R W
R D L Q A L C C R E G C S M K E L S T L C
S R Q T V K L C G L D Y V R T V I Y I C A S S A A
(1)
R D L Q A L C C R E G C S M K E L S T L C
S R Q T V A L C G L D Y V A T V I A I C A S S R W
(2)

Figure 2


a
23.111
0
10
20
30
Time (min)
b
20.582
0
10
20
30
Time (min)
c
18.207
0
10
20
30
Time (min)
d
19.962
0
10
20
30
Time (min)
e
5071.194
[M+H]+Calcd. = 5071.08
Intens. [a.u.]
4000
5000
6000
m/z

f


23.910
0
10
20
30
Time (min)

g

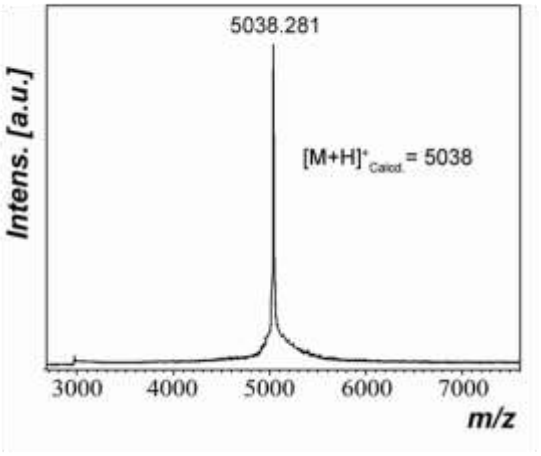
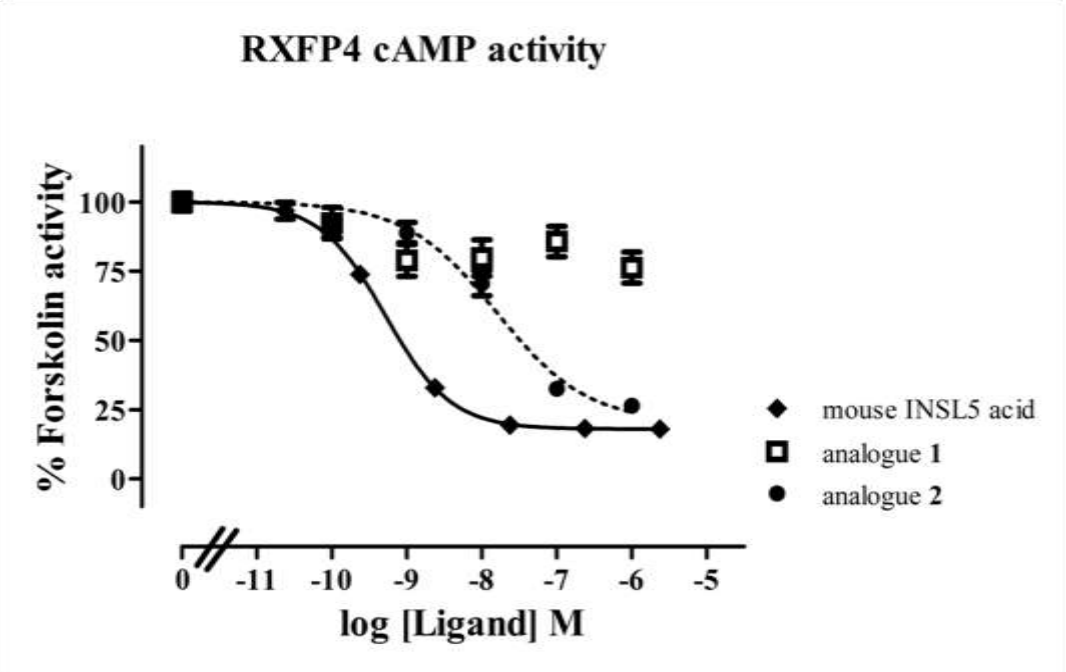
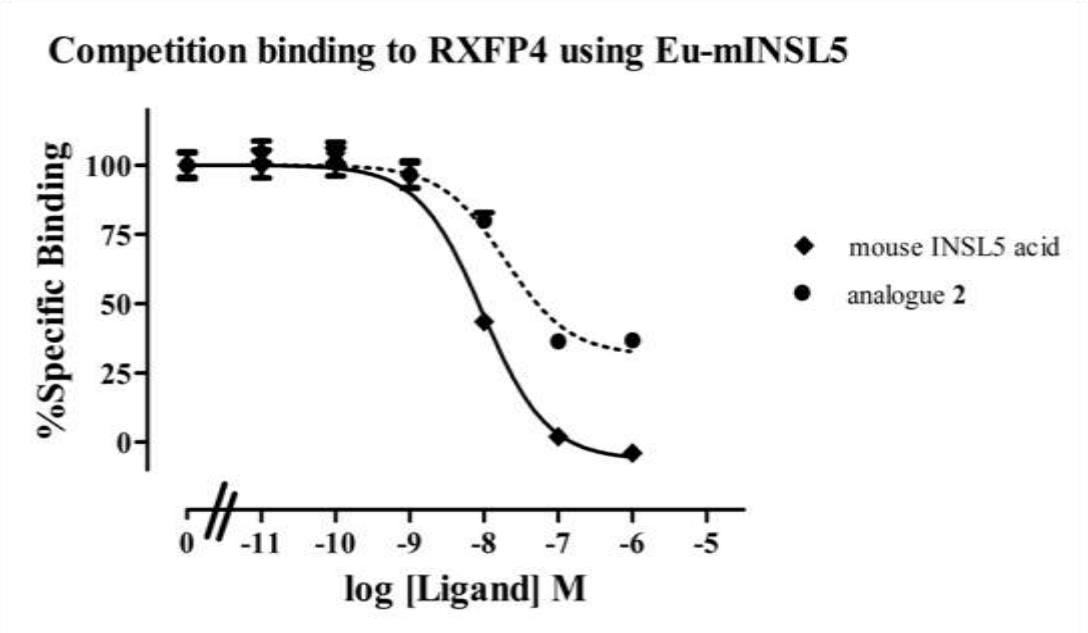

5038.281
[M+H]+Calcd. = 5038
Intens. [a.u.]
3000
4000
5000
6000
7000
m/z

Figure 3

a


RXFP4 cAMP activity
% Forskolin activity
100
75
50
25
0
0
-11
-10
-9
-8
-7
-6
-5
log [Ligand] M
mouse INSL5 acid
analogue 1
analogue 2

b


Competition binding to RXFP4 using Eu-mINSL5
%Specific Binding
100
75
50
25
0
0
-11
-10
-9
-8
-7
-6
-5
log [Ligand] M
mouse INSL5 acid
analogue 2