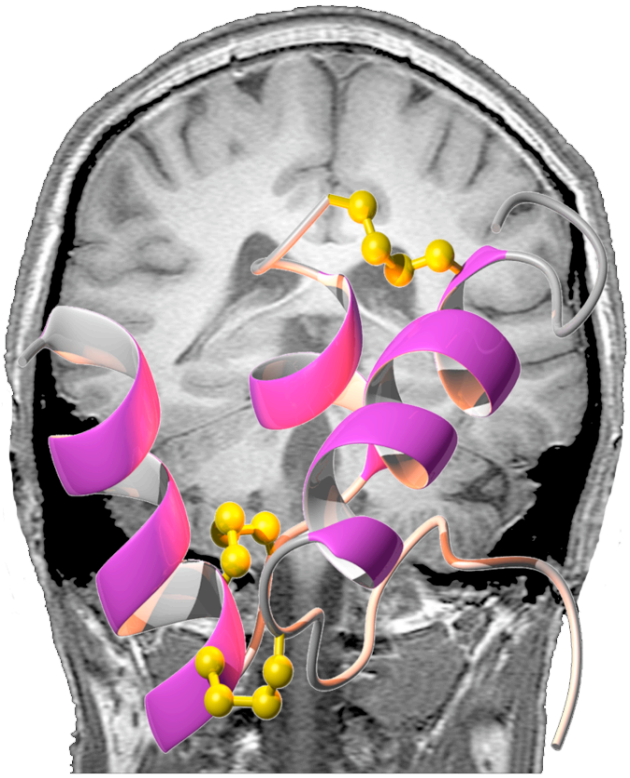
DESIGN, SYNTHESIS AND STRUCTURAL STUDIES OF RELAXIN RECEPTOR MODULATORS



Iselin Sørstad Elvheim

Institute for Molecular Bioscience
University of Queensland

Institutt for farmasi
Universitetet i Tromsø

May 2011

Abstract

The relaxin family of peptides consists of seven structurally related peptides, with a wide variety of biological functions, ranging from involvement in reproductive processes to functions as neuroendocrine modulators.

Because of the complex, two-chained structure of the relaxins, and their lack of specificity for the various relaxin family receptors (RXFPs), design of simpler, more selective analogues is important for further investigation of their biological functions and as potential drug leads.

The aim of this project was to introduce a helical structure around the receptor-binding region of single-chain relaxin analogues. This was approached by utilising helix capping sequences on truncated relaxin B-chains and by grafting of residues important for binding onto a stable peptide scaffold.

Helix capping enhanced the helical properties compared to previous single-chain analogues, but was unable to introduce a sufficiently stable helix. Despite the increased helical tendencies, no high-affinity analogues were found. As a result, the importance of Arg^{8} was investigated, and we have demonstrated that this residue might be more involved in binding than previously thought.

We successfully synthesised a relaxin-3/chlorotoxin analogue, which appeared to fold correctly. However, the yield following oxidation was poor, and no bioactivity data or structural data confirming the correct fold was obtained within the time limits of the project.

A secondary aim was to probe for favourable mutations around the receptor-binding region by synthesising a combinational library. We were able to successfully synthesise a library of peptides with mutations in one position by inserting a mixture of amino acids at this coupling step in the solid phase peptide synthesis (SPPS) procedure. Although no significant improvement in binding was seen for the analogues generated, important methodological advances were made that will be used to scan different positions for new contact points with the receptor.

Acknowledgements

The work of my master project was done at the Institute for Molecular Bioscience at the University of Queensland, Australia, in the period from October 2010 to May 2011.

First of all, I want to thank my supervisors Dr. Richard Clark and Dr. Johan Rosengren for all their help throughout my project, both with the practical labwork and their helpful comments and guidance on writing this thesis.

I much appreciate the opportunity I was given to undertake these studies, and write my master thesis at the University of Queensland. This was made possible by Dr. David Craik, for letting me do the work on my thesis in his lab, and Dr. Jon Våbenø, for agreeing to be my internal supervisor at the University of Tromsø.

I am sincerely grateful to everyone in the Craik research group for making me feel welcome in the lab, and being patient with all of my questions. Especially I would like to thank Chia Chia Tan, Crystal Huang and Alysha Elliot for helping me with various lab equipment and procedures, and Phill Walsh and Phil Sunderland for helping me with my peptide synthesis. Also, thanks to Muharrem Akcan for sharing his experience with synthesis and folding of chlorotoxin.

I also have to thank Ross Bathgate and his lab at the Florey Neuroscience Institutes in Melbourne for performing the cell based binding and activity assays.

And last, but by no means least, I would like to thank Katrine Køller and Jørgen Østerpart. Katrine for sharing this experience with me, for keeping me company and cheering me up when my experiments didn't go according to plan, and being a great friend thorughout our stay. My wonderful boyfriend Jørgen Østerpart for all his support and help, for coming to Australia, and for putting up with me in general.

Table of contents

Abbreviations:

1D	one-dimensional
2D	two-dimensional
3D	three-dimensional
Boc	*tert*-butoxycarbonyl
Buffer A	0.05 % trifluoracetic acid in water
Buffer B	90% acetonitrile, 10% water, 0.05% trifluoracetic acid
Buffer A/B	50/50 mixture of buffers A and B
DIPEA	N,N-Diisopropylethylamine
DMF	N,N-Dimethylformamide
cAMP	Cyclic adenosine monophosphate
CHO	Chinese hamster ovary
CPRG	Chlorophenol red-β-D-galatopyranoside
CRE	cAMP response element
DODT	3,6-dioxa-1,8-octanedithiol
ES-MS	Electrospray mass spectrometry
Fmoc	9H-fluoren-9-ylmethyloxycarbonyl
GPCR	G protein-coupled receptor
HBTU	O-benzotriazol-1-yl-1,1,3,3-tetramethyluronium hexafluorophosphate
HEK	Human embryonic kidney
HPLC	High performance liquid chromatography
ICV	Intracerebroventricular
IGF	Insulin-like growth factor
INSL	Insulin-like peptide
iPVN	Intraparaventricular
LDL-a	Low-density lipoprotein class A
LGR	Leucine-rich repeat-containing G protein-coupled receptor
LRR	Leucine-rich repeat
MALDI-TOF-MS	Matrix-assisted laser desorption/ionization time of flight mass spectrometry
MS	Mass spectrometry

MRNA	Messenger ribonucleic acid
NI	Nucleus incertus
NMR	Nucleic magnetic resonance
NOE	Nuclear Overhauser effect
NOESY	Nuclar Overhauser effect spectroscopy
RP-HPLC	Reverse phase high performance liquid chromatography
RXFP	Relaxin family peptide receptor
SPPS	Solid phase peptide synthesis
TFA	Trifluoracetic acid
TOCSY	Total correlation spectroscopy
TIPS	Tri-isopropylsilane
UV	Ultraviolet
VEGF	Vascular endothelial growth factor

1. Introduction

1.1. Bioactive peptides

Peptides are defined as compounds composed of two or more amino acids, coupled together by amide bonds. Peptides generally comprise of combinations of the 20 different gene encoded amino acids. There are exceptions however, with peptides being post-translationally modified.

A bioactive peptide is a peptide that shows some sort of biological activity. Bioactive peptides can be found in all parts of the natural world, and have a range of different functions. One function well known in humans, is the role peptides play as hormones. The most famous of the peptide hormones might be insulin, and its regulation of blood glucose levels.

In other organisms peptides with more exotic functions are found; for instance peptides with antimicrobial properties, or peptide toxins produced to aid in defence, or in capture of prey. Worth mentioning here is cyclosporine, a peptide metabolite from the soil fungus *Trichoderma polysporum*, which is now widely used as an immunosuppressant drug.

Peptides have many properties making them attractive drug leads; high activity and specific binding to their target, and generally low non-specific binding to molecules other than the target. Toxicity is also generally low, as is accumulation in tissue [1, 2]. However, peptides also have some major drawbacks as pharmaceuticals, including poor stability and unfavourable properties such as low oral availability and rapid clearance from the body, as well as a potential immunogenicity [1-3]. Because of these problems, much work is focused on improving peptide stability and developing delivery systems for peptide drugs [2, 3].

1.2 The relaxin family of peptides

1.2.1. Overview

The family of relaxin peptide hormones consists of seven different peptides; relaxin 1-3 and insulin like peptide (INSL) 3-6. These peptides are structurally related to insulin and insulin-like growth factors (IGF) I and II, and together these ten peptides are referred to as the insulin/relaxin superfamily [4, 5]. One common

denominator for the relaxin peptides is their two-chained structure, where an A and a B chain are held together through cross-bracing by two interchain disulfide bonds, and one disulfide bond within the A-chain [4] (Figure 1).

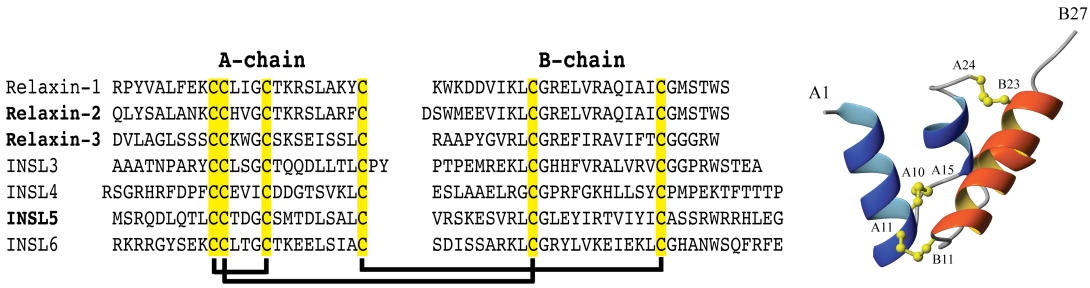


Figure 1. Sequence alignment of the peptides that make up the relaxin family. The concerved cysteine residues are highlighted in yellow, and their connectivity indicated by brackets. The overall fold of the family is expamplified by the crystal structure of relaxin-2, shown with the A-chain in blue, and the B-chain in red. Disulfide bonds are shown in yellow.

These disulfide bonds stabilise the overall structure, which is conserved in all of the relaxins. In the common structure, the A-chain folds into two helical segments, separated by a central β-strand. The helices are positioned anti-parallel to each other in a U-shaped structure. The B-chain consists of one central helix positioned across the U of the A-chain and a second β-strand forming an inter-chain β-sheet [6-9].

The focus of this project was to design new analogues of relaxin-2, relaxin-3 and INSL5, and the rest of the introduction will focus on these members of the relaxin family.

1.2.2. Biological roles and clinical applications

Relaxin-2 was discovered in 1926, and the first observed biological role was in relaxing the pubic ligaments in pregnant guinea pigs, hence the name “relaxin” [10]. The sequence of porcine relaxin was first characterized in the 1970’s [4, 11, 12].

From originally being viewed as a reproductive hormone, its functional roles have later been expanded to include regulating the cardiovascular and renal systems, inhibiting collagen biosynthesis and promoting collagen breakdown, as well as regulating allergic responses [5, 13, 14].

Relaxin-2 is involved in the cardiovascular adaptations that occur during pregnancy, including increased plasma volume, cardiac output and heart rate. It has also been found to reduce blood pressure and vascular resistance [4, 13, 14]. Studies on pregnant rats have also shown glomerular filtration rate and effective renal plasma

flow to be proportional to plasma levels of relaxin [13]. Because of these effects on cardiovascular and renal systems, relaxin-2 has a potential therapeutic use in congestive heart failure. Relaxin-2 has passed phase II clinical trials for treatment of acute heart failure with promising results [15], and entered phase III trials on this indication.

The use of relaxin-2 in women with severe preeclampsia is also being investigated in the clinic [16]. The pathophysiology of preeclampsia involves severe vasoconstriction, resulting in reduced blood flow to the placenta, hypertension, renal dysfunction and multiple organ failure [16]. Relaxin-2 is thought to improve this condition by increasing blood flow to the placenta, in addition to improving systemic heamodynamics and renal function in a way similar to the rationale for treatment of heart failure [16].

Relaxin-2's effects on connective tissue include inhibition of collagen synthesis and promotion of collagen breakdown in both reproductive and non-reproductive tissues. This effect suggests a potential use as an anti-fibrotic agent [4, 5, 13, 14].

Relaxin-2 induces cervical ripening in rats and pigs, and clinical trials on cervical ripening have been conducted on women with post-date pregnancies. In different trials relaxin-2 was administered intravaginally and intravenously, but neither trial showed any evidence of cervical ripening [4, 5].

Since the 1980's over 500 human subjects have been treated with relaxin-2 in clinical trials on different indications, and relaxin-2 has shown an excellent safety profile [14].

INSL5 was first reported in 1999, and is expressed in the colon and uterus in humans [17]. In mice it has been found to be expressed in the enteroendocrine-like cells in the crypts in colonic epithelial layer [4, 18]. Its biological function is still being investigated. Work on INSL5 knock-out mice has suggested that INSL5 may play a role in regulation of fat and glucose metabolism [4, 19].

Relaxin-3 is the latest addition to the relaxin family, discovered in 2002 [20]. Relaxin-3 is primarily expressed in the brain, but is also found in the spleen, thymus and testis [4, 20, 21].

In the brain, the main area where relaxin-3 is found is the nucleus incertus (NI) in the brainstem [20-22]. This area is close to other regions involved in behaviour control and regulation of stress responses [23, 24]. Relaxin-3 neurons also project

from the NI to other areas in the forebrain where relaxin family peptide receptor (RXFP) 3-positive neurons are present. These areas are involved in processes such as circadian rhythms, arousal and attention, learning and memory, stress responses and associated cognitive processes [23, 24]. In forced swim tests on rats, relaxin-3 neuronal activity and relaxin-3 mRNA transcription has been shown to increase in response to stress [22]. These findings suggest that relaxin-3 is involved in regulation of stress responses.

Relaxin-3 might also be involved in appetite regulation [25-27] and in long-term control of food intake [27]. Acute injections of relaxin-3 intracerebroventricularly (ICV) or intraparaventricularly (i-PVN) increase food intake and feeding behaviour in rats [25]. Chronic ICV or i-PVN administration of relaxin-3 has been shown to increase cumulative food intake [26, 27], leading to increased body weight in rats [26].

Because of the wide range of physiological roles, relaxin-3 is under investigation as a drug lead for a variety of disease states. The observation that an RXFP3 agonist is able to block fear-associated "freezing" in rats (unpublished observation), in addition to the role relaxin-3 plays in stress regulation, suggests a potential in treatment of anxiety disorders such as phobias or post traumatic stress disorder.

An RXFP3 antagonist has been reported to inhibit spatial memory in rats when injected into the medial septum [28], suggesting that relaxin-3 has a role in cognition. RXFP3 agonists may be relevant in treatment of dementia or other diseases associated with impaired cognition.

The effect on food intake and appetite, and the role in stress regulation, gives RXFP3 receptor modulators potential as drug leads in treatment of mood-related metabolic disorders, such as anorexia. RXFP3 antagonists might also be leads for anti-obesity drugs [26].

It was recently reported that relaxin-3 has antifibrotic properties similar to those of relaxin-2. These properties are mediated via relaxin-3's action on RXFP1. It was suggested that relaxin-3 may enhance the collagen-inhibitory effects of relaxin-2 [29].

1.2.3. Receptors

Whereas insulin and IGFs bind to tyrosine kinase receptors, the structurally similar relaxins bind to G protein-coupled receptors (GPCR). Four GPCRs have been identified to date as relaxin receptors; RXFP 1-4 [21, 30-34].

The RXFP3 and 4 are classic peptide ligand GPCRs [4, 34] whereas the receptors RXFP1 and 2 belong to a subgroup of family A GPCRs termed leucine-rich repeat-containing G protein-coupled receptors (LGR) [4, 35]. These receptors are characterised by a large extracellular N-terminal leucine-rich repeat (LRR) domain. More specifically, RXFP1 and 2 belong to class C in the LGR subfamily [4, 35]. The extracellular domain of RXFP1 and 2 contains a low-density lipoprotein class A (LDL-a) at their N-termini, in addition to ten LRRs [4, 35]. The B-chain of the relaxin family peptides interacts with sites in the LRR domain, while features of the A-chain are believed to be interacting with the short extracellular loops to induce activation [4, 36, 37].

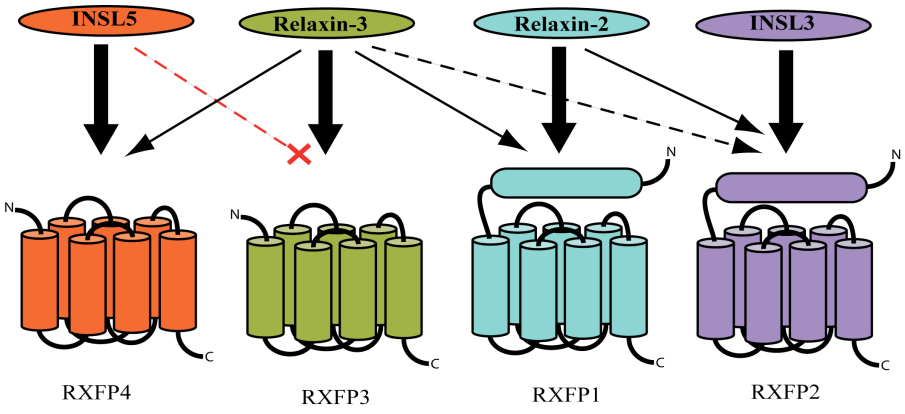


Figure 2. The RXFP receptors and their ligands. Cognate ligand-receptor pairs indicated by bold arrows, and additional interactions by fine arrows. Dotted lines denote a weak interaction, and red line an antagonistic interaction.

RXFP3 is the endogenous receptor for relaxin-3 [21], but relaxin-3 also has high affinity to, and the ability to activate, RXFP1 and RXFP4 [32, 33]. Relaxin-3 also activates RXFP2 to a lesser extent [8]. Relaxin-3 is thus able to bind to and activate receptors in both of the different subgroups of RXFP-receptors, and is the only relaxin peptide able to do so [4, 13] (Figure 2).

INSL5 binds to, and activates, RXFP4 as its endogenous receptor [34], but is also able to weakly antagonise RXFP3 [4, 7, 34]. Relaxin-2 primarily binds and

activates RXFP1, but also has high affinity for RXFP2 [30]. INSL3 is only known to bind to its endogenous receptor [4, 31, 37].

The native receptors for INSL4 and INSL6 have to date not been found [4], and little is known about their biological significance.

1.2.4. Structure/activity relationship

In 1992 the importance of the arginines B13 and B17 for relaxin-2's biological activity was reported by Büllesbach and Schwabe [38]. Later, the binding site was expanded to include isoleucine B20 [39]. The triangular contact region Arg-xxx-Arg-xx-Ile sequence is often referred to as the Relaxin binding motif or casette, or the RXXXRXXI cassette (Figure 3A). This sequence is essential for binding to relaxin-2's endogenous receptor RXFP1 [38, 39].

Relaxin-2 is also dependent on the A-chain to be able to activate the receptors RXFP1 and 2 [36, 40]. As previously mentioned, the primary binding occurs on the B-chain, involving the residues Arg^{13}, Arg^{17} and Ile^{20} [40], but the A chain residues Thr^{16}, Lys^{17} and Phe^{23} play a cooperative part in binding together with the B-chain binding region [40-42]. In addition, Trp^{B28} is involved in primary binding to RXFP2, but not to RXFP1 [40, 43].

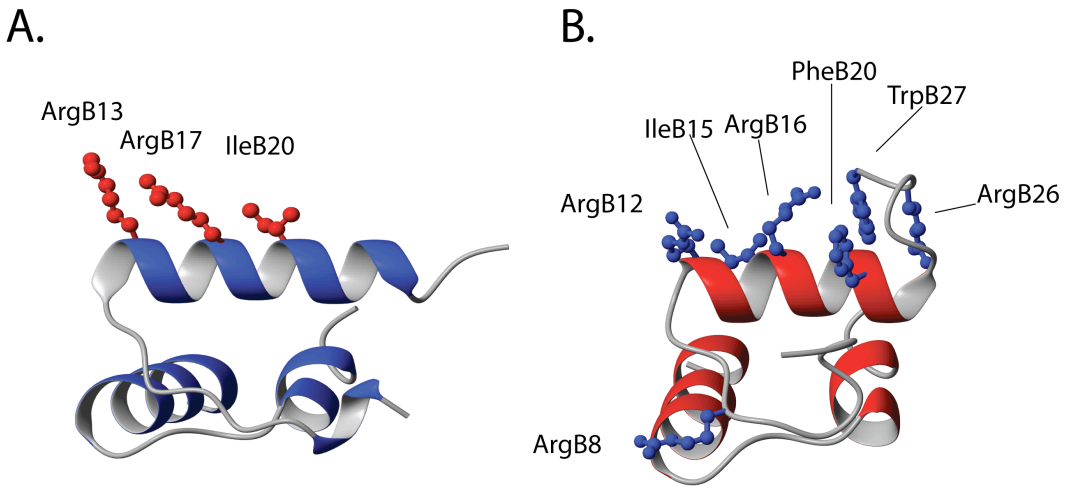


Figure 3. Crystal strucure of relaxin-2 (A.) and NMR solution structure of relaxin-3 (B.) A.) Residues in the relaxin-binding casette which are essential for binding to RXFP1 are labeled and side-chains shown in red. B.) Residues important for binding to RXFP3 are labeled, and side-chains shown in blue.

Relaxin-3 also contains a relaxin-binding cassette exposed on the B-chain helix, and these residues together with some additional neighbouring residues were suggested to be involved in relaxin-3's interaction with its endogenous receptor, RXFP3 [44] (Figure 3B).

Kuei et al. conducted mutational studies replacing either Arg^{12}, Ile^{15}, Arg^{16}, Ile^{19} or Phe^{20} with different combinations of Ala, Ser, Lys, Gly, Arg, Tyr or Val. The results confirmed the importance of the residues for binding to the receptors; Arg^{12}, Arg^{16} and Ile^{19} seem to be especially important for relaxin-3 binding to RXFP1, which fits with the earlier reported relaxin binding RXXXRXXI motif [39]. Phe^{20} and Ile^{15} were shown to be important for relaxin-3 binding to RXFP3 and 4 [8, 44]. Phe^{20} is not important for RXFP1-binding [40, 41, 44].

Kuei et al. also found that changing Arg^{8} to Ala or Ser resulted in lower potency and affinity for RXFP3, 4, and 1 [44]. It was suggested that Arg^{8} could be involved in stabilizing the overall structure of relaxin-3, as it is positioned close to the A-chain and may interact with it [44] (Figure 3B).

Altering Arg^{26} to Ala or Thr, or Trp^{27} to Ala, Phe or Arg gave a dramatic reduction in agonist activity on both RXFP3 and 4 while the activity on RXFP1 was unchanged. Interestingly, binding affinity was not lost for either RXFP3 or 4, indicating that Arg^{26} and Trp^{27} are important for receptor activation, but not essential for binding [44].

RXFP1: RAAPYGVRLCGREEIRAVIFTCGGSRW
RXFP3: RAAPYGVRLCGREEIRAVIFTCGGSRW
RXFP4: RAAPYGVRLCGREEIRAVIFTCGGSRW

Figure 4. Residues on the relaxin-3 B-chain important for binding and activation of the various RXFP receptors. Residues important for binding are shown in red, and residues crucial for activation of RXFP3 and 4 in green.

The N-terminus was shown not to be important for binding as seven residues could be removed without significant loss of binding or activity. In all of these peptides, the A-chain was left unchanged [44]. The B-chain residues important for binding and activation of the various RXFPs are summarised in Figure 4.

Although relaxin-3 and analogues containing both the A- and B-chains are able to bind to and activate all the RXFP-receptors, removing the A-chain changes this situation. Binding and activation of RXFP1 and 2 requires both the A- and B-chains of the ligands [36, 40], while the B-chain alone can activate RXPF3 and 4 [36,

40, 45]. This, along with the unchanged activity on RXFP1 on altering Arg^{26} and Trp^{27}, suggests that the activation domain for RXFP1 and 2 is located on the A-chain [36, 44]. As the primary binding site of RXFP1 and 2 is located in the leucine-rich extracellular domain of the receptor, it has been suggested that this secondary interaction, which activates the receptor involves the A-chain binding to the short extracellular loops [4, 32, 46].

Not much is known about the binding of INSL5 to RXFP4, and as it is difficult to synthesise [47], no analogues have been made. Based on expected similarity to the binding of relaxin-3 to RXFP4, it is assumed that Arg^{5}, Ile^{12}, Arg^{13}, Ile^{16} and Tyr^{17} in INSL5 are important for binding to RXFP4 [7, 8].

INSL5 has the same C-terminal residues Arg and Trp as relaxin-3. These residues are in relaxin-3 the activation domain for both RXFP3 and 4. Interestingly, INSL5 activates RXFP4, but works as an antagonist on RXFP3 [7]. It has been suggested that this is because of structural differences between relaxin-3 and INSL5; in INSL5 the B-chain α-helical segment continues to the C-terminal Trp^{24}, whereas it in relaxin-3 ends five residues earlier, with Cys^{22}, placing the key Arg-Trp motif in a different position [7].

1.2.5. Relaxin peptide analogues

Due to the important physiological roles of relaxins there is a clear potential for RXFP-receptor modulators as novel drugs. However, this requires issues both relating to the complex structure and the lack of selectivity of the native peptides to be addressed. Developing selective agonists and antagonists for the RXFP-receptors has become an intriguing field of research, both in order to get new tools to further investigate physiological roles, and to find potential drug leads [45, 48].

Relaxin-2 can be made specific for RXFP1 by truncation of TrpB28, which is required for RXFP2 binding [40, 43]. Making selective relaxin-3 analogues on the other hand, is more complex. The need for selective relaxin-3 analogues is especially important for further studies of its functional roles. Because of relaxin-3's ability to activate all known RXFP-receptors, and because RXFP3 and RXFP1 are both expressed in the brain, selective ligands are important to study its effects without the results being confounded by cross binding to RXFP1 [44, 45, 49].

Another aim in the design of new analogues is designing simpler peptides, to make synthesis easier and cheaper [45].

Liu et al. [48] developed a chimeric peptide that selectively activated RXFP3 over RXFP1. The chimera R3/I5 consisted of the A-chain from relaxin-3 and the B-chain from INSL5 [44, 48].

In 2007 Kuei et al. reported the discovery of a selective antagonist for RXFP3 [44]. The antagonist R3(BΔ23-27)R/I5 is a chimeric peptide consisting of the B-chain from relaxin-3 and the A chain from INSL5. In addition to having an extra C-terminal Arg, the B-chain Gly^{23} was replaced by Arg, and the C-terminus truncated (Gly^{24}-Trp^{27} deleted) [44].

The additional C-terminal Arg was not a planned design; it was discovered by chance that a non-native C-terminal Arg created a higher affinity antagonist, presumably by creating an additional interaction point [44, 45, 50]. Through *in vitro* pharmacological studies it was shown that R3(BΔ23-27)R/I5 binds to human RXFP3 with high affinity, but lacks affinity for human RXFP1. *In vivo* studies on rats showed that R3(BΔ23-27)R/I5 administered intracerebroventricularly blocks food intake stimulated by the selective RXFP3-agonist, R3/I5 [44].

Building on this research Haugaard-Kedström et al. recently discovered that using only the B-chain of relaxin-3, they were able to develop a selective, single-chain peptide antagonist for RXFP3 [45]. The antagonist R3 B1-22R is a truncated relaxin-3 B-chain, with the same C-terminal arginine [45] as in the R3(BΔ23-27)R/I5 peptide discovered by Kuei et al. [44]. The native cysteines have been replaced by serine residues in this single-chain antagonist [45]. The new peptide R3 B1-22R was shown to bind to RXFP3 with similar affinity as native relaxin-3, and testing on functional cyclic adenosine monophosphate (cAMP) assays showed that R3 B1-22R was indeed fully selective for RXFP3 [45].

Structural analysis by two-dimensional (2D) nucleic magnetic resonance (NMR) spectroscopy analysis suggested that the R3 B1-22R peptide lacks secondary structure, adopting a random coil conformation in solution [45]. This finding is consistent with previous assumptions that the A-chain supports the structure of the relaxin-3 B-chain, and removing the A-chain leaves the isolated B-chain with a compromised structure [44, 45, 49].

1.3. Helix-inducing strategies

To further develop the R3 B1-22R peptide towards clinical use, it is critical to incorporate modifications that reintroduce its native helical structure in order to improve stability, and possibly binding affinity. Several strategies have previously been applied to other peptides, including the use of helical capping sequences and grafting of key residues into stable scaffolds [51, 52].

1.3.1. Helix capping

Helices are stabilised by hydrogen bonds between *i* and *i-4* residues, but at the helix termini, this hydrogen bond pattern ends. Some amino acids are favoured in these end-positions because they can form interactions that compensate for the disrupted hydrogen bond pattern, and particular sequences of these particular amino acids are referred to as helix caps [53].

It was recently shown for the vascular endothelial growth factor (VEGF) peptide that incorporating capping sequences at the start and end of a helical region that displayed key residues for receptor binding was able to induce a stable helical structure, without need for the rest of the peptide. This analogue retained the high affinity of the native peptide for the VEGF receptor [51]. In this study, single-chain relaxin analogues using a similar capping sequence as in the VEGF peptide were designed to investigate whether this approach is applicable for relaxins.

1.3.2. Using chlorotoxin as a scaffold

A promising approach for designing bioactive peptides with desirable properties is to identify key residues for the function of a particular peptide, and then graft these residues onto a different peptide with similar structure but more desirable chemical properties, such as high stability [52].

Chlorotoxin is a 36 amino acid neurotoxic peptide from the venom of the scorpion *Leiurus quinquestriatus* [54]. Upon discovery it was thought to selectively bind to chloride channels, hence the name chlorotoxin [54]. Synthetic chlorotoxin is also known as TM601 [55]. Recently it has been discovered that the molecular target for the chlorotoxin peptide is Annexin A2, and not chloride channels as originally thought [56].

Chlorotoxin has four disulfide bonds, giving it a highly folded, stable structure [57, 58]. It has an overall structure of a small three-stranded antiparallel β-sheet packed against an α-helix. Three of the disulfide bonds cross link the α-helix to the β-sheet, while the fourth disulfide links the N-terminal to the rest of the molecule [57] (figure 5).

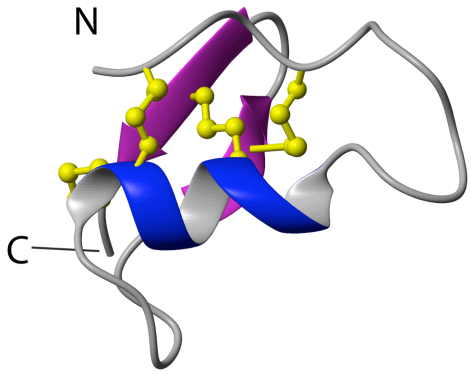


Figure 5. Solution NMR structure of chlorotoxin. Disulfide bonds are shown in yellow.

Chlorotoxin has tumour-binding properties, and is able to bind specifically to various tumour cells. Because of this, chlorotoxin has been combined with various molecules to create probes to aid in localisation of tumours during surgery [55]. Among the tumour imaging tools currently under investigation is a ^{131}I –radiolabeled chlorotoxin analogue, and a CTX:Cy5.5 dye bioconjugate [55].

As the putative drug target RXFP3 is mainly expressed in the brain, chlorotoxin's abiliy to cross the blood-brain barrier makes it highly desirable as a scaffolding molecule. In an alanine scan of chlorotoxin, all mutants were assumed to cross the blood-brain barrier as they showed similar activity to the native peptide (Muharrem Akcan, unpublished data). An analogue with chlorotoxin used as a scaffold might therefore still be able to penetrate the blood brain barrier. In addition to the physiological properties, chlorotoxin has an α-helix with a suitable size for incorporating the important residues for binding to RXFP4 [57]. All of these properties make chlorotoxin an intriguing molecule to use as a scaffold for a relaxin-3 analogue.

1.4. Combinational libraries

The finding that introducing a non-native arginine at the C-terminus of a truncated relaxin-3 B-chain increases the affinity of this antagonist analogue > 100-fold [44, 45] highlights that small changes, including replacement of single amino acids can have a major impact on the affinity of a ligand if a favourable new contact point with the receptor is introduced. Methods have been described to rationally search for such favourable mutations, including the synthesis of combinational libraries where mixtures are coupled at one or several positions in the sequence [59]. By testing the mixtures, or fractions thereof, the identity of high affinity analogues can be identified if present.

Several residues are exposed adjacent to the primary binding residues in relaxin-3 and its single-chain antagonist, and there is potential to further improve the binding to RXFP3 by modifying these. In this study we investigated the effect of replacing Glu^{13} in the R3 B1-22R antagonist with 16 different amino acids (excluding cysteine, methionine, glycine and proline).

1.5. Aims for the project

In this project we aimed to design relaxin peptide analogues where a helical structure is induced around the receptor-binding region. The hypothesis is that a stable helical structure resembling the native peptide will give higher affinity analogues.

- Capping sequences were used to induce helices in short single-chain analogues. Capped single-chain analogues were designed as relaxin-3 agonist, relaxin-3 antagonist, relaxin-2 antagonist, INSL5 agonist and INSL5 antagonist.
- Chlorotoxin was used as a scaffold, and an analogue with the amino acids important for relaxin-3 binding to RXFP3 were incorporated into the existing helical region of chlorotoxin.

Additional aims were:

- To investigate the importance of Arg^{8} in a single-chain analogue by replacing Arg^{8} in the R3 B1-22R antagonist with alanine, and to probe for other mutations that might give an increase in affinity similar to the non-native Arg^{23} in R3(BΔ23-27)R/I5 and R3 B1-22R by using combinational libraries.

2. Results

2.1. Capped relaxin analogues

2.1.1. Design

The design of the capped relaxin analogues was based on a capped analogue of the VEGF peptide, in which a stable helix was successfully introduced by incorporating helix capping sequences [51]. In this study, we have investigated whether this approach might also be applicable to single-chain analogues of relaxins. Our hypothesis was that if a stable helical structure can be achieved, it will give a receptor modulator with higher affinity than previous single-chain analogues, because it will more closely resemble the structure of the native peptides.

As shown in Figure 6, the N-terminal capping sequence KLT was introduced in all of the analogues, based on the success of the VEGF peptide analogue. The C-terminal capping sequence KGI is used as in the VEGF peptide for the relaxin-2 and INSL5 antagonists, whereas it is offset by one residue for the relaxin-3 and INSL5 agonists. This was done to utilise the existing glutamic acid's position in the sequence which would be expected to further support the capping sequence. In the relaxin-3 antagonist analogue, the C-terminal capping sequence was left out to avoid affecting the non-native C-terminal arginine known to provide an extra interaction point with the receptor [44, 45, 50]. The introduced glutamic acid in position six, as well as the leucine and the glutamine, are also found to favour helix stabilisation, due to the common occurrence of these residues in the positions relative to the ends of helical structures.

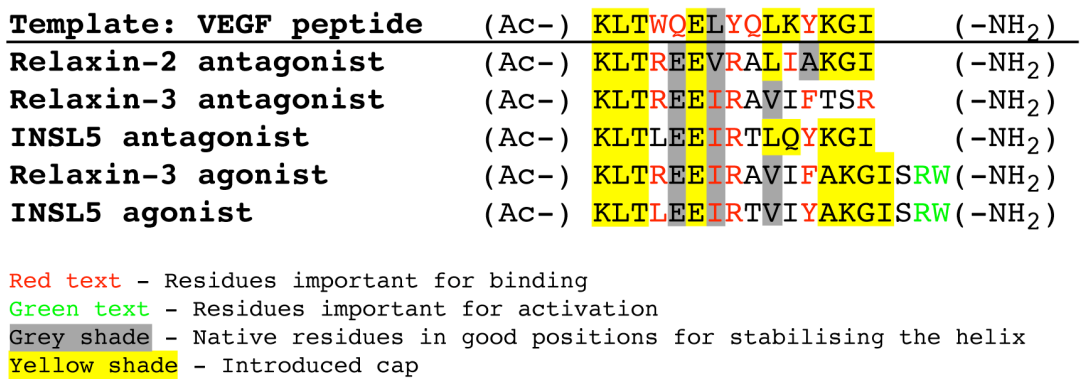


Figure 6. Amino acid sequences of capped analogues designed to act as agonists and antagonists on RXFP receptors. Capping sequences are adapted from a capped analogue of the VEGF peptide.

2.1.2. Synthesis and cleaving

Synthesis of all capped relaxin analogues was done using 9H-fluoren-9-ylmethyloxycarbonyl (Fmoc) chemistry on a Rink-Amide MBHA resin. All peptides except the INSL5 agonist were synthesised on an automated CEM Liberty microwave synthesiser. The INSL5 agonist was synthesised manually with an average coupling yield of approximately 99.4 %.

The peptides were cleaved from the resin using Trifluoracetic acid (TFA), and Electrospray mass spectrometry (ES-MS) was used following cleavage to confirm that the correct peptide was present from each synthesis.

For the relaxin-2 and INSL5 antagonists, the major peaks in the mass spectrum corresponded to the desired product. For the relaxin-3 antagonist, relaxin-3 agonist and the INSL5 agonist the major peaks belonged to side products. As the peptides were successfully purified with acceptable yield, and due to time constraints, the identity of the side products was not determined. For the INSL5 agonist the yield after purification was low, which might be a result of low synthesis yield due to side product formation.

The INSL5 antagonist, relaxin-2 antagonist, relaxin-3 antagonist analogues were synthesised and cleaved from the resin by Chia Chia Tan.

2.1.3. Purification

Using reverse phase high performance liquid chromatography (RP-HPLC), all peptides were successfully purified with the following results:

Relaxin-2 antagonist

The peptide eluted after approximately 36 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column (see section 5.4.1). The mass of pure peptide was 56.69 mg, and the yield 33.7 %. ES-MS spectra and HPLC chromatogram for the relaxin-2 antagonist are shown in Figure 7. The spectra and chromatograms for the other capped analogues are found in the appendix, Figure A1-A2.

Relaxin-3 antagonist

The peptide eluted after approximately 38 minutes with on a 1% gradient, 8 ml/min program, using a preparative C18 column. The mass of pure peptide was 3.08 mg, and the yield 1.7 %

INSL5 antagonist

The peptide eluted after approximately 41 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column. The mass of purified peptide was 25.7 mg, and the yield 12.7 %.

Relaxin-3 agonist

ES-MS after cleaving of the relaxin-3 agonist showed that masses corresponding to the peptide were present in both the A/B phase and the ether phase. Both the A/B- and the ether phase were purified by RP-HPLC. The ether phase gave the purest fractions after the first purification, and was further purified.

The peptide eluted after approximately 44 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column. The mass of purified peptide was 2.68 mg, a yield of 1.4 %.

INSL5 agonist

The peptide was purified by RP-HPLC, and needed two purifications to be sufficiently pure. The peptide eluted after approximately 47 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column. The mass of pure peptide was only 0.04 mg, a yield of 0.012 %.

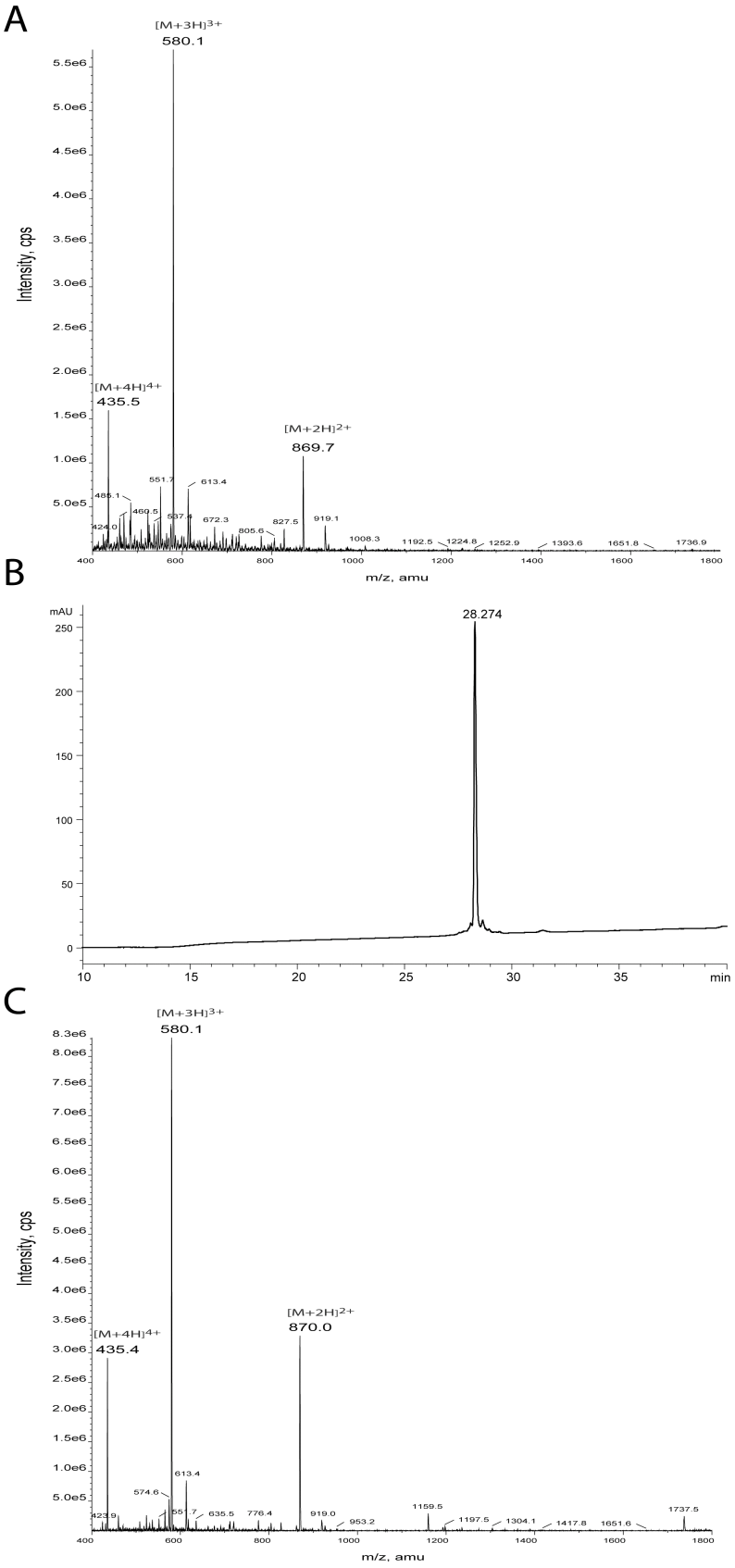


Figure 7. Purification and characterisation of the capped relaxin-2 antagonist. A) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. B) RP-HPLC trace of the purified peptide. C) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.

2.1.4. Structural studies

Structural analysis of the capped relaxin analogues was performed by recording and assigning 2D NMR spectra of the purified peptides. The spectra were assigned using the sequential assignment procedure [60]. Figure 8 shows the total correlation spectroscopy (TOCSY) and nuclear Overhauser effect spectroscopy (NOESY) spectra recorded for the relaxin-2 antagonist. The individual amino acids were identified in the TOCSY spectrum (figure 8, A), and the connections between the residues determined by the NOESY cross-peaks (figure 8, B).

Non-sequential cross-peaks were observed in the NOESY spectrum; 7Hα to 10HN and 3Hα to 5HN are labelled in figure 8. Additional non-sequential peaks were observed for the relaxin-2 antagonist, as well as in the spectra of some of the other analogues (data not shown).

To get an indication of the structure of the peptides, secondary Hα chemical shifts were calculated. Secondary Hα chemical shifts are the difference between the observed Hα-shift of a residue and the chemical shift of the same residue in a peptide with a random coil structure [61]. A sequence of negative shifts is indicative of an α-helix [62]. Figure 9 shows the calculated secondary Hα chemical shifts for the capped relaxin analogues.

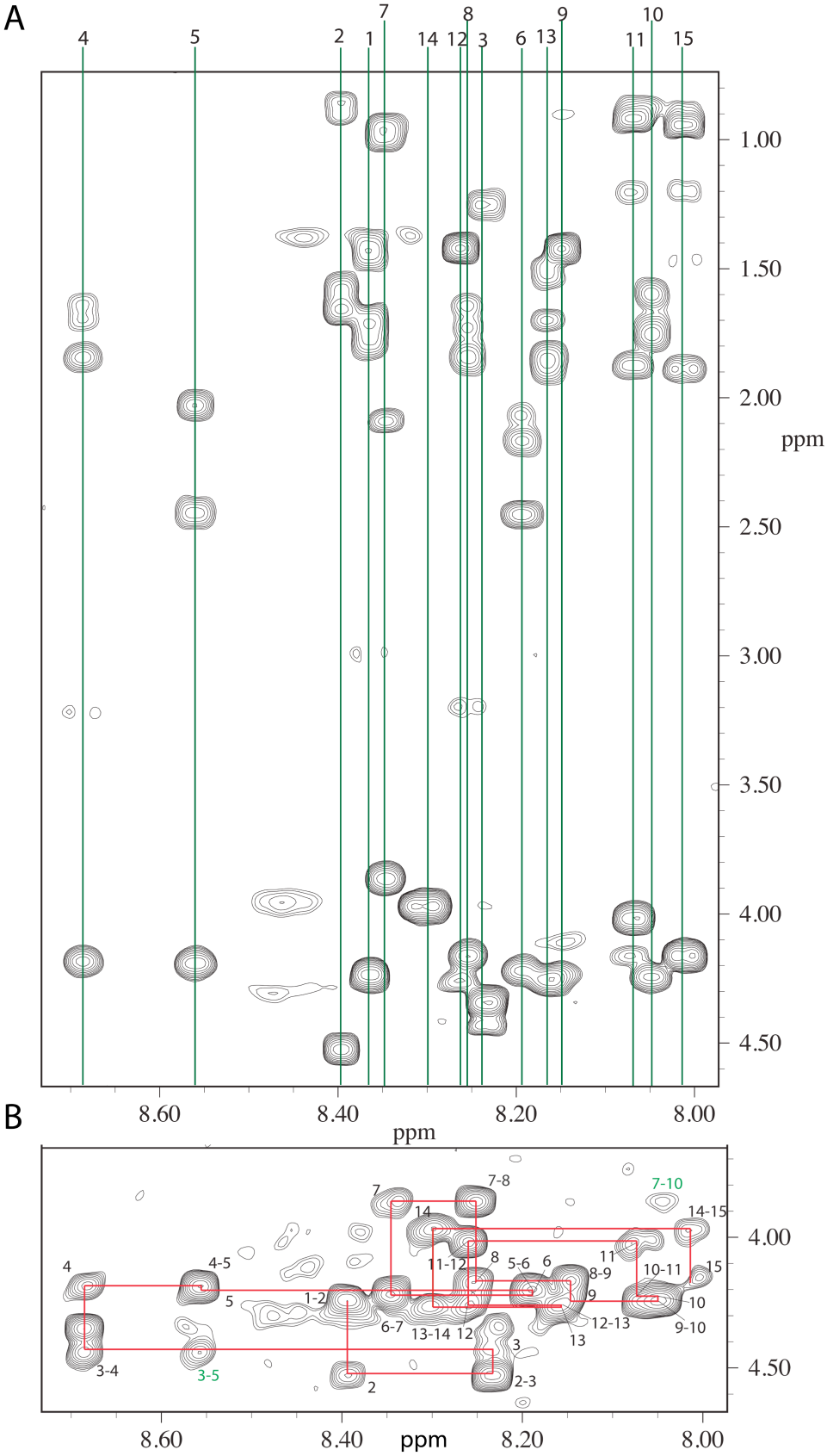


Figure 8. NMR spectra of the relaxin-2 antagonist. A) TOCSY spectrum with the individual spin systems indicated by green lines. Residue numbers are given by numbers above the lines. B) NOESY spectrum showing the sequential cross-peaks used in the spectral assignment.

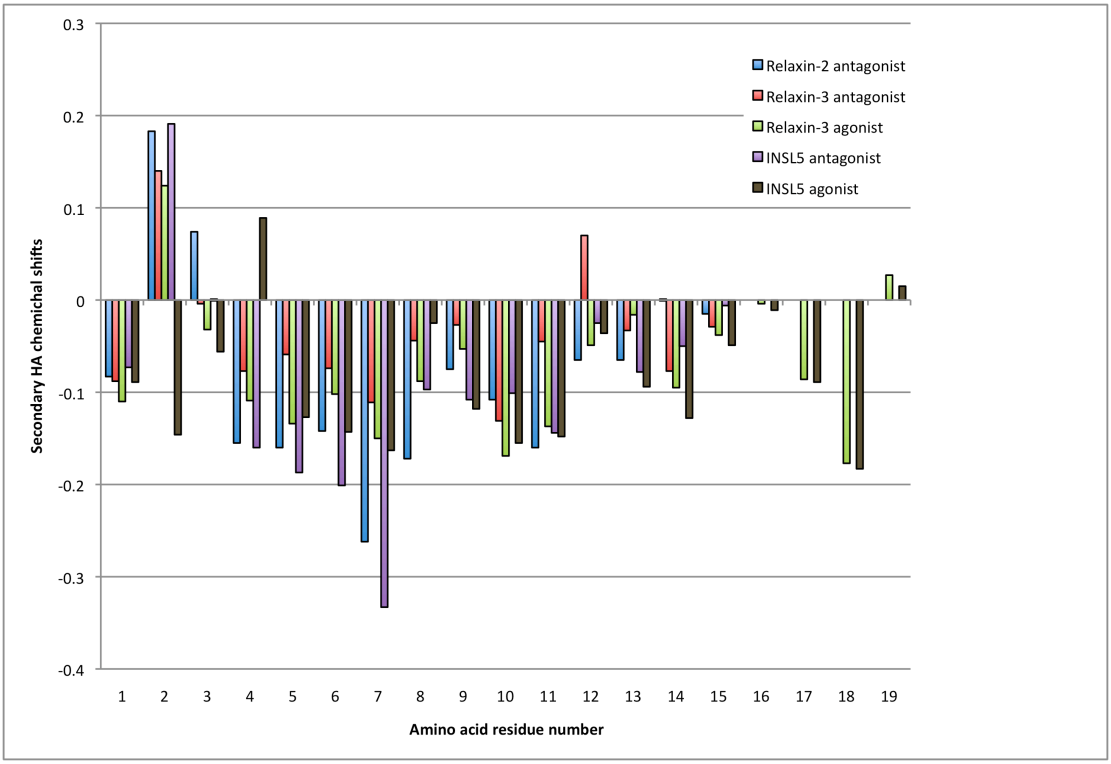


Figure 9. The secondary Hα chemical shifts of the capped relaxin analogues. The secondary Hα shifts were calculated by subtracting the random coil shifts from the experimental Hα shifts.

From the secondary Hα chemical shifts shown in Figure 9, the capped analogues display a general tendency towards negative shifts, indicating a helical structure. The relaxin-2 and INSL5 antagonists display particularly negative shifts from residue four to seven.

The secondary structure of the relaxin-2 antagonist was calculated (Figure 10) in order to determine if the helical tendencies seen from the secondary Hα chemical shifts correlated with a well defined helical structure. The structure did show two helical turns close to the N-terminus, whereas the C-terminus is mainly unstructured.

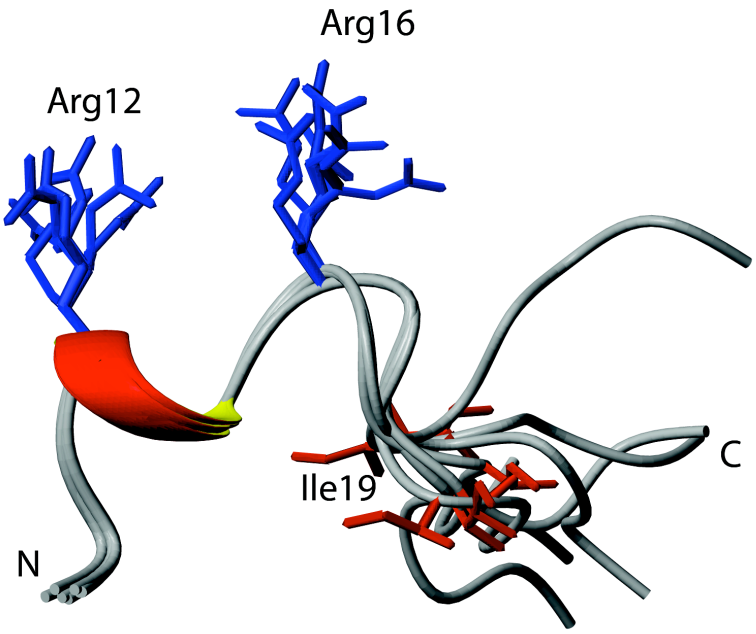


Figure 10. NMR structure of the capped relaxin-2 antagonist. The arginines 13 and 17 (blue) and the isoleucine 20 (red) in the relaxin binding motif are labeled according to their position in native relaxin-2

2.1.5. Binding and activity

In competitive binding assays on cells expressing RXFP-receptors, all of the capped analogues showed little or no binding to the receptors (Figure 11). Consistent with the lack of binding, the INSL5 analogues also failed to show any activity, measured by the ability to reduce forskolin induced cAMP activity in a cell based activity assay (Figure 12). Unfortunately, the activity assay on the relaxin-3 agonist was not completed in time to be included in this thesis.

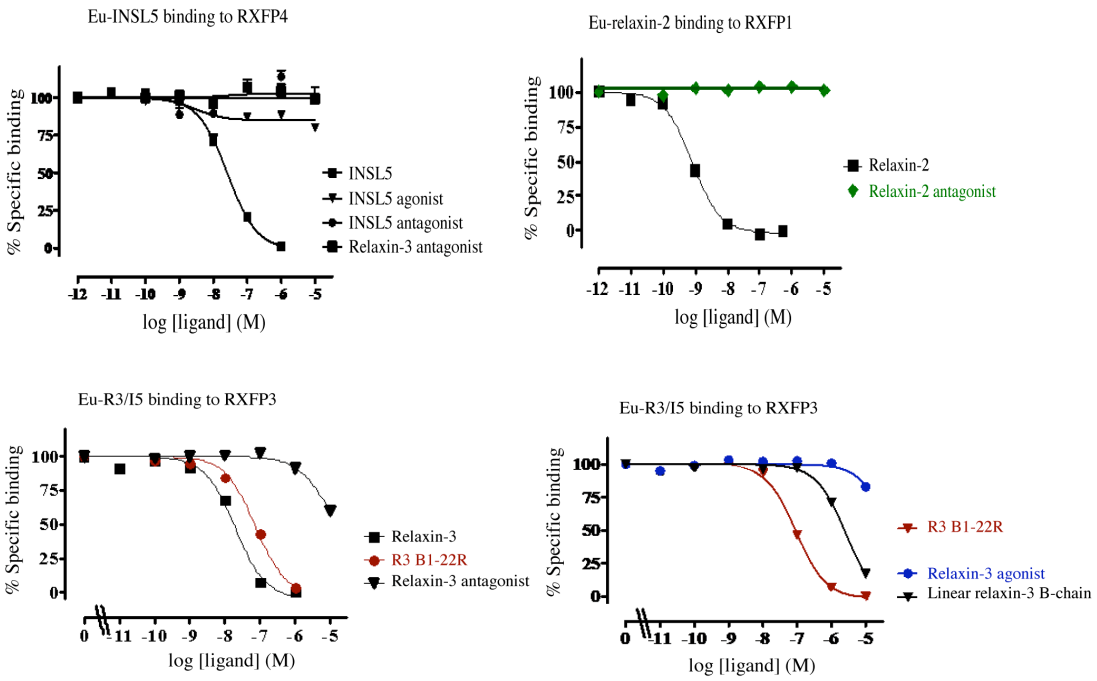


Figure 11. Competition binding curves for capped relaxin B-chain analogues in comparison to Eu-R3/I5, Eu-relaxin-2 or Eu-INSL5.

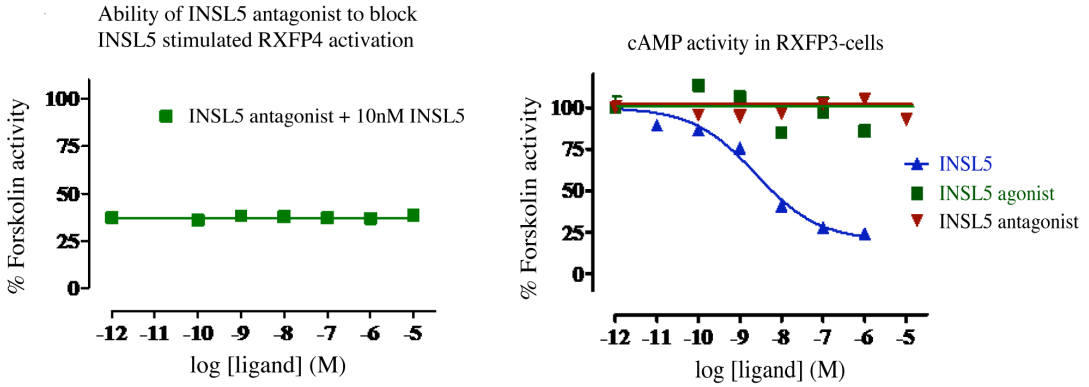


Figure 12. Agonist or antagonist activites of capped INSL5 analogues, measured by changes in cAMP activity.

2.2. R3 B1-22R R8A analogue

2.2.1. Synthesis and purification

In light of the lack of activity of the capped analogues, we considered whether additional features outside the helical segments were involved in the receptor interaction of relaxin-3. Arg^{8} is a conserved residue, and appears to play a role in the activity of native relaxin-3. Therefore we decided to make an alanine-mutated variant testing this hypothesis.

The synthesis of the R8A mutant of the R3 B1-22R analogue was performed on an automated CEM Liberty microwave synthesiser using Fmoc chemistry. The peptide was cleaved from the resin using TFA.

ES-MS after cleaving showed that the masses corresponding to the peptide were present in both the A/B phase and the ether phase. Both phases were freeze-dried, and purified by RP-HPLC. The ether phase yielded sufficiently pure fractions after one round of purification. The mass of pure peptide was 23.64 mg, resulting in a yield of 12.7 %. ES-MS spectra and HPLC chromatogram are shown in the appendix, figure A3. The peptide eluted after approximately 46 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column.

2.2.2. Binding

Competitive binding assay on the R8A analogue of R3 B1-22R showed a loss of affinity in the alanine mutant. However, it still has affinity for RXFP3 (Figure 13).

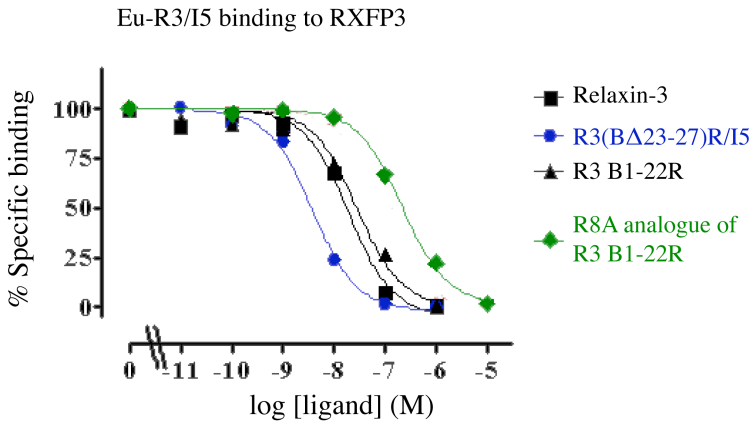


Figure 13. Competition binding curves for R8A analogue of R3 B1-22R in comparison to the parent peptide and native relaxin-3. Competitive binding was measured using Eu-R3/I5 in RXFP3 expressing cells

2.3. Relaxin-3/chlorotoxin analogue

2.3.1. Design

The relaxin-3/chlorotoxin analogue was designed by incorporating the residues known to be important for relaxin-3 binding to RXFP3 into the existing helical segment of chlorotoxin (figure 14). In light of our results showing that the R8A mutant of the R3 B1-22R peptide had decreased affinity compared to the parent peptide, an arginine was incorporated in a position outside the helix, similar to that

found in position eight of native relaxin-3. The sequence of native chlorotoxin compared to the relaxin-3/chlorotoxin analogue is shown in Figure 14.

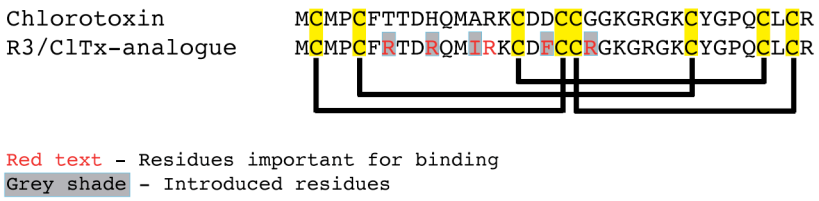


Figure 14. The amino acid sequence of native chlorotoxin and an analogue where the residues important for binding to RXFP3 have been grafted into chlorotoxin. Disulfide bonds in the native fold of the peptide are indicated by connecting lines.

2.3.2. Synthesis and purification

The synthesis of the chlorotoxin analogue was performed on an automated CEM Liberty microwave synthesiser using Fmoc chemistry. The peptide was cleaved from the resin using TFA. ES-MS on the crude peptide showed major peaks in the mass spectrum that corresponded to the mass of the peptide.

The peptide was purified by RP-HPLC on a preparative C18 column, using a 1% gradient before folding. The peptide eluted after approximately 38 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column. The mass of reduced, purified peptide was 86.38 mg in total. This gives a yield of 8.7% for the reduced peptide.

2.3.3. Folding

A small-scale folding trial was done, investigating a range of folding conditions. The peptide concentration in all buffers was 0.1 mg/mL. The following conditions were used:

1 0.1 M Tris-HCl
0.2 M NaCl, 5 mM reduced glutathione
0.5 mM oxidized glutathione
pH 7.8
2 0.1 M NH_4OAc
pH 8.3
3 0.1 M $(NH_4)_2CO_3$
pH 8.2
4 0.33 M NH_4OAc
0.5 M GnHCl
5 0.1 M Na_3PO_4
pH 7.4
6 Buffer 1/isopropanol, 1:1
7 Buffer 2/isopropanol, 1:1
8 Buffer 3/isopropanol, 1:1
9 Buffer 4/isopropanol, 1:1
10 Buffer 5/isopropanol, 1:1

After 24 hours, analytical HPLC was used to analyse the samples. All samples showed the same general pattern with a sharp, early eluting peak, followed by a broad collection of peaks presumably containing misfolded peptide (figure 15). The buffer/isopropanol mixtures showed the lowest intensity of the earliest eluting peak. After 48 hours, analytical HPLC was used to analyse the five samples without isopropanol. Results and conditions are summarised in figure 15. Again, the same pattern was seen, and the extra oxidation time did not seem to have affected the overall folding significantly.

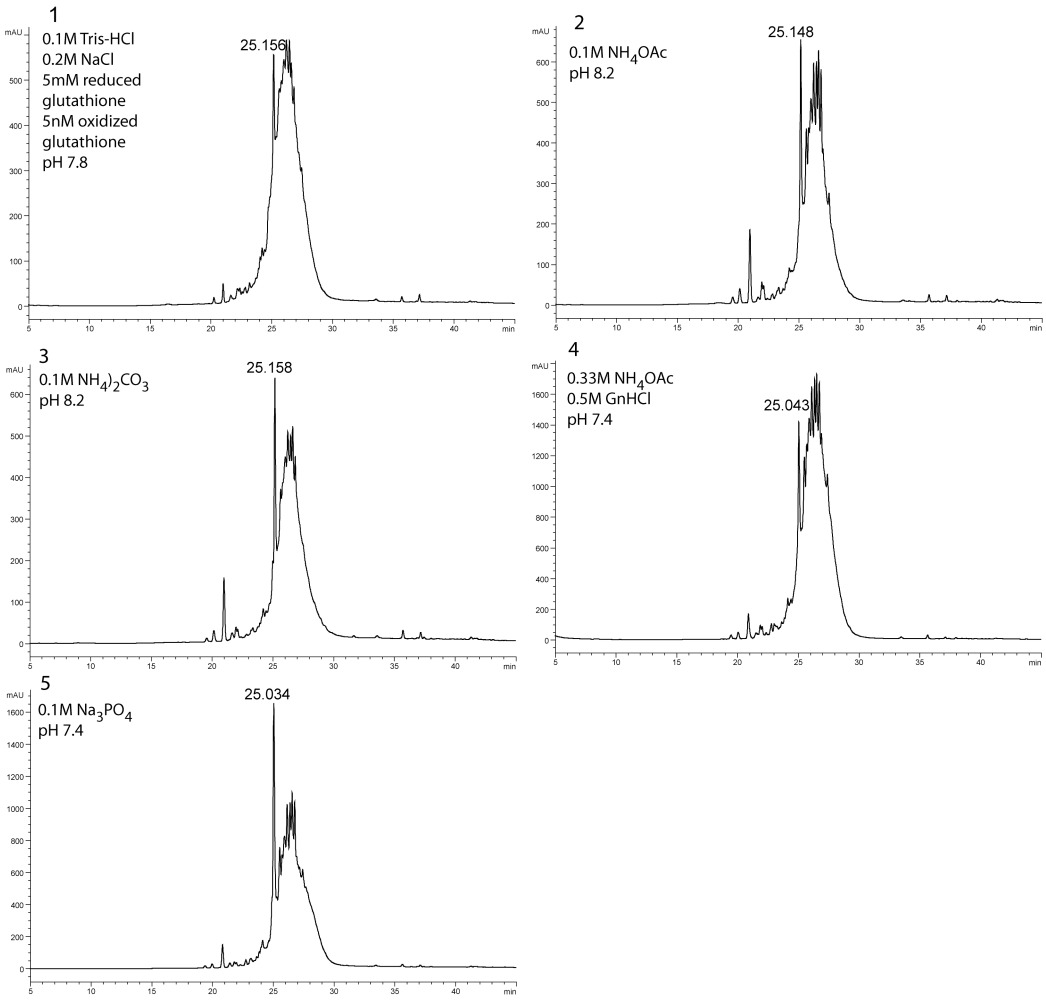


Figure 15. The oxidation conditions examined for the relaxin-3/chlorotoxin analogue. RP-HPLC chromatogram after a small scale oxidation trial.

In the small scale folding trial the 0.1 M sodium phosphate buffer (pH 7.4) seemed to give the most intense and cleanest peak in the HPLC-trace (figure 15).

A buffer with 0.1 M Tris-HCl, 0.2 M NaCl, 5 mM reduced glutathion, 0.5 mM oxidized glutathion (pH 7.8) has previously been successfully used on native chlorotoxin, and both these buffers were used for further folding of the peptide.

After folding, the peptide was purified using RP-HPLC on a 3 ml/min column and a 0.5 % gradient. After five rounds of oxididation and purification of 5-18 mg reduced, pure peptide at a time, a mass of 1.52 mg pure, oxidized peptide was obtained. A total of 62.65 mg of R3/CTX was oxidized. This gives a 2.43 % yield after oxidation, and a total yield of 0.2 %. Analytical RP-HPLC chromatograms and ES-MS spectra from the folding and purification are shown in Figure 16.

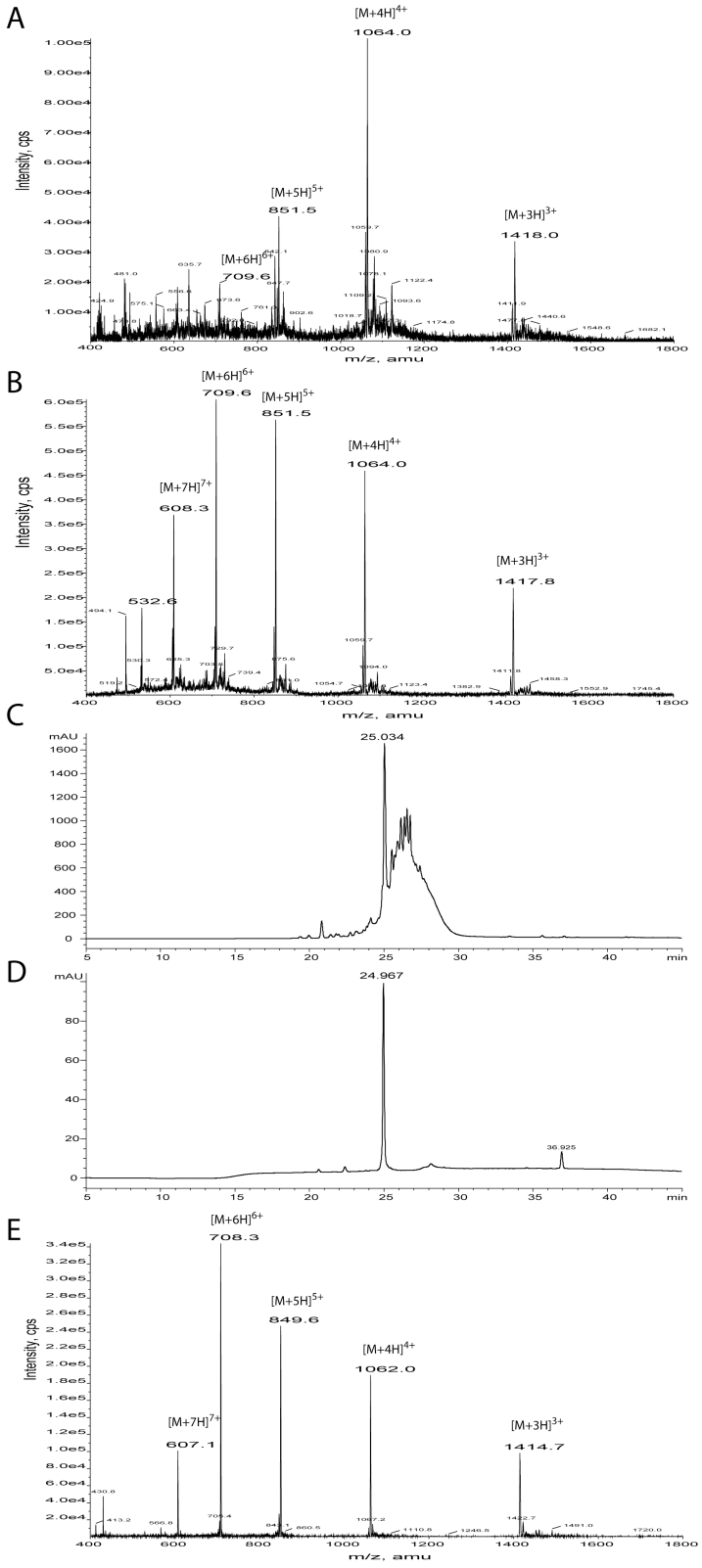


Figure 16. Purification and folding of relaxin-3/chlorotoxin analogue. A) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. B) ES-MS spectrum of purified reduced peptide. C) RP-HPLC chromatogram from oxidation trial using 0.1 M sodium phosphate buffer (pH 7.8). D) RP-HPLC chromatogram of the purified, oxidized peptide. E) ES-MS spectrum of purified, oxidized peptide, confirming that the correct peptide is present.

2.3.4. ^{1}H NMR of oxidized R3/CTX analogue

To check whether the major isomer of the peptide was correctly folded, the purified peptide was analysed by ^{1}H NMR (600 MHz). The recorded spectrum compared to that of native chlorotoxin is shown in Figure 17.

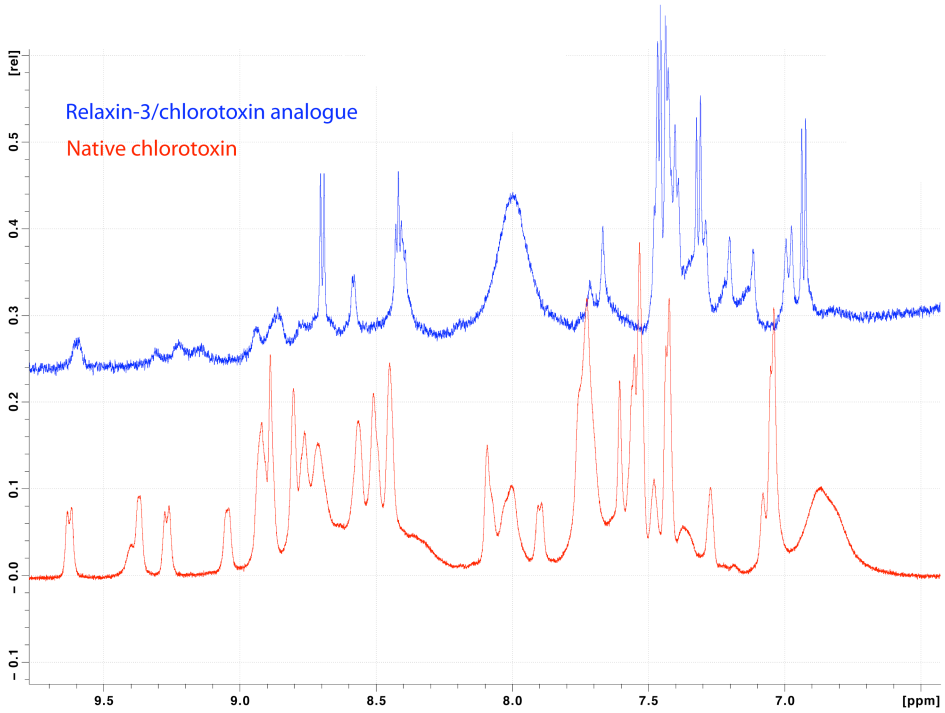


Figure 17. Comparison of the 1D NMR spectra of the relaxin-3/chlorotoxin analogue (blue) and native chlorotoxin (red). The spectra indicate that the relaxin-3/chlorotoxin analogue has adopted a chlorotoxin-like fold.

The sample analysed was small, giving very low intensity peaks, but through a large number of scans the signal-to-noise ratio was sufficiently increased to give an indication of the fold. As shown in Figure 19, the signals appear to be well dispersed, with several down field shifted peaks indicating a folded peptide.

2.4. E13 combinational library of R3 B1-22R

2.4.1. Design

Based on the accidental discovery that a non-native C-terminal arginine can create an additional contact point with the RXFP3 receptor [44, 45, 50], it was decided to do a targeted probing to potentially reveal other positions and mutations that could enhance receptor binding.

Glu[13] in the relaxin-3 B-chain is in a position close to other residues that are important for binding to RXFP3, and this residue was chosen as a position that might be used to create additional interaction points.

A combinational library of R3 B1-22R was synthesised using a mixture of 16 different amino acids (cysteine, methionine, proline and glycine excluded) in the position of Glu[13]. Cysteine and methionine were left out to simplify the process by eliminating potential problems with disulfide formation by cysteines in different peptide molecules, and oxidation of the methionine. Glycine and proline were excluded from the mixture because these residues would break any helical tendencies shown by the peptide, which would most likely cause a loss of affinity.

2.4.2. Synthesis and bioactivity guided fractionation

The E13 mix was synthesised by Richard Clark on an automated CEM Liberty microwave synthesiser, using Fmoc chemistry. The synthesis of the E13 analogues was done by using a mixture of amino acids at position 13, while following standard solid phase peptide synthesis procedures on the rest of the peptide. This resulted in a mixture of peptides differing in residue 13. The peptides were cleaved from the resin using TFA.

A bioactivity guided fractionation approach was used to separate the crude mixture. The results from this approach are summarised in figure 18.

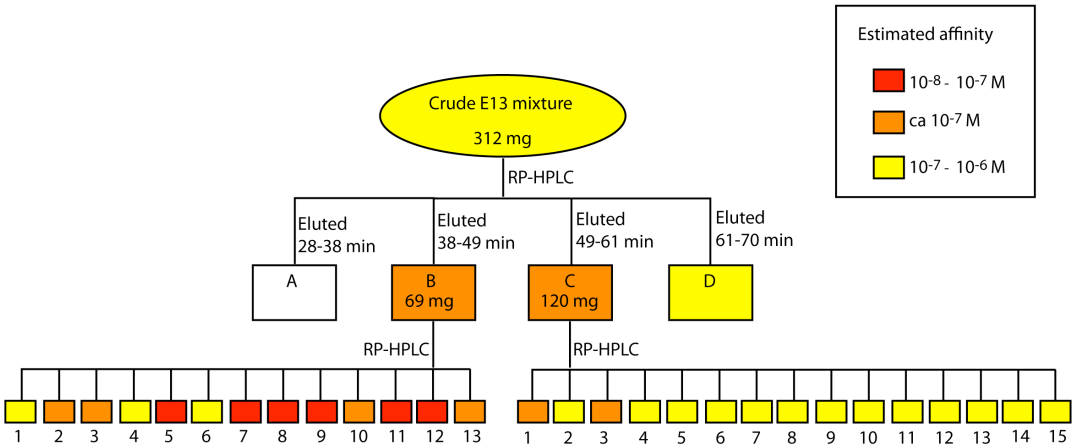


Figure 18. Summary of procedure and results from the bioactivity guided fractionation approach of the mixture of E13 mutants of R3 B1-22R.

Fractionation was based on retention times on RP-HPLC, and bioactivity determined by competitive binding assays (see section 5.7.2).

The binding assay on the mixture showed loss of binding compared to R3 B1-22R, but there was still affinity for RXFP3 (Figure 19). To determine whether this affinity was due to one high affinity analogue, or several weakly binding analogues it was decided to divide the mixture into rough fractions to do further binding assays. This was done by RP-HPLC. The RP-HPLC chromatogram showed a broad peak between 28-60 minutes (figure 20, A) and fractions collected in this period were roughly divided into four pools. The fraction collected between 28 and 48 minutes did not have any significant mass after freeze drying, and was not used any further. Binding assays were done on the other three fractions, and all three fractions showed increased affinity compared to the original mixture. However, fractions collected between 49 and 61 minutes, and between 61 and 70 minutes bound with slightly higher affinity (Figure 19).

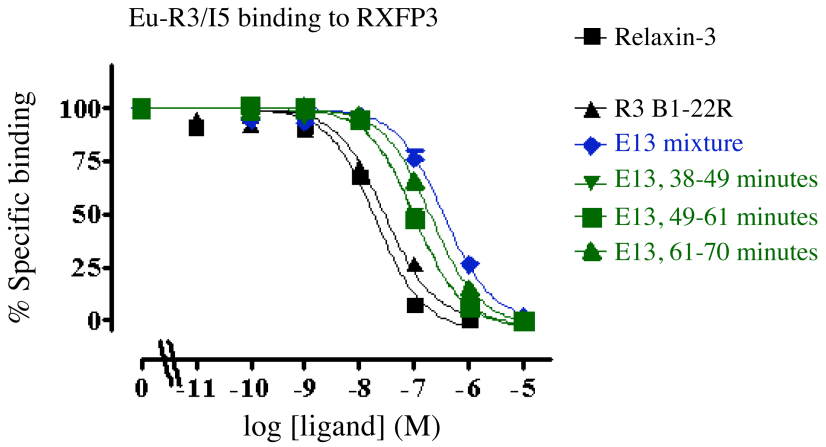


Figure 19. Competition binding curves for fractions after RP-HPLC on a mixture of R3 B1-22R analogues with mutations at position 13. Affinity shown in comparison to the parent peptide and the original crude E13 mixture. Competitive binding measured using Eu-R3/I5 in RXFP3 expressing cells

These fractions were then further divided by RP-HPLC on a 0.25 % gradient (Figure 20, B-C), and the fractions analysed by ES-MS, and combined according to the masses each fraction contained. Through ES-MS analysis we were also able to get an indication as to which analogue of R3 B1-22R that was found in the different fractions.

The fractions found to contain peptides were then tested for binding to RXFP3, and fractions B.5 and B.7 showed similar or slightly increased affinity compared to R3 B1-22R (figure 21).

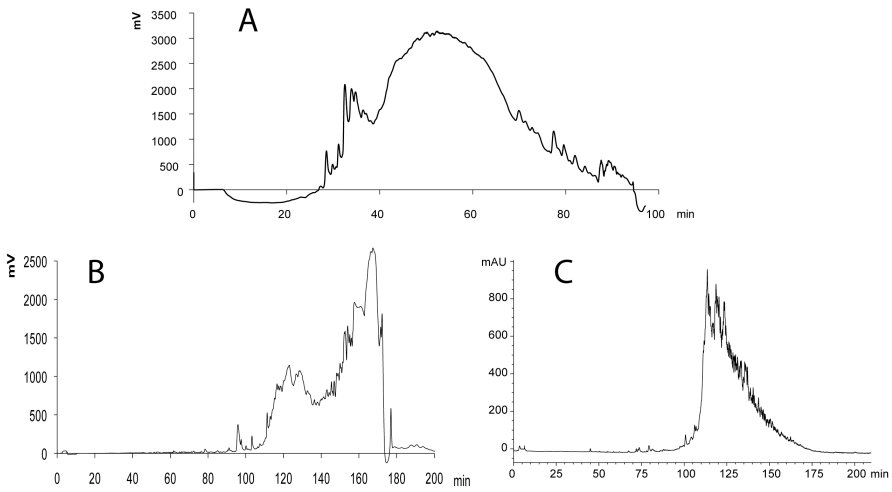


Figure 20. RP-HPLC traces of the combinational library of the R3 B1-22R peptide with mutations at the position of E13. A) The original E13 mixture B) fractions collected between 38 and 49 minutes, and C) fractions collected between 49 and 60 minutes during RP-HPLC of the original mixture.

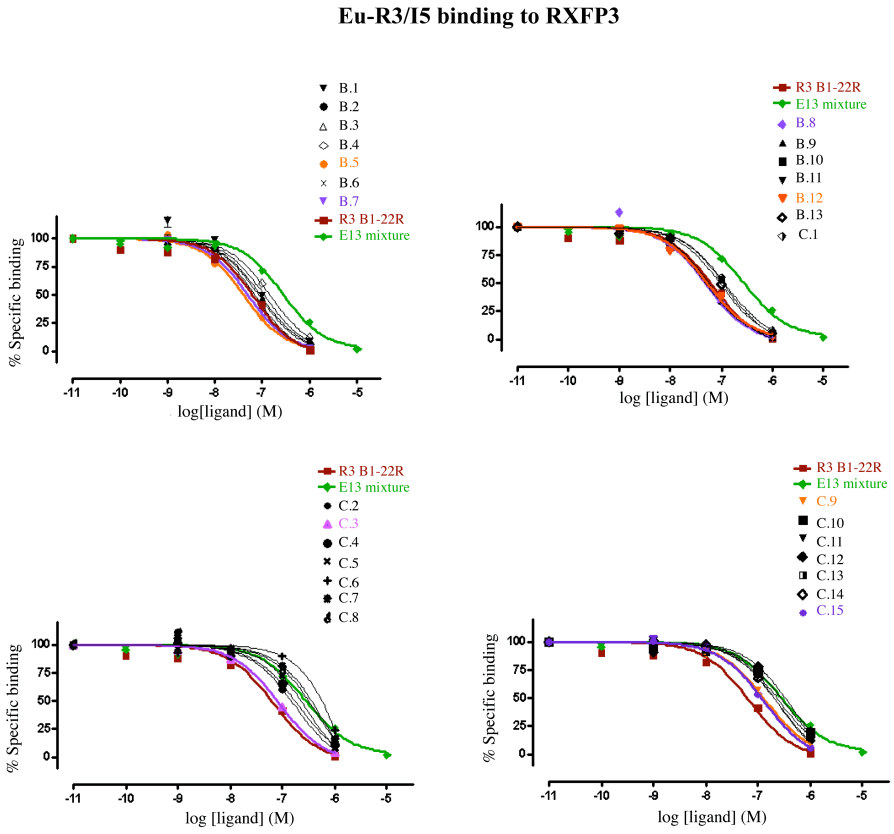


Figure 21. Competition binding curves for final fractions after RP-HPLC on a mixture of R3 B1-22R analogues with mutations on position 13. Affinity shown in comparison to the parent peptide and the original crude E13 mixture. Competitive binding measured using Eu-R3/I5 in RXFP3 expressing cells

2.4.3. MALDI-TOF-MS

Matrix-assisted laser desorption/ionization-time of flight (MALDI-TOF) MS was performed on the two fractions that showed a possible increase in affinity compared to R3 B1-22R, in order to determine which analogues were in those samples. Fraction B.7 showed two major peaks with monoisotopic masses of 2593.7 and 2620.7 (figure 22, D). These masses correspond to E13T and E13Q or E13K, respectively. Fraction B.5 had one major peak with a monoisotopic mass of 2562.7 (figure 22, B). This monoisotopic mass is one mass unit off compared to the theoretical monoisotopic mass of the E13A analogue, but as there are no other analogues with mass close to E13A (see table 1), this peptide is likely to be the E13A analogue.

Table 1

Peptide	Monoisotopic mass (amide)	Monoisotopic $[M+1H]^{1+}$	Average mass (amide)
E13G	2548,45	2549,45	2550
E13A	2562,46	2563,46	2564,02
E13S	2578,46	2579,46	2580,02
E13P	2588,48	2589,48	2590,06
E13V	2590,49	2591,49	2592,08
E13T	2592,47	2593,47	2594,05
E13L	2604,51	2605,51	2606,1
E13I	2604,51	2605,51	2606,1
E13N	2605,47	2606,47	2607,05
E13D	2606,45	2607,45	2608,03
E13Q	2619,48	2620,48	2621,07
E13K	2619,52	2620,52	2621,12
E13H	2628,48	2629,48	2630,08
E13F	2638,49	2639,49	2640,12
E13R	2647,53	2648,53	2649,13
E13Y	2654,49	2655,49	2656,12
E13W	2677,5	2678,5	2679,16
R3 B1-22R	2620,47	2621,47	2622,06

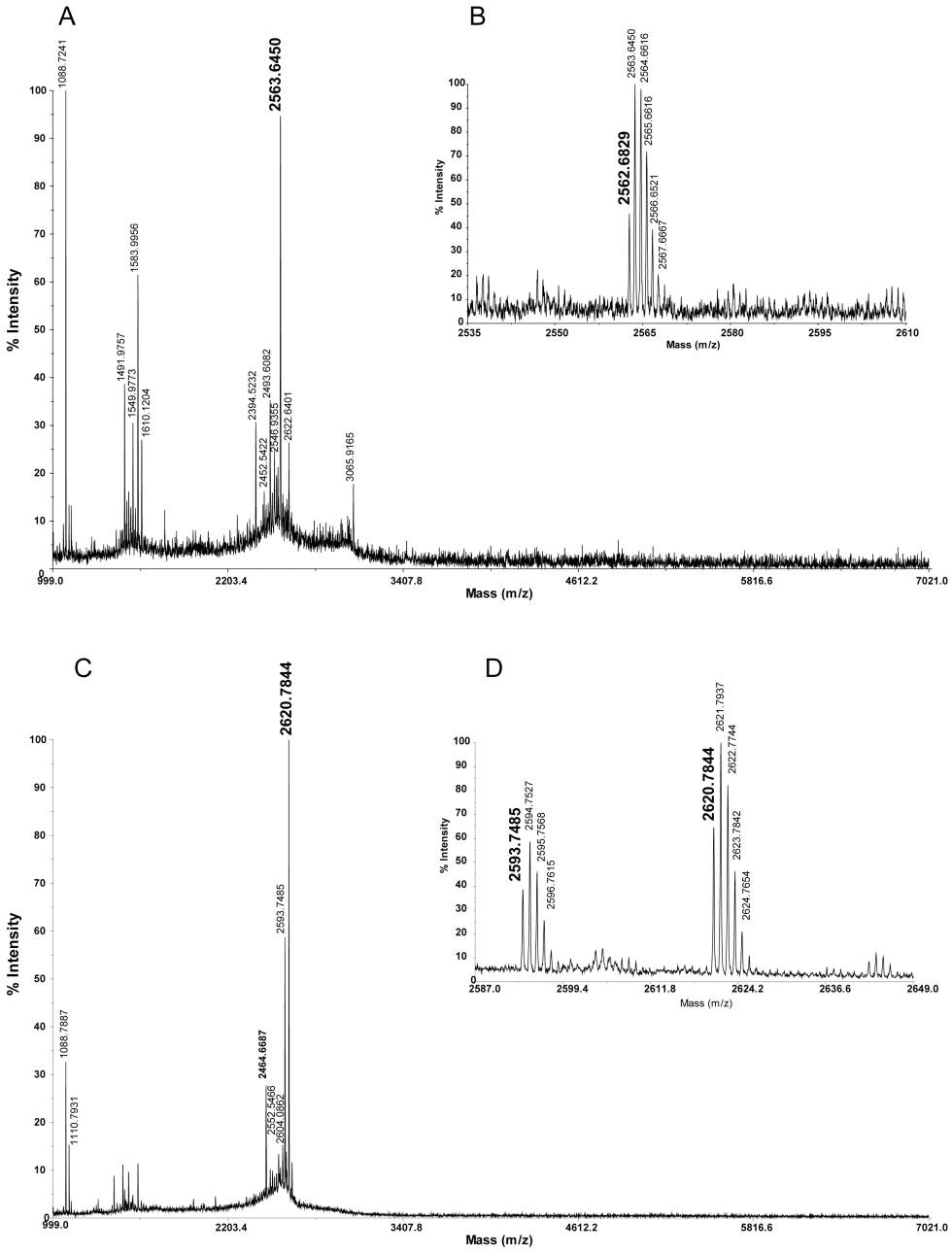


Figure 22. Spectra from MALDI-TOF-MS of fraction B.5 (panels A-B) and fraction B.7 (panels C-D). Panels A and C show the full spectrum, and panels B and C a selected range showing the isotope pattern of the E13 analogues present in the samples.

Both the samples we used for MALDI-TOF MS were impure, and as there were insufficient amounts for repurification, it was decided to synthesise one of the analogues in order to see if a pure sample would increase the affinity compared to the impure sample.

2.5. Synthesis, purification and binding of the E13A analogue of R3 B1-22R

The synthesis of the E13A mutant of the R3 B1-22R analogue was performed on an automated CEM Liberty microwave synthesiser using Fmoc chemistry. The peptide was cleaved from the resin using TFA.

Extraction of the peptide from the ether phase was not complete and ES-MS on both phases showed major peaks corresponding to the mass of the peptide. Both phases were freeze-dried and the peptide was successfully purified by RP-HPLC.

The peptide eluted after approximately 44 minutes on a 1% gradient, 8 ml/min program, using a preparative C18 column. After purification the mass of pure peptide was 13.68 mg, giving a yield of 4 %. ES-MS spectra and HPLC chromatogram are shown in the appendix, figure A3.

As shown in Figure 23, the competitive binding assay on the E13A analogue of R3 B1-22R seemed to bind with slightly higher affinity compared to the parent peptide. The affinity is very similar to that shown by fraction B.5 (Figure 21).

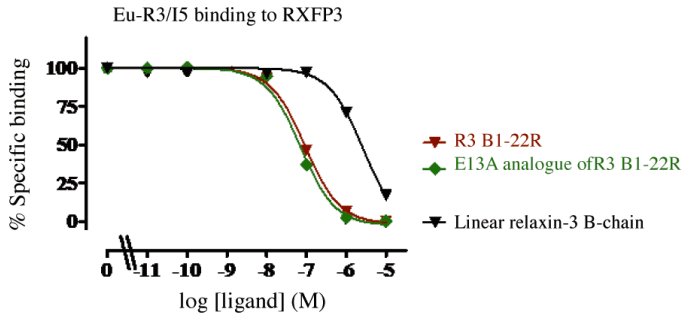


Figure 23. Competition binding curves the E13A mutant of R3 B1-22R. Competitive binding measured using Eu-R3/I5 in RXFP3 expressing cells

2.6. Synthesis and purification of R3 B1-22R

The R3 B1-22R analogue was synthesised to use as a positive control in the binding assays for RXFP3. The synthesis was performed on an automated CEM Liberty microwave synthesiser using Fmoc chemistry. The peptide was cleaved from the resin using TFA.

ES-MS on the crude peptide showed major peaks corresponding to the mass of the peptide. The peptide was purified by RP-HPLC, and three purification runs were needed to get a pure peptide. The peptide eluted after approximately 48 minutes on a

1% gradient, 8 ml/min program, using a preparative C18 column. The mass of purified peptide was 3.92 mg, giving a yield after purification of 2 %. ES-MS spectra and HPLC chromatogram are shown in the appendix, figure A4.

3. Discussion

Relaxins have a complex structure with two chains interlinked by disulfide bonds [4, 5, 40], making them difficult to synthesise [45]. They also show a complicated binding pattern to the relaxin family receptors, lacking high selectivity for individual receptors [8, 21, 30-34, 37]. As the RXFPs are promising drug targets for a range of diseases [4, 5, 13], designing simpler and selective analogues is desirable. One particularly promising analogue is the selective antagonist R3 B1-22R, which binds to RXFP3 with high affinity. This analogue is however mainly unstructured in solution [45]. To further develop single-chain analogues of the relaxin family peptides as drug leads, being able to reintroduce the helical structure of the native peptides is paramount. A structured peptide will have increased stability compared to a random coil analogue [63, 64], and reintroducing a native helical structure around the receptor-binding region is likely to improve affinity.

The receptor-binding site of relaxin-3 and INSL5 is believed to be located in the B-chain alone, but without the structural support of the A-chain these molecules have poor affinity for their receptors [36, 40, 45]. Therefore, in this study we have explored the possibility of introducing a stable helical structure around the receptor-binding region. We investigated the use of helix capping sequences for stabilising the B-chain key helical region in the absence of an A-chain. This strategy did induce some helicity, however it was not sufficient to generate high affinity analogues for any of the receptors. We have shown that this might in part be explained by the removal of Arg^8, which appears to be more involved in receptor interaction than previously believed. In a second approach, the key receptor binding residues of relaxin-3, including Arg^8, were grafted onto a stable disulfide-rich peptide; chlorotoxin. We were able to successfully synthesise the relaxin-3/chlorotoxin analogue, but the yield after oxidation was low and we were not able to get any information about structure or bioactivity within the time limits of the project.

Finally, by synthesising libraries of the RXFP3 antagonist R3 B1-22R, we investigated the possibility of identifying additional contact points close to the residues known to be involved in binding. We were able to successfully synthesise a library of peptides with mutations in position 13, and by a bioactivity guided fractionation approach and MS we were able to separate out the more active analogues. However, we did not find any mutants with especially high activity.

3.1. Capped relaxin analogues

Capping sequences that have previously been successfully applied on the VEGF peptide [51] were used as a template in designing capped analogues of relaxin-2, -3 and INSL5. For the VEGF peptide these caps were sufficient to introduce a fully stable helical conformation [51], suggesting that a similar strategy may be applicable to relaxins. The capping sequences were adapted around the residues present in the binding region of the B-chains of the selected peptides. 2D NMR spectroscopy experiments were used to get an indication of secondary structure in the peptides. All of the capped analogues displayed generally negative secondary shifts, indicating a helical structure. Compared to the secondary Hα chemical shifts seen in the R3 B1-22R peptide [45], there is a general tendency towards increased helical properties in the capped analogues. However, the helical tendency is not as strong in the native relaxins. A few non-sequential nuclear Overhauser effects (NOE) were observed in the spectra, however not to the extent expected for a stable helical conformation, and what is seen in the native two-chain relaxin peptides.

Of the analogues showing the most helical structure, the spectra of the capped relaxin-2 antagonist appeared to be of the highest quality. The spectra recorded for the relaxin-2 antagonist were therefore used to calculate a three dimentional (3D) structure of the peptide, in order to further investigate the effect of the capping sequences on the secondary structure. From the structure of the relaxin-2 antagonist, we learned that the N-terminal capping sequence had introduced two helical turns to the peptide. This led to the two arginines in the relaxin-binding motif being positioned correctly; exposed next to each other on the same face of the helix. The C-terminal cap, on the other hand, was not able to introduce structure. This resulted in disruption of the helix, with the position of the isoleucine essential for binding being positioned significantly different compared to the native peptide.

All analogues were tested in cell-based assays for binding and activation of their respective receptors. However, in the competitive binding assays performed on Chinese hamster ovary (CHO) and Human embryonic kidney (HEK2293) T cells stably transfected with the receptors, only the relaxin-3 antagonist and agonist showed any affinity for its receptor, and even for this analogue the binding was very poor. Functional cAMP assays on the INSL5 analogues showed no response of either the

agonist or the antagonist. Unfortunately, the functional cAMP assay was not completed in time to be included in this thesis.

Relaxin-2 is known to have residues involved in binding in both the A and the B-chain [36, 39, 41, 65], making design of single-chain analogues more complicated. The capped single-chain analogue was therefore never expected to bind with similar affinity as the native peptide. However, if a stable helix is successfully introduced, it should in theory have measurable affinity, and be a good starting point for further modifications towards a drug lead.

For relaxin-3 and INSL5, all residues important for binding are located in the B-chain [7, 8, 36, 40, 41, 45]. Based on this, a single-chain analogue containing the residues important for binding to the receptors should be able to bind at least weakly, even if it lacks secondary structure. Consequently, the extent of the loss of binding in the capped analogues was somewhat surprising.

The main difference between the R3 B1-22R peptide and our capped analogues, apart from the capping sequences, was an N-terminal truncation of residues 1-8. The only residue in the removed sequence that has been shown to affect binding of relaxin-3 is Arg^8. The binding of the relaxin-3 agonist and antagonist is even poorer than that seen for the relaxin-3 B-chain alone. This further supports the hypothesis that the lack of binding is due to more than a lack of structure. Arg^8 has previously been thought to affect binding mainly by stabilising the fold of the native peptide, not by binding to the receptor [44]. This led to a discussion of whether Arg^8 is more involved in binding of a single-chain analogue than previously thought. To further investigate, an R8A mutant of R3 B1-22R was synthesised as described in previous sections. The results from the competitive binding assay showed that the R8A mutant had lower affinity for RXFP3 compared to the R3 B1-22R peptide. However, it still retained higher affinity than the capped relaxin-3 antagonist. The loss of binding in removing the Arg^8 does not account for all of the lost affinity in the capped analogue, and since the capped analogue did have more of a helical structure than R3 B1-22R, this suggests that there may be additional binding sites in the N-terminal part of the relaxin B-chain. The residues that were replaced by the capping sequences might also have been more involved in binding than assumed. Additionally, the introduction of the capping sequences might have resulted in different chemical properties at the receptor-binding site, compromising the affinity.

If capped single-chain relaxin analogues are to be further developed, different C-terminal capping sequences should be investigated, seeing as the sequence used in the VEGF peptide did not introduce stable helices in the relaxin-analogues. Additionally, future capped analogues including Arg^8 could further illuminate its role in a single-chain analogue.

Another approach in designing a stable helical analogue is the use of lactam-bridges to stabilize the helix. By bridging residues positioned on the side of the helix not involved in the receptor interaction, the structure could be stabilised.

Future studies should also include mutational studies to further investigate the importance of the various residues in the B-chain, especially around the N-terminus.

3.2. Relaxin-3/chlorotoxin analogue

In light of the findings that Arg^8 appears to be involved in receptor binding, and that the helix capping sequences used were insufficient to introduce a stable helical structure, we decided to investigate the use of stable scaffold molecules. The amino acids involved in binding, including Arg8, were grafted onto the helical region of the peptide chlorotoxin.

3.2.1. Synthesis and oxidative folding

RXFP3, the endogenous receptor for relaxin-3, is mainly expressed in the brain [21], and a potential drug lead targeting this receptor needs to be able to penetrate the blood-brain barrier to reach its proposed site of action. Native chlorotoxin is able to pass the blood-brain barrier, and an alanine-scan of chlorotoxin showed that all mutants were still able to penetrate the blood-brain barrier (Muharrem Akcan, unpublished data), suggesting that chlorotoxin used as a scaffold might still maintain this property.

Structurally, chlorotoxin has a helix of a similar size to that found in the relaxin-3 B-chain [6, 57], making grafting of the residues important for RXFP3 binding possible while still maintaining the helical structure. In addition, its structure allows for Arg^8 to be grafted into a position similar to that in native relaxin-3, outside of the helix [6].

Chlorotoxin is also a very stable molecule, showing high stability both *in vivo* and *in vitro* [58]. Because of these structural and physiological properties, a successful relaxin-3/chlorotoxin grafted analogue would make an excellent drug lead.

The relaxin-3/chlorotoxin analogue was synthesised and oxidized as described in previous sections. From the folding trial the buffers containing isopropanol gave very low yield of the isomer thought to be the correctly folded peptide. Isopropanol favours folding of hydrophobic peptides, and as the more hydrophilic, correctly folded, peptide elutes earlier in RP-HPLC, these results are to be expected.

The yield after oxidation and purification by RP-HPLC was very low, but ^{1}H-NMR on a small sample of purified, oxidized peptide showed a well-dispersed spectrum, correlating well with that seen in native chlorotoxin, suggesting the peptide was indeed correctly folded. A pure sample of 1.5 mg peptide was obtained, and the intention was to test for bioactivity and do structural analysis by 2D NMR. A one-dimentional (1D) NMR however, showed very low intensity peaks, and performing any 2D NMR spectroscopy was not possible. ES-MS on the sample showed peaks with masses corresponding to the peptide, confirming that there was peptide present in the sample. Because of the low signal in the NMR spectroscopy experiments, this led us to believe that the peptide might be aggregating, preventing further structural characterisation. Further studies will have to be undertaken to confirm that a native disulfide arrangement had been achieved, and to generate sufficient material for functional testing.

3.3. E13 combinational library of R3 B1-22R

A different strategy to improve binding of single-chain analogues is to utilise residues that are not part of the native interaction with the receptor, but are close enough to the receptor-binding site to be able to form new contacts if modified.

The selective RXFP3 antagonist R3 B1-22R is able to bind with high affinity despite its lack of structure because of an additional contact point with the receptor provided by the non-native C-terminal arginine [44, 45, 50]. This introduced residue improves affinity ≥100-fold compared to the relaxin-3 B-chain alone [44, 45]. Studies have shown that rapid identification of favourable “mutations” is possible by synthesising chemical libraries of analogues that include mixtures of different amino

acids at one or several points. Among other things this strategy has been used to improve the affinity of conotoxins for their target receptor [59].

We identified E13 as a position that may be used to improve receptor binding in the R3 B1-22R analogue. A combinational library was successfully synthesised, with a mixture of amino acids inserted at position 13. LC-MS analysis revealed the presence of all expected masses confirming that the coupling of the mixture had been successful. After fractionation by RP-HPLC we were able to identify analogues with varying degrees of affinity for the RXFP3 receptor. None of the identified analogues were significantly more potent than the parent peptide, but some of the analogues apppeared to have slightly improved binding. Thus, an E13A variant that appeared to bind with the highest affinity was synthesised and tested. The synthesised analogue had comparable binding to that shown by the B.5 fraction. Thus, a pure analogue did not yield a higher affinity analogue.

While it was disappointing that we were not able to identify any analogues of especially high activity, these results were encouraging in that we have proven the use of peptide mixtures in automated synthesis to be a possible method for synthesis of combinational libraries. We were also able to readily fractionate and test these mixtures to identify components with varying degree of binding. Future studies should continue probing for favourable mutations at other positions using this method.

4. Conclusions

By utilising capping sequences, the helical character of single-chain analogues was successfully increased compared to native relaxin B-chains alone, and previous single-chain analogues. However, the capping sequences were not sufficient to introduce a stable helix.

The capped relaxin-3 antagonist was found to bind RXFP3 with lower affinity than the parent peptide R3 B1-22R, despite a more helical structure. This finding implies that further residues than those known are involved in receptor binding. An R8A analogue of R3 B1-22R was synthesised, and binding assays confirmed that Arg^8 is likely to be more involved in binding than previously thought.

We were able to successfully synthesise a relaxin-3/chlorotoxin analogue, incorporating the residues important for binding to RXFP3 into the stable scaffold

created by the chlorotoxin structure. From 1D NMR data the analogue seemed to fold correctly. However, the yield following oxidation was poor, and no bioactivity data or structural data confirming the correct fold was obtained within the time limits of the project.

A combinational library of R3 B1-22R analogues with mutations in position 13 was synthesised using standard solid phase synthesis, by inserting a mixture of amino acids at this position. We were able to successfully separate the more active analogues from the bulk of the mixture by a bioactivity guided fractionation approach. However, we found no mutants with especially high activity.

5. Materials and methods

5.1. Peptide Synthesis

5.1.1. Solid phase peptide synthesis

Solid phase peptide synthesis (SPPS) is a common method of synthesising peptides where the peptide chain is built on a solid support consisting of resin beads. The general principle of the synthesis is to repeat cycles with coupling of amino acids, and then deprotecting; removing the protecting group from the peptide chain N-terminus to reveal a free NH_2-group, ready for coupling of the next amino acid (Figure 24).

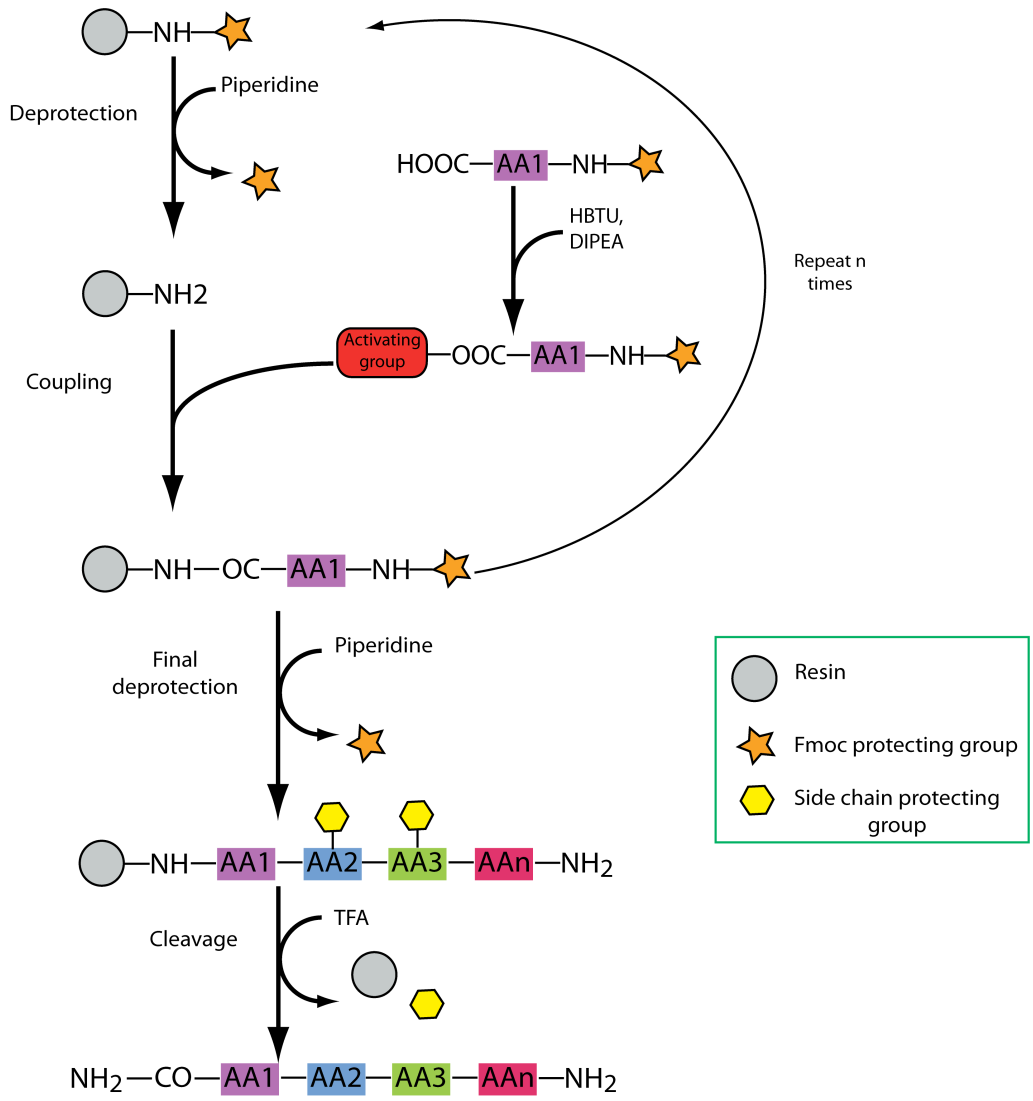


Figure 24. The basic steps in Fmoc solid phase peptide synthesis, using Rink-Amide resin.

There are two major types of chemistry used in SPPS, based on the protecting groups on the amino acids; *tert*-butoxycarbonyl (Boc) and Fmoc. In these two methods different resins are used and the amino acids have different protection groups, however, the main practical difference between these methods is in the deprotection step and the cleavage reaction. Whereas Fmoc-chemistry uses a base (piperidine) to deprotect the amino acids, Boc chemistry uses an acid (typically TFA). When cleaving the finished peptide from the resin, a peptide synthesised using Boc-chemistry requires hydrogen fluoride for the cleaving. With Fmoc chemistry the peptide can be cleaved with TFA, which is much easier and safer.

In both types of SPPS, different chemicals can be used for the coupling reaction. Peptide synthesis in this project was done using O-benzotriazol-1-yl-1,1,3,3-

tetramethyluronium hexafluorophosphate (HBTU) and diisopropylethylamine (DIPEA) in the coupling reactions.

The amino acids are added to the resin together with the appropriate amount of HBTU and DIPEA. Synthesis on a 0.25 mmole scale uses 1 mmole of amino acid, 2 ml of 0.5 M HBTU in N,N-dimethylformamide (DMF), and 174 μL DIPEA per coupling reaction.

5.1.1.1. Manual synthesis

The resin was placed in a reaction vessel with a sintered glass filter and a tap at the bottom that is connected to a vacuum line via a solvent trap. The filter allows reagents to be drained away while the resin, and the growing peptide, is held back in the vessel.

After each coupling step a ninhydrin test was performed to get an estimate of the reaction yield [66]. The ninhydrin test was done by drying a small amount of resin and weighing out approximately 3-5 mg to a test tube. To the resin two drops of 76 % w/w phenol in ethanol, four drops of 0.2 M potassium cyanide in pyridine and two drops of 0.28 M ninhydrin in ethanol was added. The test tube was put on a heating block at 100 °C for five minutes together with a blank sample containing no resin. 2.8 ml 60% ethanol in water was added, and the absorbance at 570 nm was measured using an ultraviolet (UV) spectrophotometer and an estimate of the yield was calculated using the following formula:

Equation 1. Calculation of coupling percentages based on the measured absorbance from the ninhydrin test

$$Coupling = 100 \times (1 - (\frac{A_{570} \times 200}{SV \times m}))\ ,$$

A_{570} = Absorbance measured at 570 nm

SV = Substitution value for the coupling in mmol/g

m = mass of resin

If needed, the coupling step was repeated in order to get sufficient couplings. The ideal coupling percentage of ≥ 99.6 % was not always practically feasible, and in some cases a percentage over or close to 99 % had to suffice.

After an amino acid had been added, and an acceptable coupling yield had been achieved, the resin was washed with 3x15 ml DMF, and then treated twice with 50/50 DMF/piperidine; first for one minute, and then three minutes. The resin was then thoroughly rinsed with DMF to remove all the piperidine, and was ready for addition of the next amino acid.

5.1.1.2. Automated synthesis

The automated synthesis was done on a CEM Liberty microwave synthesiser operated via PepDriver software.

Double couplings were carried out on the following residues: arginine, isoleucine, leucine, phenylalanine, threonine, tryptophane, valine and cysteine.

5.1.2. The synthesis of relaxin B-chain analogues

5.1.2.1. Manual synthesis

The INSL5 agonist analogue was synthesized manually using Fmoc-chemistry on a 0.25 mmol scale following the procedure described above (5.1.1). The resin used was Rink-amide MBHA (Merck), which results in an amidated C-terminus after peptide cleavage.

5.1.2.2. Automated synthesis

The synthesis of all relaxin analogues except the INSL5 agonist was done on a CEM Liberty microwave synthesiser using Fmoc chemistry. Synthesis of all the peptides except the R3 B1-22R was done on a 0.25 mmol scale. The R3 B1-22R was synthesised on a 0.5 mmol scale. The resin used for all these analogues was Rink-amide MBHA (Merck).

The synthesis of the capped relaxin analogues INSL5 antagonist, relaxin-2 antagonist and relaxin-3 antagonist was done by Chia Chia Tan.

5.1.2.3. Acetylation

The capped relaxin analogues (see section 2.1.1) were acetylated N-terminally. After deprotecting the final amino acid and rinsing with DMF an acetylation mixture was added. The acetylation mixture was made of 15 ml DMF, 870 μL acetic

anhydride and 470 µL DIPEA. This mixture was split in two and the resin treated with the mixture for 2 x 5 minutes.

5.1.3. Synthesis of the relaxin-3/chlorotoxin analogue

The Relaxin-3/Chlorotoxin-analogue was synthesised on a CEM Liberty microwave synthesiser on a 0.25 mmoles scale using Fmoc amino acids and a Fmoc-L-Arg(Pbf) resin (Applied biosystems). The synthesis was carried out with four times excess of amino acids.

5.1.4. Synthesis and bioactivity guided fractionation of E13 mixture

The E13-mixture was synthesized by Richard Clark on a CEM Liberty microwave synthesiser. A mixture of 16 amino acids (methionine, cysteine, proline and glycine excluded) was used in the coupling step on position 13 of the R3 B1-22R-peptide. The amino acids in the mixture were added in different ratios according to their coupling efficiencies (Table 2). (Mixture composition provided by Chris Armishaw, personal communication)

Table 1

Amino Acid	Coupling ratio	Amino Acid	Coupling ratio
Ala	0.95	Asn	1.2
Asp	0.9	Gln	1.2
Glu	0.95	Arg	1.42
Phe	0.81	Ser	1.3
Ile	1.16	Thr	1.6
His	0.85	Val	1.14
Lys	1.05	Trp	0.89
Leu	1.08	Tyr	1.26

Purification was done following a bioactivity guided fractionation approach. After testing, fractions eluted between 38-49 and 49-60 minutes in the HPLC on the original mixture seemed to have some affinity. These fractions were purified by RP-HPLC, using a 8 ml/min column and a 0.25 % gradient. Further fractionation based on retention times was performed. Using ES-MS, the fractions were divided according to the masses each fraction contained.

Of the final fractions, two fractions showed increased affinity compared to R3 B1-22R, which was used as a positive control.

These fractions were analysed by MALDI-TOF-MS to determine which analogues the fractions contained.

5.2. TFA cleavage

All peptides except the chlorotoxin analogue were cleaved using 90 % TFA/5 % tri-isopropylsilane (TIPS)/ 5 % distilled water. 50 ml cleaving mixture was used for 400 mg of resin.

The chlorotoxin analogue was cleaved from the resin using 96 ml TFA/2 ml TIPS/2 ml 3,6-dioxa-1,8-octanedithiol (DODT)/1 ml distilled water. 100 ml of cleaving mixture was used per 1 g of resin. For all peptides, the resin was left to react on a magnetic stirrer, in room temperature, for two hours. The resin was then filtered off, and TFA removed on a rotary evaporator, followed by the addition of approximately 100 ml cold diethyl ether to precipitate the peptide. The precipitated peptide was extracted with 50/50 buffer A/B using a separation funnel. Any residual ether was removed on the rotary evaporator and the peptide solution was lyophilised.

5.3. Oxidative folding of the relaxin-3/chlorotoxin analogue

Before oxidizing larger amounts of peptide, a small scale folding trial was conducted. 100 μg of purified peptide was added to 200 μL of the following buffers:

Table 3

Buffer number	Constituents	pH
1	0.1 M Tris-HCl 0.2 M NaCl, 5 mM reduced glutathione 0.5 mM oxidized glutathione	7.8
2	0.1 M NH_4OAc	8.3
3	0.1 M $(NH_4)_2CO_3$	8.2
4	0.33 M NH_4OAc 0.5 M GnHCl	7.8
5	0.1 M Na_3PO_4	7.4
6	Buffer 1/isopropanol, 1:1	
7	Buffer 2/isopropanol, 1:1	
8	Buffer 3/isopropanol, 1:1	
9	Buffer 4/isopropanol, 1:1	
10	Buffer 5/isopropanol, 1:1	

The samples in the folding trial were analysed by analytical HPLC after 24 and 48 hours, and buffers 1 and 5 were subsequently used to oxidize larger amounts of peptide.

Reduced, purified peptide was oxidized in buffer 1 or 5 (Table 3) to form the disulfide bonds. The reaction was left for 24 hours at room temperature, under continuous stirring. A concentration of 0.1 mg/ml peptide was used.

1D NMR spectroscopy was used on purified peptide after folding in buffer 1 to confirm that the correct folding was achieved.

5.4. HPLC

5.4.1. Preparative

The peptides were purified using RP-HPLC.
All purification was done on one of the following systems:

- Waters 600 E System Controller with a Waters 484 Tunable Absorbance Detector and Waters Empower software
- Waters 600 MS System Controller with a Waters 484 Tunable Absorbance Detector and Delta 5.5 Chromatography Data System software
- Shimadzu LC 20 AT Prominence Liquid Chromatograph with a Shimadzu SPD-20A Prominence UV/VIS Detector and LC Solutions software
- Agilent 1100 series with an ICI Instruments LC 1500 HPLC pump and LC Solutions software

Columns used:

- Preparative:
    - Phenomenex Jupiter C18 (250 x 21.20 mm, 300 Å pore size, 15 µm particle size)
    - VYDAC Protein & Peptide C18 (250 x 22 mm, 300 Å pore size, 10 µm particle size)
- Semi-preparative:
    - Phenomenex C18 (250 x 10,00 5 micron, 300 Å pore size, 5 µm particle size)

- VYDAC Protein & Peptide C18 (250 x 10 mm, 300 Å pore size, 5 µm particle size)

- Analytical:
  - Phenomenex Jupiter C18 (150 x 4.60 5 micron, 300 Å pore size, 5 µm particle size)

Gradients used in purification:

- 8 ml/min:
  - 1 % gradient of 0-80 % buffer B in 80 minutes. UV detection at 230 or 280 nm
  - 0.5 % gradient of 0-80 % buffer B in 160 minutes. UV detection at 230 or 280 nm
  - 0.25 % gradient of 0-80 % buffer B in 160 minutes. UV detection at 230 or 280 nm
- 3 ml/min:
  - 0.5 % gradient of 0-80 % buffer B in 160 minutes. UV detection at 230 or 280 nm
- 1 ml/min:
  - 1 % gradient of 0-80% buffer B in 80 minutes. UV detection at 215, 254 and 280 nm.

5.4.2. Analytical

Analytical HPLC was done on one of the following:

- Agilent 1100 series. The column used was a Phenomenex Jupiter C18 (150 x 2.00 mm 5 micron, 300 Å pore size, 5 µm particle size). A flow rate of 0.3 ml/min, and a 2 % gradient of 0-80 % buffer B in 40 minutes was used. UV-detection at 215, 254 and 280 nm.
- Shimadzu LC 20 AT Prominence Liquid Chromatograph with a Shimadzu SPD-20A Prominence UV/VIS Detector and SIL-20A HT prominence autosampler, and LC Solutions software. The column used was a Phenomenex Jupiter C18 (150 x 4.60 5 micron, 300 Å pore size, 5 µm particle size) or a VYDAC Protein & Peptide C18 (250 x 4.60 mm 5 micron, 300 Å pore size, 5 µm particle size)

5.5. MS

5.5.1. ES-MS

The identification of HPLC-fraction constituents was based on mass. All mass data were obtained using ES-MS on an Applied Biosystems Sciex API 2000 system.

5.5.2. MALDI-TOF-MS

Identification of monoisotopic masses present in the two most active fractions of the E13 mixture was done on a 4700 Proteomics Analyzer MS-MS TOF-TOF from Applied Biosystems.

Samples from the two fractions were made in two different dilutions by mixing 1 μl of the fractions re-dissolved in buffer A, with 1 or 2 μl matrix. 1 μl was then taken from each dilution and placed on a MALDI-plate. The sample showing the best signal-to-noise signal in the final spectrum was then chosen to use for mass determination.

5.6. NMR spectroscopy

0.5-1 mg purified, freeze dried sample was dissolved in distilled H_2O/D_2O (500 μL/50 μL) and subjected to solution NMR. For the experiments done at 285 K, 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS) was added as a reference in the samples.

A 1D ^1H spectrum was recorded of the relaxin-3/chlorotoxin analogue on a Bruker 600 MHz spectrometer. Due to the very low concentration of the sample this was recorded with 20000 scans to ensure a good enough signal-to-noise ratio to clearly observe the resonance peaks.

2D NMR experiments included TOCSY and NOESY spectra, and were recorded with a mixing time of 200 ms at on a Bruker 500 MHz or a Bruker 600 MHz spectrometer.

The spectra for the relaxin-3 agonist and the relaxin-2 antagonist were acquired at 285 K. All other spectra were acquired at 298 K. The pH in all samples was approximately 4.

All spectra were analysed on Silicon Graphics Indigo Workstations using Topspin (Bruker) software, and analysed using CARA.

Spectra were assigned using the sequential assignment technique [60], and secondary shifts were calculated using the random coil chemical shifts described by Wishart et al. [61].

5.6.1. Structural calculations

Structural calculations were done using CYANA software. NOE cross-peaks were translated into distance restraints based on peak intensities.

The three dimensional structure was then calculated using torsion angle dynamics within CYANA. 20 structures were calculated, and the ten lowest energy structures were used to get the 3D structure of the peptide. Structures were visualised using the program MOLMOL [67].

5.7. Activity and binding assays

All binding and activity assays were done by our collaborators at the Florey Neuroscience Institutes in Melbourne.

Binding and activity assays for RXFP1 was done on HEK293T cells expressing the receptor, and for RXFP3 and 4 on CHO cells.

5.7.1. Cyclic Adenosine Monophosphate assay

CHO cells were transfected with either human RXFP3 or RXFP4 receptor, together with a cAMP response element (CRE)-β-galactosidase reporter plasmid as described by Haugaard-Kedström et al. [45].

After incubation, the cells were stimulated with 5 μM forskolin (RXFP3 expressing cells) or 1 μM forskolin (RXFP4 expressing cells). Activation of RXFP3 or 4 inhibits production of cAMP, and the activity of ligands for these receptors is measured by the ability to reduce the forskolin induced cAMP activity.

The cells were stimulated with peptide, and then treated with lysis buffer and the β-galactosidase substrate chlorophenol red-β-D-galactopyranoside (CPRG). The yellow-orange CPRG is hydrolysed by β-galactosidase to the chromophore chlorophenol red and galactose, resulting in a dark red solution. The absorbance of chlorophenol red is proportional to cAMP production, and used as a measure of the activity of the receptor. The absorbance was measured at 570 nm.

5.7.2. Competitive binding assay

Binding properties were determined by competitive binding assays on CHO cells stably transfected with RXFP3 or RXFP4, or HEK193T cells stably transfected with RXFP1. Cells were incubated with Europium(III) labelled relaxin-2, R3/I5 or INSL5 for assays on RXFP1, RXFP3 and RXFP4 respectively. The affinity of the tested ligands was assessed based on their ability to compete with the labelled ligand [45].

6. References

1. Bellmann-Sickert, K. and A.G. Beck-Sickinger, *Peptide drugs to target G protein-coupled receptors.* Trends Pharmacol Sci, 2010. **31**(9): p. 434-41.
2. Sato, A.K., et al., *Therapeutic peptides: technological advances driving peptides into development.* Curr Opin Biotechnol, 2006. **17**(6): p. 638-42.
3. Lien, S. and H.B. Lowman, *Therapeutic peptides.* Trends Biotechnol, 2003. **21**(12): p. 556-62.
4. Kong, R.C., et al., *Membrane receptors: structure and function of the relaxin family peptide receptors.* Mol Cell Endocrinol, 2010. **320**(1-2): p. 1-15.
5. Chan, L.J., et al., *The Relaxin Peptide Family - Structure, Function and Clinical Applications.* Protein Pept Lett, 2010.
6. Rosengren, K.J., et al., *Solution structure and novel insights into the determinants of the receptor specificity of human relaxin-3.* J Biol Chem, 2006. **281**(9): p. 5845-51.
7. Haugaard-Jonsson, L.M., et al., *Structure of human insulin-like peptide 5 and characterization of conserved hydrogen bonds and electrostatic interactions within the relaxin framework.* Biochem J, 2009. **419**(3): p. 619-27.
8. Rosengren, K.J., et al., *Structural insights into the function of relaxins.* Ann N Y Acad Sci, 2009. **1160**: p. 20-6.
9. Eigenbrot, C., et al., *X-ray structure of human relaxin at 1.5 A. Comparison to insulin and implications for receptor binding determinants.* J Mol Biol, 1991. **221**(1): p. 15-21.
10. Hisaw, F.L., *Experimental relaxation of the pubic ligament of the guinea pig.* Proc Soc Exp Biol Med, 1926. **23**: p. 661-62.
11. Schwabe, C., J.K. McDonald, and B.G. Steinetz, *Primary structure of the A chain of porcine relaxin.* Biochem Biophys Res Commun, 1976. **70**(2): p. 397-405.
12. Schwabe, C. and J.K. McDonald, *Primary structure of the B-chain of porcine relaxin.* Biochem Biophys Res Commun, 1977. **75**(2): p. 503-10.
13. van der Westhuizen, E.T., et al., *Relaxin family peptide receptors--from orphans to therapeutic targets.* Drug Discov Today, 2008. **13**(15-16): p. 640-51.
14. Samuel, C.S., et al., *Drugs of the future: the hormone relaxin.* Cell Mol Life Sci, 2007. **64**(12): p. 1539-57.

15. Teerlink, J.R., et al., *Relaxin for the treatment of patients with acute heart failure (Pre-RELAX-AHF): a multicentre, randomised, placebo-controlled, parallel-group, dose-finding phase IIb study.* Lancet, 2009. **373**(9673): p. 1429-39.
16. Unemori, E., B. Sibai, and S.L. Teichman, *Scientific rationale and design of a phase I safety study of relaxin in women with severe preeclampsia.* Ann N Y Acad Sci, 2009. **1160**: p. 381-4.
17. Conklin, D., et al., *Identification of INSL5, a new member of the insulin superfamily.* Genomics, 1999. **60**(1): p. 50-6.
18. Jaspers, S., Lok, S., Lofton-Day, C.E., Haldeman, B.A, Whitmore, T.E., Foley, K.P., Conklin, D. *The genomics of INSL5*. in *Relaxin 2000: the Third International Conference on Relaxin and Related Peptides*. 2001.
19. Aparicio, S., Dixon, J., Hendrick, A., Horwood, J. M., Zahn, D., *Use of Gpr100 receptor in diabetes and obesity regulation*. 2008: United States.
20. Bathgate, R.A., et al., *Human relaxin gene 3 (H3) and the equivalent mouse relaxin (M3) gene. Novel members of the relaxin peptide family.* J Biol Chem, 2002. **277**(2): p. 1148-57.
21. Liu, C., et al., *Identification of relaxin-3/INSL7 as an endogenous ligand for the orphan G-protein-coupled receptor GPCR135.* J Biol Chem, 2003. **278**(50): p. 50754-64.
22. Tanaka, M., et al., *Neurons expressing relaxin 3/INSL 7 in the nucleus incertus respond to stress.* Eur J Neurosci, 2005. **21**(6): p. 1659-70.
23. Ma, S., et al., *Relaxin-3 in GABA projection neurons of nucleus incertus suggests widespread influence on forebrain circuits via G-protein-coupled receptor-135 in the rat.* Neuroscience, 2007. **144**(1): p. 165-90.
24. Goto, M., L.W. Swanson, and N.S. Canteras, *Connections of the nucleus incertus.* J Comp Neurol, 2001. **438**(1): p. 86-122.
25. McGowan, B.M., et al., *Central relaxin-3 administration causes hyperphagia in male Wistar rats.* Endocrinology, 2005. **146**(8): p. 3295-300.
26. Hida, T., et al., *Chronic intracerebroventricular administration of relaxin-3 increases body weight in rats.* J Recept Signal Transduct Res, 2006. **26**(3): p. 147-58.
27. McGowan, B.M., et al., *Effects of acute and chronic relaxin-3 on food intake and energy expenditure in rats.* Regul Pept, 2006. **136**(1-3): p. 72-7.
28. Ma, S., et al., *Modulation of hippocampal theta oscillations and spatial memory by relaxin-3 neurons of the nucleus incertus.* Learn Mem, 2009. **16**(11): p. 730-42.
29. Hossain, M.A., et al., *H3 relaxin demonstrates antifibrotic properties via the RXFP1 receptor.* Biochemistry, 2011. **50**(8): p. 1368-75.
30. Hsu, S.Y., et al., *Activation of orphan receptors by the hormone relaxin.* Science, 2002. **295**(5555): p. 671-4.
31. Kumagai, J., et al., *INSL3/Leydig insulin-like peptide activates the LGR8 receptor important in testis descent.* J Biol Chem, 2002. **277**(35): p. 31283-6.
32. Sudo, S., et al., *H3 relaxin is a specific ligand for LGR7 and activates the receptor by interacting with both the ectodomain and the exoloop 2.* J Biol Chem, 2003. **278**(10): p. 7855-62.
33. Liu, C., et al., *Identification of relaxin-3/INSL7 as a ligand for GPCR142.* J Biol Chem, 2003. **278**(50): p. 50765-70.
34. Liu, C., et al., *INSL5 is a high affinity specific agonist for GPCR142 (GPR100).* J Biol Chem, 2005. **280**(1): p. 292-300.

35. Hsu, S.Y., et al., *The three subfamilies of leucine-rich repeat-containing G protein-coupled receptors (LGR): identification of LGR6 and LGR7 and the signaling mechanism for LGR7.* Mol Endocrinol, 2000. **14**(8): p. 1257-71.
36. Hossain, M.A., et al., *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, 2008. **283**(25): p. 17287-97.
37. Bullesbach, E.E. and C. Schwabe, *LGR8 signal activation by the relaxin-like factor.* J Biol Chem, 2005. **280**(15): p. 14586-90.
38. Bullesbach, E.E., S. Yang, and C. Schwabe, *The receptor-binding site of human relaxin II. A dual prong-binding mechanism.* J Biol Chem, 1992. **267**(32): p. 22957-60.
39. Bullesbach, E.E. and C. Schwabe, *The relaxin receptor-binding site geometry suggests a novel gripping mode of interaction.* J Biol Chem, 2000. **275**(45): p. 35276-80.
40. Tregear, G.W., et al., *Structure and activity in the relaxin family of peptides.* Ann N Y Acad Sci, 2009. **1160**: p. 5-10.
41. Hossain, M.A. and J.D. Wade, *The Roles of the A- and B-Chains of Human Relaxin-2 and -3 on their Biological Activity.* Curr Protein Pept Sci, 2010.
42. Park, J.I., et al., *Regulation of receptor signaling by relaxin A chain motifs: derivation of pan-specific and LGR7-specific human relaxin analogs.* J Biol Chem, 2008. **283**(46): p. 32099-109.
43. Bullesbach, E.E. and C. Schwabe, *Tryptophan B27 in the relaxin-like factor (RLF) is crucial for RLF receptor-binding.* Biochemistry, 1999. **38**(10): p. 3073-8.
44. Kuei, C., et al., *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, 2007. **282**(35): p. 25425-35.
45. Haugaard-Kedstrom, L.M., et al., *Design, Synthesis, and Characterization of a Single-Chain Peptide Antagonist for the Relaxin-3 Receptor RXFP3.* J Am Chem Soc, 2011.
46. Kern, A., A.I. Agoulnik, and G.D. Bryant-Greenwood, *The low-density lipoprotein class A module of the relaxin receptor (leucine-rich repeat containing G-protein coupled receptor 7): its role in signaling and trafficking to the cell membrane.* Endocrinology, 2007. **148**(3): p. 1181-94.
47. Akhter Hossain, M., et al., *Synthesis, conformation, and activity of human insulin-like peptide 5 (INSL5).* Chembiochem, 2008. **9**(11): p. 1816-22.
48. Liu, C., et al., *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, 2005. **67**(1): p. 231-40.
49. Haugaard-Jonsson, L.M., et al., *Structural properties of relaxin chimeras.* Ann N Y Acad Sci, 2009. **1160**: p. 27-30.
50. Hossain, M.A., et al., *The structural and functional role of the B-chain C-terminal arginine in the relaxin-3 peptide antagonist, R3(BDelta23-27)R/I5.* Chem Biol Drug Des, 2009. **73**(1): p. 46-52.
51. D'Andrea, L.D., et al., *Targeting angiogenesis: structural characterization and biological properties of a de novo engineered VEGF mimicking peptide.* Proc Natl Acad Sci U S A, 2005. **102**(40): p. 14215-20.

52. Gunasekera, S., et al., *Engineering stabilized vascular endothelial growth factor-A antagonists: synthesis, structural characterization, and bioactivity of grafted analogues of cyclotides.* J Med Chem, 2008. **51**(24): p. 7697-704.
53. Aurora, R. and G.D. Rose, *Helix capping.* Protein Sci, 1998. **7**(1): p. 21-38.
54. DeBin, J.A., J.E. Maggio, and G.R. Strichartz, *Purification and characterization of chlorotoxin, a chloride channel ligand from the venom of the scorpion.* Am J Physiol, 1993. **264**(2 Pt 1): p. C361-9.
55. Wu, X.S., et al., *Development of the research on the application of chlorotoxin in imaging diagnostics and targeted therapies for tumors.* Chin J Cancer, 2010. **29**(6): p. 626-30.
56. Kesavan, K., et al., *Annexin A2 is a molecular target for TM601, a peptide with tumor-targeting and anti-angiogenic effects.* J Biol Chem, 2010. **285**(7): p. 4366-74.
57. Lippens, G., et al., *NMR sequential assignments and solution structure of chlorotoxin, a small scorpion toxin that blocks chloride channels.* Biochemistry, 1995. **34**(1): p. 13-21.
58. Akcan, M., et al., *Chemical Re-engineering of Chlorotoxin Improves Bioconjugation Properties for Tumor Imaging and Targeted Therapy.* J Med Chem, 2011.
59. Armishaw, C.J., et al., *A synthetic combinatorial strategy for developing alpha-conotoxin analogs as potent alpha7 nicotinic acetylcholine receptor antagonists.* J Biol Chem, 2010. **285**(3): p. 1809-21.
60. Wütrich, K., *NMR of proteins and nucleic acids*. 1986, New York: Wiley-Interscience.
61. Wishart, D.S., et al., *1H, 13C and 15N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J Biomol NMR, 1995. **5**(1): p. 67-81.
62. Wishart, D.S., B.D. Sykes, and F.M. Richards, *The chemical shift index: a fast and simple method for the assignment of protein secondary structure through NMR spectroscopy.* Biochemistry, 1992. **31**(6): p. 1647-51.
63. Dincbas-Renqvist, V., et al., *Thermodynamics of folding, stabilization, and binding in an engineered protein--protein complex.* J Am Chem Soc, 2004. **126**(36): p. 11220-30.
64. Hubbard, S.J., *The structural aspects of limited proteolysis of native proteins.* Biochim Biophys Acta, 1998. **1382**(2): p. 191-206.
65. Bullesbach, E.E. and C. Schwabe, *The trap-like relaxin-binding site of the leucine-rich G-protein-coupled receptor 7.* J Biol Chem, 2005. **280**(14): p. 14051-6.
66. Sarin, V.K., et al., *Quantitative monitoring of solid-phase peptide synthesis by the ninhydrin reaction.* Anal Biochem, 1981. **117**(1): p. 147-57.
67. Koradi, R., M. Billeter, and K. Wuthrich, *MOLMOL: a program for display and analysis of macromolecular structures.* J Mol Graph, 1996. **14**(1): p. 51-5, 29-32.

7. Appendix

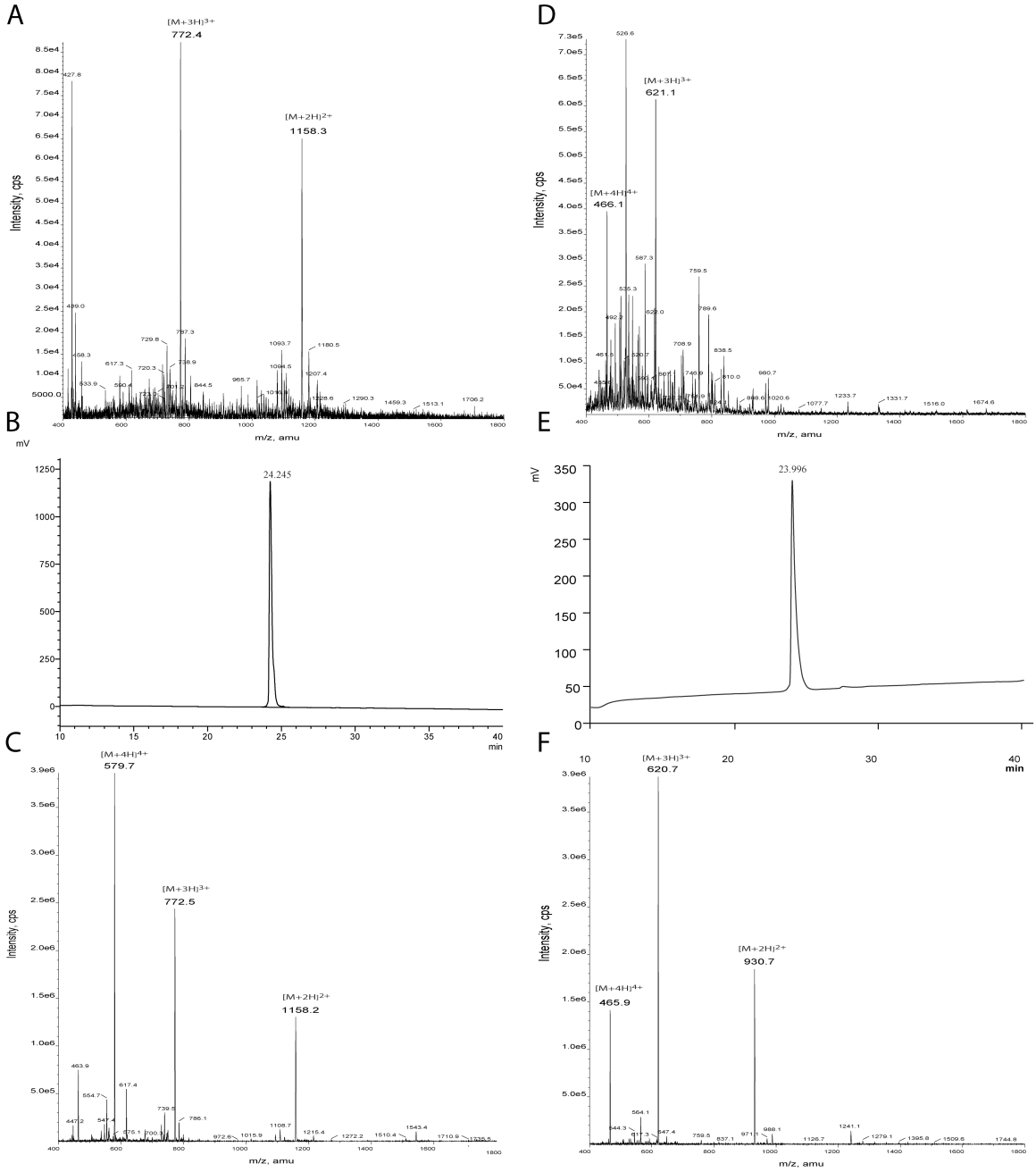


Figure A1. Purification and characterisation of the capped relaxin-3 analogues.
Relaxin-3 agonist: A) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. B) RP-HPLC trace of the purified peptide. C) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.
Relaxin-3 antagonist: D) ES-MS spectrum after TFA cleavage showing masses belonging to the peptide. E) RP-HPLC trace of the purified peptide. F) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.

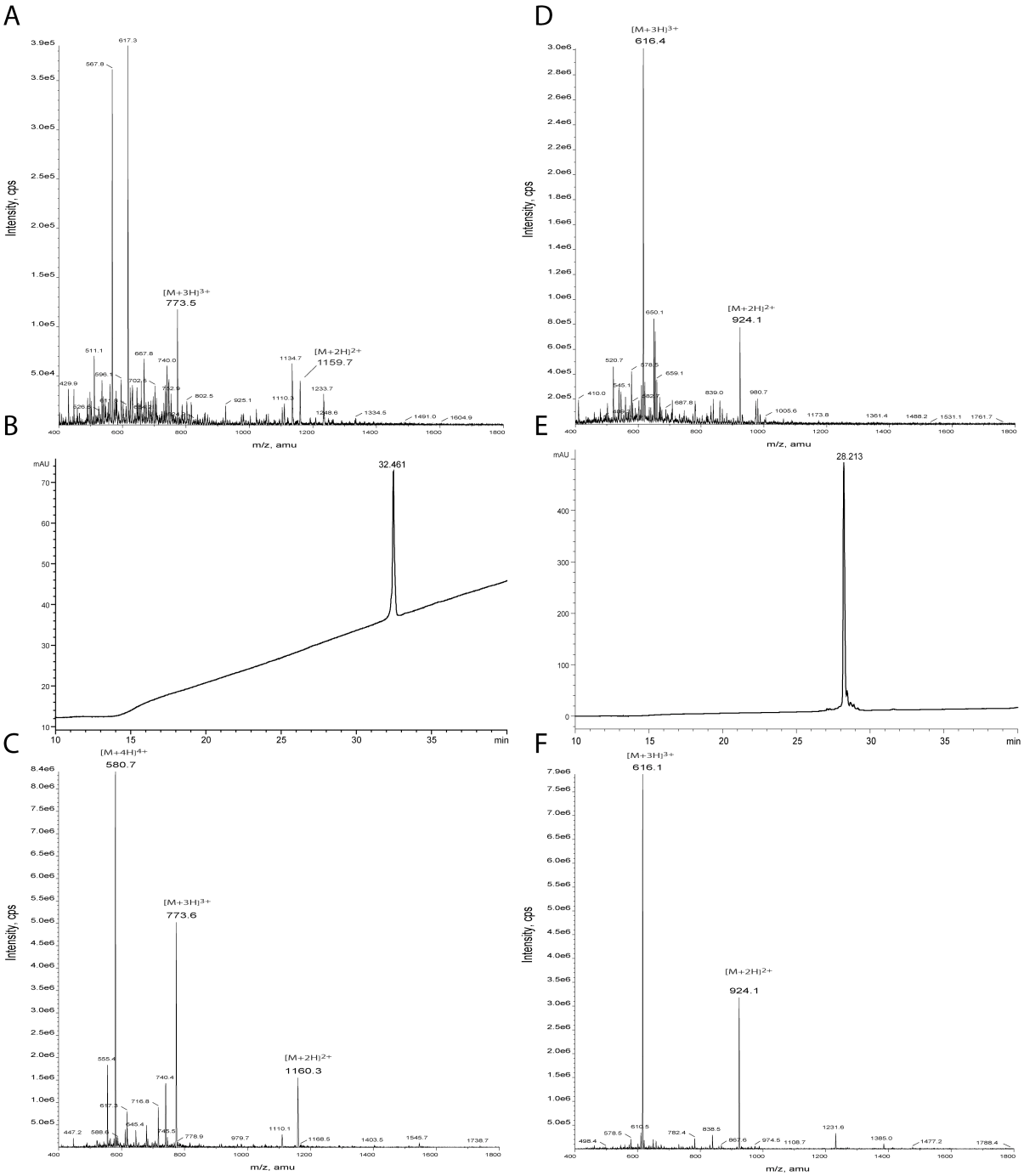


Figure A2. Purification and characterisation of the capped INSL5 analogues.
INSL5 agonist: A) ES-MS spectrum after TFA cleavage showing masses belonging to the peptide. B) RP-HPLC trace of the purified peptide. C) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.
INSL5 antagonist: D) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. E) RP-HPLC trace of the purified peptide. F) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.

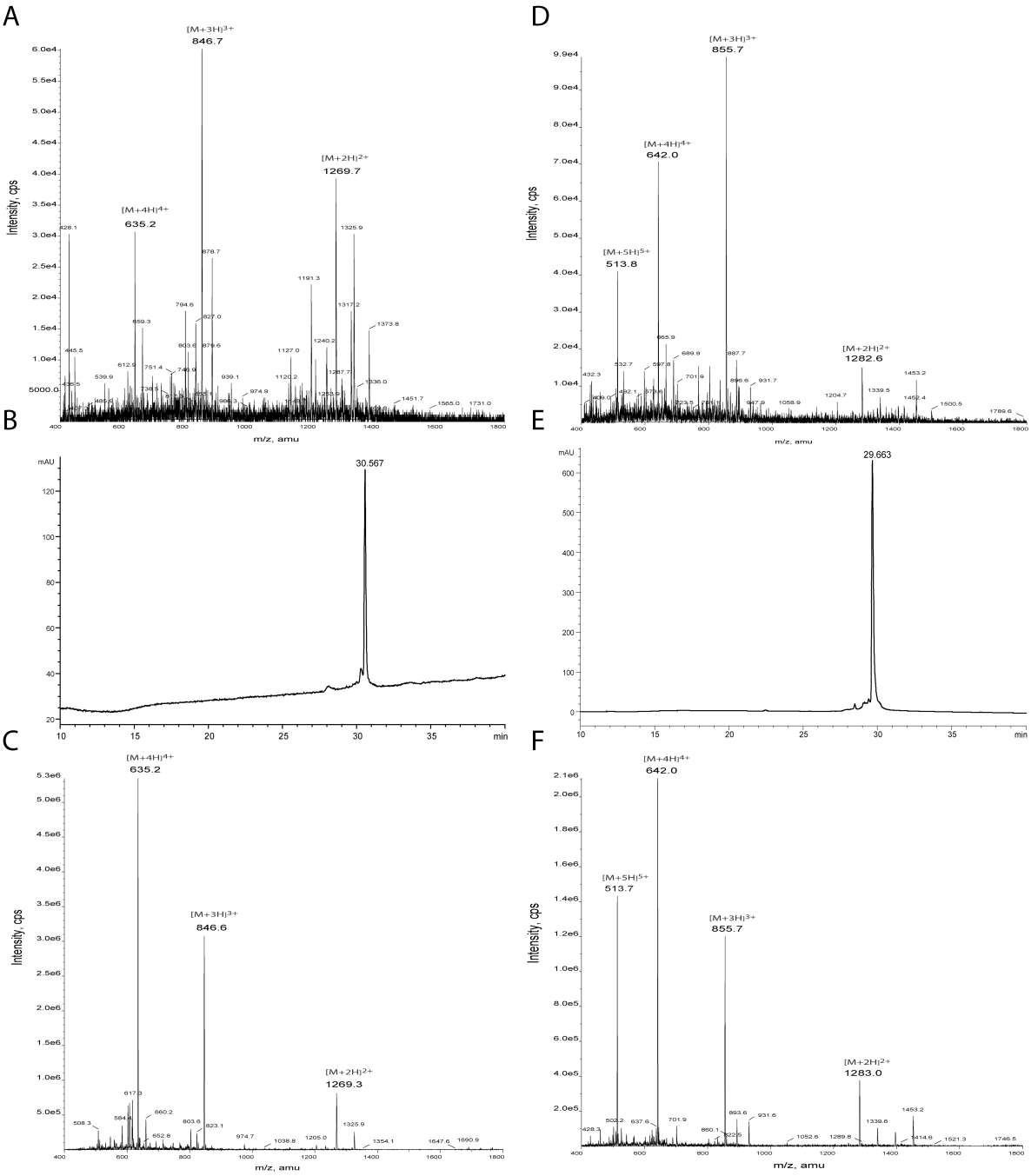


Figure A3. Purification and characterisation of R8A and E13A analogues of R3 B1-22R.
R8A analogue A) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. B) RP-HPLC trace of the purified peptide. C) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.
E13A analogue: D) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. E) RP-HPLC trace of the purified peptide. F) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.

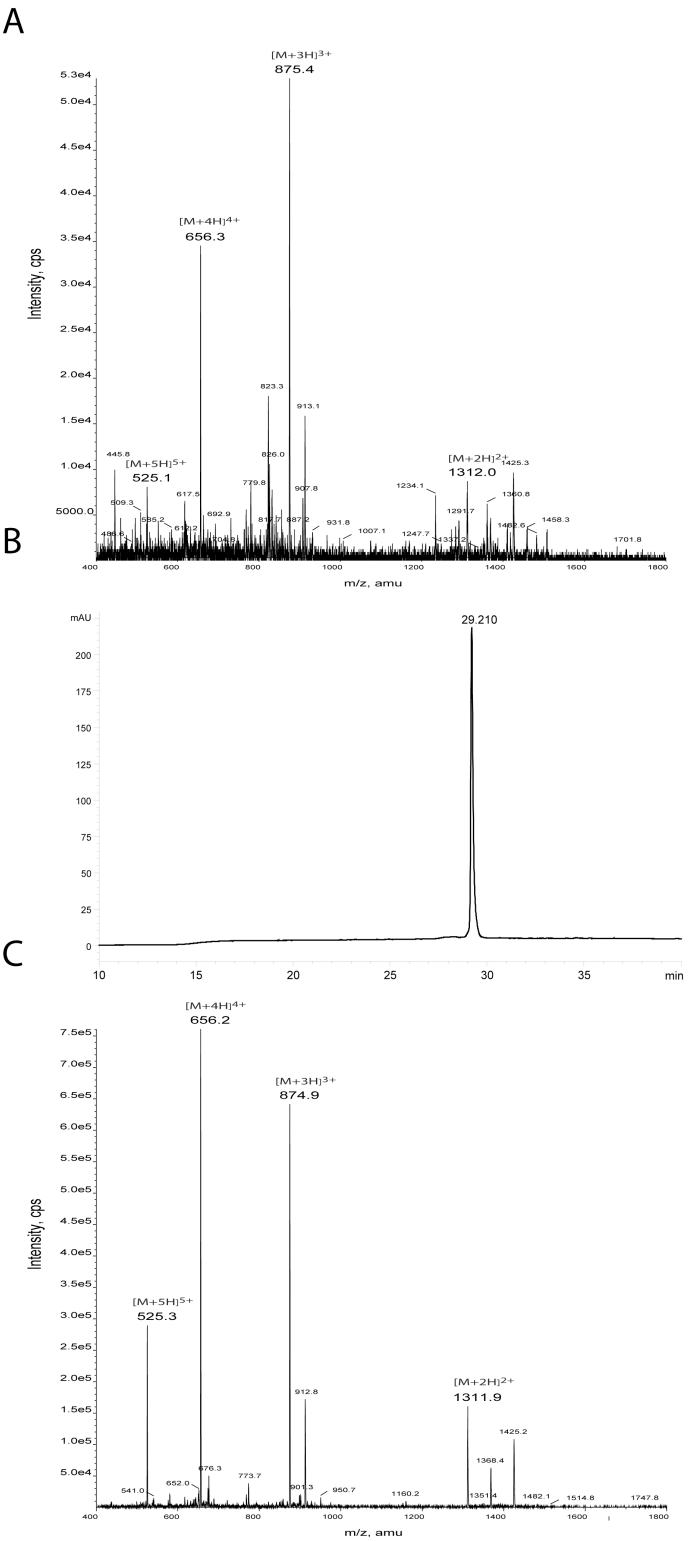


Figure A4. Purification and cahracterisation of R3 B1-22R. A) ES-MS spectrum after TFA cleavage showing major masses belonging to the peptide. B) RP-HPLC trace of the purified peptide. C) ES-MS spectrum of purified peptide, confirming that the correct peptide is present.