Furan-modified peptides for crosslinking to GPCRs

Emiel Ongenae

Promotor: Prof. Dr. Annemieke Madder
Supervisor: Dr. Laia Miret Casals
Examiners: Prof. Dr. Marleen Van Troys
                    Prof. Dr. Sandra Van Vlierberghe

Acknowledgments

I've always been fascinated by the chemistry that fuels the wonders of the world, life itself in particular. Throughout my 5-year educative journey, the mysteries surrounding the world at a molecular level have largely been unraveled, yet its magic remains. Ghent University treated me with fantastic courses taught by passionate professors. I will always look back at my student years with a nostalgic feeling of pure joy and excitement.

A single professor did stand out, partially because the course material carried my biggest interest, but more importantly, because of the unequalled enthusiasm for her own field of study. It was a true honor to perform both my bachelor thesis and my master thesis at her research group, I felt at home. Professor Madder, thank you so much for all the opportunities I received to perform scientific research at a top level in the branch of chemistry that interests me the most.

Every final year student is guided throughout their master thesis project by a thesis supervisor, which was Laia Miret Casals in my case. If this thesis were a water skiing competition, Laia would be the speedboat driver. Laia was always thinking two steps ahead, pulling me along on her fast-paced trajectory, powered by scientific dedication and Catalan fury instead of diesel. Without the speedboat driver, there is simply no way I could have covered the same distance. Laia, I learned so much from you, working together in the lab was a delight and my greatest gratitude goes to you. I could not have completed this thesis without you.

Obviously it wasn't just Laia and me in the lab, the whole corridor was always filled with people from the OBCR group who were more than willing to help out with whatever issue I was encountering. A special thanks to: Nathan, always ready to jam together to some song on the below-average sounding lab radio, Dorien and Jan for the cheerful chats, Jos for the technical support and everyone else for just being in an overall good mood all the time. Apart from people affiliated to the OBCR group, I want to thank people from the VIB, especially Professor Marleen Van Troys and Professor Christophe Ampe for allowing me to perform the cell experiments in their lab.

I would also like to thank all of my fellow last year students, parents, brother, family, friends and girlfriend for the fun times and support when needed, not only during my final year but

also during the four wonderful years in advance. Each and every one of you contributed to shaping me into the person I am proud to be today. In addition, the COVID-19 quarantine weighed rather heavily on me and I am very happy that I could rely on any of you for cheering me up, especially my girlfriend Laura.

Finally, I would like to thank the members of the reading commission for spending some of their valuable time on reading and grading my master thesis dissertation. I hope the reading process will be as enjoyable as possible.

Enjoy reading!

Emiel Ongenae (a.k.a. Chemiel)

Table of Contents

1 Introduction

1.1 Peptide Therapeutics

As from the 1920s, peptide therapeutics have been playing part in the quickly evolving field of medical practice. The well-known first example of a peptide therapeutic is insulin, which treats high glucose concentrations in the blood of diabetic patients. A century later, insulin is still a true necessity to the health of about 150-200 million individuals worldwide[1]. Apart from insulin, over 140 peptide therapeutics are currently available for use in general medicine, of which 11 got approved by the United States food and drug administration (FDA) during the period 2016-2019[2]. Although featuring a low toxicity in general, peptide therapeutics have some drawbacks compared to alternatives, which translates into their currently rather modest share in the pharmaceutical industry[3]. Nevertheless, they play a key role in the treatment of a wide variety of diseases, as shown in Figure 1.
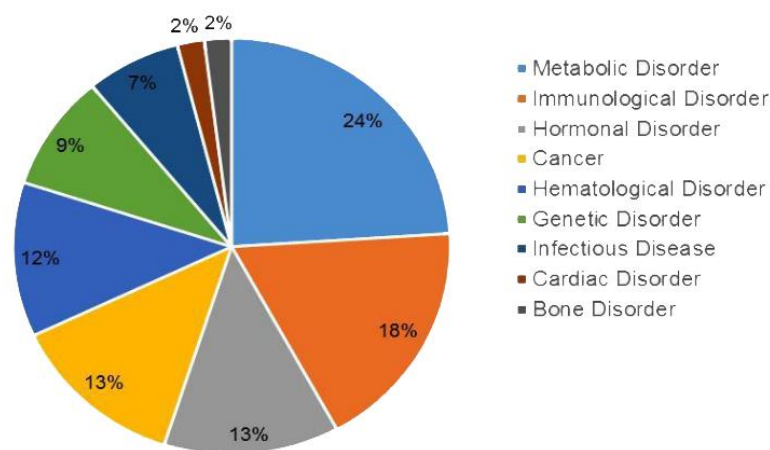
Figure 1: Diseases treated by FDA-approved peptide and protein drugs (December 2016)[3].

The most fundamental hurdle for peptide therapeutics to overcome, is their rapid degradation inside the human body, most specifically under oral administration. The digestive system contains plenty of enzymes (peptidases) specialized in cleaving the amide bonds of proteins and thus the peptide drugs. Scientists have been performing extensive research to overcome this challenge, as well as other challenges like improving absorption, distribution, metabolism, excretion and enhancing activity in general. This has yielded promising results by making use of modified amino acids[4], peptide cyclization[5], conjugation[6], stapling[7] and more[8]. However, despite these solutions to overcome the challenges peptide drugs typically face, their market share has simply been overshadowed by rapid evolutions in the field of small

molecule drugs thanks to combinatorial chemistry and high-throughput screening technologies, benefiting from higher interest from the industry's big corporations[9].

Although being more prevalent, the small molecule approach has some drawbacks of its own. Compared to peptides that typically contain a specific complementary 3D-structure to the targeted receptor, small molecules often lack the necessary chemical complexity required for such high specificity interaction with the receptor[9]. In this case, a receptor might only be druggable by more complex chemical entities. On top of that, it is known that small molecules tend to act as antagonists more frequently as compared to fulfilling an agonist role and given their intrinsic less specific and complex structure, they can give rise to undesired side effects[10].

In order to increase the size of the druggable genome and thus maximizing the possibilities of treating current and future diseases, research on peptide therapeutics has received increased interest from the pharmaceutical industry over the past recent years and novel ways of tackling peptides' intrinsic chemical and pharmacological challenges has been the topic of numerous detailed publications.

1.2   Covalent drugs

Drugs that form a covalent bond with their target are not the first thing that comes to mind in the process of new drug development in the pharmaceutical industry. This category of therapeutics has obtained a rather bad reputation due to the fear for off-target binding, something that was studied for example for bromobenzene in the 1970s[11]. It was observed that reactive metabolites of bromobenzene generated in the liver, covalently bound to liver proteins resulting in hepatotoxic phenomena. Apart from side effects observed in the liver, direct tissue damage as well as immune reactions were reported, resulting in wide-spread aversion to therapeutics with a covalent mode of action[12].

Despite off-target binding being a major concern, it should not be forgotten that many covalent drugs play a daily role in the health of numerous individuals. Aspirin, the most broadly used medication worldwide, has a covalent mode of action. Other FDA-approved examples are found in a wide variety of applications, ranging from oncology to anti-infectives, cancer treatments and treatments for gastrointestinal and cardiovascular disorders[13]. The majority of covalent drugs target enzymes in the human body, yet examples like clopidogrel

(Figure 2) are known to target receptors. In this case, the target is a G protein-coupled receptor (GPCR), more specifically P2Y12[14]. Examples like clopidogrel prove the applicability and potential of a covalent mode of action for targeting receptors and more specifically GPCRs.
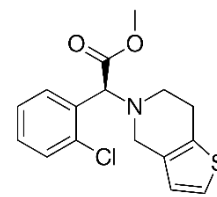
Figure 2: Clopidogrel

Apart from their mode of action, covalent drugs share another feature. They have mostly been discovered serendipitously through screening experiments of some sort. Their covalent mechanism only became apparent at some point in time after their pharmaceutical role had long been established. This is something that could change in the near future, with drug discovery research increasingly focusing on a structure-based approach, with a lesser role for serendipity. By employing a covalent mode of action, drugs obtain a higher potency and prolonged activity. This is caused by the irreversible covalent bond, which allows for the drug interaction to persist longer, even beyond drug clearance. This on its turn leads to less frequent dosing and less exposure to the drugs. On top of that, covalent drugs could prove their value in targeting known shallow binding sites [15][16].

1.3 Peptide-protein crosslinking

While peptide-protein interactions are very dynamic by nature, chemically crosslinking peptides to proteins by means of forming a covalent bond allows to freeze these interactions in place (Figure 3). Chemical crosslinking has been heavily explored over the past decade in the field of structural biology and proteomics[17]–[19]. Typically, new information on the structure of the formed complex, binding partner identity, binding site topology and more is gathered by performing a crosslink experiment followed by enzymatic digestion and advanced mass spectrometric analysis. This strategy has successfully led to numerous new insights in the complex worlds of structural and chemical biology in the recent past and will continue to elucidate nature's biochemical cellular processes in the coming decades[20][21].

Figure 3: Principle of peptide-protein crosslinking

3

The two most widely used classes of crosslinking reagents for the strategy described above are amine- or thiol-reactive crosslinkers (e.g. carbodiimides, reactive esters, isocyanates…) and photoreactive cross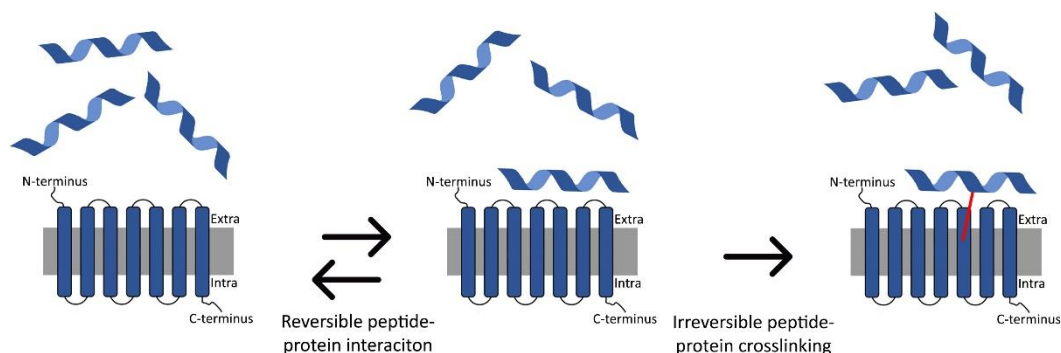linkers (e.g. azides, benzophenones…)[22]. While the first class relies on specific reactivity from certain amino acids (e.g. lysine, cysteine, …) present in the target protein, the latter can react with a larger variety of amino acids to form the covalent crosslink-bond upon activation of the photoreactive chemical moiety. The two can also be combined, where the introduced bifunctional molecule possesses for example an amide-sensitive group on one side of the molecule in combination with a photoreactive moiety on the other side[23].

Apart from different classes of chemical reagents, several different implementations of these reactive groups exist in the crosslinking field. Reactive esters for example can be used in the form of a homobifunctional molecules, serving as glue to covalently bind many interacting peptide-protein pairs at once (often called 'shotgun approach')[24]. Photoreactive crosslinkers on the other hand can be incorporated by means of an unnatural amino acid in the sequence of peptides or proteins in order to enable a more selective crosslinking experiment. This unnatural amino acid can be introduced either during chemical peptide synthesis, or through genetic incorporation.

In order to extract accurate information from crosslinking experiments, it is of high importance that the crosslinking event occurs in a natural cellular environment, as altered conditions like addition of certain chemical reagents or UV-light can have detrimental effects on cells and can even yield false positive results[25].

1.4   Furan-based crosslinking

Over the past ten years, the Organic and Biomimetic Chemistry Research group (OBCR) at the University of Ghent has built up considerable expertise in furan-based modification of biomolecules and subsequent interstrand crosslinking of DNA[26]–[30], RNA[31] and PNA (peptide nucleic acid)[32][33], peptide-peptide ligations[34][35] or ligand-receptor crosslinking[36]. Incorporation of a furan residue in sequences of nature's most relevant molecules, followed by oxidation and subsequent crosslinking to the biological binding partners has led to many noteworthy publications.
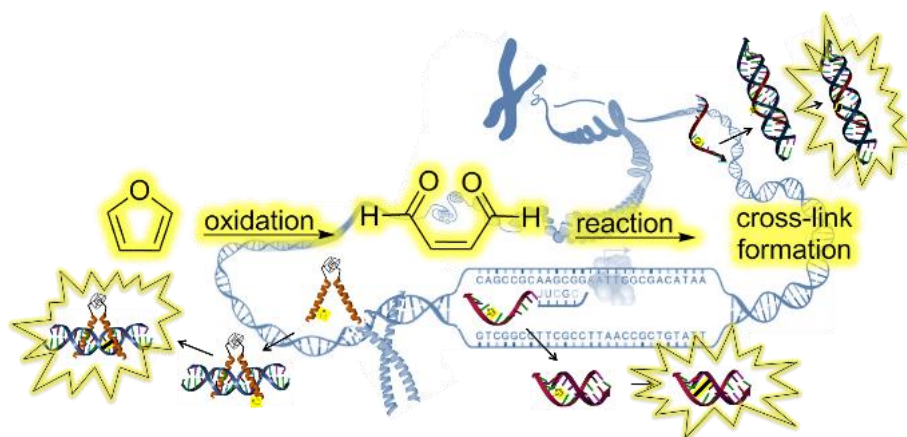
Figure 4: Applicability of furan-modified crosslinking on nucleic acid systems[30]

The reactivity of the initially stable furan moiety is based upon its oxidation to the corresponding keto-enal, which is receptive for nucleophilic attack by proximal amine or sulfhydryl groups present in the binding partner (Figure 5). The furan oxidation can be performed in several ways: *in vitro* by selective N-bromosuccinimide (NBS) oxidation[26] or by singlet oxygen produced by photosensitizers when irradiated with a visible light source[29]. Alternatively, oxidation can even occur spontaneously in living cells by endogenously produced reactive oxygen species (ROS)[36].
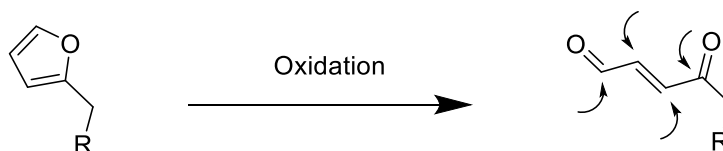
Figure 5: Furan oxidation to keto-enal, arrows indicate possible nucleophilic attack sites.

The potential for endogenous oxidation of furan-containing molecules when applied on living cells, was serendipitously discovered through a control experiment during previous research. It was observed that furan-crosslinking unexpectedly occurred without the addition of the NBS oxidizing reagent to the cell culture when crosslink experiments on the KISS-receptor were performed with a biotin-furan-modified kisspeptin-10 (BFK) ligand. It was hypothesized that ROS produced by NADPH oxidases (NOX) contained in the cell membrane were responsible for this oxidation. This hypothesis was supported by the fact that crosslinking occurred to lesser extent when NOX inhibitors were added to the cell culture[37]. In addition, when a

5

scrambled kisspeptin-10 peptide containing a furan at the same position was added, no crosslinking was observed, highlighting the very high selectivity of the crosslinking process.

In the field of structural biology and proteomics, this endogenous oxidation in cells has substantial benefits over the widely used alternatives. Experiments can be performed with complete omission of the formerly required addition of chemicals or administration of UV-light, which are known to have an influence on the cellular environment and thus interfere with the results of experiments[25].

In addition, this discovery even offers the possibility to further expand the applicability of furan-based crosslinking beyond fundamental research in structural biology and proteomics. The furan moiety is intrinsically stable prior to oxidation and is easily incorporable in peptide sequences through the commercially available Fmoc protected 3-(2-furyl)alanine, which is suitable for solid phase peptide synthesis (SPPS). These two attractive characteristics of furan-based crosslinking could potentially pave the road for novel pharmaceutical applications.

1.5 G protein-coupled receptors (GPCRs)

One of the necessities for the existence of multicellular living organisms is communication between cells and subsequent reaction to signals. This cellular communication goes under the name of cell signaling and plays a crucial role in homeostasis, metabolism, tissue function and repair, immunity and much more. The signal carriers in the world of cellular biology are called hormones and can be chemically categorized as ranging from amino acid derivatives to full peptides and more lipid-like molecules (e.g. steroids).

Signal reception of these hormones is achieved by so-called receptors, of which the G protein-coupled receptors are a major class, with over 750 known GPCRs found in human cells. The role of these receptors is to sense the hormone's presence at the outer side of the cell membrane and to initiate a cellular response at the inner side of the cell membrane. Receptors of this class are able to sense a wide variety of signals: peptides/proteins, lipids, sugars and even energy in the form of light.

A signal is transduced over the cell membrane by means of a conformational change in the receptor. A GPCR consists of a single, long peptide chain, but possesses a distinct three-dimensional folded structure, just like many proteins. This folded structure consists of a series of loops that cross the cell membrane seven times, giving rise to three distinct zones in the receptor: the extracellular part responsible for hormone sensing (ligand binding), the seven transmembrane helices and the intracellular part triggering a response inside the cell. The N-terminus is located on the outside of the cell, the C-terminus on the inside (Figure 6).
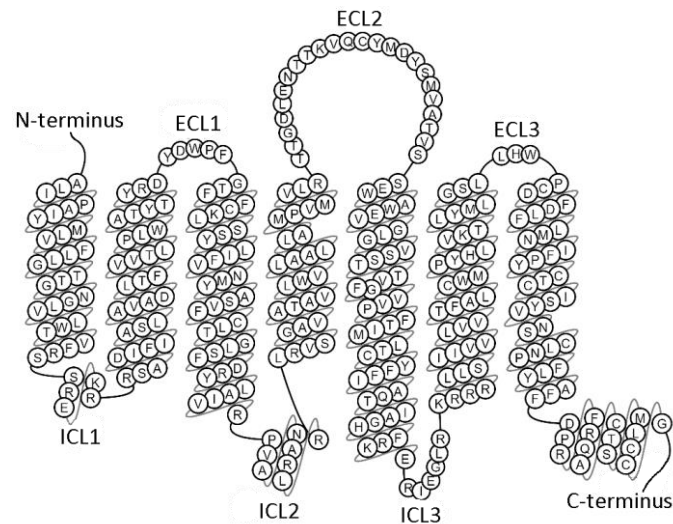
Figure 6: Schematic representation of Apelin receptor (GPCR) ECL = extra-cellular loop, ICL - intra-cellular loop.

The N-terminus on the outside of the cell together with the extracellular loops are responsible for ligand specificity and are known to vary a lot between different types of GPCRs. The N-terminus plays a particular key role when the ligands are larger molecules like peptides[38], as small molecules are known to interact more often with the transmembrane part[39].

As GPCRs are known to regulate many cellular processes in the human body, they are a popular drug target. A third of all FDA-approved drugs are focused on this type of receptors, targeting 108 different members of this receptor family[40]. Nevertheless, there are a considerable number of GPCRs that currently still have undiscovered ligands. These so-called orphan GPCRs are a great source of future drug targets and are being heavily investigated.

2 Aims and Objectives

The OBCR group has developed a novel crosslinking strategy that involves endogenous, *in situ* oxidation of a furan moiety incorporated into a peptide ligand, in an attempt to bridge the gap between peptide therapeutics and covalent drugs. Peptide therapeutics as well as covalent drugs present intrinsic challenges, which could be mitigated by combining the two in one in the design of peptides with the potential to form covalent bonds towards their targets. The challenging pharmacokinetic properties of peptide therapeutics on one side can be countered by a very high potency resulting from a covalent mode of action. On the other side, off-target covalent binding should be prevented by the furan moiety only being reactive after *in situ* oxidation on living cells. Peptide therapeutics also possess a very specific molecular structure which grants them their typical high selectivity towards a specific target. This combination could contribute to a possible solution for the increasing problem of drug resistance that society is facing today, and tomorrow[41][42].

This thesis is focused on the synthesis and study of a variety of furan-modified peptide ligands for the apelin and μ-opioid receptor. To introduce the furan moiety into the peptide ligand, the commercially available Fmoc protected 3-(2-furyl)alanine, an unnatural amino acid containing a furan moiety (Figure 7), is utilized. The peptide ligands will be synthesized using solid phase peptide synthesis (SPPS), a procedure

Figure 7: Fmoc-3-(2-furyl)alanine.

that has nowadays become the go-to approach for synthesizing peptides of moderate size (less than 50 amino acids). The Fmoc protected 3-(2-furyl)alanine (Fua) is compatible with standard SPPS, yet a cleavage cocktail optimization will be performed to maximize the overall synthetic yield of the furan-modified peptides.

Once this goal is achieved, the aim of the thesis is to evaluate the general applicability and robustness of the furan crosslinking strategy based on endogenous oxidation, a method which has previously been demonstrated on the KISS-receptor - kisspeptin-10 system in a proof of concept study[37]. For this, two therapeutically relevant receptors, both being GPCRs, will be targeted: the apelin receptor and the μ-opioid receptor. The applicability of the novel crosslinking strategy will be evaluated by cell incubation tests using the in-house synthesized furan-modified peptides, followed by Western blot experiments. Several variables in the cell

incubation experiments will be investigated in order to obtain a more detailed understanding of the factors influencing the crosslinking process and efficiency, which are currently not yet entirely understood. A schematic overview of the aims and objectives of this thesis can be found in Figure 8.

In addition to performing peptide synthesis and crosslinking experiments, *in-silico* visualization and simulation of the apelin ligand-receptor complex will be attempted, based on a confirmed molecular crystal structure[43]. Inside the binding pocket, distances between the furan moiety (incorporated in the ligand) and possible nucleophiles in the receptor in close proximity to the furan, will be calculated. The specific position of the furan in the ligand sequence might result in this distance being shortened or lengthened, which could possibly be reflected in the crosslinking efficiency of the furan-modified apelin peptide with the receptor. Combining results from the crosslinking experiments with structural insights obtained from this *in-silico* visualization and simulation could lead to a deeper understanding of the actual covalent bond being formed. This part of the thesis was performed as a supplementary assignment to compensate for the loss of lab time caused by the COVID-19 outbreak.
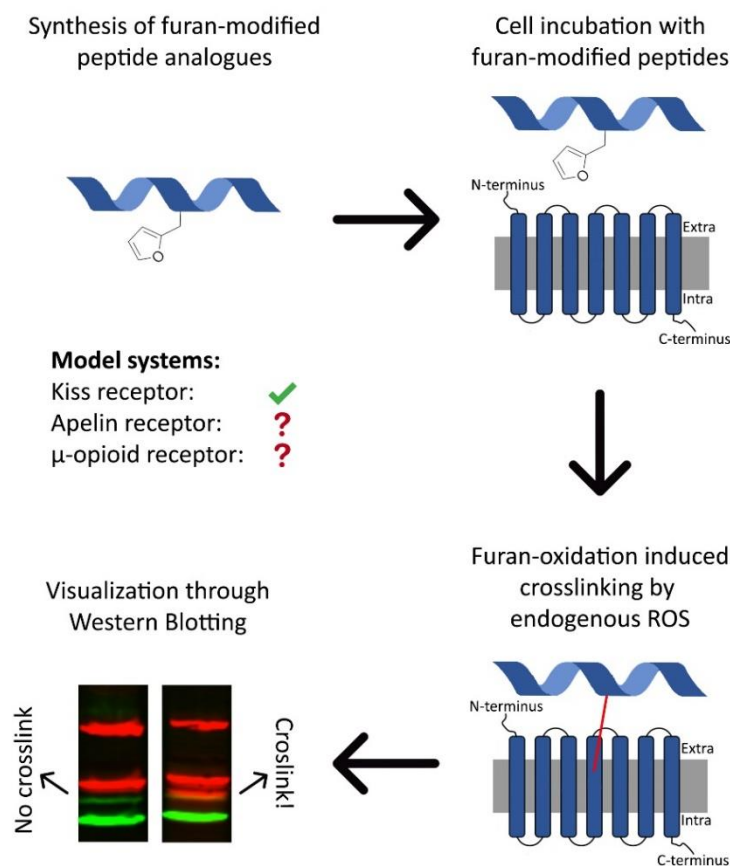

Figure 8: Schematic representation of aims and objectives.

3 Relevance of the chosen model systems

3.1 Model system 1: apelin receptor

The apelin receptor, also known as the APJR receptor, is a G protein-coupled receptor (GPCR) belonging to class A, also known as the rhodopsin-like GPCRs. The apelin receptor is consistently expressed in human cells, most notably in the central nervous system, lungs and mammary gland[44]. The apelin receptor has several known natural apelin peptide ligands: apelin 36, apelin 17 and apelin 13 being the shortest (the number index indicates the number of amino acids in the peptide ligand, for complete structures *vide infra* section 4.1). On top of this, Elabela has recently been proven as a second type of endogenous peptide ligand of the apelin receptor[45]. The apelin receptor has been associated with cardiovascular regulations and fluid homeostasis. It also plays a role in the inhibition of HIV infection and has an influence on the glucose and energy metabolism, making it an interesting drug target. The different apelin peptide ligands have non-identical potencies in regulating the functions above, the longest ligand (apelin 36) plays a more effective role in the HIV infection inhibition, while the shortest peptide (apelin 13) has a more pronounced effect on the cardiovascular system[46][47]. Although some mysteries regarding the apelin receptor and its ligands remain unsolved, furan-modified apelin ligands could prove to be very useful covalent drugs for cardiovascular and metabolic diseases.

3.2 Model system 2: μ-opioid receptor

Opioid receptors, just like the apelin receptor, can be categorized as GPCRs belonging to class A. Opioid receptors can be categorized in three main categories: μ-opioid receptors (MOR), δ-opioid receptor (DOR) and the κ-opioid receptor (KOR)[48]. The opioid receptors play their part in the mechanism of the most effective analgesics, which is the scientific term for compounds that relief pain[49][50]. As opioid receptors are some of the oldest drug targets, responsible for the physiological effects of e.g. opium, intensive research has resulted in a wide variety of known ligands, ranging from small molecules to endogenous or synthetic peptides[51]. These different ligands all bind to the different opioid receptors with varying affinity and specificity. Opioid pain relief therapy is an established practice but can come with harmful side effects (nausea, dizziness, vomiting) and possible addiction to the administered substances. On top of that, a typical phenomenon is desensitization of the receptor, something which is a known

issue for GPCRs[52]. Given the wealth of information concerning these GPCR systems and the existence of relevant peptide ligands, it was decided to also investigate furan-modified versions thereof as a third potential proof of concept system.

3.3 Scientific relevance of the model systems

Both systems are therapeutically interesting G protein-coupled receptors belonging to class A. The choice for these systems was based on previous collaborations with Confo Therapeutics (for the apelin receptor system) and with Professor Steven Ballet from the University of Brussels (for the μ-opioid receptor system). Some digging inside Web of Science showed that the μ-opioid receptor has been studied extensively for many years, while the apelin receptor is a rather novel receptor, first discovered in 1993[53], where a lot remains to be studied (Figure 9).
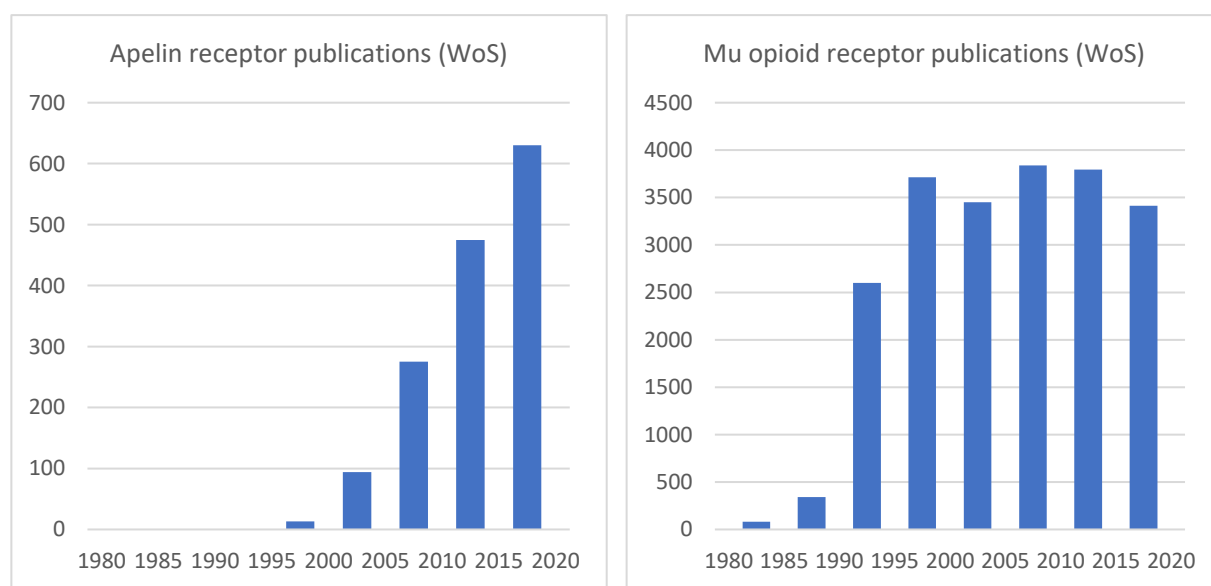

Figure 9: Web of Science publications on the apelin and μ-opioid receptors over the last 40 years.

4 Synthesis of furan-modified peptides

4.1 Peptide sequence design

In the following part, the sequence design of the synthesized furan-modified peptides will be addressed. All synthesized peptides include the 3-(2-furyl)alanine amino acid (except for the "native(Nle)" apelin analogue) and a biotin moiety. The furan-moiety will be responsible for forming the covalent bond in the ligand-receptor crosslinking event. A biotin moiety is attached to the peptides to allow detection of the crosslinked ligand-receptor complex through Western blotting experiments. The biotin moiety is attached via a PEG-linker to avoid as much as possible any steric interference in the interaction between the peptide ligand and the receptor. In addition, the PEG-linker increases accessibility of the biotin moiety after crosslinking, facilitating its recognition by streptavidin for visualization purposes. More information on this can be found in section 5.3. In total 9 different apelin peptides and 6 different opioid peptides were synthesized, all sequences can be found in appendix A.

4.1.1 Apelin receptor ligands

The synthesized furan-modified apelin sequences are based upon insights originating from the crystal structure of the apelin receptor obtained by Ma et al.[43] and earlier preliminary work performed by Dr. Khoubaib Ben Haj Salah within the OBCR group. The crystal structure reported in the paper by Ma et al. describes a modified apelin 17 peptide ligand interacting with the apelin receptor. In addition, the paper reports work performed using molecular dynamics focused on the interaction between natural apelin 13 and the apelin receptor. The combined results of the work performed by Ma et al. can be found in Figure 10, where an overlay is presented of the obtained crystal structure (yellow and blue) and the molecular dynamics results (pink and light blue).
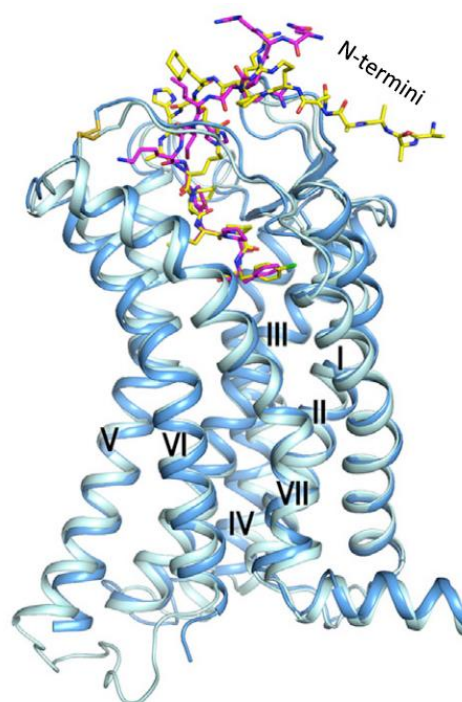

Figure 10: Overlay of the crystal structure of the modified apelin 17 - apelin receptor complex (yellow - blue) and molecular dynamics results of the apelin 13 - apelin receptor complex (pink - light blue)[43].

When examining Figure 10 in detail, it is remarkable how the C-terminal parts of both peptide ligands occupy almost the exact same available space. However, the position of the N-terminal part differs a lot, which indicates a less stringent interaction of these parts with the receptor. It was hypothesized that the N-terminus of the peptide ligands play a lesser role in the ligand-receptor interaction and positioning of residues in that part is more difficult to predict or rationalize. Therefore it was decided to use the shorter natural apelin 13 as a template for ligand design. As illustrated in Table 1, synthesizing apelin 13 peptides instead of apelin 17 peptides only results in a loss of N-terminal amino acids that seem to play a less important role in binding to the receptor, no C-terminal amino acids are lost.

Apelin peptide	Sequence (N-terminus ➜ C-terminus)																
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17
Apelin-17	Lys	Phe	Arg	Arg	Gln	Arg	Pro	Arg	Leu	Ser	His	Lys	Gly	Pro	Met	Pro	Phe
Modified-17	Lys	Phe	Arg	Arg	Gln	Arg	Pro	hArg	Cha	Glu	His	Lys	Lys	Oic	Nle	Pro	4-Cl-Phe
Apelin-13					Gln	Arg	Pro	Arg	Leu	Ser	His	Lys	Gly	Pro	Met	Pro	Phe
					1	2	3	4	5	6	7	8	9	10	11	12	13

Table 1: Apelin peptide sequences, modified-17 is the peptide synthesized by Ma et al.[43] for the crystal structure determination, blue residues are unnatural amino acids, red residues represent mutations used to form a lactam ring. The yellow and pink represent the colors of the peptide ligands depicted in Figure 10.

In previous work performed by Dr. Khoubaib Ben Haj Salah (at the OBCR group) and Confo Therapeutics, it was studied at which positions the 3-(2-furyl)alanine could be inserted in the apelin 13 peptide, without losing the interaction with the apelin receptor. In addition, Dr. Khoubaib Ben Haj Salah studied if having a carboxylic functionality or an amide functionality at the C-terminus of the peptides made a difference towards receptor interaction (both functionalities are synthetically accessible by simply varying the linker in SPPS, see section 4.2.2). Lastly, the influence of N-terminal cyclization of the glutamine to pyroglutamic acid was also examined (Pyr(1) apelin 13), as this is a stabilized isoform largely present in the human cardiovascular system[54].

These studies were carried out with furan-modified apelin 13 peptides with amino acids 1, 5, 7, 11, 12 or 13 counting from the N-terminus being replaced by 3-(2-furyl)alanine in the peptide sequence. All sequences can be found in Table 2. As a next step, indirect radioligand binding assays (RBA) were performed to determine the binding affinity of the various

synthesized furan-modified apelin peptide ligands. RBA is the most sensitive quantitative approach to measuring receptor binding parameters *in vitro*.

Code	Sequence (N-terminus ➜ C-terminus)
Fur(13) Apelin 13:	H-GlnArgProArgLeuSerHisLysGlyProMetProFua-OH
Fur(13) Apelin 13-NH$_2$	H-GlnArgProArgLeuSerHisLysGlyProMetProFua-NH$_2$
Fur(12) Apelin 13	H-GlnArgProArgLeuSerHisLysGlyProMetFuaPhe-OH
Fur(12) Apelin 13-NH$_2$	H-GlnArgProArgLeuSerHisLysGlyProMetFuaPhe-NH$_2$
Pyr(1) Apelin 13	H-PyrArgProArgLeuSerHisLysGlyProMetProPhe-OH
Apelin 13	H-GlnArgProArgLeuSerHisLysGlyProMetProPhe-OH
Apelin 13-NH$_2$	H-GlnArgProArgLeuSerHisLysGlyProMetProPhe- NH$_2$
Fur(11) Apelin 13	H-GlnArgProArgLeuSerHisLysGlyProFuaProPhe-OH
Fur(7) Apelin 13	H-GlnArgProArgLeuSerFuaLysGlyProMetProPhe -OH
Fur(5) Apelin 13	H-GlnArgProArgFuaSerHisLysGlyProMetProPhe-OH
Fur(1) Apelin 13	H-FuaArgProArgLeuSerHisLysGlyProMetProPhe-OH

Table 2: (Furan)-modified apelin 13 peptide sequences used in RBA.

There are three main ways of performing radioligand binding assays based on saturation binding, indirect binding, or kinetic binding[55]. The furan-modified apelin analogues were tested with an indirect binding assay, which is based on binding competition. A fixed concentration of apelin receptor in membrane extracts was incubated with a fixed concentration of radioactively labeled apelin 13 ligand (125-I apelin 13, an apelin peptide with radioactive ^{125}I bound to the lysine using iodinated Bolton & Hunter reagent[56]) and varying concentrations of the unlabeled ligand of interest (furan-modified apelin peptides in this case). The radioligand binding in the absence of unlabeled ligand typically delivers the maximal binding in the assay (B$_0$) and increasing concentrations of unlabeled ligand progressively inhibit radioligand binding through competition. In most cases, a semilogarithmic plot of bound radioligand against the ^{10}log of the concentration of unlabeled ligand gives an inverse sigmoid curve with a maximum equal to the B$_0$ and a minimum equal to nonspecific binding (NSB), while the unlabeled ligand concentration that inhibits binding by 50% (IC50) may be used to estimate the binding affinity of the unlabeled ligand of interest (Figure 11).

In the assays performed by Confo Therapeutics, unlabeled apelin 13 (without a furan) was also included as one of the 'unlabeled ligands of interest' as a positive control, allowing quantitative comparisons with the furan-modified apelin analogues. In case the inverse

sigmoid curve (signal output) for a certain unlabeled furan-modified apelin ligand is displaced to the left compared to the inverse sigmoid curve of the unlabeled non-furan-modified apelin 13 (positive control), this means that the unlabeled furan-modified apelin has a more pronounced interaction with the receptor (stronger competition at lower concentrations directly relates to higher binding affinity). The opposite is also true, when the inverse sigmoid curve of a certain unlabeled furan-modified apelin peptide is located to the right compared to the unlabeled non-furan-modified apelin 13 (positive control), the furan-modified ligand has less binding affinity. The results of the radioligand binding assays can be found in Figure 11.
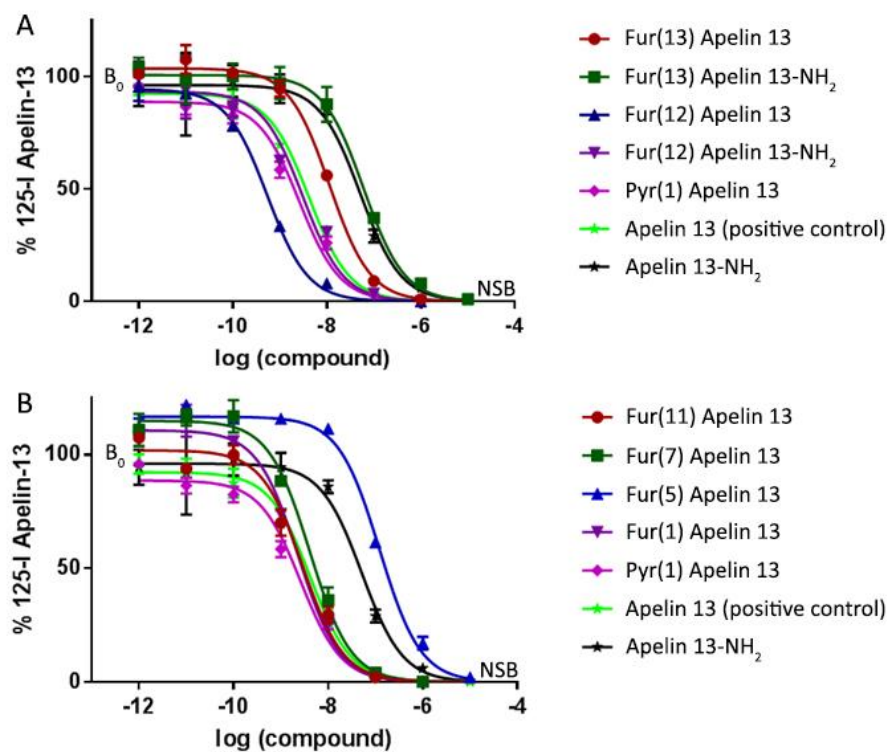
Figure 11: Results of indirect radioligand binding assay on (furan-) modified apelin 13 peptides, performed by Confo Therapeutics.

Although some of the inverse sigmoid curves of furan-modified apelin analogues are clearly located to the right compared to the non-furan-modified apelin 13 ligand (meaning lower binding affinity compared to the non-furan-modified apelin 13 ligand), these results still demonstrate interaction with the receptor at higher concentrations. The experiments prove that incorporation of the 3-(2-furyl)alanine in apelin 13 on positions 1, 5, 7, 11, 12 or 13 is allowed with conservation of biological activity. In addition, it can be observed in Figure 11A that non-furan-modified apelin 13 and two different furan-modified apelin peptides (furan at position 12 or 13) containing a carboxylic acid at the C-terminus have higher binding affinities

compared to their C-terminal amide equivalents. From these results it was decided to only synthesize the remaining furan-modified apelin peptides with a C-terminal carboxylic acid, which explains the lack of data for other C-terminal amide furan-modified apelin peptides in Figure 11B. In addition, it can be observed that the binding affinity of the Pyr(1) apelin 13 is very similar to the non-furan-modified apelin 13.

After completion of the experiments performed by Confo Therapeutics, furan-modified apelin peptides with the 3-(2-furyl)alanine at position 11, 12 or 13, containing a carboxylic acid at the C-terminus, were selected for further studies. Based upon the analysis of the crystal structure reported by Ma et al.[43] (vide infra, section 6.1) it can be noticed that these three modification sites are in close proximity to nucleophilic amino acids (Lys, Cys and His) and to amino acids that can engage in electrophilic aromatic substitutions (Tyr and Trp) present in the targeted apelin receptor. Furan-modified apelin peptides with the 3-(2-furyl)alanine at positions 1, 5 or 7 were discarded as it can be observed in the crystal structure that these positions at the N-terminal part of the apelin 13 peptide have a less stringent interaction with the receptor. It was hypothesized that it would be more difficult to form a crosslinked product if the furan moiety was placed in these positions.

The next step in the sequence design was to include a biotin moiety in the selected furan-modified apelin 13 peptides for easy visualization by a fluorescent-streptavidin tag during Western blotting. It was chosen to couple the biotin via a PEG4 linker to the N-terminus of the furan-modified apelin 13 peptides, as this side of the peptide is located in a more solvent exposed area of the receptor upon interaction, as can be seen in the crystal structure. Attachment of the biotin and linker was achieved by performing a regular peptide coupling step between biotin-PEG4-propionic acid and the free N-terminal amine of the peptide, using the same coupling reagents as during the synthesis of the peptide, which will be described in more detail later in this chapter.

An additional element in the sequence design of the apelin peptides was introduced after synthesis of the first two apelin analogues, as troublesome methionine (Met) oxidation was observed during cleavage of the peptides. To solve this issue, all remaining apelin peptides were synthesized with norleucine (Nle) as a methionine replacement (Figure 12). This is a fairly common modification that has been proven to retain binding affinity[57]–[59]. This replacement of methionine by norleucine was also included in the scrambled sequences and the native

apelin 13 ligand (containing no furan). It is worth noting that this modification was also introduced by Ma et al. in their modified apelin 17 ligand to obtain the crystal structure[43].
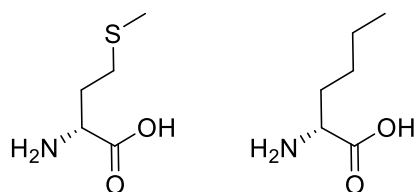

Figure 12: Methionine (left), Norleucine (right), note the sterical similarity.

In total, 9 different furan-modified apelin 13 peptides were synthesized. All sequences can be found in Table 3. Looking at the codes in Table 3, AP stands for apelin peptide, the number 1,2 or 3 indicates the position of the furan moiety counting from the C-terminus, Met and Nle indicate which amino acid was included at the third position starting from the C-terminus. Figure 13 shows the detailed chemical structure of the first synthesized apelin 13 analogue: AP1(Met), which contains the 3-(2-furyl)alanine at the first position from the C-terminus, a methionine at the third position from the C-terminus and has a biotin-PEG4 at the N-terminus (every synthesized apelin analogue has a biotin-PEG4 attached to the N-terminus). An apelin peptide label starting with the letter S (e.g. SAP2(Nle)) means the apelin 13 peptide sequence was scrambled for later control experiments, as this should nullify the ligand-receptor interaction. The 'Native' apelin 13 peptide does not contain the 3-(2-furyl)alanine.

| Code	Sequence (N-terminus ➜ C-terminus)
AP1(Met):	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProMetProFua-OH
AP2(Met):	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProMetFuaPhe-OH
AP1(Nle):	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProNleProFua-OH
AP2(Nle):	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProNleFuaPhe-OH
AP3:	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProFuaProPhe-OH
SAP1(Nle):	Biotin-PEG4-GlyProLysLeuNleArgProGlnHisArgProSerFua-OH
SAP2(Nle):	Biotin-PEG4-GlyProLysLeuNleArgProGlnHisArgProFuaSer-OH
SAP3(Nle):	Biotin-PEG4-GlyProLysLeuNleArgProGlnHisArgFuaProSer-OH
Native AP(Nle):	Biotin-PEG4-GlnArgProArgLeuSerHisLysGlyProNleProPhe-OH

Table 3: Furan-modified apelin 13 peptide sequences.
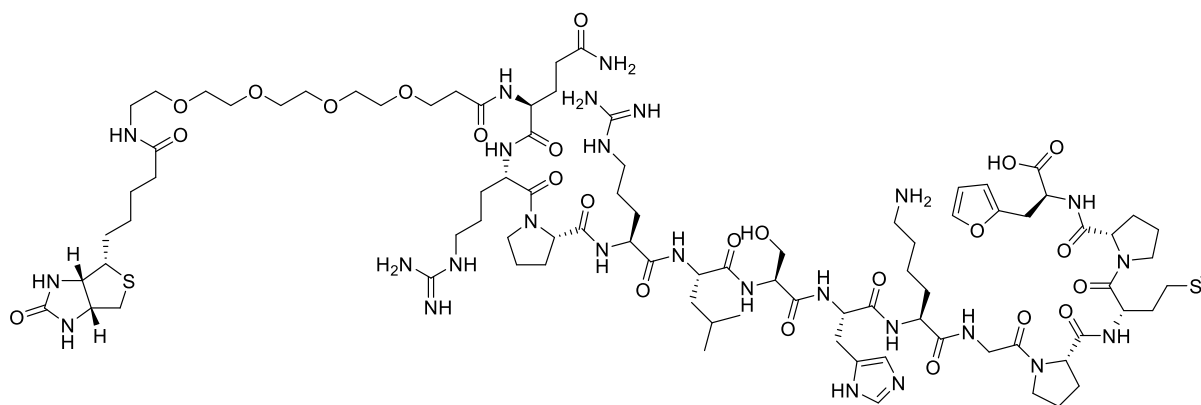
Figure 13: Chemical structure of AP1(Met).

4.1.2 μ-opioid receptor ligands

Currently, two crystal structures for the μ-opioid receptor interacting with a ligand have been reported[51][60]. However, both crystal structures contain a non-peptide ligand, making design of the envisaged furan-modified opioid peptides a bit more challenging. As a consequence, design of the μ-opioid peptides relies on published research concerning structure-activity relationships and additional expertise in this area of Professor Steven Ballet from the University of Brussels[61], [62]. As a starting point for the sequence design of opioid peptides for interaction with the μ-opioid receptor, the tetrapeptides endomorphin-1 and -2 (Figure 14) and the heptapeptide dermorphin (Figure 15) were used.

Endomorphin-1 is, just like endomorphin-2, considered a natural opioid tetrapeptide neurotransmitter involved in pain relief. Their C-terminal amide peptide sequences give rise to a folded tertiary structure which grants them their very strong and exclusive binding properties towards the μ-opioid receptor. Endomorphin-1 can be found in the brain and the upper brainstem, whereas endomorphin-2 is more concentrated in the lower brainstem and the spinal cord[63].
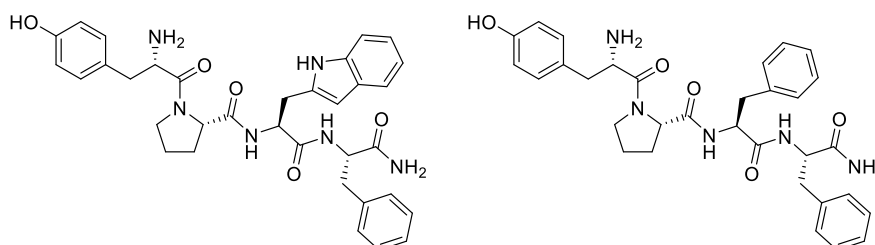

Figure 14: Structure of Endomorphin-1 (H-Tyr-Pro-Trp-Phe-NH$_2$, left) and Endomorphin-2 (H-Tyr-Pro-Phe-Phe-NH$_2$, right).

Dermorphin is a natural opioid agonist first isolated from frogs (Phyllomedusa)[64]. This C-terminal amide peptide is known to bind with high potency and selectivity to the μ-opioid receptor, being 30-40 times more potent than morphine[65]. This peptide cannot be found in the human body or other mammals, as it contains a D-amino acid (D-alanine), which is not included in the human genetic code. This means that the peptide cannot be synthesized naturally. However, bacteria, amphibians and mollusks are known to produce D-amino acid containing peptides through a posttranslational modification by amino acid isomerases.
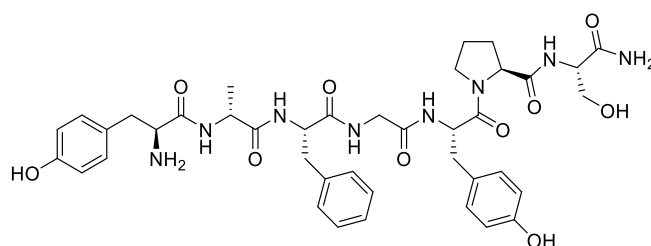

Figure 15: Chemical structure of Dermorphin (H-Tyr-D-Ala-Phe-Gly-Tyr-Pro-Ser-NH$_2$).

Extensive research over many years has led to the development of small synthetic opioid peptides with high bioactivity. Replacing the N-terminal tyrosine by a 2', 6' methylated analogue (2', 6'-dimethyl-L-tyrosine, Dmt) proved to be a successful approach in enhancing the biological activity[66] and will therefore be incorporated in the sequence design of the opioid peptides for this thesis. This knowledge was combined with the state of the art and expertise from Professor Steven Ballet from the University of Brussels, which has led to the six sequences of the furan-modified opioid peptides for this thesis (Table 4).

The furan-modified opioid peptides synthesized in this thesis are five to seven amino acids long and have an amide function at the C-terminus in accordance with the dermorphin and endomorphin-1 and endomorphin-2 structures. All sequences contain a Dmt-residue at the N-terminus to increase biological activity. D-alanine was introduced at position two to prevent protease cleavage (in accordance with the dermorphin sequence). A furan moiety was introduced as 3-(2-furyl)alanine at position three or four counting from the N-terminus to establish the covalent crosslink bond with the receptor. Phenylalanine was introduced at position three or four (except for OP5) as it is known that an aromatic amino acid at these positions increase the biological activity (in accordance with both the dermorphin and the endomorphin-1 sequence). A ß-alanine was included at position 5 to act as a spacer to prevent

steric hindrance between the pharmacophore part (positions 1-4) and the biotin-PEG3 moiety at the C-terminus.

As the N-terminal Dmt is known to be highly important for the opioid peptide bioactivity, attaching the biotin at the N-terminus in a similar way as performed with the apelin peptides was not an option.
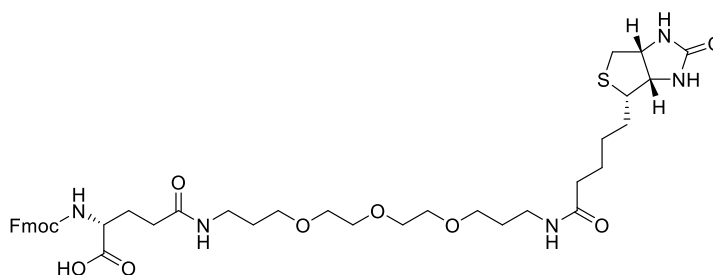

Figure 16: Commercially available Fmoc-Glu(biotinyl-PEG)-OH introduced at the C-termini of the furan-modified opioid peptides.

Attaching the biotin at the N-terminal side of the peptide would most likely cause steric interference and thus lower the ligand-receptor interaction. The solution for this was to couple the commercially available Fmoc-Glu(biotinyl-PEG)-OH at the C-terminus instead (Figure 16). Incorporation of this unnatural biotin-containing amino acid in the regular solid phase peptide synthesis described later in this chapter, was realized without problems. The biotin-PEG3 moiety at the C-terminus will be used for easy visualization of the crosslinked receptor-ligand complex in Western blot assays. The opioid peptides were arbitrary labeled as opioid peptide (OP) 1-6, the sequences can be found in Table 4. The chemical structure of the opioid peptide that has most in common with dermorphin, OP6, can be found in Figure 17.

Code	Sequence (N-terminus ➜ C-terminus)
OP1:	Dmt - D-Ala - Fua - Phe - ß-Ala - Glu - Glu(Biot) - NH$_2$
OP2:	Dmt - D-Ala - Phe - Fua - ß-Ala - Glu - Glu(Biot) - NH$_2$
OP3:	Dmt - D-Ala - Phe - Fua - ß-Ala - Glu(Biot) - NH$_2$
OP4:	Dmt - D-Ala - Fua - Phe - ß-Ala - Glu(Biot) - NH$_2$
OP5:	Dmt - D-Ala - Fua - ß-Ala - Glu(Biot) - NH$_2$
OP6:	Dmt - D-Ala - Phe - Fua - Glu(Biot) - NH$_2$

Table 4: Furan-modified opioid peptide sequences.
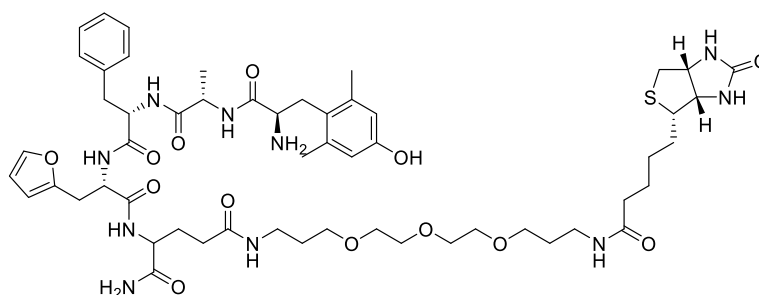

Figure 17: Chemical structure of OP6.

4.2   Solid phase peptide synthesis

As the furan-modified apelin peptides are 13 amino acids long and the furan-modified opioid peptides are 5 to 7 amino acids long, solid phase peptide synthesis (SPPS) is the most effective approach to the synthesis. SPPS allows synthesis of peptides with full control over the sequence up to 50 amino acids, a technique that has matured for over 50 years[67] into a very robust practice with established protocols and lots of expertise present within the OBCR group. The apelin peptides were synthesized using an automated peptide synthesizer. The opioid peptides on the other hand were synthesized manually, as they are relatively short and a number of expensive unnatural amino acids were used in the sequences.

4.2.1   General principle

Solid phase peptide synthesis combines the ease of in-solution organic chemistry with simple filtration washing steps. This is achieved by coupling the first amino acid of the desired sequence via its carboxyl group to an insoluble, yet swellable, polymer bead through a cleavable linker. After this initial step, the entire synthesis is based upon a repeated cycle of deprotecting the N-terminal amine group of the growing peptide chain attached to the insoluble resin, followed by performing the coupling reaction between the deprotected N-terminal amine and the carboxylic acid group of the next N-terminal protected amino acid in line, resulting in the desired peptide bond. This way, the peptide is built up from the C-terminus of the sequence to the N-terminus by repeating the deprotection-coupling cycle until the final amino acid is introduced. At this point, the peptide is ready to be removed from the polymer bead by cleaving off the linker and resin. A schematic overview of the general SPPS procedure can be found in Figure 18.
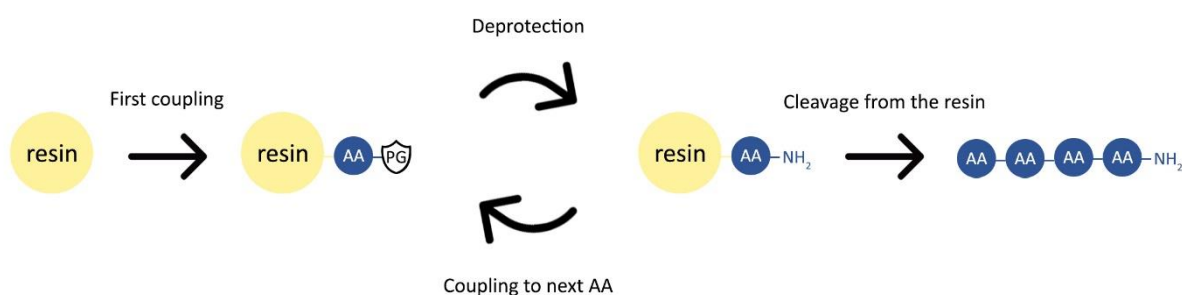

Figure 18: Schematic overview of the general SPPS procedure, 'AA' stands for amino acid, 'PG' stands for protecting group.

The key benefit of this approach to peptide synthesis is the unidirectional growth of the peptide chain, allowing easy sequential control. In addition, equimolarity between the added amino acids and the growing peptide chain is not required, allowing coupling reactions to be driven to full conversion by utilizing an excess of amino acid followed by extensive washing with dimethylformamide (DMF) and dichloromethane (DCM) of the polymer beads on a filter.

4.2.2   Polymer resin and linker

The choice of an adequate polymer resin to attach the peptides to is essential. Preferably, the polymer bead should exhibit a high degree of swelling in the solvent that will be used throughout the synthesis. When the polymer that contains the attached growing peptide chains is swollen properly, the required reagents for both the deprotection and coupling reactions can easily reach the N-terminal site, leading to higher efficiency of the reactions. In this thesis, polystyrene polymer beads were used containing 1% of the divinylbenzene crosslinker. Higher crosslinker concentration would cause a lower degree of swelling, yet some crosslinker is required to ensure mechanical rigidity of the polymer beads.

The first amino acid of the sequence (counting from the C-terminus) is attached to the resin using a pre-installed chemical moiety on the polymer bead, called the linker. A wide variety of linkers exist, allowing for tunable cleavage conditions and different C-terminal functionalities of the finalized peptide after cleavage. Some more exotic linker molecules can only be cleaved when activated in a prior step (safety catch linker) or can even be cleaved by enzymes or upon UV irradiation.

In this thesis, the 2-chlorotrityl chloride linker was used for the apelin peptides. This is an acid labile linker that results in a carboxylic acid functionality at the C-terminus after cleavage. For the opioid peptides, the Rink amide linker was used in order to obtain an amide functionality at the C-terminus after acidic cleavage (Figure 19). The resins were purchased with the linker molecule pre-attached. Coupling the first amino acid to the 2-chlorotrityl chloride resin is achieved through an SN_1 reaction in alkaline conditions, while the Rink amide linker is bought in a Fmoc protected form, which allows for the immediate start of the regular peptide synthesis (after initial Fmoc deprotection).
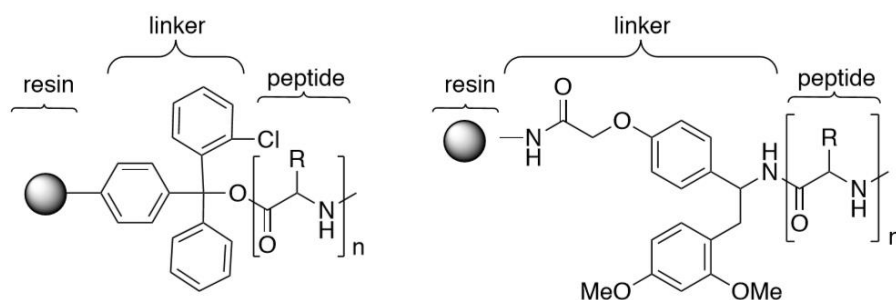
Figure 19: 2-Chlorotrityl chloride linker (left) and Rink amide linker (right).

4.2.3 Fmoc strategy and protecting groups

As amino acids are intrinsically at least bifunctional, preventing uncontrolled polymerization resulting in a random polyamide is key to allow sequential control during the peptide synthesis. The way this is achieved is by utilizing protecting groups, which temporarily shield the amine functionality of the amino acids. The two most frequently used protecting groups for this purpose are the tert-butyloxycarbonyl (Boc) group, which is removed by addition of an acid (e.g. TFA) and the fluorenylmethoxycarbonyl (Fmoc), which is removed by addition of a base (e.g. piperidine).

On top of protecting the amine functionality of the amino acid, many sidechains of amino acids contain reactive moieties that can interfere in the peptide synthesis process. These sidechains also require protecting groups. As these protecting groups are only supposed to be removed at the very end of the peptide synthesis, they should withstand the conditions used for deprotecting the Boc or Fmoc group during the repeated deprotection-coupling cycles. For this reason, orthogonal chemistries are used. In case of the acid-labile Boc, a common sidechain protecting group is benzyl, which requires higher acidity (e.g. HF) to be removed and thus withstands the acidic conditions used to remove the Boc. In the case of the base-labile
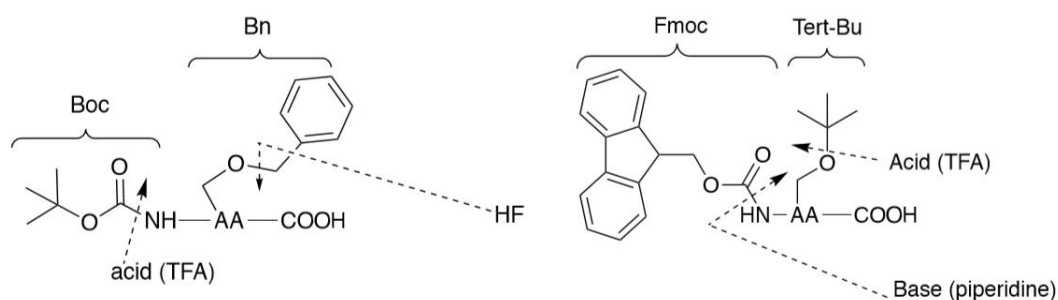

Figure 20: The two main protection strategies in SPPS: Boc/Bn (left) and Fmoc/tert-Bu (right).

23

Fmoc, tert-butyl groups are often used, which are removed by acid treatment (e.g. trifluoroacetic acid (TFA)) (Figure 20).

The more pronounced chemical orthogonality and more efficient cleavage reaction of the Fmoc/tert-butyl strategy (base/acid) compared to the Boc/benzyl strategy (acid/stronger acid) is the reason why, during this thesis, the Fmoc/tert-butyl strategy was utilized for the synthesis of all peptides. All amino acids used in the synthesis were purchased in a Fmoc protected, sidechain protected form.

During the synthesis, the base-labile Fmoc protecting groups were removed by addition of piperidine, the reaction mechanism for deprotection is shown in Figure 21. The sidechain protecting groups used during this thesis (Boc, tBu, Trt and Pbf) are removed by the same acidic conditions required to cleave the peptides from the resin. Cleavage and deprotection of the sidechains is performed simultaneously using a cleavage cocktail containing mainly TFA, more info on cleavage cocktails can be found in section 4.2.5.
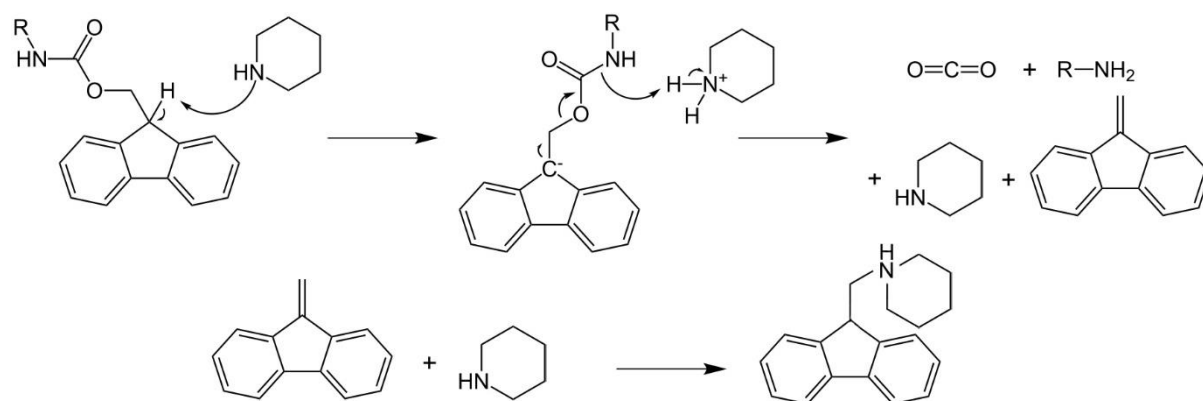
Figure 21: Fmoc deprotection mechanism (by addition of piperidine).

Coupling reagents (HBTU + DIPEA)

In order to form the desired amide bond between the N-terminal amine of the growing peptide chain and the carboxylic acid of the added Fmoc-protected amino acid, activation of the carboxylic acid is required to ensure an efficient coupling reaction. The coupling reaction can be categorized as a nucleophilic attack of the N-terminal amine on an activated carboxylic acid, which results in the formation of an amide bond. If the carboxylic acid would not be activated, an ordinary acid-base reaction would occur, yielding an undesired ionic bond between the protonated amine and the deprotonated carboxylic acid. For the activation of the carboxylic acid, so-called coupling reagents are commonly used.

Many coupling reagents were developed over the past thirteen decades of peptide synthesis research[68]. For this thesis hexafluorophosphate benzotriazole tetramethyl uronium (HBTU) was used as this coupling reagent benefits from a low degree of racemization of the activated amino acids. It fulfills its activating role after the carboxylic acid is being deprotonated by a non-nucleophilic base, for which di-isopropylethylamine (DIPEA) was used. After this deprotonation, it forms an activated ester which on its turn can efficiently react with the N-terminal amine to form the amide bond. After the coupling, any excess of coupling reagents and all side-products are efficiently washed away. The reaction mechanism of an HBTU assisted coupling reaction can be found in Figure 22.
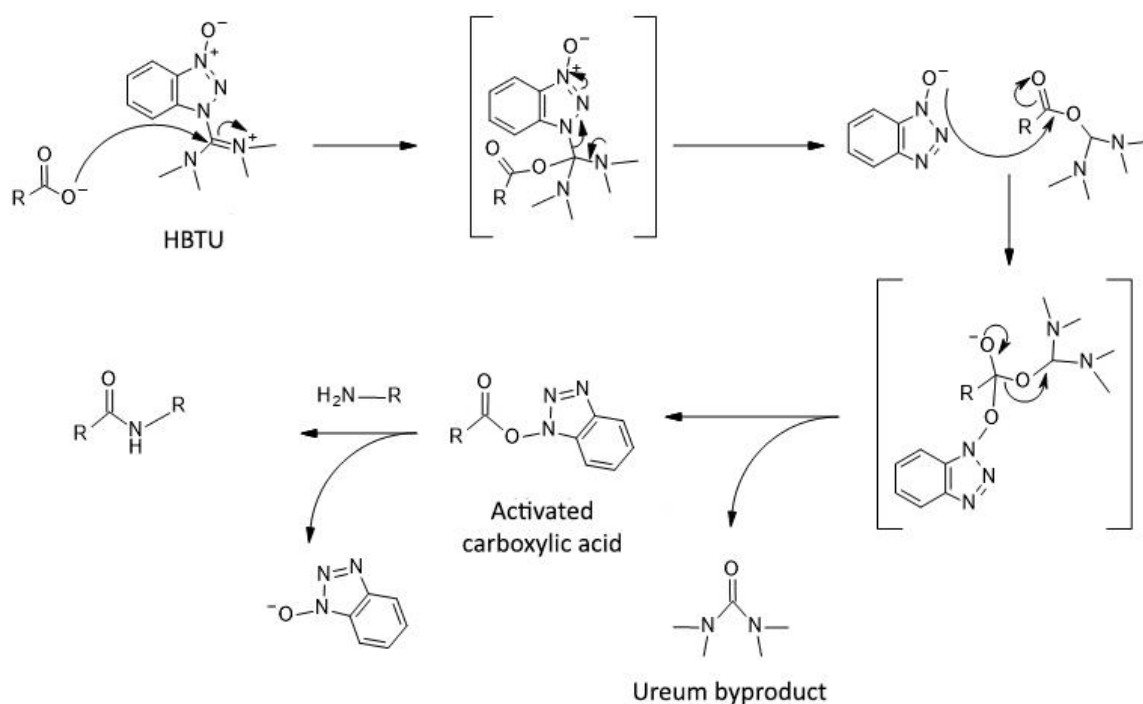
Figure 22: HBTU coupling mechanism.

4.2.4 TNBS test

When synthesizing a peptide manually, it is essential to be able to evaluate the conversion of the coupling steps throughout the entire synthesis. When a next coupling is performed before full conversion of the previous one, two different peptide chains are obtained: one with the intended sequence and one with a deletion of an amino acid. As these peptides only differ one residue, separation is nearly impossible.

For assessing the conversion of the amide coupling reactions, the trinitrobenzene sulphonic acid (TNBS) test was utilized, which is a qualitative color test where TNBS acts as coloring

agent. When free primary amines are present in the solution (or on the swollen beads in this case), they will react with the TNBS through nucleophilic aromatic substitution (Figure 23). This reaction product has a distinct bright red color. In practice, a very small amount of polymer beads is used to perform the test. When the coupling has not reached full conversion, the beads turn bright red. If the beads remain pale yellow, the coupling is finished. When red beads are observed, the previous coupling step and the subsequent TNBS-test are typically repeated, until no more red colored beads are observed.
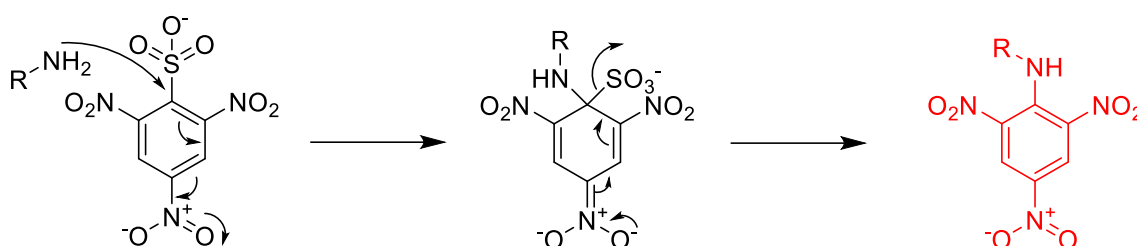

Figure 23: TNBS-test mechanism.

4.2.5 Cleavage cocktail optimization

Cleavage of the peptide from both the 2-chlorotrityl chloride and the Rink amide linkers is generally performed with cleavage cocktails containing 90% or more TFA with addition of certain scavengers. In the Fmoc/tert-butyl strategy used in this thesis, cleavage of the peptide happens simultaneously with the deprotection of the sidechain protecting groups. This deprotection generates stabilized carbocation compounds, which can subsequently react with electron rich sidechains leading to undesired byproducts. To mitigate this problem, scavenger molecules are typically used in the TFA-cleavage cocktails, which irreversibly capture these stabilized carbocations.

A typical cleavage cocktail for the linkers used in this thesis is a mixture of 95% TFA, 2.5% triisopropylsilane (TIS) and 2.5% H_2O, the last two being the scavengers. However, the furan residue contained in the peptide is known to degrade under harsh acidic conditions, most notably when the furan is close to one of the ends of the peptide chain. Introduction of the furan in the middle part of the peptide protects the furan to some extent from the degrading conditions. As for the apelin peptides, the furan is located at the C-terminus of the peptides, previously published work on this specific degradation issue from within the OBCR group was consulted[69].

In order to optimize the yield of the cleavage reactions, it was decided to perform test cleavages for comparing the classic 95% TFA, 2.5% TIS and 2.5% H_2O cocktail to four different cleavage cocktails (Table 5), which contained two commonly used scavengers that were not examined before in the published paper. In practice, a small number of beads from the AP1(Nle) synthesis was subjected to the different cleavage cocktails for 45 minutes and analyzed via reversed phase high performance liquid chromatography (RP-HPLC) and liquid chromatography followed by mass spectrometry (LC-MS) for determination of the yield and identification of by-products respectively. Not all by-products could be identified, yet the most prevalent by-products could be identified as the peptide with a reduced furan moiety, or peptide-protecting group adducts (+tBu, +Trt). The obtained HPLC and LC-MS spectra of all test cleavages can be found in the experimental data section (section 10.1.3.12-10.1.3.18).

Cleavage cocktail	Concentrations
Classic	95% TFA, 2.5% TIS, 2.5% H_2O
CC1	95% TFA, 2.5% Thioanisole, 2.5% TIS
CC2	92.5% TFA, 2.5% Thioanisole, 2.5% TIS, 2.5% H_2O
CC3	95% TFA, 2.5% thioanisole, 2.5% m-cresol
CC4	97.5% TFA, 2.5% m-cresol

Table 5: Tested cleavage cocktails

To summarize, both CC1 and CC3 initially showed promising results, resulting in higher yields compared to CC2, CC4 and the classic cocktail. CC1 allowed to obtain the highest yield of the two, as can be seen in Figure 24. CC3 led to a slightly lower yield, but the resulting crude mixture did not contain the reduced furan by-product which is located in close proximity
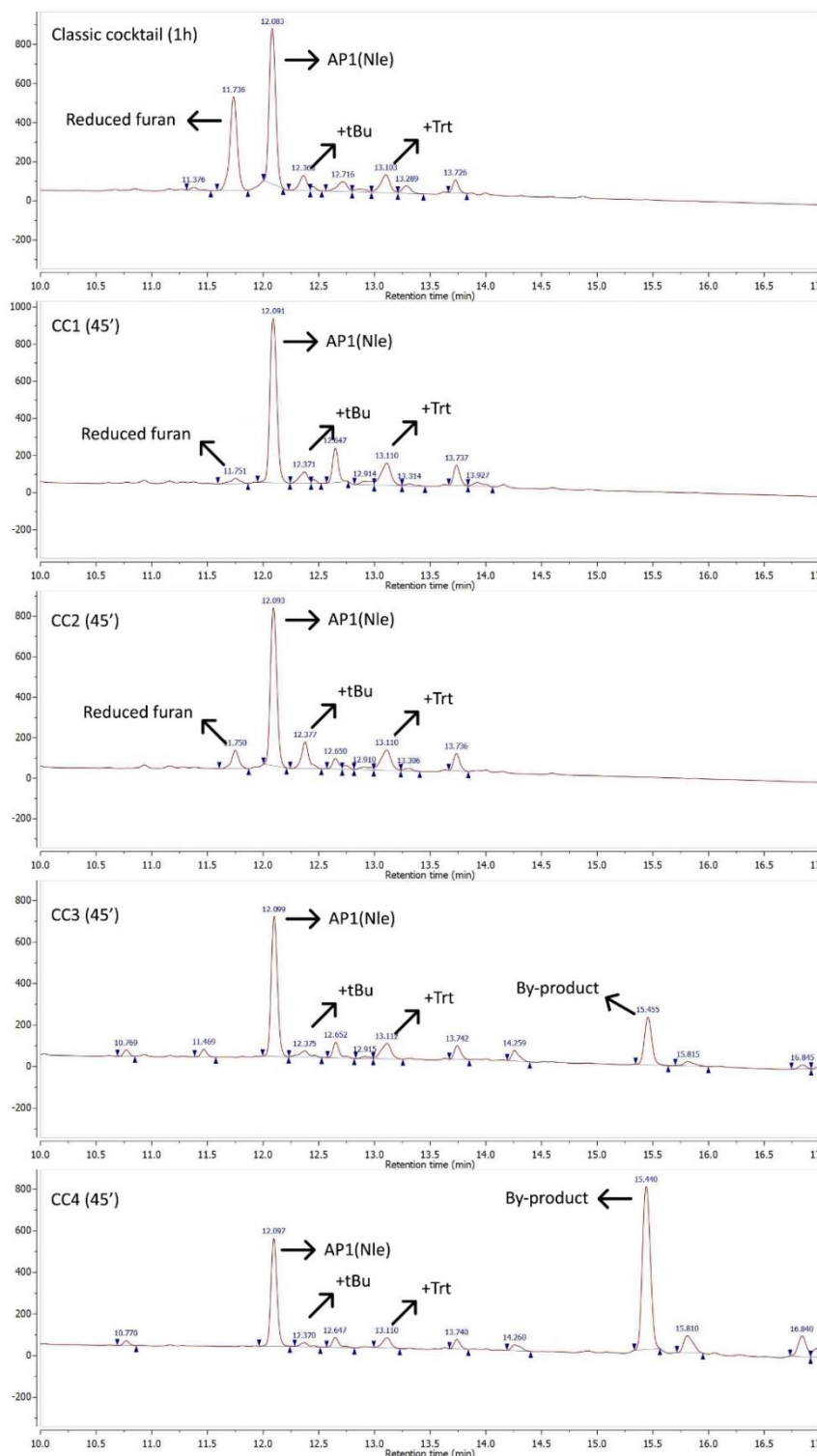

Figure 24: Close-up of most relevant test cleavage HPLC data, identification based on LC-MS data, more details can be found in section 10.1.3.

(HPLC-wise) to the main product in the RP-HPLC chromatogram, which is typically difficult to remove during purification. In CC3 and CC4, an unidentified by-product was formed (peak at minute 15.5), which is not present after cleavage with the classic cocktail, CC1 or CC2.

A second test cleavage was performed with a longer cleavage time of one hour and 45 minutes in order to determine more clearly the winner between CC1 and CC3. CC1 came out on top, as upon CC3 treatment, an increased amount of an unidentified by-product was formed after this longer cleavage time, as can be seen in Figure 25. It has to be noted that these spectra have a slight shift in elution time (x-axis) compared to the spectra from Figure 24. This is caused by usage of a larger loop in the HPLC system, resulting in a larger dead volume and therefore later peak elution. In addition, the CC3.2 sample was more diluted than the CC1.2 sample, as can be deduced from the y-axis displaying UV absorption. Nevertheless, relative peak size does not depend on sample concentration and therefore the conclusions drawn are valid.
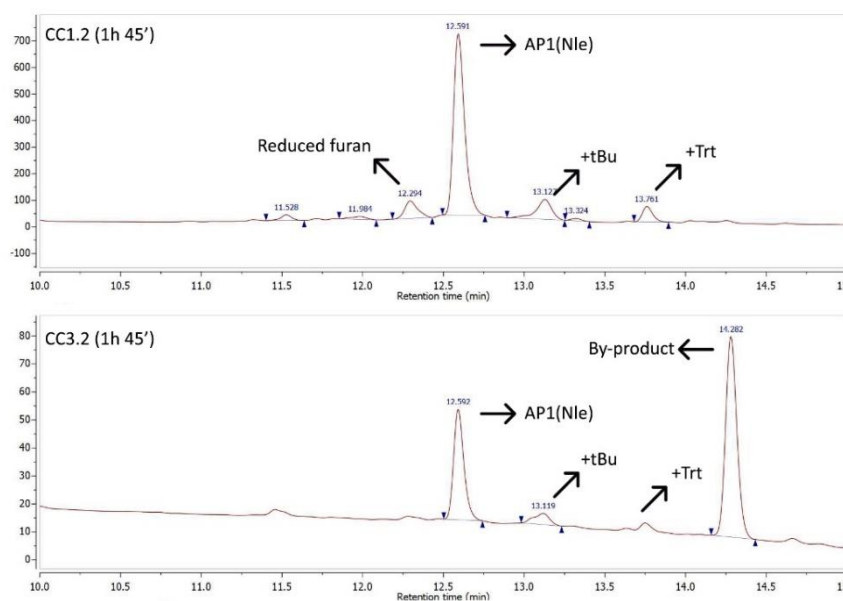

Figure 25: Close-up of second test cleavage HPLC data, identification based on LC-MS data.

In conclusion, CC1, consisting of 95% TFA, 2.5% thioanisole and 2.5% TIS, yielded the best results. Therefore, it was used for all cleavages with exclusion of AP1(Met) and AP2(Met), as these were cleaved before this cleavage cocktail optimization was performed.

4.2.6 Work-up and purification

The last stage in obtaining high purity peptides consists of the work-up and purification process after cleavage. Work-up of all peptides happened in a very traditional manner, namely using methyl-tert-butylether (MTBE) precipitation. The concept of this process is simple, after cleavage, the peptide is present in the cleavage cocktail solution, together with the TFA, scavengers and cleaved protecting groups. To separate this mixture, the TFA is evaporated under a flux of nitrogen, followed by addition of cold MTBE to the crude mixture. The peptide precipitates, while the residual TFA, scavengers and cleaved protecting groups remain dissolved in the MTBE. After centrifugation at cold temperatures, the supernatant is discarded and new MTBE is added for further washing. This is repeated three times.

Final purification is performed utilizing a preparative reversed phase HPLC. In Figure 26, the purity of the crude after MTBE work-up is compared to the purity after RP-HPLC for SAP1(Nle). For all furan-modified peptides, a purity exceeding 95% was achieved.
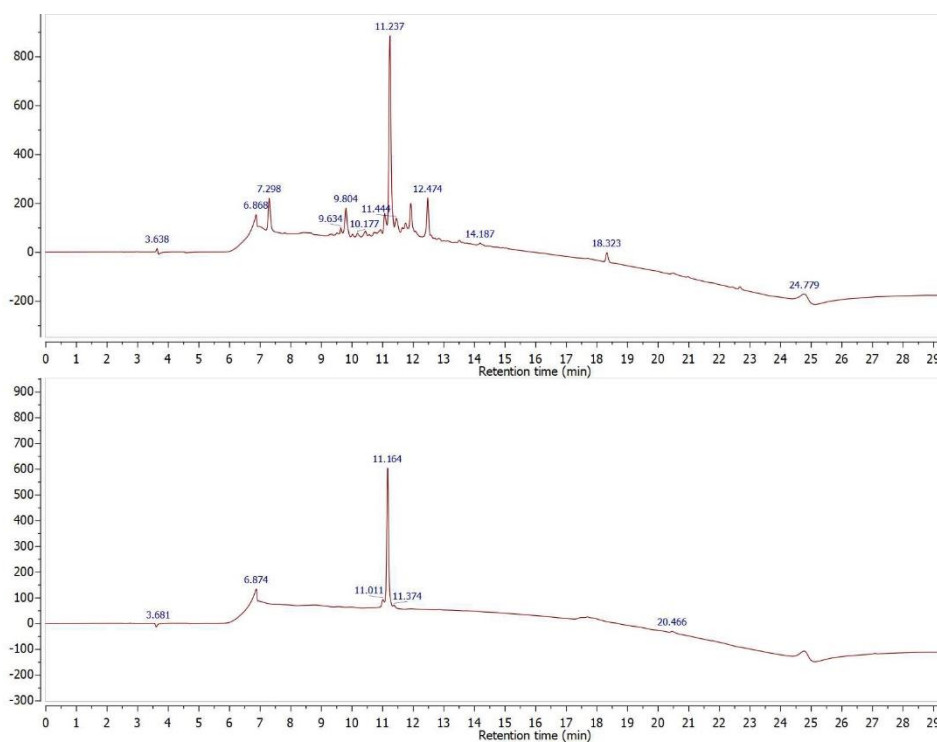

Figure 26: RP-HPLC chromatogram of crude SAP1(Nle) mixture after MTBE work-up (top),
RP-HPLC chromatogram of SAP1(Nle) after purification (bottom).

4.2.7   SPPS applied to furan-modified apelin peptides

As mentioned before, a C-terminal carboxylic acid functionality was desired for the furan-modified apelin peptides. For this reason, the first amino acid was linked to the 2-chlorotrityl chloride resin, an acid-labile resin for solid phase immobilization of carboxylic acids. To do so, a simple SN_1 reaction was performed with 1 equivalent of the desired amino acid using DIPEA as a base, before capping the resin with methanol to block remaining reactive sites on the resin that have not reacted with the first amino acid. After coupling of the first amino acid to the resin through the linker, the remaining part of the apelin peptides was synthesized using an automated peptide synthesizer, performing the deprotection-coupling cycles fully autonomous, using 4 equivalents of amino acid and 4 equivalents of HBTU. After completion of the sequence, the peptides were manually coupled to biotin-PEG4-propionic acid using standard coupling conditions. The only steps remaining at this point are the cleavage (which simultaneously leads to deprotection of the sidechains), followed by work-up and purification. Some apelin peptides were difficult to purify, fractions of the first purification where sufficient
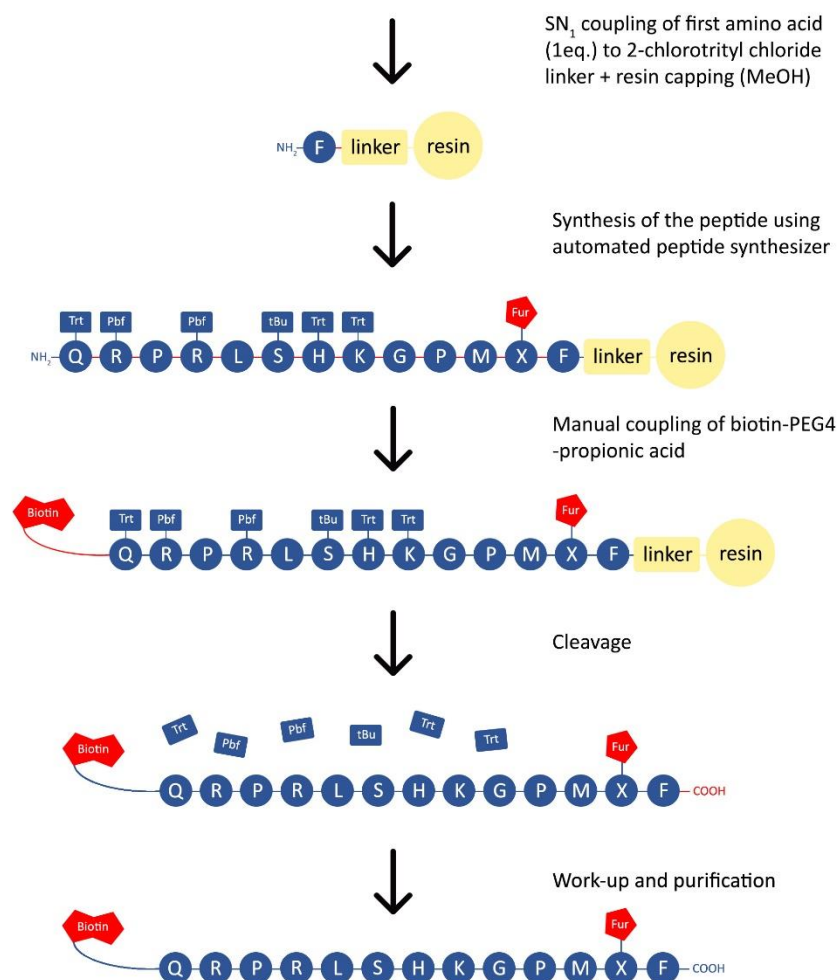

Figure 27: Schematic overview of the AP2(Met) synthesis procedure.

separation was not achieved, were repurified after lyophilization. In Figure 27, a schematic overview of the AP2(Met) synthesis procedure is given as an example.

4.2.8   SPPS applied to furan-modified opioid peptides

For the opioid peptides, a C-terminal amide functionality was desired. In order to obtain this, a polystyrene resin with the Rink amide linker was utilized. This linker contains a Fmoc protected amine functionality which is used for attaching the first amino acid after Fmoc deprotection using piperidine. This first coupling is performed just like all other peptide couplings in the synthesis. As mentioned in the part on sequence design, the first amino acid starting from the C-terminus is the unnatural Glu(biotinyl-PEG). For all furan-modified opioid peptides, the entire peptide was synthesized manually, as expensive unnatural amino acids are incorporated in the sequence. Cleavage, work-up and purification was performed in an identical way to the apelin peptide syntheses. In Figure 28, a schematic overview of the synthesis of OP2 is given as an example.
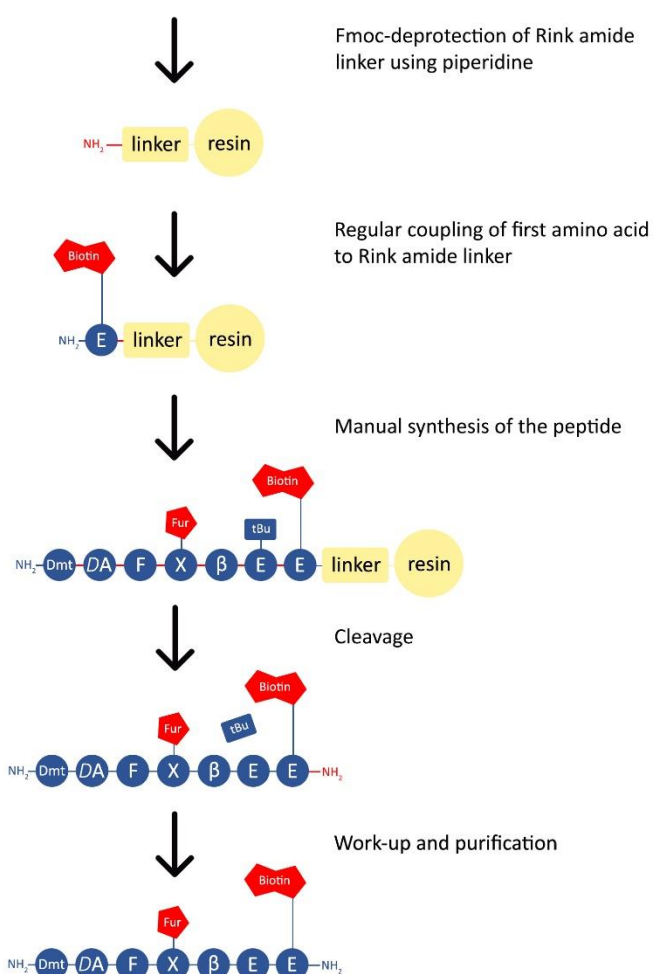

Figure 28: Schematic overview of the OP2 synthesis.

4.3   Results & discussion

Synthesis of all envisaged furan-modified peptides was successful. The commercially available Fmoc protected 3-(2-furyl)alanine was included seamlessly in regular SPPS procedures, both in automated peptide synthesis and in manual synthesis. The only challenging hurdle to overcome was the acid-induced furan degradation, which occurred in the cleavage step. Cleavage cocktail optimization allowed partial mitigation of this issue. Yields of all synthesized peptides can be found in Table 6.

Of the nine apelin peptides, seven were obtained with high purity and moderate yields ranging from 2.9% to 28.1% with an average yield of 14.5%. When the furan is inserted at the first position counting from the C-terminus, the yield is substantially lower (indicated in red in Table 6) Two of the envisaged apelin peptide sequences have not been finished in view of the COVID-19 outbreak.

Of the six opioid peptides, four were obtained with high purity and moderate yields ranging from 13.8% to 45.6% with an average yield of 34.6%. These yields are remarkably higher compared to the apelin peptides. This can be attributed to the smaller number of amino acids in the opioid sequences and thus less reactions required in the synthesis. A second explanation for the higher yield is the fact that less side-reactions occurred with the furan during cleavage, as it is protected by the peptide backbone (with the apelin peptides, the furan was closer to the end of the peptide). Two of the envisaged opioid peptide sequences have not been finished in view of the COVID-19 outbreak

Furan-modified apelin peptides			Furan-modified opioid peptides		
Code	Yield (%)	Yield (mg)	Code	Yield (%)	Yield (mg)
AP1(Met):	2.9	5.9	OP1:	unfinished	
AP2(Met):	14.4	29.9	OP2:	unfinished	
AP1(Nle):	7.6	14.5	OP3:	42.8	42.3
AP2(Nle):	28.1	55.1	OP4:	45.6	37.5
AP3:	15.4	30.1	OP5:	36.3	26.2
SAP1(Nle):	11.8	22.7	OP6:	13.8	10.7
SAP2(Nle):	unfinished				
SAP3(Nle):	unfinished				
Native AP(Nle):	21.1	40.6			

Table 6: Yields of the performed furan-modified apelin and opioid peptide syntheses.

5 Crosslinking experiments

The crosslinking experiments encompass two distinct phases. The first phase covers the treatment of the living cells with the synthesized furan-modified peptides at different concentrations during different incubation periods. During this phase, the endogenously activated crosslinking event takes place. The second phase consists of a Western blot experiment in order to examine if crosslinking of the furan-modified peptide to the GPCRs has occurred. In order to perform the Western blot, the cells need to be lysed, which is achieved by addition of a lysis buffer to the cells. Western blotting covers a series of sample processing steps, starting with gel electrophoresis to separate the different cell lysate components by size. This is followed by a blotting step (transfer of the proteins from the gel to a membrane) and finally the separated compounds are visualized by means of dual immunodetection.

Due to the COVID-19 outbreak, only a single crosslinking experiment using three different apelin peptides and subsequent Western blotting was performed by me. Results of a crosslinking experiment performed by my thesis supervisor, Laia Miret Casals, with 4 different opioid peptides are also included. Many planned experiments were cancelled due to the COVID-19 lockdown measures.
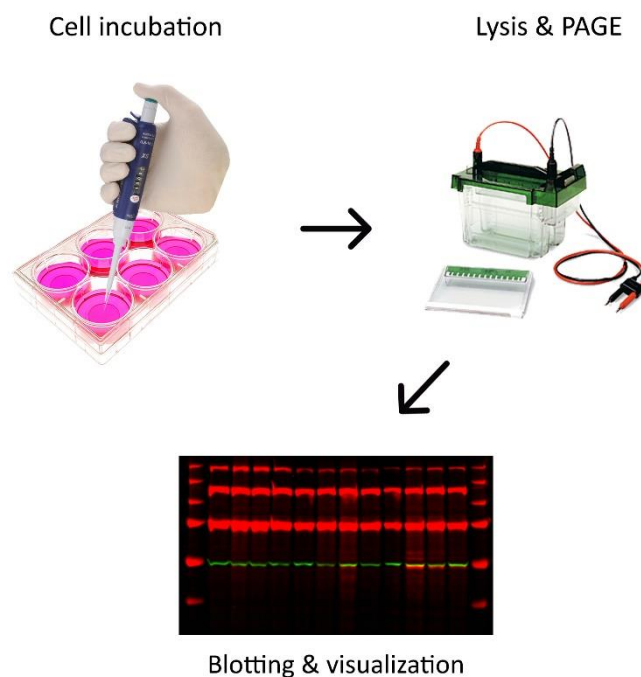

Figure 29: Workflow of a typical crosslinking experiment.

5.1   Crosslinking in living human cells

The human cells used to perform the crosslinking experiments originate from the MDA-MB-231 cell line, a late-stage breast cancer cell line. The MDA-MB-231 cells were used in view of the fact that they were available in the cell lab and it is known that this cell line expresses the apelin receptor. Detailed protocols for optimized crosslinking experiments through the endogenously oxidized furan methodology had already been established within the group[37] and were adopted for this thesis. The cells were seeded on six-well plates (825.000 cells per well) before treatment with the furan-modified apelin peptides the next day, when the cells were about 80% confluent. The cells were treated with three different furan-modified apelin peptides (AP1(Nle), AP2(Nle) and AP3) at different concentrations (25, 50 or 100µM) during different periods of time (30 minutes, 1 hour, 2 hours).

The furan-modified peptides were dissolved after synthesis and purification in DMSO at 25mM as a stock solution for storage. To obtain cell medium with different concentrations of furan-modified peptide, this stock solution got dissolved to the desired concentration using Dulbecco's Modified Eagle Medium (DMEM) without fetal bovine serum (FBS), Penicillin or Streptomycin (antibiotics used to prevent bacterial contamination during cell culturing). It was observed during experiments prior to this thesis that the cell medium used for a successful crosslink experiment should not contain FBS: when the medium contains FBS, crosslinking was only observed when higher concentrations of the peptide were added to the cells. As every peptide solution contains DMSO from the stock solution, it is important to be sure that the amount of DMSO present is not toxic for the cells. For this reason, a control sample was included in the crosslinking experiment: cells were treated with cell medium containing 100µM DMSO (without peptide) at the same conditions. It is known that DMSO is commonly used as a solvent for pharmacological substances, as well as a control group for testing products in cultured cells

The aim of the first experiment was to optimize the crosslinking conditions (peptide concentration and incubation time) and to evaluate if notable differences could be observed between the three added apelin peptides, containing the furan moiety at the first, second or third position from the C-terminus respectively.

5.2   Cell lysis and SDS-PAGE

After incubation of the cells with the furan-modified peptides for a desired period of time, the cells need to be lysed to obtain a lysate solution that contains all proteins, including the crosslinked ligand-receptor complex of interest. Upon addition of the lysis buffer, which contains a protease-inhibitor, the cell lysate is meticulously kept on ice to further reduce protease activity. Preventing proteases from digesting proteins in the cell lysate solution plays an essential role in the process, as digested proteins can't be detected through Western blotting and too much digestion would result in an unsuccessful experiment.

The cell lysate solution at this point contains cell debris consisting of cell membrane and coagulated cell organelles, which would interfere with the Western blot experiment. After careful homogenization of the cell lysate solution by forcing it through a thin needle for ten times and subsequent centrifugation, the obtained supernatant no longer contains any cell debris and is a clear solution. To this solution, Laemmli sample buffer 5X is added which denatures the proteins and gives them a uniform negative charge distribution. Bromophenol blue present in the Laemmli sample buffer serves as a tracking dye. Before proceeding to the next step, the samples were heat-treated at 95 degrees Celsius for 5 minutes.

Separation of the proteins in the lysate solution is achieved by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE), a well-established technique based upon separation of components by size. As the proteins are denatured and uniformly negatively charged by addition of the sample buffer, separation by size can in practice be considered as a separation by molecular weight of the proteins. Proteins with higher molecular weight take up more volume after denaturation and migrate at a slower rate through the gel. Additionally, a protein marker is added to one of the electrophoresis gel lanes. This marker is a solution containing a set of stained proteins of known molecular weight, which serves to estimate the molecular weight of unknown proteins in other lanes. For the performed experiment, cell lysates corresponding to a single apelin peptide were run together on a gel. This way, three different SDS-PAGE separations were performed, one for every apelin peptide (AP1(Nle), AP2(Nle) and AP3).

5.3 Blotting and visualization

After the SDS-PAGE, the polyacrylamide gel contains the separated proteins. In order to allow for an easier identification of the proteins currently contained in the gel, the proteins are transferred to a membrane, a practice which is called 'blotting'. This is essential as identification of proteins relies on interactions with antibodies, for which the proteins need to be easily accessible. The membrane is placed on the gel and the proteins migrate from the gel to the membrane surface under the influence of an electric potential. The positions of the proteins inside the gel (based upon molecular weight), are retained on the membrane: the separation achieved by gel electrophoresis is preserved.

Once transferred to the membrane, the aim is to identify proteins of interest based on specific antibody interactions. As the membrane used in the blotting step is especially efficient at binding proteins, unspecific binding of the antibody to the membrane would result in very high background signal. To prevent this, the membrane is typically saturated in a so-called blocking step, by addition of a protein-rich solution to the membrane. This way, all possible unspecific binding sites are blocked by proteins that will not interfere in any further steps in the visualization process.

For the actual visualization, fluorescent antibodies are utilized. A dual immunodetection approach is used, where two different antibodies having fluorescent properties at different wavelengths are added to the membrane. For the performed experiment with the apelin peptides, the aim was to visualize the apelin receptor using a primary apelin receptor antibody, amplified by a secondary antibody which features green fluorescence. On top of that, biotin was visualized using red fluorescent streptavidin (Figure 30). This dual immunodetection should in theory allow for easy identification of the crosslinked ligand-protein
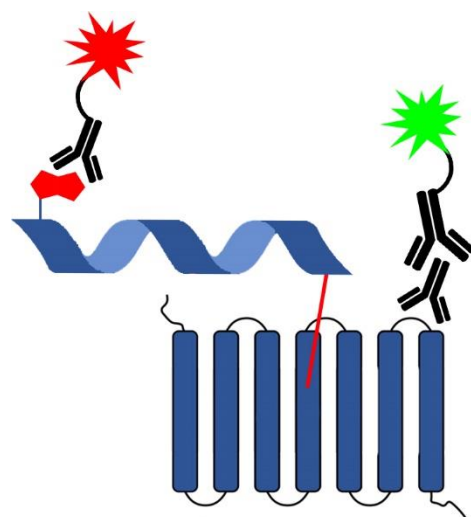

Figure 30: Principle of dual immunodetection using red fluorescent Streptavidin (left) and green fluorescent secondary antibody (recognizing the primary antibody, right)

complex. The apelin receptor itself should be clearly visible as a green fluorescent signal. If crosslinking occurred, a red signal should be observable in close proximity on the membrane (or a resulting yellow band if both signals overlap), as the covalently bound apelin peptide

ligand was synthesized to contain a biotin moiety (which interacts with the red fluorescent streptavidin).

Incubation of the membrane with the apelin antibody was performed overnight, before washing the membrane with phosphate buffered saline (PBS) containing 0.1% TWEEN the next day (3x). Incubation with the green fluorescent secondary antibody and the red fluorescent streptavidin was performed over a period of 45 minutes, before washing the membranes with PBS containing 1% SDS (3x), regular PBS (2x) and a final washing step with water. As the membranes now contain fluorescent moieties bound to the proteins of interest, the Western blot procedure is finished by scanning the membranes.

5.4 Results & conclusion

5.4.1 Apelin receptor system

From the crosslinking experiment performed with three apelin peptides, promising results were obtained. The results are based upon the scan of the three membranes obtained after blotting and antibody incubation. Each membrane contains experiments with only one single type of apelin peptide. The three membrane scans are displayed in Figure 31.

When analyzing the scanned membranes, the first thing to notice when starting from the left side, are the red signals belonging to the protein marker which was added in the very first lane of every gel during SDS-PAGE. The marker was also added to a second lane in the gel, with varying position for every tested apelin peptide, which aided in easy identification of the gels after many experimental manipulations. As mentioned before, these markers are a guide to estimate the molecular weight of visualized proteins in other lanes.

Figure 31: Western Blot for three different administered apelin peptides at different concentrations (100, 50, or 25 µM) and different incubation times (2 hours, 1 hour or 30 minutes). Observed crosslinking signal is indicated with a yellow ellipsoid.

Apart from the red signals that can be attributed to the protein marker, every lane contains at least three intense red signals (±250 kDa, ±130 kDa and ±75 kDa). These red signals can be attributed to the fluorescent streptavidin that was used to incubate the membranes and aims at visualizing the biotin attached to the crosslinked ligand-receptor complex. Streptavidin specifically binds to biotin, which implies that multiple natural proteins containing a biotin moiety are present in the cell lysate mixture. Human cells are known to contain multiple endogenously biotinylated proteins, which explains these red signals: coA carboxylase at 220 kDa, pyruvate carboxylase at 130 kDa, 3-methylcrotonyl coA carboxylase at 75 kDa and
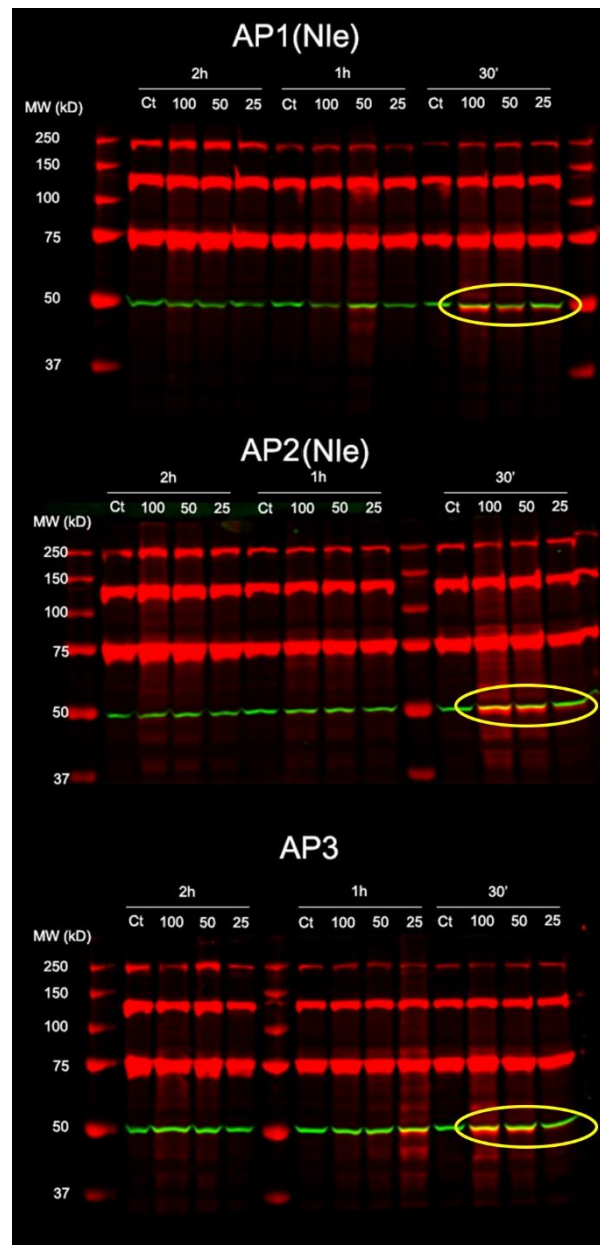
propionyl coA carboxylase at 72 kDa[70]. These last two give rise to a single signal due to overlap, as the resolution of the performed SDS-PAGE is limited. These signals originating from the endogenously biotinylated proteins mentioned above confirm that the fluorescent streptavidin incubation works and the signals from these proteins can serve as a loading control.

Around 50 kDa, one or two different signals can be observed: a green signal and a red signal (a focused extract from the Western blot of AP1(Nle) around 50 kDa can be found in Figure 32). As the apelin receptor has a mass of 43 kDa and the furan-modified apelin peptides have a mass around 2 kDa, the crosslinked apelin ligand-receptor complex should have a mass of 45 kDa. This means that the signals around 50 kDa in the Western blot carry the most valuable information about the crosslinking experiment. The green signal originates from the fluorescent secondary antibody bound to the apelin receptor antibody, making its interpretation very straightforward: a green signal represents the presence of the apelin receptor. However, when the cells were incubated with the furan-modified apelin peptides (AP1(Nle), AP2(Nle) or AP3) for 30 minutes, a red signal can be observed very close to the green apelin antibody signal. This red signal, just like the red signals mentioned in the previous paragraph, originates from the fluorescent streptavidin bound to a biotin moiety.
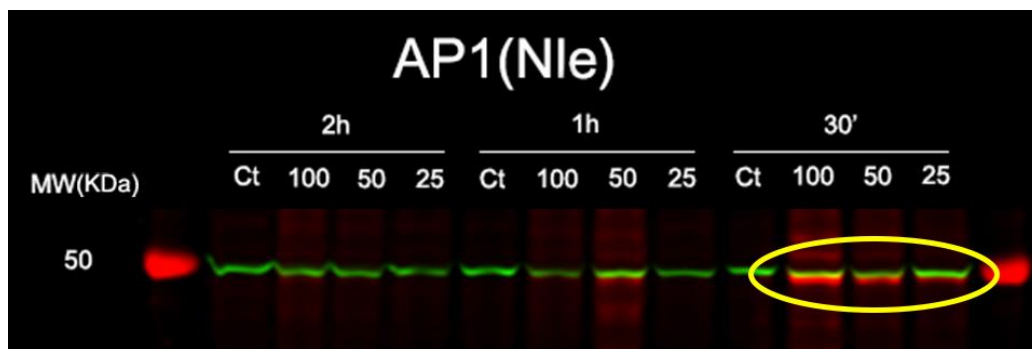

Figure 32: Extract from Western Blot (Figure 30) focused at a molecular weight of ±50 kDa for AP1(Nle).

As the apelin receptor is not reported to be biotinylated, these red signals should originate from the crosslinked ligand-receptor complex, as a biotin moiety was attached to the furan-modified apelin peptides for easy detection in Western blot experiments. No red signal around 50 kDa is observed in the control experiment, where no furan-modified peptide was added to the cells. This also strengthens the conclusion of the red signal around 50 kDa representing the apelin receptor crosslinked to the synthesized furan-modified apelin peptide.

The crosslinked product also does not appear after incubation with the peptide over a period of two hours. When AP1(Nle) is added to the cells for 1 hour at 50 µM, a crosslink signal is observable, but not at 100 or 25 µM. These results raise some questions about the dose- and time- dependent activity of the furan-modified apelin ligands, about the turnover of the apelin receptor and how the internalization of the apelin receptor is affected by the interaction with the furan-modified apelin ligands. A possible explanation is that the apelin receptor has a high rate of turnover after crosslinking and the receptor gets internalized over time, a common phenomenon observed with GPCRs and described for the apelin receptor[71]. The crosslinked complex is likely to be degraded by proteases upon internalization, preventing its detection by the utilized Western blotting process. This is a possible explanation for the presence of crosslinked product after 30 minutes of incubation, but not after one or two hours.

5.4.2   µ-opioid receptor system

In addition to the single crosslinking experiment performed with the three apelin peptides, my thesis supervisor performed a crosslinking experiment with 4 opioid peptides (OP3, OP4, OP5 and OP6). For this experiment, living SH-SY5Y cells were used, a human cell line originating from a bone marrow biopsy from a young female with neuroblastoma, a nerve tissue cancer. Contrary to the MDA-MB-231 cells utilized for the apelin peptide crosslinking experiments, the µ-opioid receptor is expressed sufficiently in the SH-SY5Y cells[72]. The crosslinking conditions used for this experiment were almost identical compared to the conditions described above for the apelin peptide experiment. However, only incubation during 1 hour or 30 minutes was performed. All opioid peptides were added to the cells at 50 µM concentration, only OP3 was added at two different concentrations of 50 µM and 100 µM. The results can be found in Figure 33.

Interpretation of the results is very similar to the Western blot for the apelin peptides. In the scan, signals for naturally occurring endogenously biotinylated proteins are present in red (pyruvate carboxylase at 130 kDa, 3-methylcrotonyl coA carboxylase at 75 kDa and propionyl coA carboxylase at 72 kDa[70]). In contrast to the scan from the crosslinking experiment with the apelin peptides, there are multiple green signals present due to background issues. This means that the antibody used to detect the μ-opioid receptor is not fully specific and also recognizes other proteins or receptor multimerization takes place and the antibody is able to recognize these higher order structures. It is known that the μ-opioid receptor has a molecular weight of 50 kDa, so crosslinking signals should be sought for in that region, where a green signal is observable.
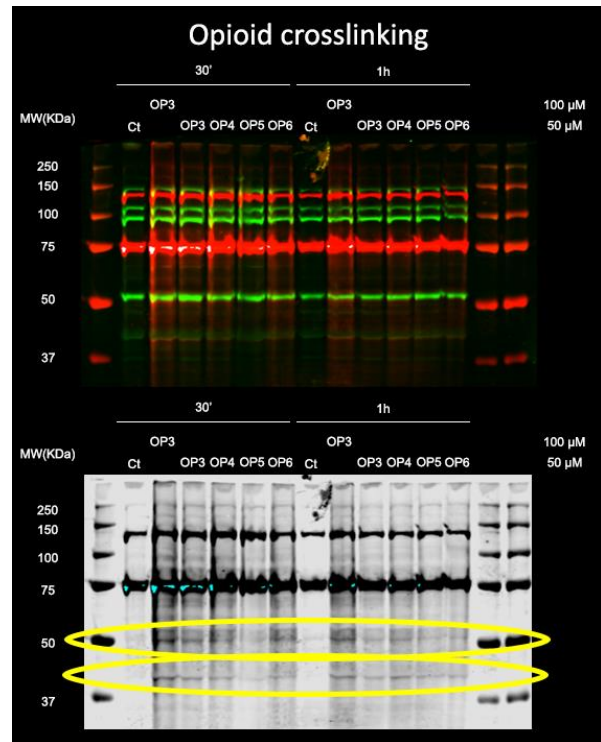

Figure 33: Western Blot for four different administered opioid peptides at two different concentrations (100 or 50 μM) and two different incubation times (1 hour or 30 minutes). Bottom blot is an inverted grey-scale version of the blot that only contains the red-colored signals. Observed crosslinking signal is indicated with a yellow ellipsoid.

The red signals belonging to the crosslinked complex (opioid peptide - μ-opioid receptor) coming from the fluorescent streptavidin interacting with the biotin incorporated in the opioid peptide, are less pronounced compared to the scan of the apelin peptide experiment. Yet, when solely the red signals are scanned for and the scan is inversed and grey-scaled, the signals can be observed for all furan-modified opioid peptides, except for OP5. This visual enhancement of the scan is found in the lower half of Figure 33 and a focused extract can be found in Figure 34. The crosslinking signals are more pronounced after thirty minutes of incubation, similar to the apelin peptide experiment. As expected, no crosslink signal is observable in the control samples, in which no peptide was added to the cells. The signal observed around 37 kDa (with lower intensity) can be attributed to a crosslinked receptor-ligand complex with a different glycosylation state. It is known that GPCRs can occur in different glycosylated forms: the upper band (around 50 kDa) is most probably the matured

glycosylated form of the receptor, which is present at the cell-surface. The band around 37 kDa can be the deglycosylated crosslinked complex after internalization. Further experiments are required to confirm this hypothesis.
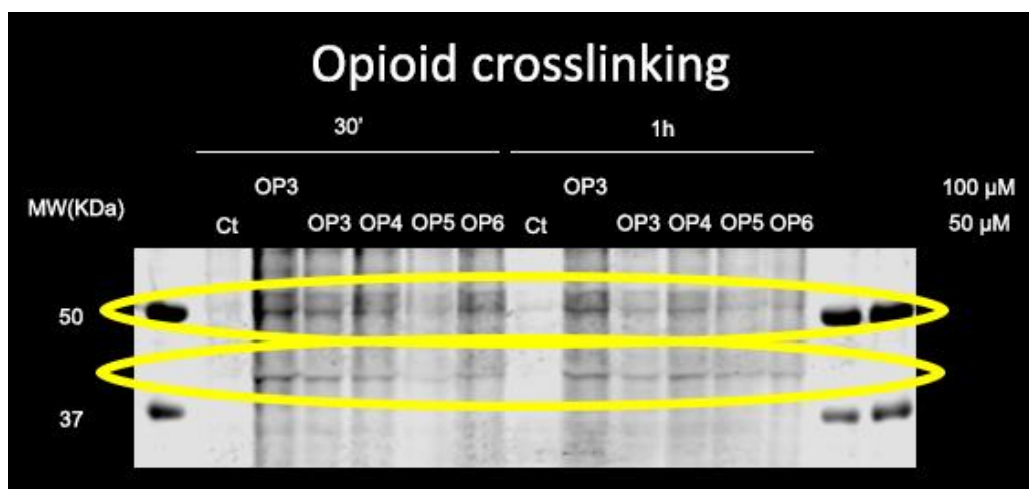

Figure 34: Extract from Western Blot (Figure 32) focused at a molecular weight of ±37-50 kDa for opioid peptides.

5.4.3   Conclusion

As a conclusion, the spontaneous oxidation of the furan moiety and subsequent crosslink reaction between the furan-modified ligand and the G protein-coupled receptor in living cells appears to work in both the apelin receptor and the µ-opioid receptor model systems. Using the described methods, it is possible to crosslink the synthesized furan-modified peptide ligand to the respective receptor selectively and subsequently visualize the crosslinked product via Western blotting. The lack of crosslinked product signal in the Western blot when the cells were incubated with the synthesized peptides for one hour or longer can be attributed to internalization of the receptor over time after the occurrence of the crosslinking event. When an incubation time of 30 minutes is utilized, the crosslinking signal is most pronounced. Further investigation of influencing factors still needs to be performed in order to fully understand and further develop the furan crosslinking technology on living cells.

6 Visualization and simulation of ligand-receptor complexes

As a supplementary assignment to complement the practical work that has been performed and to compensate the lost lab work due to COVID-19 restrictions, an *in-silico* visualization and simulation of the apelin ligand-receptor complex was performed. The visualization and simulation is based on the apelin ligand-receptor crystal structure obtained by Ma et al.[43].

For the second model system studied in this thesis, the μ-opioid system, no crystal structure of the receptor interacting with a peptide ligand is currently available, obstructing any *in-silico* analysis. Starting from a crystal structure of the μ-opioid receptor interacting with a small-molecule ligand, as reported by Huang et al.[51] or Manglik et al.[60], and modifying the ligand to a peptide *in-silico,* would lead to a very unreliable structure. Alternatively, a docking approach could be considered, by fitting the peptide ligand into the crystal structure of the active receptor after removing the small-molecule ligand, yet this requires expertise and specific software beyond the scope of this thesis.

6.1 Visualization of apelin receptor with AP1(Nle), AP2(Nle) and AP3

The main goal of this visualization is to understand which amino acid residues inside the receptor are likely to form a covalent crosslink bond with the furan moiety introduced in the apelin ligand peptide. Three nucleophilic amino acids were considered as possible nucleophiles for the oxidized furan moiety: lysine, cysteine and histidine. In addition, two amino acids that can potentially engage in electrophilic aromatic substitution will be taken into account: tyrosine and tryptophan. The apelin 17 peptide utilized by Ma et al. is first modified *in-silico* to natural apelin 13 followed by modification to obtain AP1(Nle), AP2(Nle) or AP3. For the visualizations, modifications and measurements, MAESTRO software developed by Schrödinger is used.

Firstly, the heavily modified apelin 17 ligand in the crystal structure was modified to the natural apelin 13 ligand using the MAESTRO software. This is achieved by deleting the four N-terminal amino acids and undoing the modifications introduced by Ma et al. in the remaining part (Figure 35). To examine which amino acids in the receptor are likely to form a covalent bond with the oxidized furan, a selection of any of the five amino acid residues mentioned above within a 10 Å radius measured from the three C-terminal α-carbons of the apelin 13 peptide ligand, was made. In this range of 10 Å from the three C-terminal α-carbons, 12

residues are observed that could potentially engage in a covalent bond formation with the oxidized furan. These residues are located on the N-terminal part, on the first or second extracellular loop or on the first, second, sixth or seventh transmembrane helix of the apelin receptor (Figure 36). It is important to note that the lysine at position 8 in the apelin ligand itself could also be included in this selection, formation of a covalent bond would in this case result in cyclization of the oxidized furan-modified apelin ligand. However, as the side-chain of this lysine is unfortunately not present in the crystal structure reported by Ma et al. for unknown reasons, this residue cannot be included in further measurements.
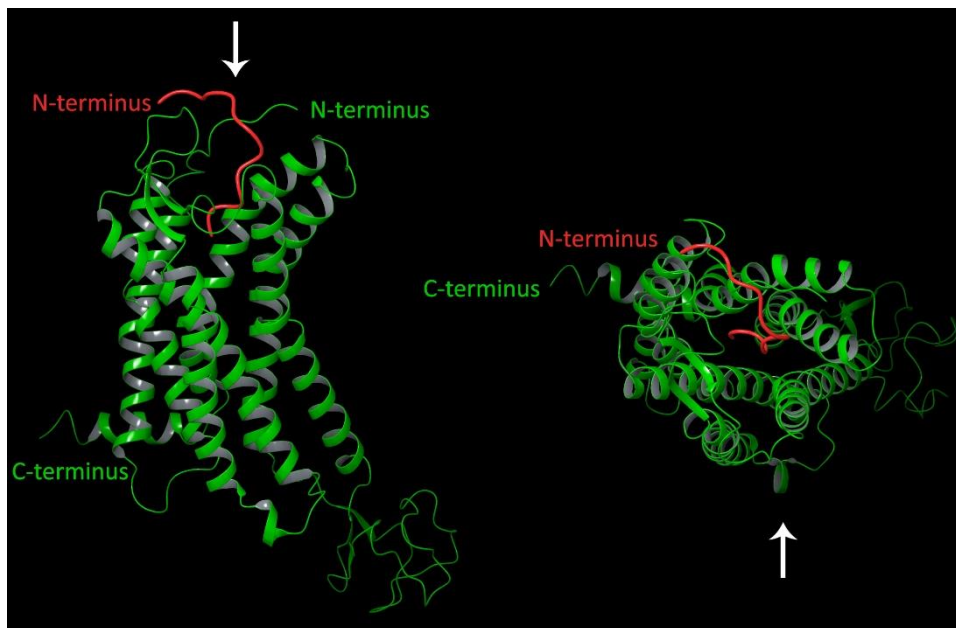

Figure 35: Apelin 13 (red) interacting with the apelin receptor (green). Front view (left) and top view (right), white arrow on the left indicates viewing angle for the right and vice versa.
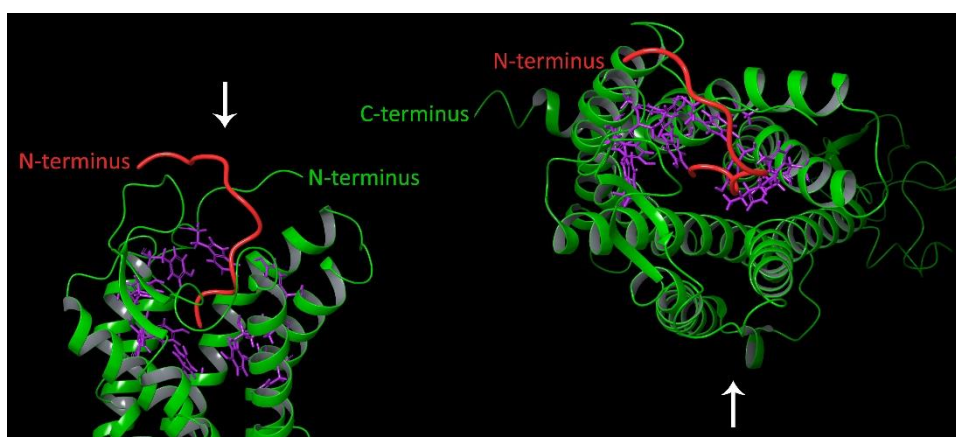

Figure 36: Apelin 13 (red) interacting with the apelin receptor (green). Receptor residues of interest are colored in purple. Front zoomed view (left) and top zoomed view (right), white arrow on the left indicates viewing angle for the right and vice versa.

After making the selection of residues that could potentially form a covalent bond with the oxidized furan, the natural apelin 13 peptide ligand was modified to AP1(Nle), AP2(Nle) or AP3 using the MAESTRO software. Using these 3 ligands, distances were measured from the relevant carbon in the introduced furan (at position 1, 2 or 3 from the C-terminus of the ligand) to the relevant atom in one of the selected residues that could potentially engage in covalent bond formation. In case this distance is 10 Å or lower, the dihedral angles in both residues were optimized manually (free rotation over σ-bonds) to minimize the distance between the two atoms. Results of these measurements can be found in Table 7.

Residue	AP1(Nle): Distance to furan	AP1(Nle): Minimized distance	AP2(Nle): Distance to furan	AP2(Nle): Minimized distance	AP3: Distance to furan	AP3: Minimized distance
W24	10.48		4.85	1.92	6.81	6.78
Y35	5.35	4.33	9.73	6.79	14	
W85	4.55	3.21	11.15		15.58	
Y88	5.09	2.34	7.37	<1.5	11.98	
Y93	4.55	2.77	4.34	<1.5	9.22	5.41
C181	DSB - C102	DSB - C102	DSB - C102	DSB - C102	DSB - C102	DSB - C102
Y264	7.87	<1.5	8.24	6.04	8.68	2.88
H265	13.84		13.87		12.68	
K268	9.38	<1.5	8.72	5.43	8.38	<1.5
Y271	13.13		8.37	6.09	3.91	<1.5
C294	12.2		12.15		8.39	4.48
Y299	4.28	<1.5	9.67	6.09	13.11	

Table 7: Results of the distance measurements (Å) performed in MAESTRO, 'DSB' stands for disulfide bridge.

When interpreting these results, it is of high importance to remember that the distance measurements and dihedral angle optimizations were performed on the non-oxidized furan. When the furan is oxidized, it becomes longer and more flexible, allowing residues further away to also potentially form a covalent bond, as shown in Figure 37.
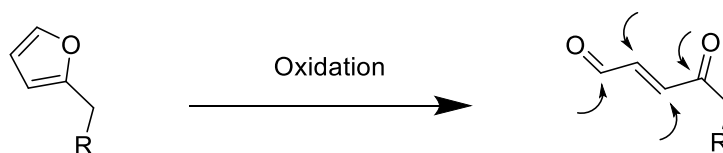

Figure 37: Furan oxidation to keto-enal, arrows indicate possible nucleophilic attack sites.

Making predictions about the residue that will actually engage in the crosslink bond formation is very difficult, as this depends on many factors. Factors like relative reactivity of the selected residues towards the keto-enal (= oxidized furan moiety), flexibility of the residue and steric interference of surrounding residues all play an important role. Nevertheless, when taking a close look at the results in Table 7, lysine 268 seems to be a very likely candidate, as it is in close proximity to the furan in all three furan-modified apelin peptides and is an extremely flexible residue. In addition, tyrosine 88, 93, 264 and 299 can also get very close to one or more furans of the furan-modified apelin 13 peptides.

As a conclusion to this visualization study, Figure 38 shows the 5 residues of interest that can come closest to the furan in AP1(Nle), AP2(Nle) or AP3 by free rotation over σ-bonds.
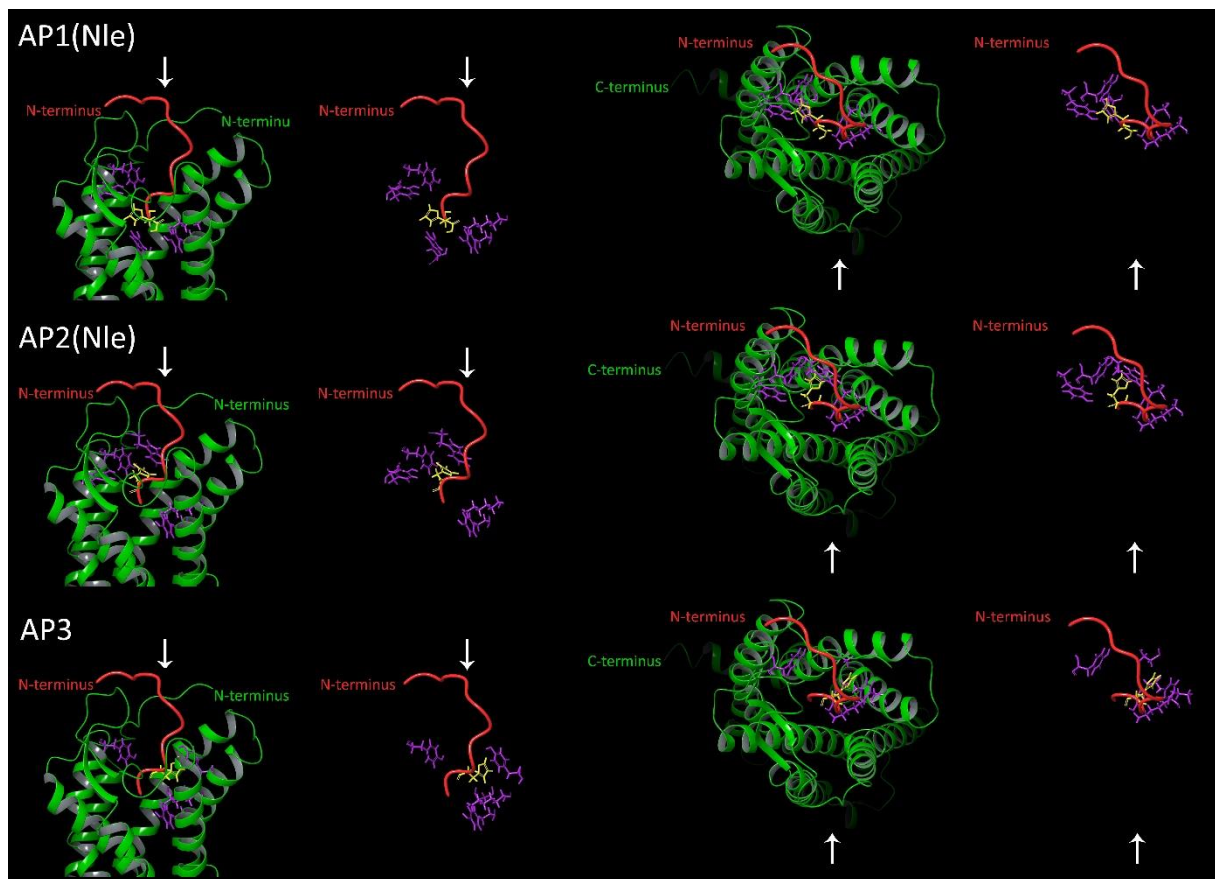

Figure 38: AP1(Nle), AP2(Nle) or AP3 in red interacting with the apelin receptor in green, the introduced furan is marked in yellow. The 5 most relevant receptor residues are colored in purple. Front zoomed view (left) and top zoomed view (right), white arrow on white arrow on the left indicates viewing angle for the right and vice versa.

6.2 Molecular simulation of apelin 13 interacting with apelin receptor

In context of the course 'Molecular Simulations of Biosystems', taught by Professor Toon Verstraelen, a molecular dynamics simulation was performed using openMM in Python, based on the published work of Ma et al.[43]. This simulation allows the apelin receptor and the apelin 13 ligand to move freely over a very short period of time. In order to keep this thesis accessible to read, details will be described in the experimental part only, a brief summary of molecular dynamics in general, the applied strategy and the acquired results will be given on the following pages.

Molecular dynamics simulations are based on the movement of atoms in a molecule under the influence of a forcefield and find their use in a.o. highly accurate *in-silico* docking applications. For docking simulations, a receptor and ligand are free to move around in hope of a binding event taking place at a certain point during the simulation. As the starting position of the ligand is outside of the receptor, it often takes long for the ligand to recognize and bind to the receptor. To perform simulations like these, supercomputers are needed to allow for long simulation, as this increases the chances of a binding event. The simulation performed for the course and this thesis was based on a different approach, it started from a situation where the ligand is already bound to the receptor, based on a crystal structure. This allows for shorter simulation times, requiring less computational power. However, the amount of extracted information is obviously a lot lower.

For the performed simulation, the AMBER forcefield was used. This forcefield is incompatible with unnatural amino acids, so the crystal structure had to be modified back to the natural apelin 13 ligand and the natural apelin receptor, as Ma et al. introduced a number of modifications in order to crystallize the ligand-receptor complex. This was achieved using an online tool called 'CHARMM-GUI-solution builder'. The simulation was performed with explicit water solvent molecules at 300 K, 50.000 simulations steps of 2 femtoseconds were run, resulting in a total simulation time of 100 picoseconds. One simulation step consists of letting the atoms evolve under the influence of the forcefield for a given period of time (2 femtoseconds in this case), followed by reporting all new coordinates of all atoms. These reported coordinates are then used to visualize the evolution of the system over time in an animated way. In addition, a number of analytic procedures can be carried out using the reported coordinates. Since a printed version of an animated visual format is quite irrelevant,

the focus in the following part will be on an analysis of the coordinates relevant in the context of this thesis.

As observed in the visualization part of this supplementary assignment (*vide supra*), lysine 268 is an interesting residue in the receptor in the context of possible crosslink bond formation. After running the simulation, the reported coordinates were used to calculate the distance between every α-carbon of the natural apelin 13 ligand and the α-carbon of this particular lysine 268 residue after every simulation step. The results are displayed in Figure 39.

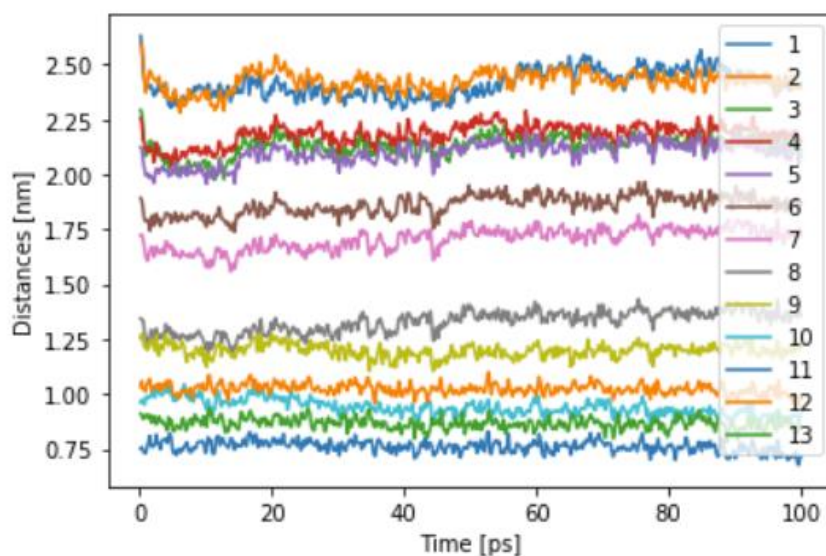*Figure 39: Distance between the different α-carbons of natural apelin 13 (counting from the N-terminus) and the α-carbon of lysine 268.*

As the simulation was run on my own personal computer, a long simulation was not possible: 100 picoseconds of simulation time took about 2 hours. However, from the 100 picoseconds of simulation, some valuable insights can already be extracted. Firstly, as was already deduced from the visualization, the lysine 268 residue is in close proximity to the C-terminal residues of the apelin 13 peptide, possibly allowing crosslink formation if a furan is inserted at any of these residues and oxidized to the reactive keto-enal. In addition, the simulation proves that the C-terminal part of the peptide is bound more firmly in place compared to the N-terminal part (distances between N-terminal α-carbons and the lysine 268 α-carbon fluctuate a lot more). However, to allow more information to be deduced from simulations like this one, a much longer simulation time would be required.

7 COVID-19 restrictions and consequences

Due to the COVID-19 outbreak, all practical work from March 13th onward was cancelled due to safety regulations aimed at preventing further spreading of the virus. For this particular thesis, this translates to a loss of about a third of all planned practical working hours in the lab, as the majority of the lab work was scheduled in the second semester. As the peptide synthesis chronologically needs to happen before any crosslinking experiments in cells are possible, the majority of the lost lab time was originally foreseen for performing the crosslinking experiments, as illustrated in Figure 40. This is the reason why only a single crosslinking experiment (preceded by extensive training) was performed and a lot of synthesized peptides have not been further investigated in this context yet. Results from a second crosslink experiment with four different opioids (OP3-6) are included in this thesis as well, but this experiment was performed by my thesis supervisor (Laia Miret Casals).

The exact content of the experimental time that was lost will be covered in section 9.1 of this thesis. To compensate for the lost lab time, a complementary assignment was performed focused on the visualization and simulation of the binding pockets for the apelin model systems, which was covered in section 6.
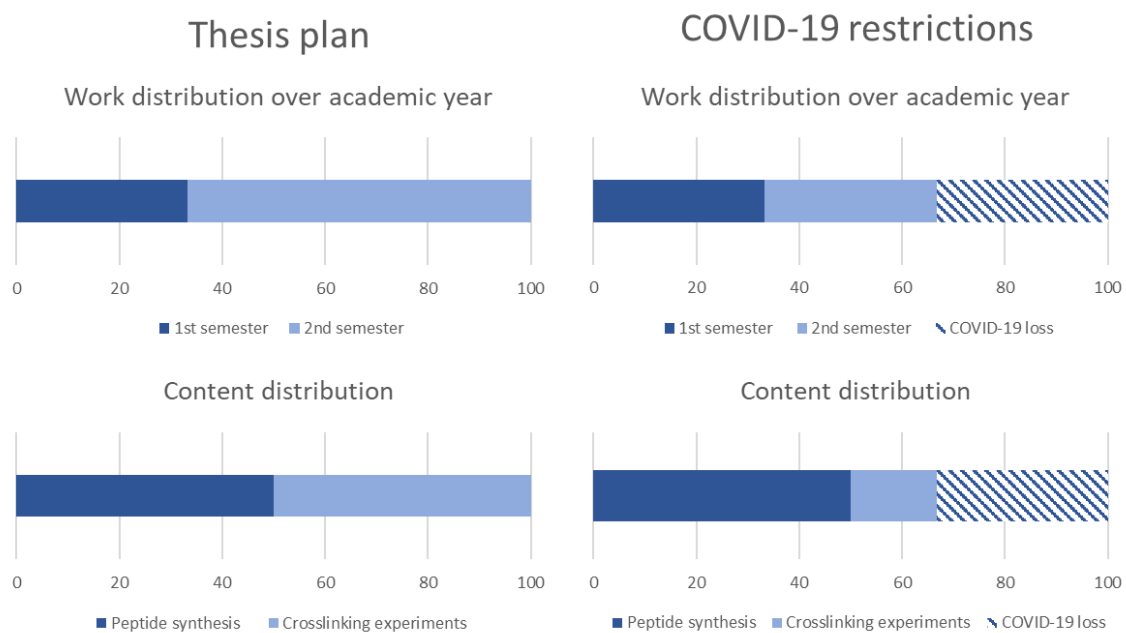

Figure 40: Schematic representation of COVID-19 work losses.

8 Conclusions

This thesis consisted of two major parts, supplemented by an additional assignment. The first part, synthesis of furan-modified ligand peptides for two model systems by inclusion of the commercially available Fmoc protected 3-(2-furyl)alanine into regular SPPS, proved successful. The second part of the thesis, crosslinking of the synthesized peptides to GPCRs in two model systems (apelin receptor and μ-opioid receptor), was achieved as well. It was proven that the furan-based crosslinking works in two new ligand-GPCR systems and the spontaneous oxidation of the furan moiety is steered by endogenous cellular activity. Nevertheless, many more experiments were planned to optimize the crosslinking conditions and to obtain a better understanding of the furan crosslink technology in living cells. The loss of experimental work due to the COVID-19 lockdown was compensated by performing a supplementary assignment on *in-silico* visualization and simulation of the apelin model system using MAESTRO software and molecular simulation tools.

The synthesis of the furan-modified peptides, both apelin peptides and opioid peptides, has been performed successfully. In total, 11 different furan-modified peptides were synthesized and obtained with a high degree of purity (7 apelin peptides, 4 opioid peptides). In addition, optimization of the TFA cleavage from the resin and deprotection of the side chains was performed to achieve higher yields in the synthesis of furan-modified peptides. This cleavage optimization will be useful in the future when more furan-modified peptides will be synthesized. Four peptides (2 apelin peptides and 2 opioid peptides) were not finished in view of the COVID-19 lockdown.

Crosslinking of the synthesized furan-modified peptide ligands to the corresponding GPCRs triggered after endogenous furan oxidation was observed in experiments on living human breast-cancer cells (MDA-MB-231) for the apelin receptor and on neuroblast cells (SH-SY5Y) for the opioid receptor. The cells were treated with the furan-modified peptides for 30 minutes at 37 °C in medium containing no FBS, the crosslinked ligand-receptor complex was detected by Western blot.

In addition to experiments performed in the lab, a visualization study and molecular dynamics simulation of the apelin 13 – apelin receptor system was performed. This highlighted the potential of lysine 268 for the formation of the crosslink bond. Nevertheless, more residues

capable of forming a covalent crosslink bond with the oxidized furan moiety, are in close proximity.

In conclusion, the applicability of the spontaneous oxidation and subsequent covalent crosslinking of a furan-containing peptide to its receptor when added to a cell culture, has been expanded successfully to two new ligand-GPCR model systems. This novel technique was first described for the Kisspeptin-10/KISS receptor model system[37] by the OBCR group, yielding promising results as a proof of concept. This thesis succeeded in demonstrating the reproducibility and robustness of the novel furan-based crosslinking strategy.
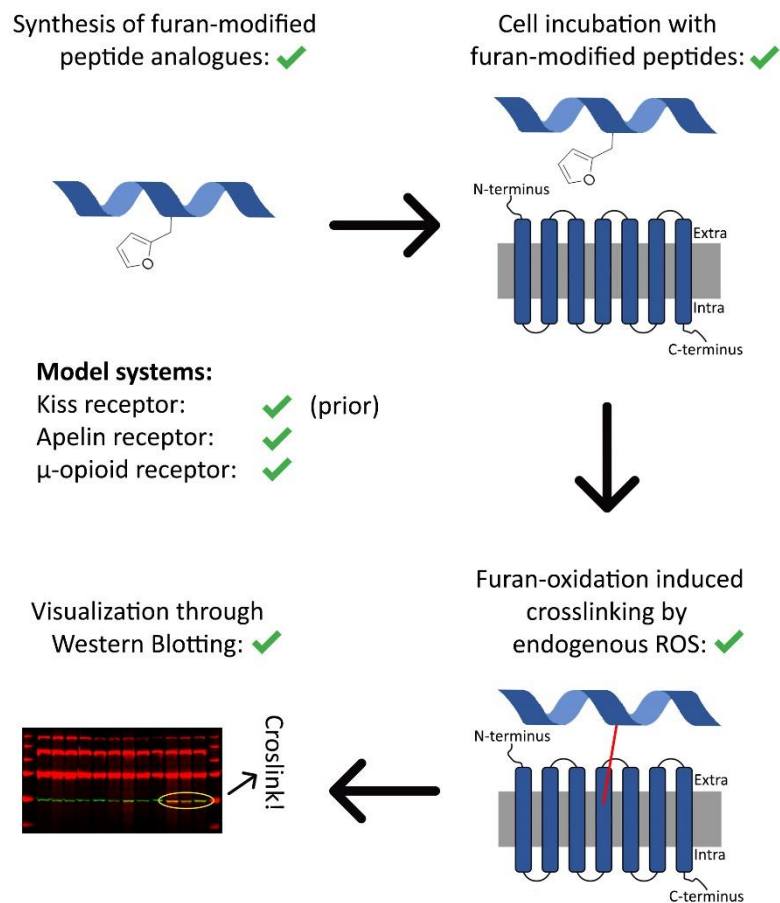

Figure 41: Schematic representation of achieved objectives

9 Future perspectives

9.1 Cancelled experiments due to COVID-19

Firstly, the crosslink experiments should be repeated twice more in both model systems to prove reproducibility of the results over a total of three experiments. If all three experiments yield similar results, the results can be considered accurate and reliable. Next, optimization of the crosslinking conditions should be achieved. It can be observed in the performed experiments that incubation time plays an important role in crosslink visualization. As GPCRs have a high turnover rate, shorter durations of furan-modified ligand treatment (less than 30 minutes) could improve the crosslinked complex signal. In addition, incubation at lowered temperatures (4 °C) could be explored in order to slow down the cellular metabolism and therefore potentially mitigate the issue of receptor internalization. Internalization is hypothesized to be the cause of the absence of crosslink signal after incubation for one or two hours in the experiment with the apelin peptides. Different cell lines could also be tested, as receptor expression levels will most definitely have an influence on the occurence of crosslinking. For the opioid peptides, expression levels of the μ-opioid receptor in the MDA-MB-231 cell line are known to be low, which motivated the choice of working with SH-SY5Y cells[72] for the opioid peptide experiment performed by my thesis supervisor.

In the performed experiments, the control sample consisted of cells that were incubated with cell medium containing only 100 μM DMSO. For future experiments with the apelin receptor, two more types of control samples could be included: a natural apelin 13 peptide (without furan-moiety) and three different scrambled apelin sequences containing the furan moiety at position 1, 2 and 3 from the C-terminus. When treating the cells with the natural apelin 13, the crosslinked complex should not be observed due to the lack of furan moiety. When treating the cells with the scrambled apelin sequences containing the furan moiety, the peptides should not be able to recognize the binding site of the apelin receptor, therefore the formation of a crosslinked complex is also not expected. The natural and scrambled apelin sequences have already been fully or partially synthesized and also contain the biotin-PEG4 for visualization purposes. These experiments would further prove that the specific three-dimensional structure of the amino acids in the apelin sequence are a requirement for binding and subsequent crosslinking.

Apart from different control samples, NOX-enzyme inhibitors such as diphenyleneiodonium, plumbagin and 4-(2-aminoethyl)benzenesulfonyl fluoride (AEBSF) could be added to the cells prior to furan-modified peptide treatment to prove the responsibility of these enzymes in producing the reactive oxygen species causing the endogenous furan oxidation triggering the crosslink. In this case a reduced crosslink signal is expected due to less furan oxidation leading to fewer crosslinked complexes to be formed.

Furthermore, a competitive assay could be very informative to confirm the specificity and efficiency of the furan crosslinking (Figure 42). When cells are simultaneously incubated with a combination of a fixed concentration of the furan-modified apelin peptides and an increasing concentration of natural apelin 13 (without furan moiety), a reduced crosslinked product signal is expected when increasing the concentration of the natural apelin 13. This can

Figure 42: Principle of competitive assays.

be attributed to competition for the binding site of the receptor. This reduction of the crosslinked product signal could possibly be less pronounced if the furan crosslink technology is specific and very fast.
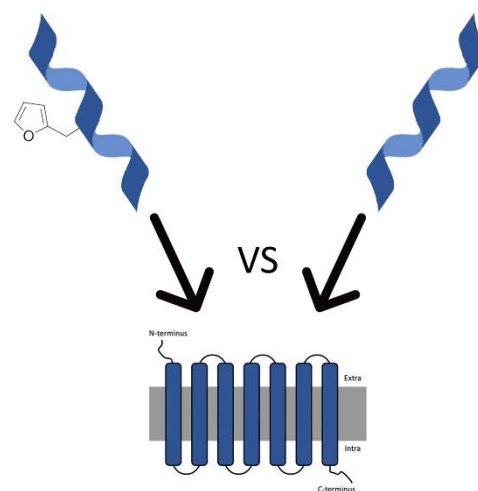
9.2   Perspectives beyond this thesis

This thesis was aimed at proving the robustness and general applicability of a novel furan crosslinking technique that offers significant benefits over other alternatives. Nevertheless, novel techniques are only viable if they allow for new applications. This novel crosslinking technique could allow for novel applications of peptide-protein crosslinking to identify and validate new drug targets and provide structural information on the ligand-receptor interface. In addition to structural biology and proteomics applications, the furan crosslink technology can have therapeutic applications. Furan modified peptides can be used as a covalent agonist and lead to sustained receptor activation. Obviously, studies on toxicity and stability in plasma of furan-modified peptides should be performed before any further steps can be taken. On paper, this first hurdle should be easy to overcome, as peptide drugs typically possess low toxicity and the furan moiety is non-reactive until it gets oxidized *in situ*. The second hurdle,

preventing the generally rapid degradation of peptides inside the human body, seems a little more troublesome. Nevertheless, extensive research is being performed on mitigating this problem for peptide drugs in general, as mentioned in the introductory part of this thesis on peptide drugs (section 1.1).

Apart from the potential of furan-modified peptides to play an agonist role by covalently binding to a receptor, other potential strategies could be applied. One example could be to use the furan-modified peptide as a carrier for a particular bioactive molecule in order to reach the lysosomes, which are typically difficult to target. By attaching the bioactive molecule to the furan-modified peptide ligand, the crosslinking process followed by internalization of the crosslinked complex could allow for the bioactive molecule to reach the lysosomes. This concept is based on internalization of the crosslinked ligand-receptor complex and subsequent endocytic trafficking to lysosomes[73]. Attaching the bioactive moiety in this case could be achieved in a similar way as the biotin moiety was attached to the furan-modified peptides in this thesis.

A second concept based on the internalization of the crosslinked ligand-receptor complex is utilizing furan-modified peptides to trigger internalization of oncogenic receptors. Internalization of these oncogenic receptors, which are overexpressed and/or overactive compared to their non-cancer counterparts, could suppress the development of a tumor. Apart from suppressing tumor development by internalization of the oncogenic receptors, a cell death-inducing compound could be attached to the furan-modified peptide ligand. Upon internalization of the crosslinked ligand-receptor complex, the cell death-inducing compound that was co-internalized could provoke cell apoptosis, resulting in a more enhanced tumor suppression strategy.

Many more valuable concepts could be invented upon utilization of both creativity and knowledge of cellular biology.

10 Procedures and experimental data

10.1 Synthesis of furan-modified peptides

10.1.1 General and reagents

All synthesis work was performed in the S4 building on Campus Sterre of Ghent University in the labs of Professor Annemieke Madder, guided by my thesis supervisor Laia Miret Casals. Organic solvents used throughout this thesis include DMF (Biosolve), DCM (Merckx), ACN (Fischer scientific) and MeOH (Merck) were used without extra drying or purification procedures. Milli-Q grade water was obtained from a Sartorius Arium 611 Dl water purification system present in the lab. HBTU, TFA and Fmoc-protected amino acids used for the peptide synthesis were obtained from Iris Biotech GmbH. Amino acids with sensitive side-chains were purchased side-chain protected: Fmoc-Lys(Boc)-OH, Fmoc-His(Trt)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH. Piperidine and DIPEA were purchased from Sigma-Aldrich. The used resins are 2-chlorotrityl chloride PS (1% DVB) and Rink amide PS (1% DVB). TNBS (5 w/v% in methanol) was obtained from Thermo Scientific.

10.1.2 Procedures and instrumentation

For the synthesis of all peptides, SPPS was used. All apelin peptides were synthesized using the fully automated SYRO Multiple Peptide Synthesizer robot, equipped with a shaker (vortex) for the reactor block accommodating a maximum of 24 reactors (polypropylene syringe tubes with a polyethylene filter at the bottom). Usually, synthesis with double coupling steps was performed as following: a mixture of 5 equivalents (eq.) amino acid in DMF (0.5 M), 5 eq. HBTU in DMF (0.5 M) and 10 eq. DIPEA in NMP (2 M) is added to the resin, with subsequent reaction for 40 minutes at room temperature. The Fmoc-3-(2-furyl)-L-alanine was incorporated as a standard amino acid. All opioid peptides were synthesized manually, employing the same filter tubes.

10.1.2.1 Apelin and μ-opioid peptides synthesis

For the apelin peptides, 2-chlorotrityl chloride polystyrene resin (1% DVB) was used. For every synthesis, 100 mg (AP1(Met), AP2(Met)) or 120 mg (all other apelin peptides) was loaded with 1 mmol/g of the first amino acid of the sequence, resulting in a synthesis scale of 0.1 or 0.12 mmol. Loading was performed through an SN_1 reaction. First, the resin gets swollen for 30 minutes in DCM and 30 minutes in DMF. Next, the resin gets washed 3 times with DMF and 3

times with DCM (1 wash consists of 2 Pasteur pipettes of solvent added to the resin and subsequent removal by vacuum). 1 mmol/g of the desired amino acid was weighed and dissolved in the minimal required amount of DCM to obtain a clear solution. 10 equivalents of DIPEA was added to the solution. The solution was added to the resin and shaken for 2 hours on an automated shaker. After the 2 hours of shaking, 10 equivalents of MeOH were added to cap the resin, the mixture was shaken for another 15 minutes. The resin got washed 3 times with DCM and 3 times with DMF. At this point, the resin is ready to start the automated peptide synthesis in the SYRO Multiple Peptide Synthesizer robot. The robot starts with deprotecting the Fmoc group and finishes after coupling of the last amino acid (no Fmoc deprotection of the last amino acid). Finally, the deprotection of the Fmoc group and the coupling of the biotin-PEG4-propionic acid is performed manually due to the high cost of the reagent. Fmoc deprotection was achieved by adding 4 mL of 40% piperidine in DMF for three times with 1, 2 and 2 minutes of stirring. Next, the resin was washed with DCM (3x), DMF (3x) and DCM again (3x). After this, a solution of 1.5 eq. of biotin-PEG4-propionic acid, 1.5 eq. of HBTU and 1.5 eq. of DIPEA in 1 mL of DMF was added to the resin. The mixture was shaken using an automated shaker for at least 2 hours (or overnight in some cases). After the coupling, the resin was washed with DCM (3x) and DMF (3x). To assess coupling completion, a TNBS test was performed.

For the opioid peptides, Rink amide polystyrene resin (Rink Amide-ChemMatrix) with a loadding of 0.69 mmol/g was used. For every opioid peptide 100 mg of resin was used, resulting in a 0.069 mmol synthesis scale. The resins were swollen for 30 minutes with DCM followed by 30 minutes of swelling in DMF. After this the resins got washed with DCM (3x) and DMF (3x). As the linker is Fmoc-protected at this point, it needs to get deprotected first. This was achieved by adding 5mL of 40% piperidine in DMF for three times with 1, 2 and 2 minutes of stirring. After this, the resin was washed with DCM (3x), DMF (3x) and DCM again (3x). 1.5 eq. of the Fmoc-Glu(biotinyl-PEG)-OH was dissolved in DMF, together with 1.5 eq. of HBTU and 1.5 eq. of DIPEA. Only 1.5 eq. were used due to the high cost of the Fmoc-Glu(biotinyl-PEG)-OH. This solution was added to the deprotected resin and shaken for 1 hour. A TNBS test was used to check completion of the resin loading (TNBS test is described in section 10.1.2.2). If the loading was not complete, the resin was capped by addition of 3 equivalents of acetic anhydride together with 3 equivalents DIPEA dissolved in 2 mL DMF. This was shaken for 15

minutes before doing a new TNBS test. The next step was to deprotect the Fmoc group using 40% piperidine in DMF, as was previously described. The rest of the amino acids were also coupled manually using 3 eq. of Fmoc-protected amino acids, 3 eq. of HBTU and 3 eq. of DIPEA. To get a clear solution when dissolving the amino acids in DMF, an ultrasound bath was used. After addition of the dissolved components to the resin, the mixture was shaken on an automated shaker for 30 minutes. After this, the resin was washed with DCM (3x) and DMF (3x). After every coupling, a TNBS test was performed to assess completion of the coupling reaction. The last amino acid, the Boc-2,6-dimethyl-L-tyrosine (Boc-Dmt-OH), was provided by Professor Steven Ballet's group from the University of Brussels and coupling of the Boc-Dmt-OH was performed using a carbodiimide, not uranium of phosphonium salts. Boc-Dmt-OH was coupled using DIC (1.5 eq.) as a coupling reagent and Oxyma (1.5 eq.) as an additive. This was added and shaken during 1 hour. The resin was washed with DCM (3x) and DMF (3x) after. Finally, a TNBS-test was performed to assess completion of the coupling reaction. The N-terminal Boc group was cleaved simultaneously with the protecting groups of the side chain and the peptide from the resin. These syntheses were performed by Laia Miret Casals, Nathalie Verbeke and myself.

10.1.2.2 TNBS test

After every coupling reaction, completion of the coupling was performed using the TNBS test. A very small amount of beads were transferred to a tiny glass tube. To this, 3 drops of 10% DIPEA in DMF was added, followed by 3 drops of 5% TNBS in MeOH. If free primary amines are present, they will turn red due to their reaction with TNBS. The glass tube was held in bright light in front of a white surface to examine. If a red color was observed on the beads, the previously performed coupling reaction was repeated, followed by a new TNBS test.

10.1.2.3 Fmoc-deprotection

Fmoc deprotections were performed automatically by the SYRO Multiple Peptide Synthesizer robot during synthesis of the apelin peptides. For all other Fmoc-deprotections the following procedure was employed: 5mL of 40% piperidine in DMF was added to the resin for three times with 1, 2 and 2 minutes of stirring using a metal spatula respectively. After this, the resin got washed with DCM (3X), DMF (3x) and DCM again (3x).

10.1.2.4 Alloc deprotection

For the synthesis of AP1(Nle), AP2(Nle), AP3, SAP1(Nle), SAP2(Nle), SAP3(Nle) and Native AP(Nle), Alloc protected lysine was used due to availability issues. The inclusion of this amino acid in the peptide requires an additional deprotection step before cleavage, as the Alloc protection group is not cleaved by the total cleavage conditions. To deprotect Alloc-protected lysine, 1.5 mL solution containing 14 mg of $Pd(PPh_3)_4$ and 150 µL of phenylsilane in DCM was added to the resins and shaken for 15 minutes, followed by washing with DCM (3x) and DMF (3x). This was repeated for 3 times. In order to remove final traces of the Pd-containing compound, the resins were washed 3 times with a solution of 0.02 M $NaS_2CN(C_2H_3)_2$ in DMF. The dark color in the resin disappears after this final washing.

10.1.2.5 Small-scale cleavage of peptide from resin

After completion of an automated synthesis sequence for the apelin peptides, test cleavages were performed to check if the SYRO Multiple Peptide Synthesizer robot accomplished the synthesis without errors. For this, the resin first got Fmoc-deprotected. After this, a very small amount of resin was added to a 1 mL Eppendorf tube. To this, 1 mL of a cleavage cocktail (95% TFA, 2.5% thioanisole, 2.5% TIS) was added. The mixture was shaken using an automatic shaker for 1 hour. After this, the TFA was removed using an N_2 gas flush (a Pasteur pipette attached to the gas source blowing N_2 gas in the Eppendorf). This was followed by MTBE work-up (section 10.1.2.8) on a miniature scale (no transfer to falcon tube, entire MTBE work-up is performed in the Eppendorf). The resulting 1:1 H_2O/ACN solution containing the peptide of interest was then checked by MALDI-TOF (section 10.1.2.9) for the desired mass

10.1.2.6 Total cleavage

For all total cleavages, both for the apelin and the opioid peptides, the same conditions were used. The optimized cleavage cocktail (section 10.1.2.7) contains 95% TFA, 2.5% TIS and 2.5% thioanisole. For every cleavage at least 60 mL of this cocktail was used. The experimental setup consisted of a round bottom flask with a stir bar, a glass piece that was attached to a vacuum pump and the syringe tube containing the peptide-carrying resin on top. First, 10 mL of cocktail is added to the flask. Then the flask is covered with the glass piece connected to the vacuum pump with the syringe tube on top. The vacuum is turned on and 5 mL of cocktail is run over the resin into the round bottom flask. Next, 5 mL of cocktail is added to the resin with the vacuum pump turned off, the resin is stirred with a metal spatula for 1 minute before

turning the vacuum on to pull the 5 mL into the flask. This is repeated twice more. After this, 5 mL of cocktail is added to the resin for three more times followed by 1 minute stirring, yet waiting 15 minutes, 15 minutes and 1.5 hours before pulling the cocktail in the round bottom flask by turning the vacuum pump on. If during the waiting time the liquid level in the syringe tube lowers, additional cleavage cocktail was added. In the end, all of the cleavage cocktail containing the cleaved peptide is inside the round-bottom flask.

10.1.2.7 Cleavage cocktail optimization

A cleavage cocktail optimization was performed to lower furan-degradation during the cleavage step. For the test cleavages, resin from the SYRO Multiple Peptide Synthesizer robot containing AP1(Nle) peptide was used (the peptide was still contained the Alloc protecting group on the lysine). 4 different cocktails were tested and compared to a classic cleavage cocktail (Table 8). The applied protocol is similar to the test cleavage protocol described in section 10.1.2.5, yet for the first run, the Eppendorfs with 1 mL cleavage cocktail were shaken for 45 minutes before MTBE work-up (section 10.1.2.8). A second round of tests was required to make a final decision on what cleavage cocktail was better, only CC1 and CC3 were examined in the second round of tests. For this test, resin containing AP1(Nle) that had an Alloc-protected lysine and was coupled to the biotin-PEG4-propionic acid was utilized. Once again a similar protocol to the test cleavage described in section 10.1.2.5 was used. This time the Eppendorfs were shaken for 1 hour and 45 minutes. All solutions containing the peptides of interest obtained after the MTBE work-up were examined using RP-HPLC and LC-MS (section 10.1.2.11). An important note has to be made: the classic cocktail was tested before performing the cleavage cocktail optimization, this cocktail was shaken for 1 hour instead of 40 minutes, which might have an influence on the results.

Cleavage cocktail	Concentrations
Classic	95% TFA, 2.5% TIS, 2.5% H_2O
CC1	95% TFA, 2.5% Thioanisole, 2.5% TIS
CC2	92.5% TFA, 2.5% Thioanisole, 2.5% TIS, 2.5% H_2O
CC3	95% TFA, 2.5% thioanisole, 2.5% m-cresol
CC4	97.5% TFA, 2.5% m-cresol

Table 8: Tested cleavage cocktails.

10.1.2.8 MTBE work-up

After cleavage of the peptides is performed, a work-up is done prior to further purification. First, the TFA gets evaporated using an N_2-gas flush. When an oily residue is obtained, it is transferred to a falcon tube (50 mL). This falcon tube is filled with cold MTBE (from the freezer), a white precipitation gets formed. The falcon tube is vortexed, followed by centrifugation at 7500 RPM for 10 minutes at 4 degrees Celsius. The supernatant is carefully discarded. This is repeated 3 times. Before dissolving the white precipitation pellet (= peptide of interest) in a mixture of 1:1 H_2O/ACN, the pellet is left at room temperature for the residual MTBE to evaporate. This solution is ready for further purification and verification.

10.1.2.9 MALDI-TOF

MALDI-TOF experiments were performed on a Sciex/Applied Biosystems 4800plus MALDI TOF/TOF analyzer equipped with a Nd-YAG solid state laser (355nm) and a pulse frequency of 200 Hz. Spotting on the carrier plate was performed by adding 0.5 µL of matrix solution (2,5-Dihydroxybenzoic acid (DHB) or α-Cyano-4-hydroxycinnamic acid (α-HCCA)) to the plate using a micropipette. 0.5 µL of the peptide containing solution was added on top, before mixing by pipetting the solution up and down. This 1 µL drop was left to crystallize before analysis.

10.1.2.10 Preparative RP-HPLC purification

Purification of the peptides after MTBE work-up was performed using a preparative RP-HPLC machine equipped with a UV-detector measuring absorbance at 214 nm (EOS, maximum loading of 200 mg, loop size 20 mL). Gradients for the purifications were composed of MQ water containing 0.1% TFA (A) and ACN containing 0.1% TFA (B). The first run of every day starts with 12 minutes washing of the column at 100% B, followed by 12 minutes equilibration at 100% A. All gradients are depicted in Table 9. Continuing the purification until 100% B was unnecessary and would have caused a big amount of wasted organic solvents and wasted time. All relevant compounds eluted from the column long before reaching 100% B. When multiple peaks were observed in the chromatogram, MALDI-TOF (section 10.1.2.9) was used to identify the peak containing the desired peptide. After the purification, fractions of the same compound were combined and freeze-dried on a lyophilizator. Purifications of the opioid peptides were performed by Laia Miret Casals, Nathalie Verbeke and myself.

	AP1(Met)	AP2(Met)	AP1(Nle)	AP2(Nle)		AP3	SAP1(Nle)	Native AP(Nle)	OP1-6
Gradient:	1	2	3	3		3	3	3	4

Gradient 1:	0% to 5% ACN in 1 minute, 5% to 15% ACN in 5 minutes, 15% to 50% ACN in 40 minutes
Gradient 2:	0% to 15% ACN in 3 minutes, 15% to 25% ACN in 5 minutes, 25% to 50% ACN in 30 minutes
Gradient 3:	0% to 70% ACN in 35 minutes
Gradient 4:	0% to 10% ACN in 2 minutes, 10% to 20% ACN in 5 minutes, 20% to 60% ACN in 40 minutes

Table 9: Gradients used for the preparative RP-HPLC purifications.

10.1.2.11 Freeze drying

The purified fractions were lyophilized using a Heto Drywinner freeze dryer in combination with a Thermoelectron corporation Savat SPD111V Speedvac concentrator or in a RVC 2-18 CDplus (Christ) in combination with an Alpha 2-4 LDplus (Christ) lyophilisator.

10.1.2.12 RP-HPLC and LC-MS Verification

Analysis of synthesized peptides was performed by RP-HPLC and LC-MS. The used RP-HPLC system is an Agilent 1100 series equipped with a Phenomenex Luna 5u C18 column (250 mm x 4.6 mm, 5 μm). The solvent for the analysis was composed of a gradient of MQ H_2O containing 0.1% TFA (A) and ACN containing 0.1% TFA (B), going from 100% A to 100% B in 20 minutes. UV absorbance was measured at 214, 254 and 280 nm. In section 10.1.3, RP-HPLC data is included. Only the chromatograms at 214 nm are included, as the amide bonds inside the peptides absorb UV-light between 190 and 230 nm.

The LC-MS system used in this thesis was not operated by me, samples were submitted for analysis. The system used is a reversed phase Agilent 1100 with diode array detector (set to 214, 254, 280, 310, 360 nm), equipped with a Phenomenex Kinetex C18 100Å (150 mm x 4.6 mm, 5 μm, 35 °C) coupled to an Agilent ESI-single quadrupole MS detector type VL. Mass detection was performed in positive mode. The system is flushed for 5 minutes with MQ water containing 0.1% TFA, followed by a gradient from 0% to 100% ACN containing 0.1% TFA in 6 minutes, ending with a 2 minute flush at 100% ACN containing 0.1% TFA. (Flow of 1.5 mL/min). In section 10.1.3, LC-MS data is included. Only the chromatograms at 214 nm is included with inclusion of the mass spectrum corresponding to the most noticeable peaks.

10.1.3 Experimental data

For all synthesized and subsequently purified peptide compounds, RP-HPLC data combined with ESI-MS data is included in this section. The displayed ESI-MS spectrum (bottom) corresponds to the compound of interest peak of the HPLC chromatogram at 214 nm (top). The same counts for the data from the cleavage cocktail optimization, yet for these, the identified by-products are indicated on the RP-HPLC spectrum. Data was processed using MestreNova 12. No data is available on peptides that haven't been finished yet (SAP2(Nle), SAP3(Nle), OP1 and OP2).
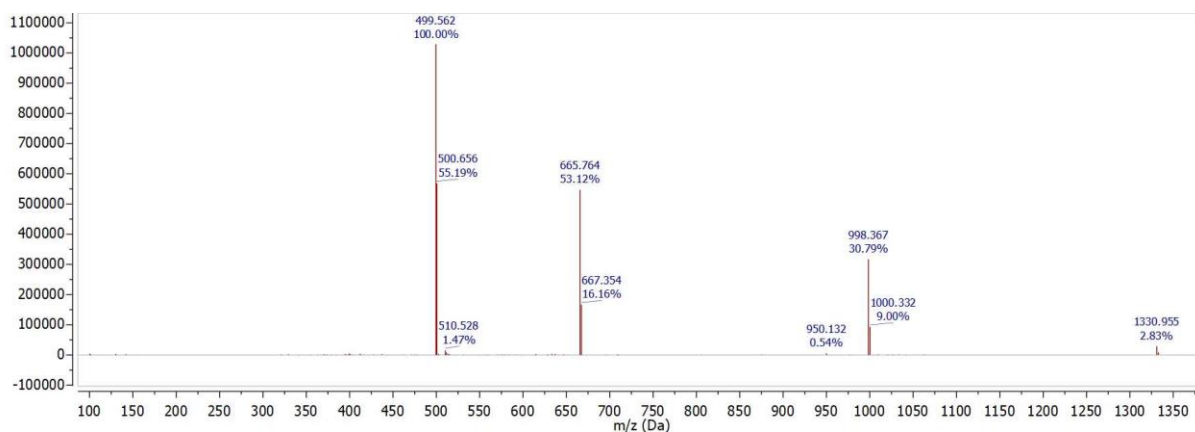
10.1.3.1 AP1(Met)

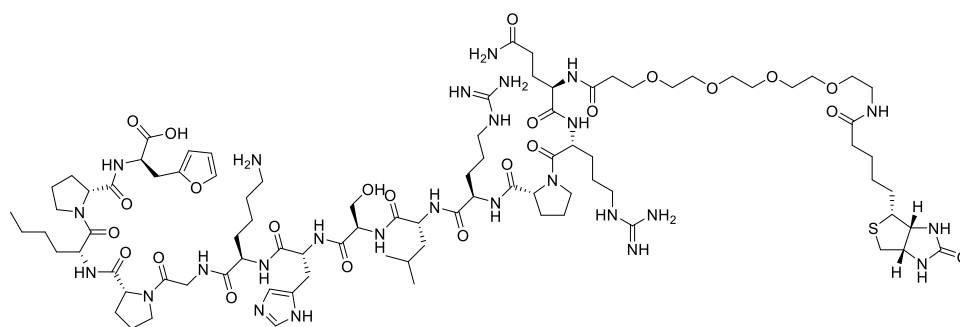Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProMetProFua-OH
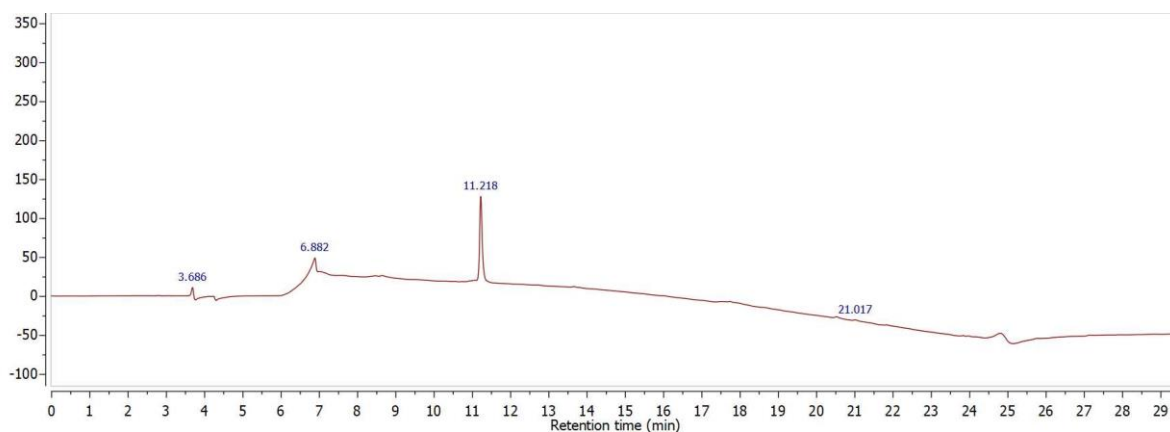
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
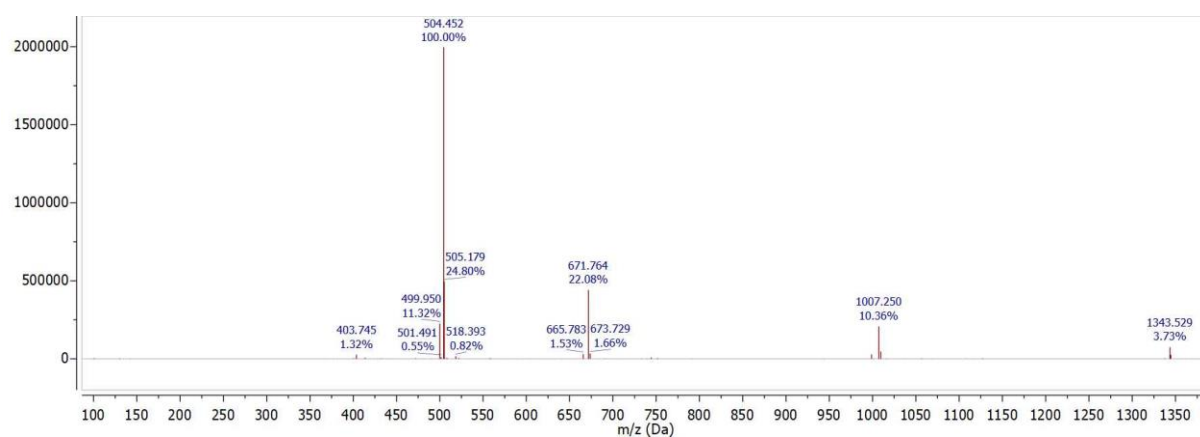
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1008.200	672.467	504.600	403.880	336.733
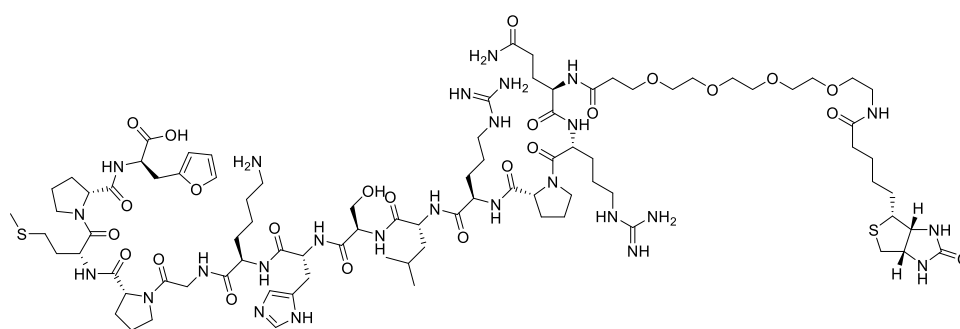
Chemical Formula: $C_{88}H_{144}N_{26}O_{24}S_2$
Exact Mass: 2013.03
Molecular Weight: 2014.40

Figure 43: Experimental data on purified AP1(Met)
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProMetFuaPhe-OH
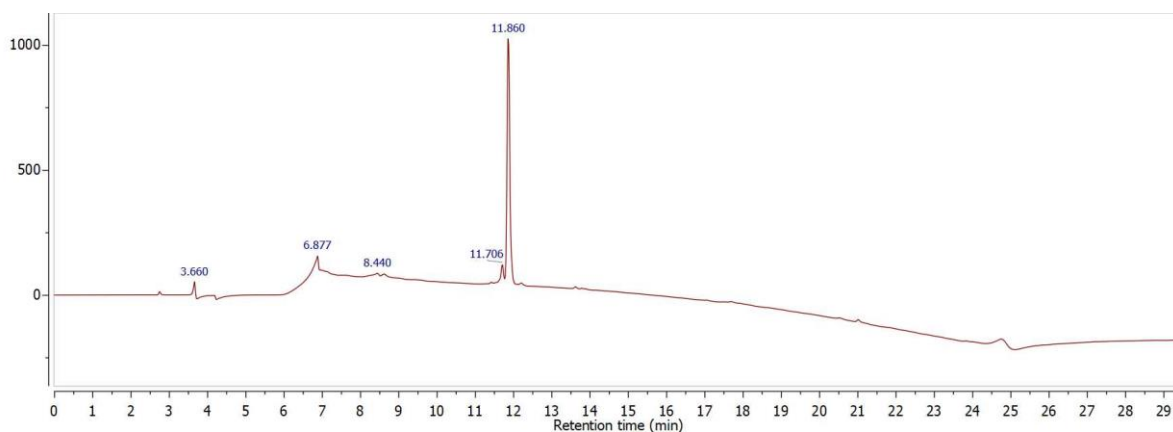
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
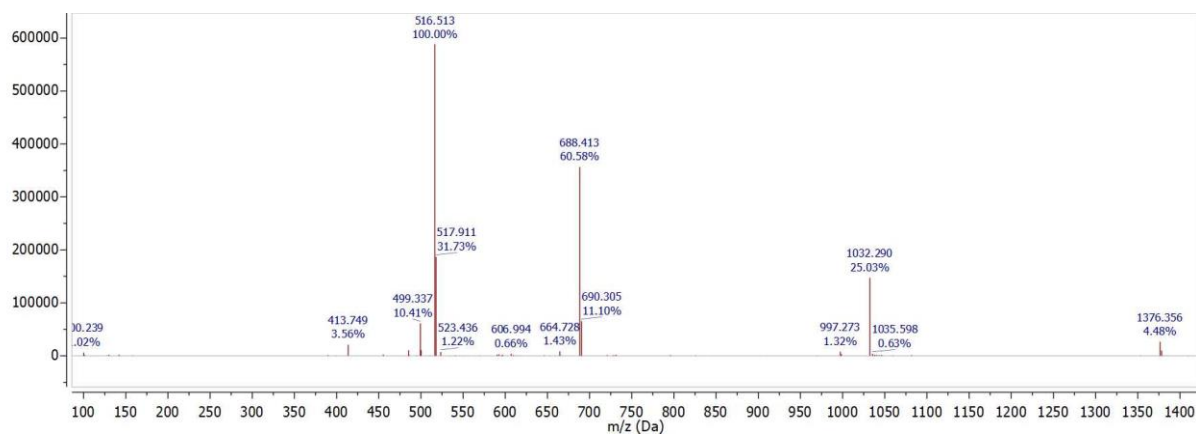
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1033.230	689.153	517.115	413.892	345.077
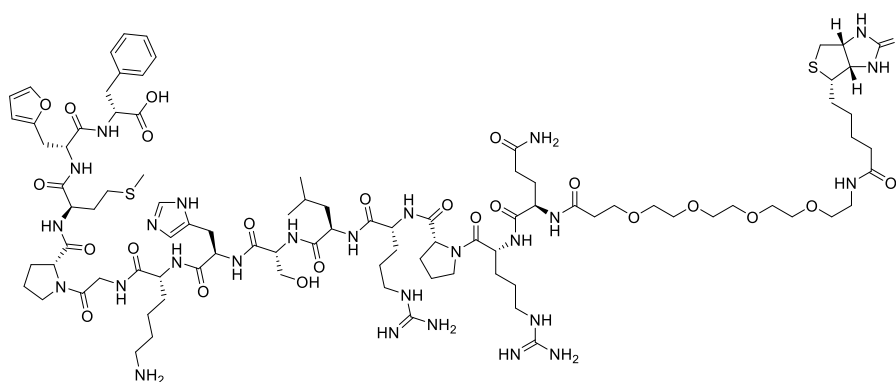
Chemical Formula: C92H146N26O24S2
Exact Mass: 2063.04
Molecular Weight: 2064.46

Figure 44: Experimental data on purified AP2(Met)
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.
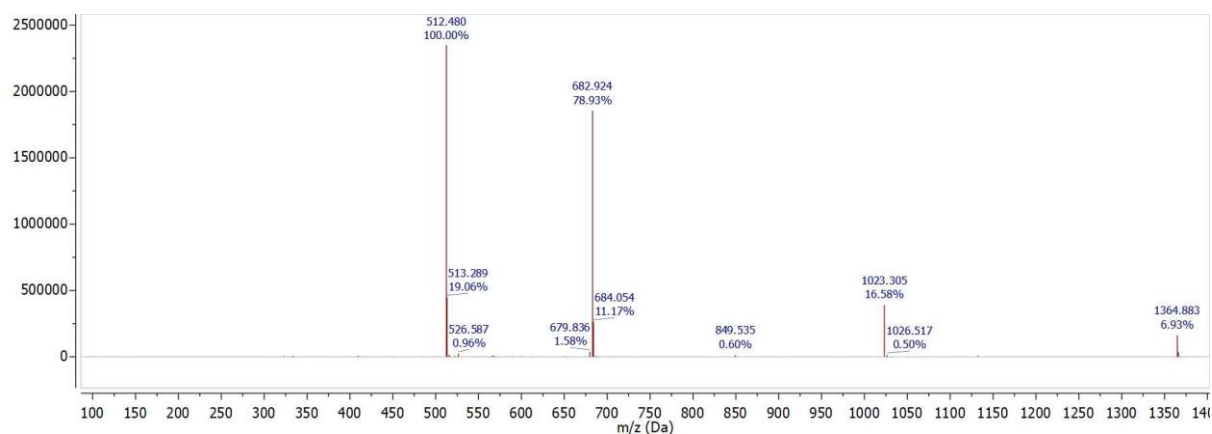
10.1.3.3 AP1(Nle)

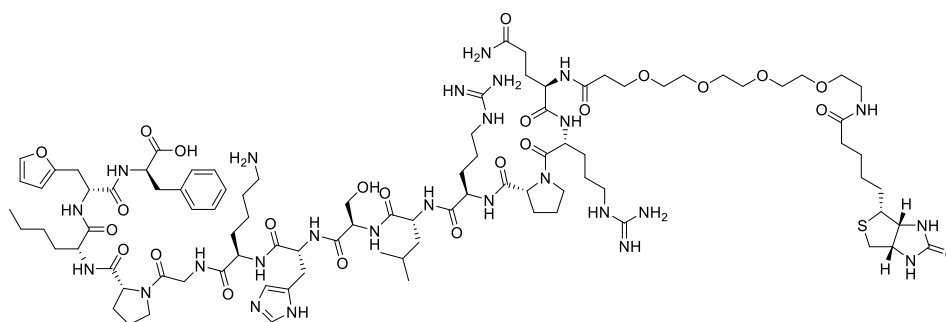Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProFua-OH
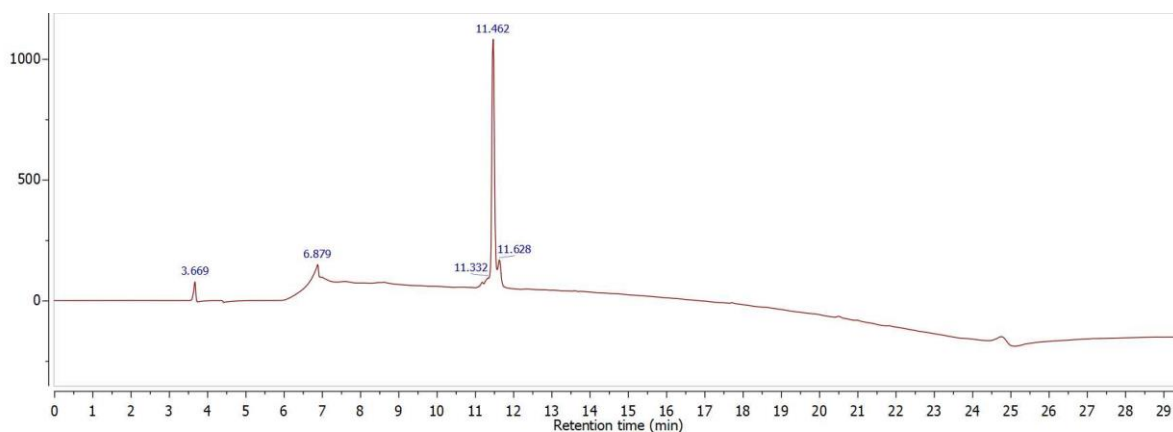
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	999.185	666.457	500.093	400.274	333.728

Chemical Formula: $C_{89}H_{146}N_{26}O_{24}S$
Molecular Weight: 1996.37

Figure 45: Experimental data on purified AP1(Nle)
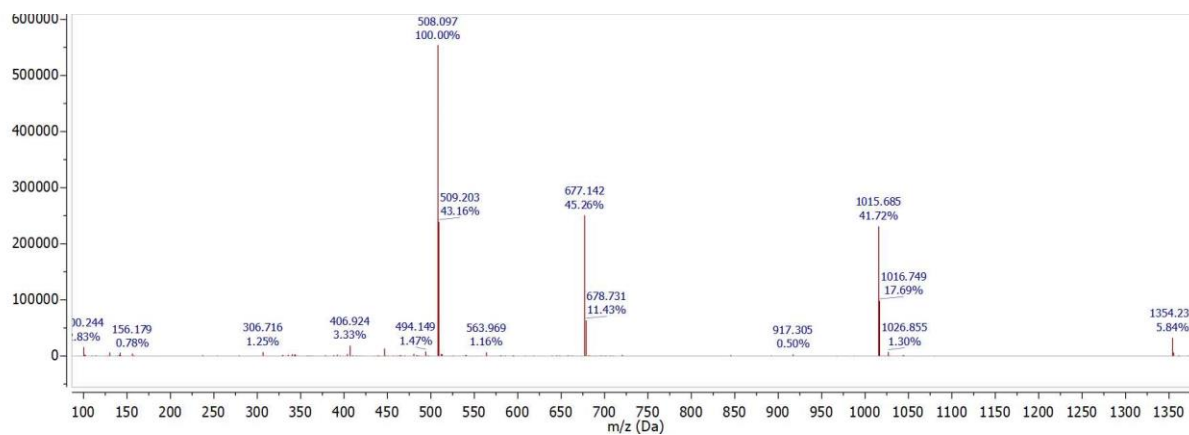*top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.*

Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleFuaPhe-OH
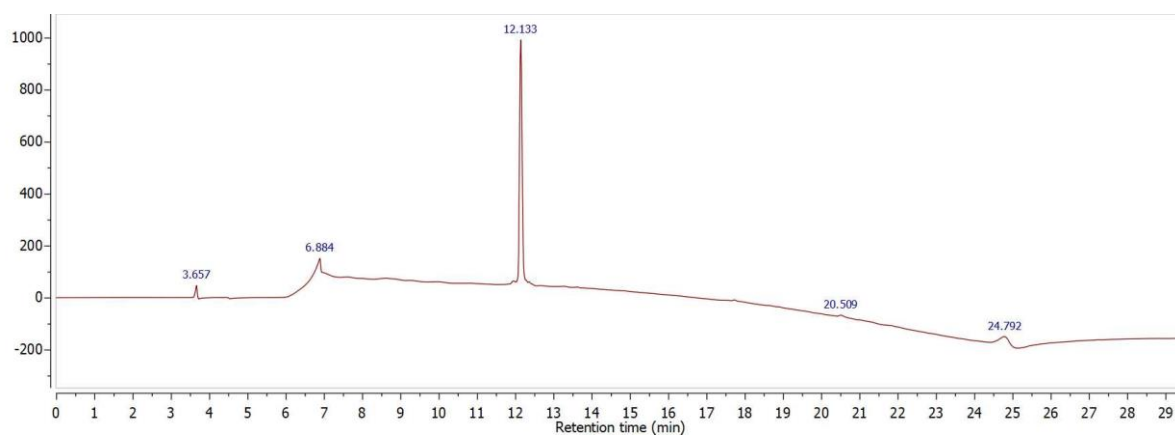
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)
RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1024.215	683.143	512.608	410.286	342.072

Chemical Formula: $C_{93}H_{148}N_{26}O_{24}S$
Molecular Weight: 2046.43

Figure 46: Experimental data on purified AP2(Nle)
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

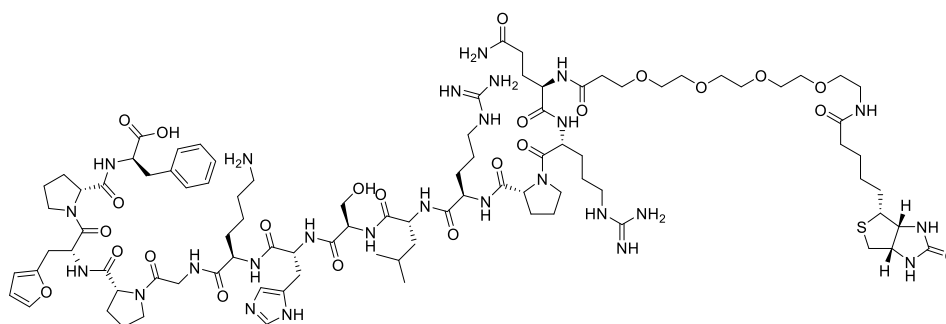Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProFuaProPhe-OH
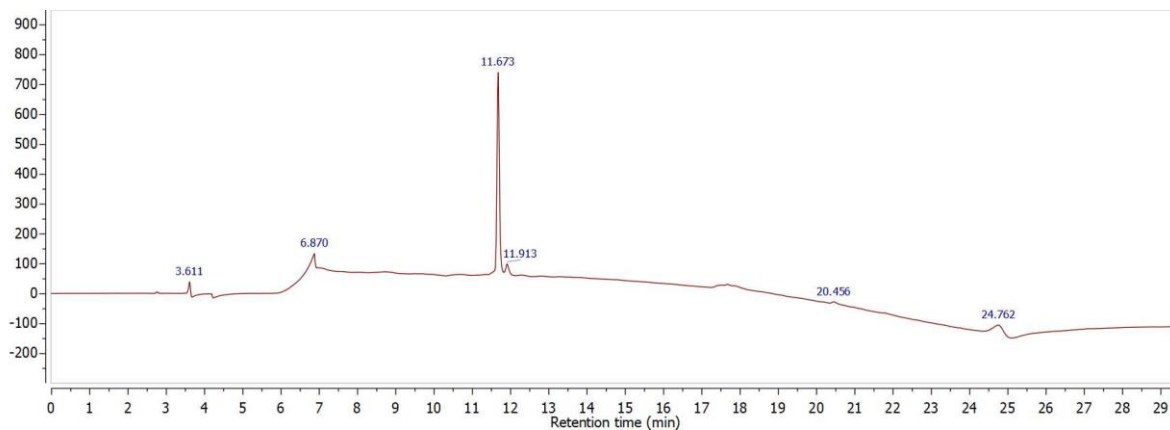
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

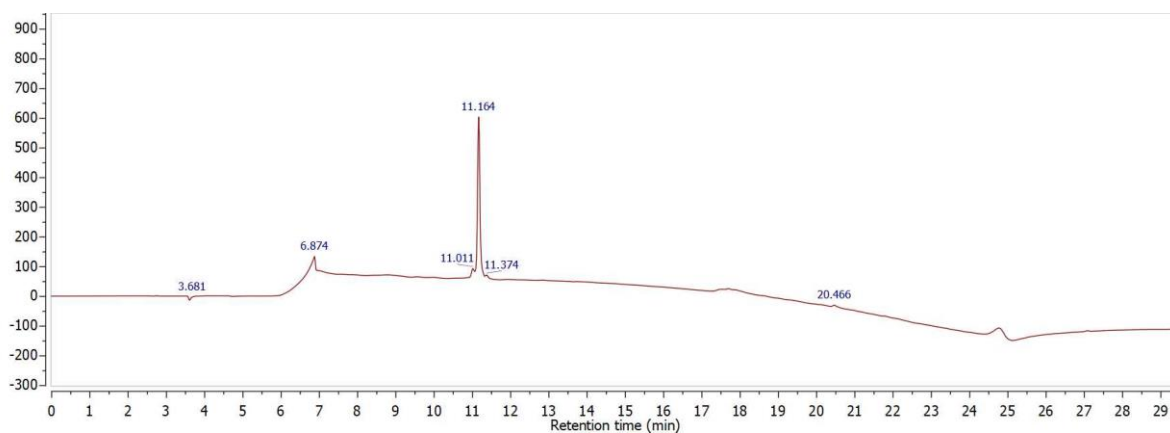| LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1016.190	677.793	508.595	407.076	339.397

Chemical Formula: $C_{92}H_{144}N_{26}O_{24}S$
Molecular Weight: 2030.38

Figure 47: Experimental data on purified AP3
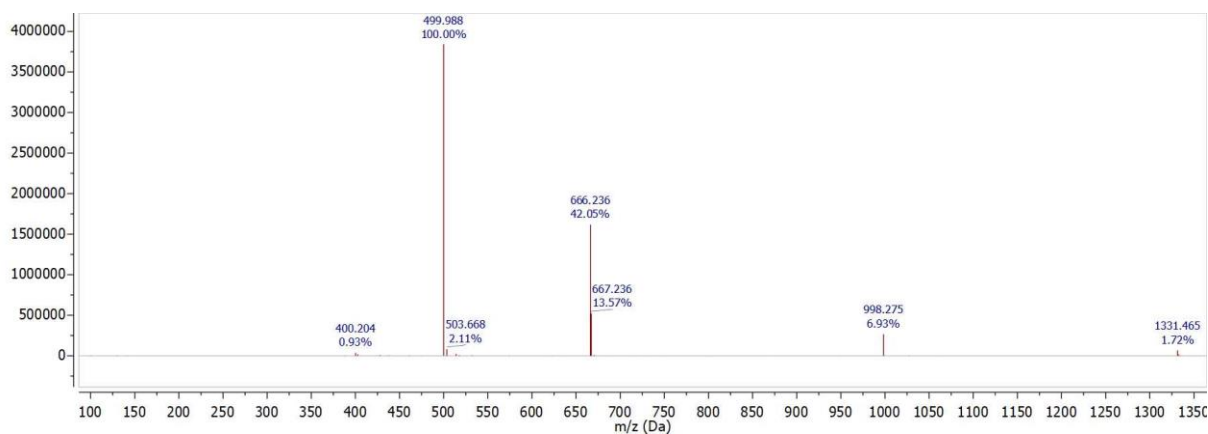*top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.*

Sequence: Biotin-PEG4-GlyProLysLeuNleArgProGluHisArgProSerFua-OH

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)
RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	999.185	666.457	500.093	400.274	333.728

Chemical Formula: $C_{89}H_{146}N_{26}O_{24}S$
Molecular Weight: 1996.37

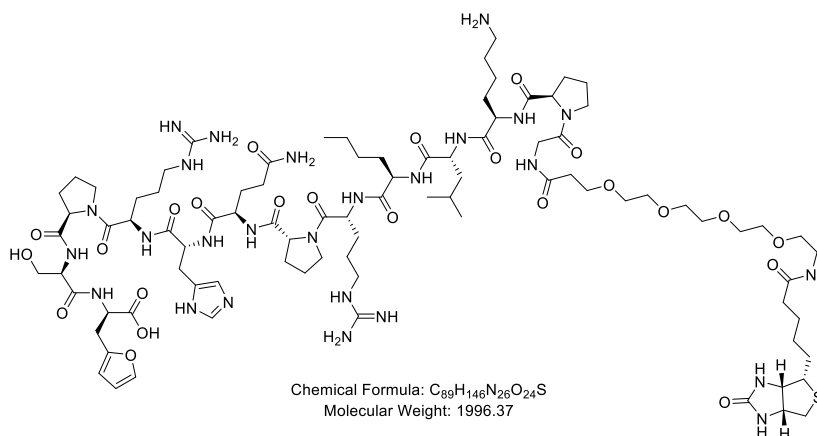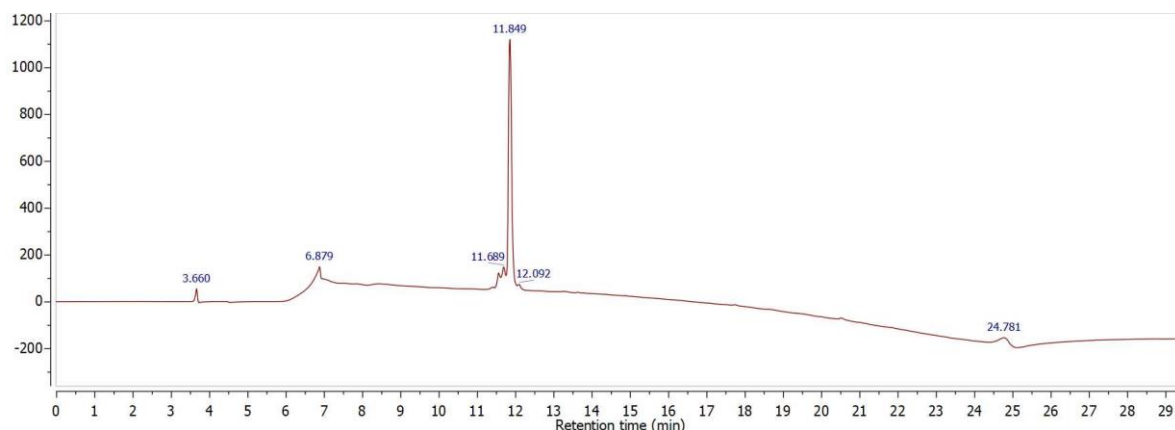*Figure 48: Experimental data on purified SAP1(Nle)*
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProPhe-OH

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
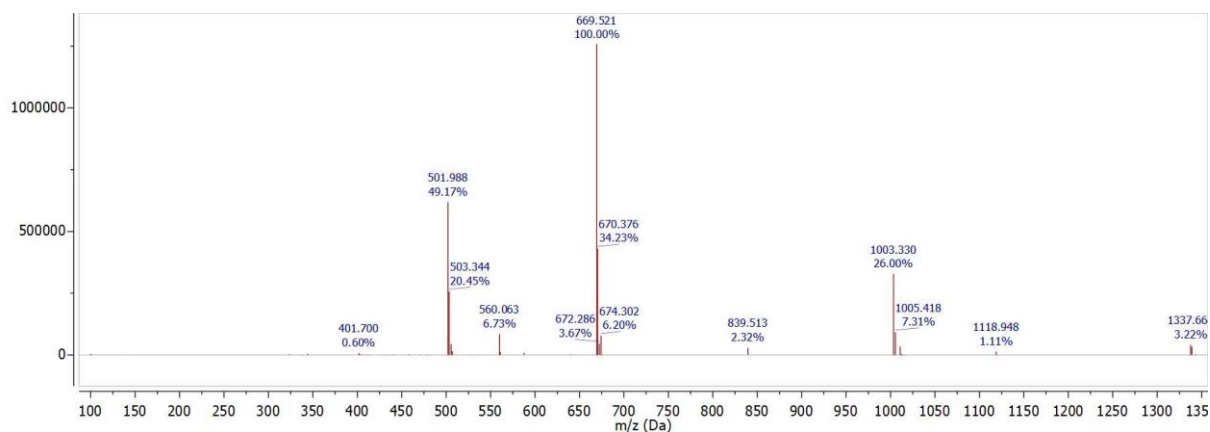
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1004.200	669.800	502.600	402.280	335.400

Chemical Formula: $C_{91}H_{148}N_{26}O_{23}S$
Molecular Weight: 2006.40

Figure 49: Experimental data on purified Native AP(Nle)
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.
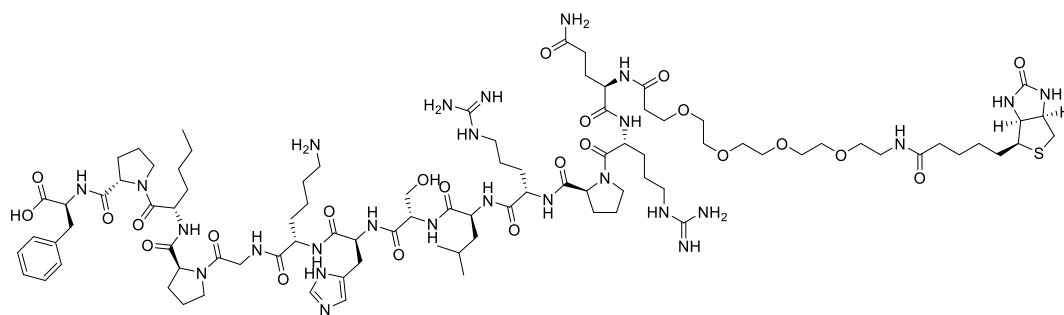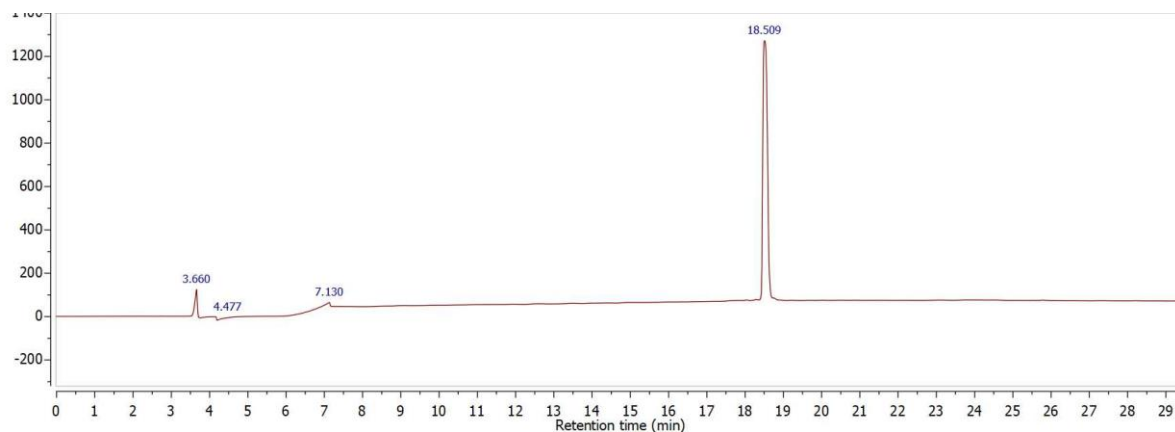
Sequence: DmtD-AlaPheFuaßAlaBio-NH2

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	597.220	398.480	299.110	239.488	199.740

Chemical Formula: $C_{58}H_{85}N_{11}O_{14}S$
Exact Mass: 1191.60
Molecular Weight: 1192.44

Figure 50: Experimental data on purified OP3
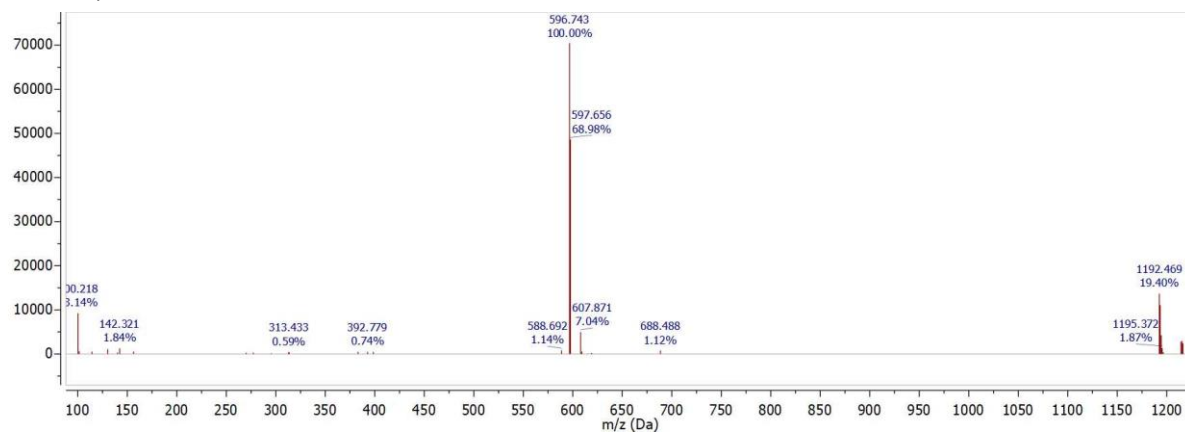*top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.*

Sequence: DmtD-AlaFuaPheßAlaBio-NH2

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	597.220	398.480	299.110	239.488	199.740

Chemical Formula: $C_{58}H_{85}N_{11}O_{14}S$
Exact Mass: 1191.60
Molecular Weight: 1192.44

Figure 51: Experimental data on purified OP4
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.
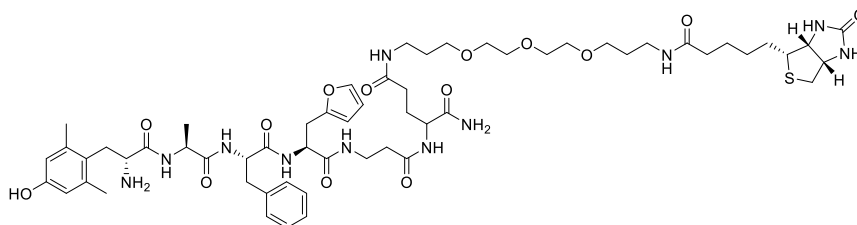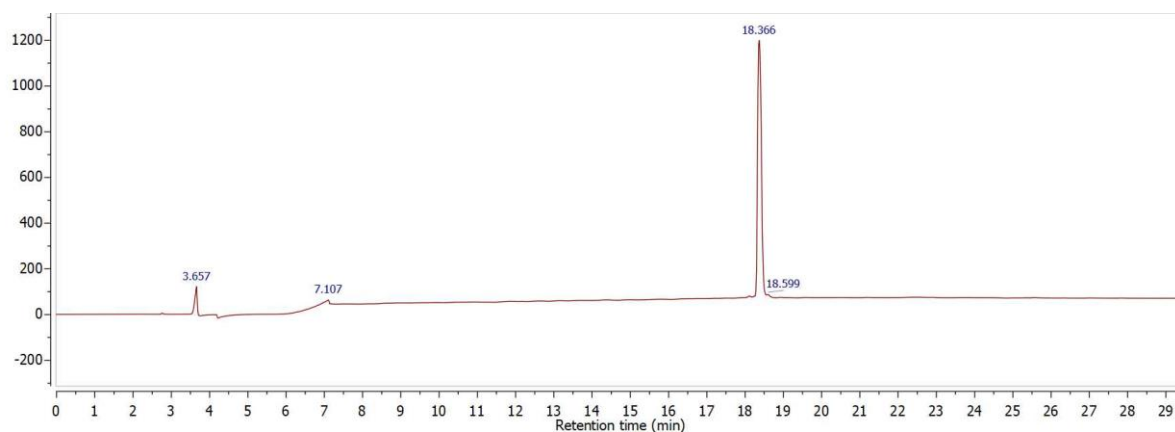
Sequence: DmtD-AlaFuaßAlaBio-NH2
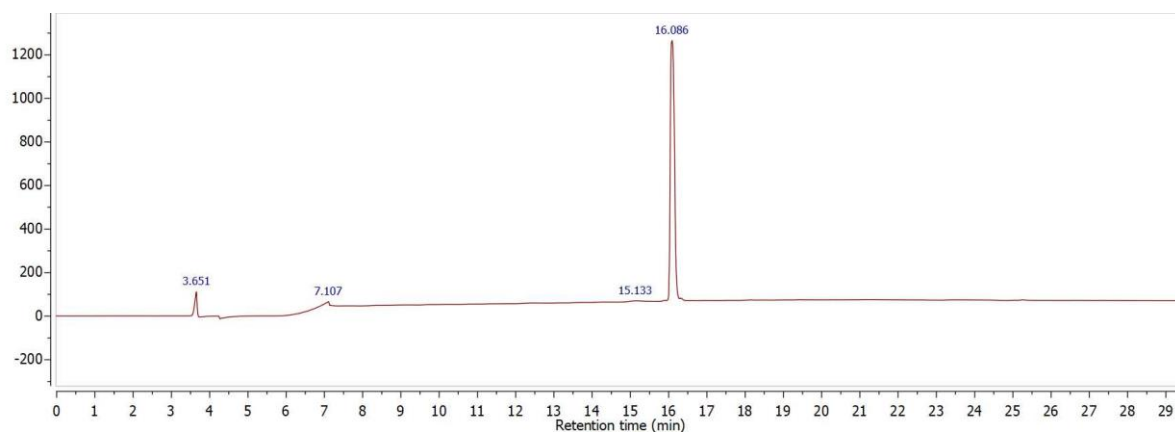
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)
RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	523.630	349.420	262.315	210.052	175.210

Chemical Formula: $C_{49}H_{76}N_{10}O_{13}S$
Exact Mass: 1044.53
Molecular Weight: 1045.26

Figure 52: Experimental data on purified OP5
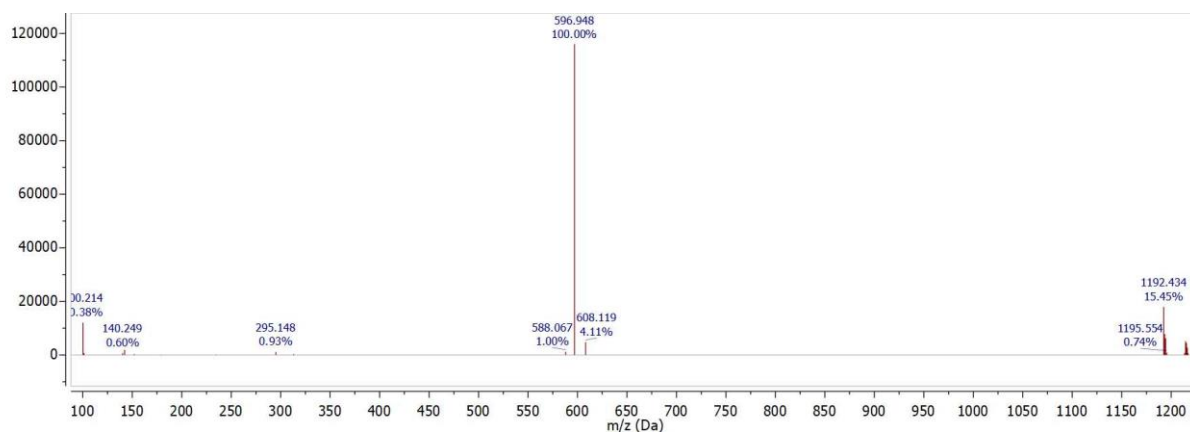*top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.*
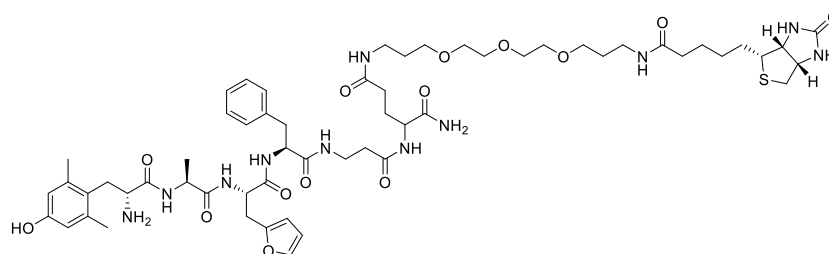
10.1.3.11 OP6

Sequence: DmtD-AlaPheFuaBio-NH2
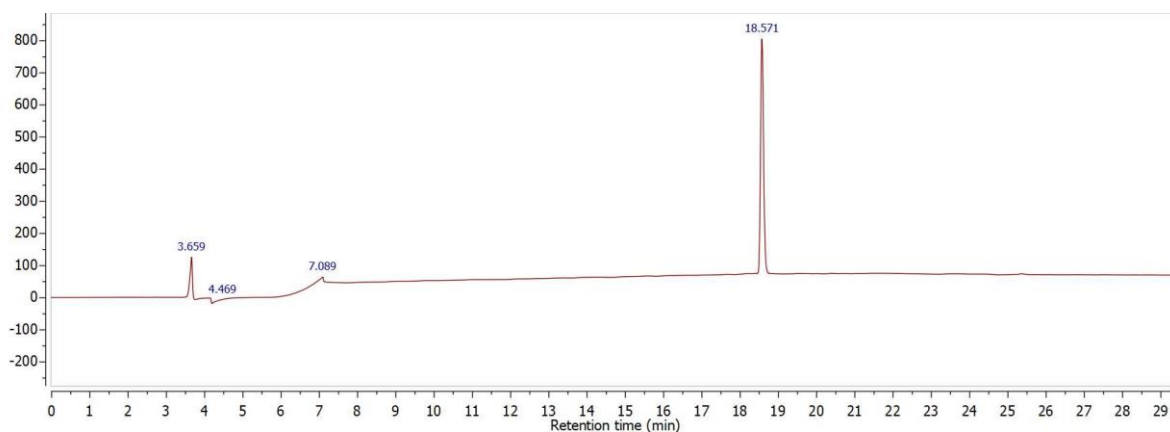
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
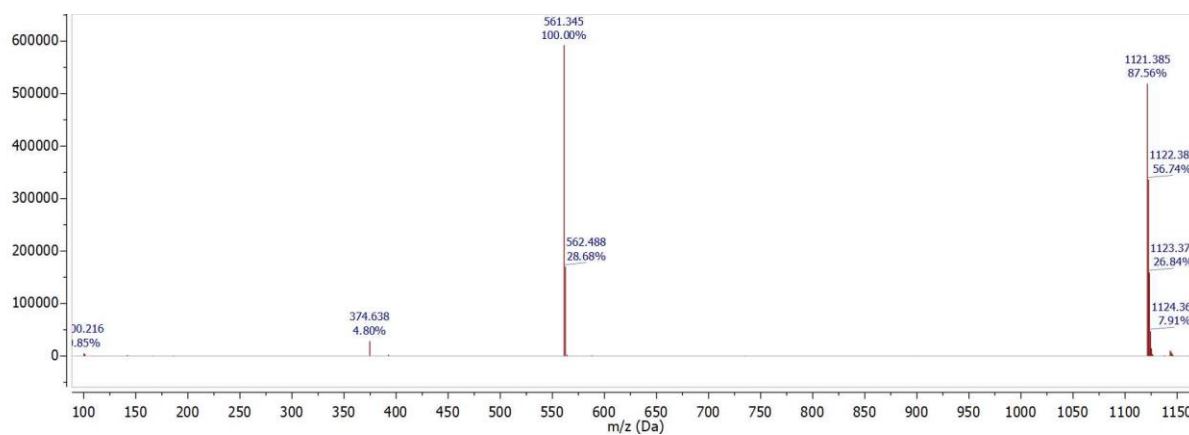
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	561.680	374.787	281.340	225.272	187.893

Chemical Formula: $C_{55}H_{80}N_{10}O_{13}S$
Exact Mass: 1120.56
Molecular Weight: 1121.36

Figure 53: Experimental data on purified OP6
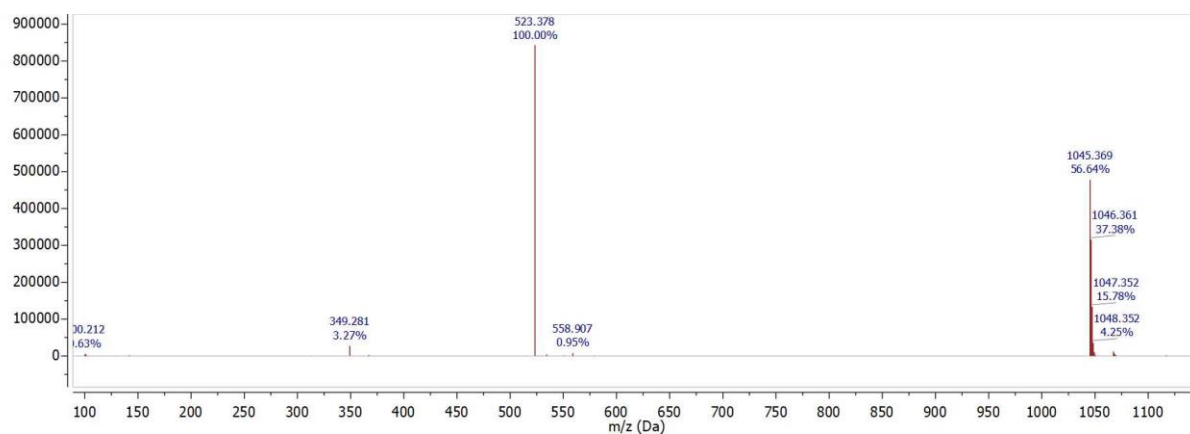*top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.*

10.1.3.12 Classic cleavage cocktail

Sequence: H-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
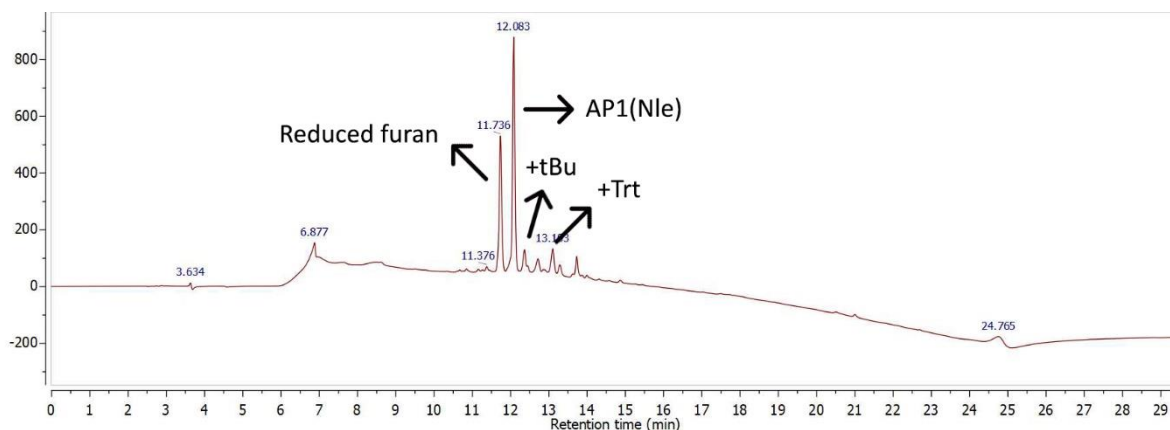
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	804.425	536.617	402.713	322.370	268.808

Chemical Formula: $C_{72}H_{115}N_{23}O_{19}$
Molecular Weight: 1606.85

Figure 54:Experimental data on classic cleavage cocktail
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

75

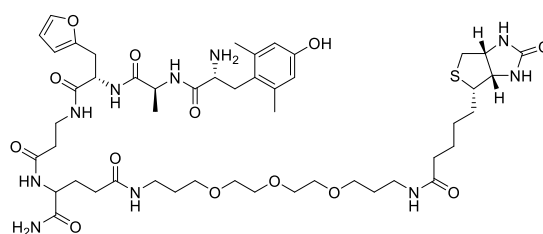Sequence: H-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
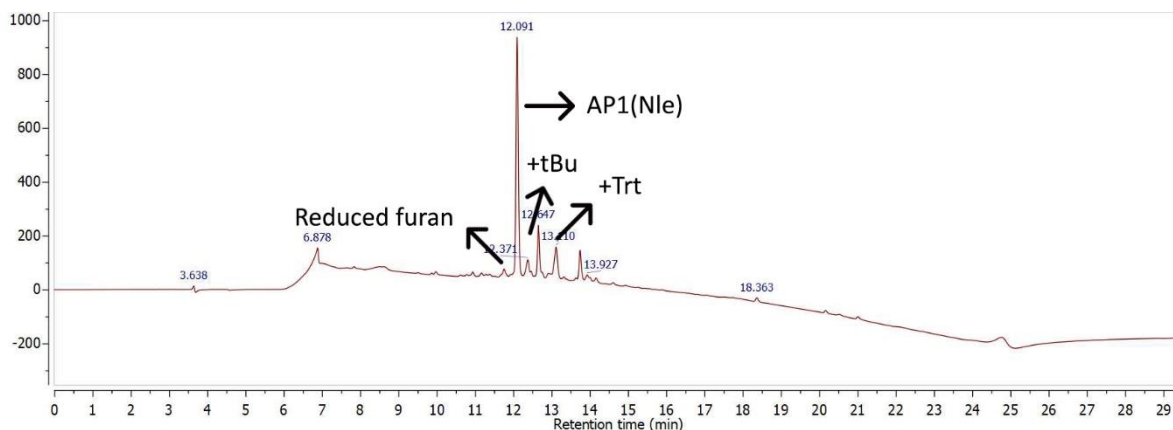
Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	804.425	536.617	402.713	322.370	268.808

Chemical Formula: $C_{72}H_{115}N_{23}O_{19}$
Molecular Weight: 1606.85

Figure 56:Experimental data on classic cleavage cocktail
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

76

Sequence: H-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
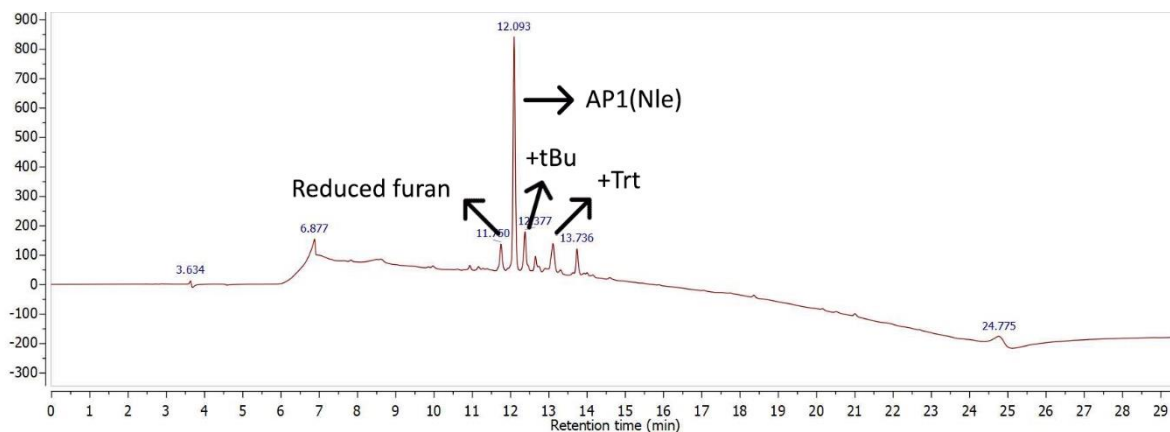
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
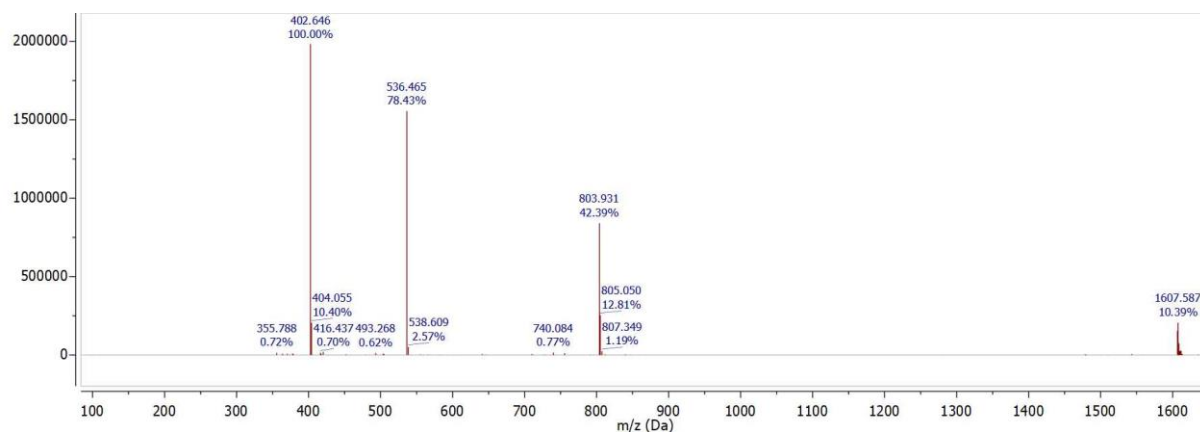
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	804.425	536.617	402.713	322.370	268.808
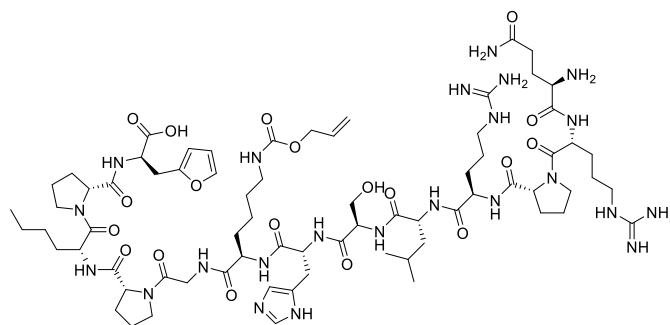
Chemical Formula: $C_{72}H_{115}N_{23}O_{19}$
Molecular Weight: 1606.85

Figure 58:Experimental data on classic cleavage cocktail
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

77

Sequence: H-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
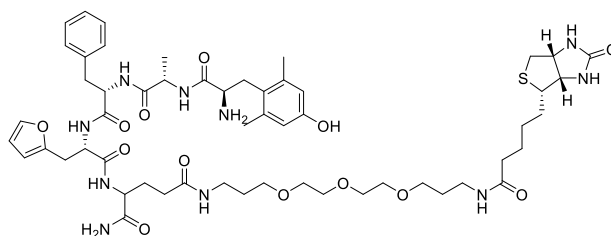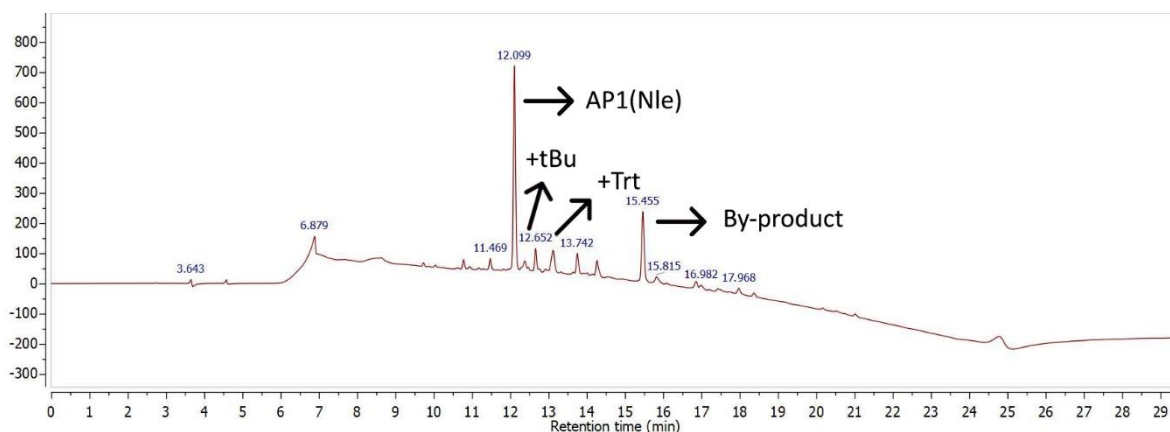

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
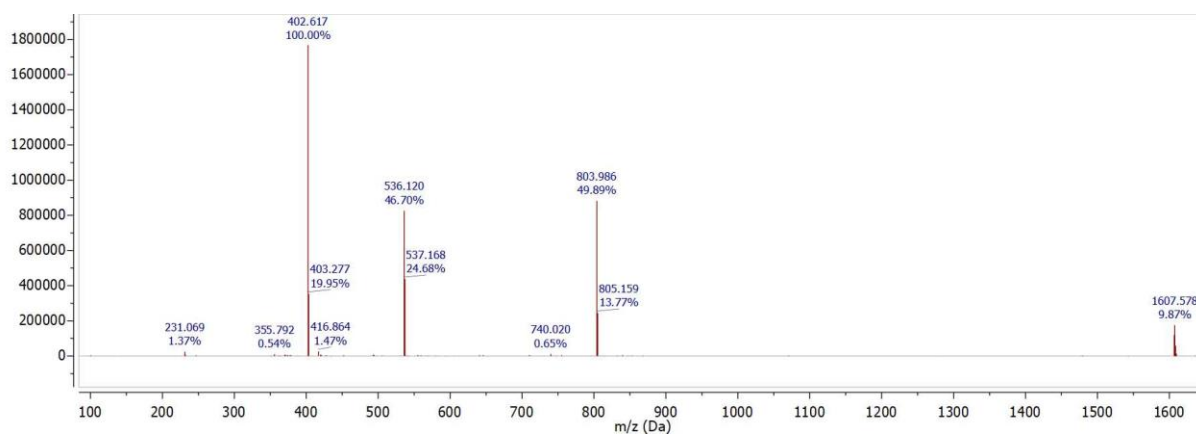
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	804.425	536.617	402.713	322.370	268.808
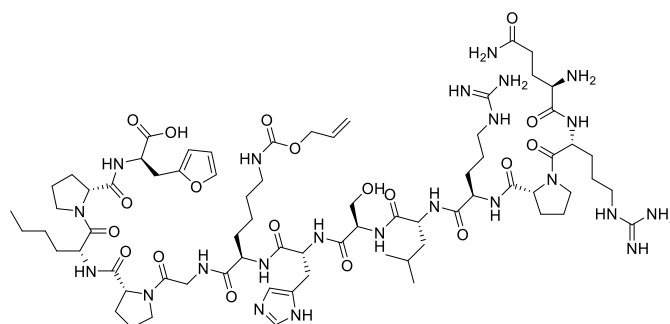
Chemical Formula: $C_{72}H_{115}N_{23}O_{19}$
Molecular Weight: 1606.85

Figure 59: Experimental data on CC3
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

78
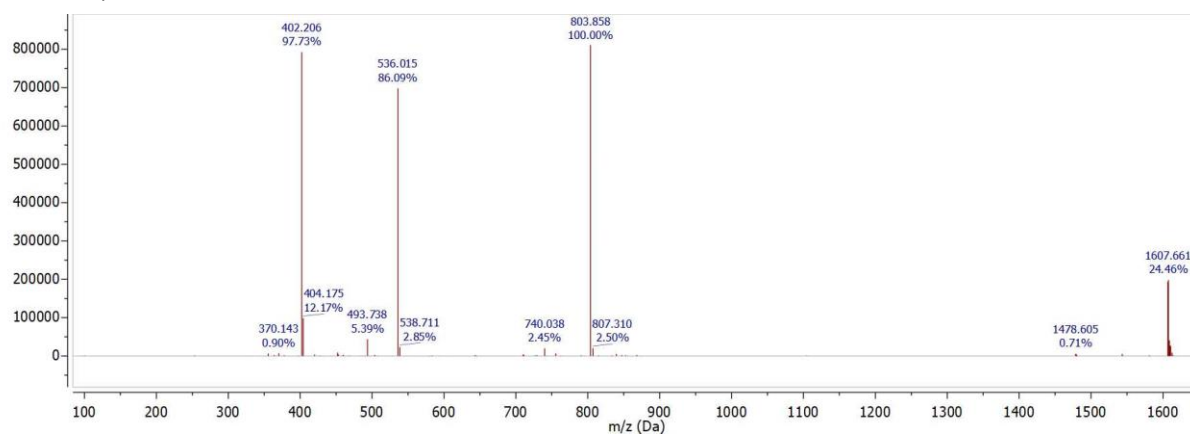
10.1.3.16 CC4

Sequence: H-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
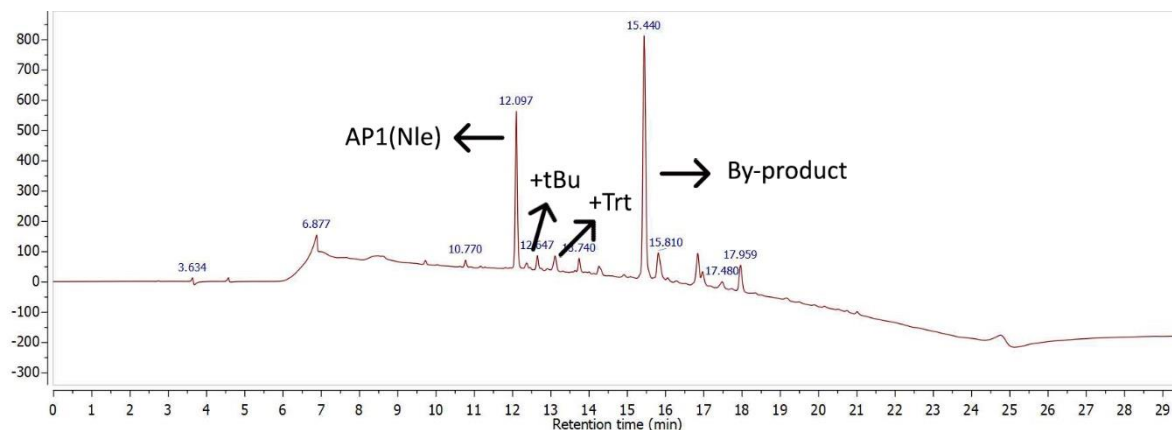
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes
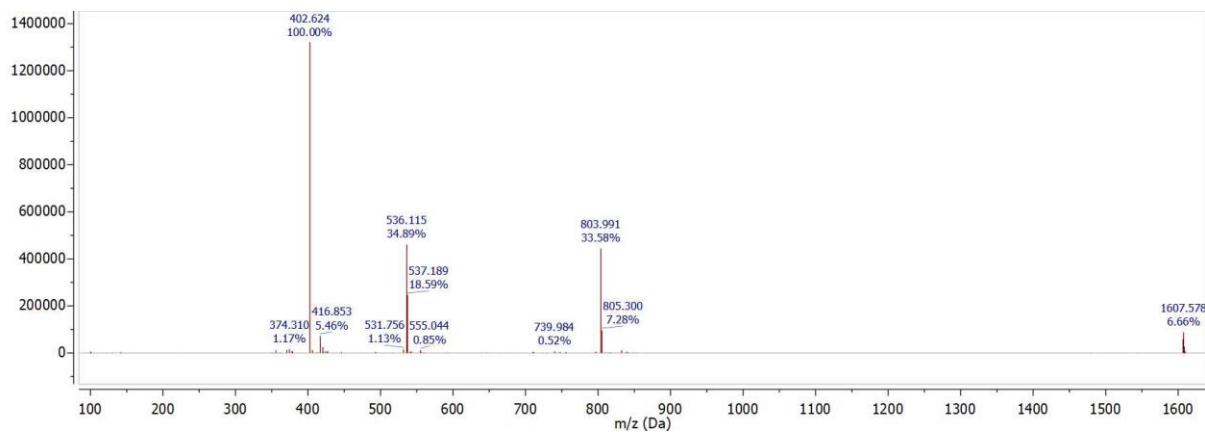
LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	804.425	536.617	402.713	322.370	268.808
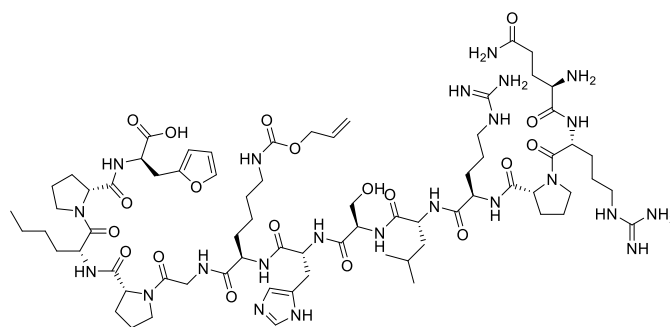
Chemical Formula: $C_{72}H_{115}N_{23}O_{19}$
Molecular Weight: 1606.85

Figure 60: Experimental data on CC4
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

10.1.3.17 CC1.2

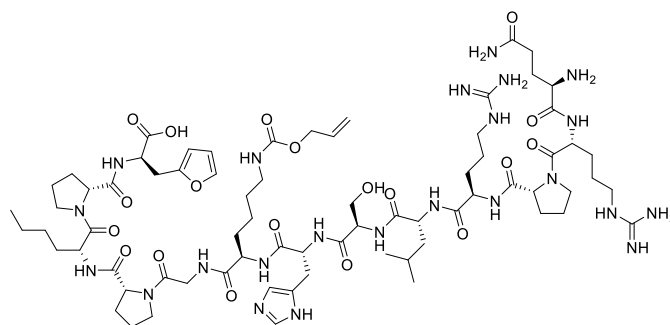Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
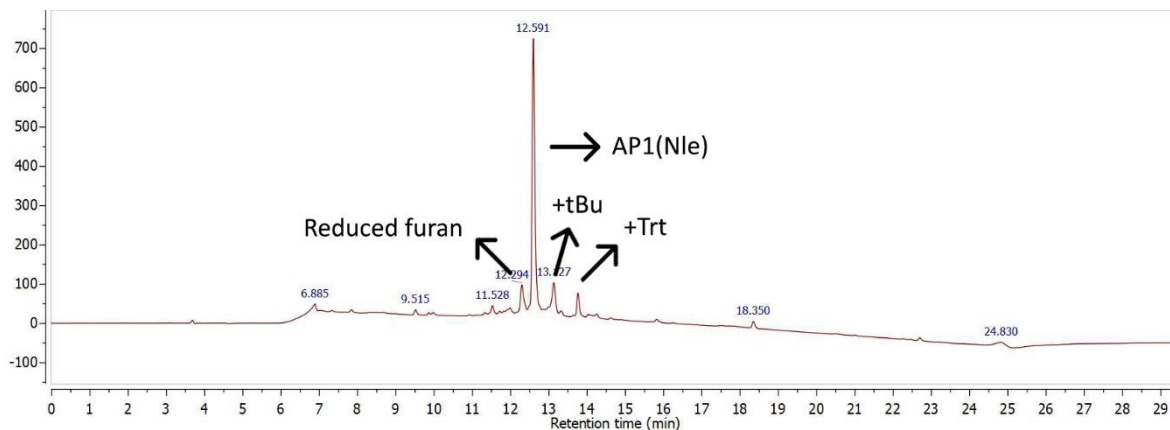
Column: Phenomenex Luna C18 (5 μm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1041.220	694.480	521.110	417.088	347.740

Chemical Formula: $C_{93}H_{150}N_{26}O_{26}S$
Molecular Weight: 2080.44
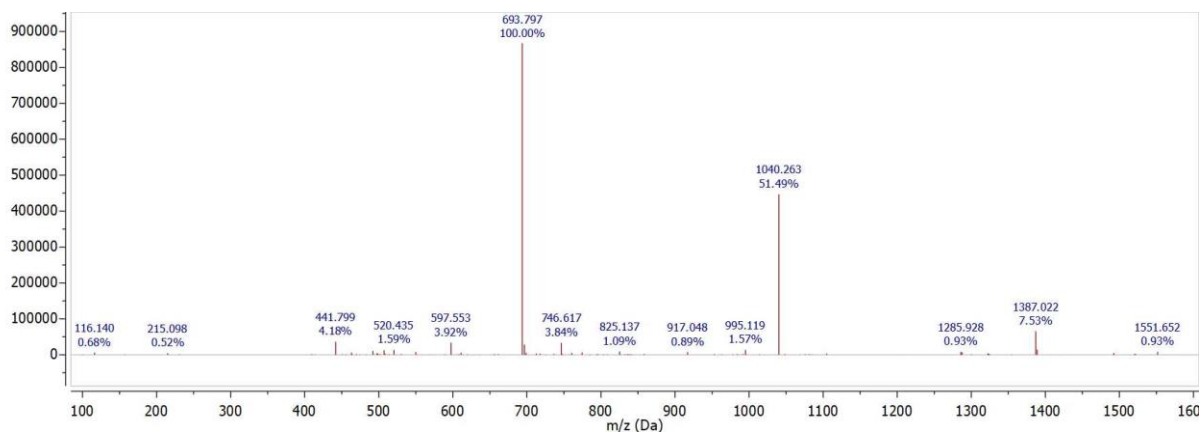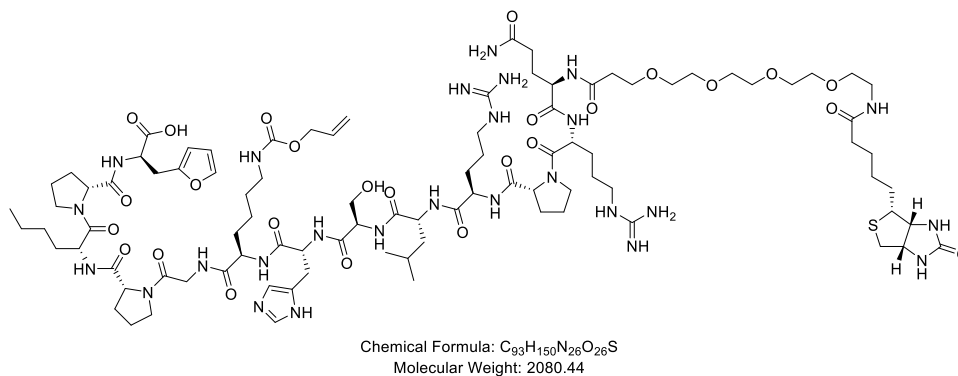
Figure 61: Experimental data on CC1.2
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.
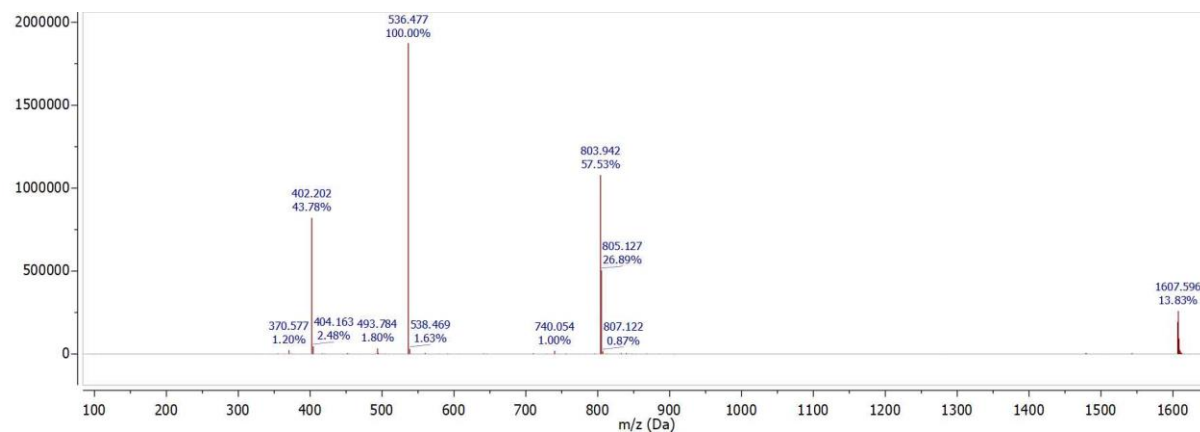
Sequence: biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProFua-OH (Alloc protected)
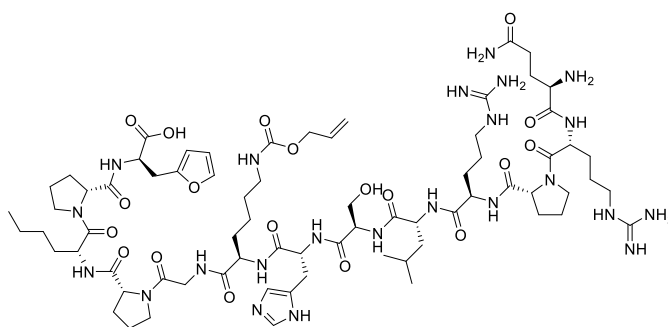

Column: Phenomenex Luna C18 (5 µm x 4.6 mm x 250 mm)

RP-HPLC, 0% to 100% ACN over 20 minutes

LC-ESI-MS	[M+2H]/2	[M+3H]/3	[M+4H]/4	[M+5H]/5	[M+6H]/6
Theoretical m/z:	1041.220	694.480	521.110	417.088	347.740

Chemical Formula: $C_{93}H_{150}N_{26}O_{26}S$
Molecular Weight: 2080.44
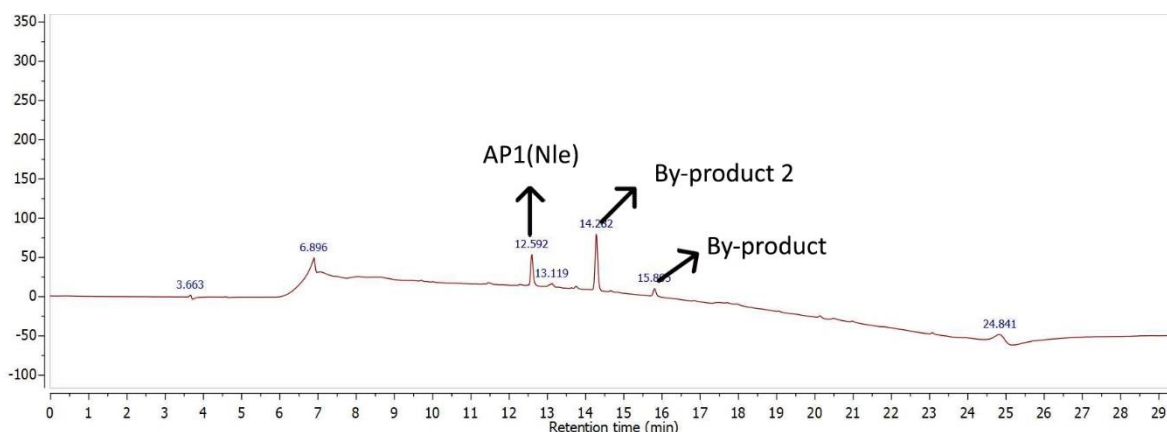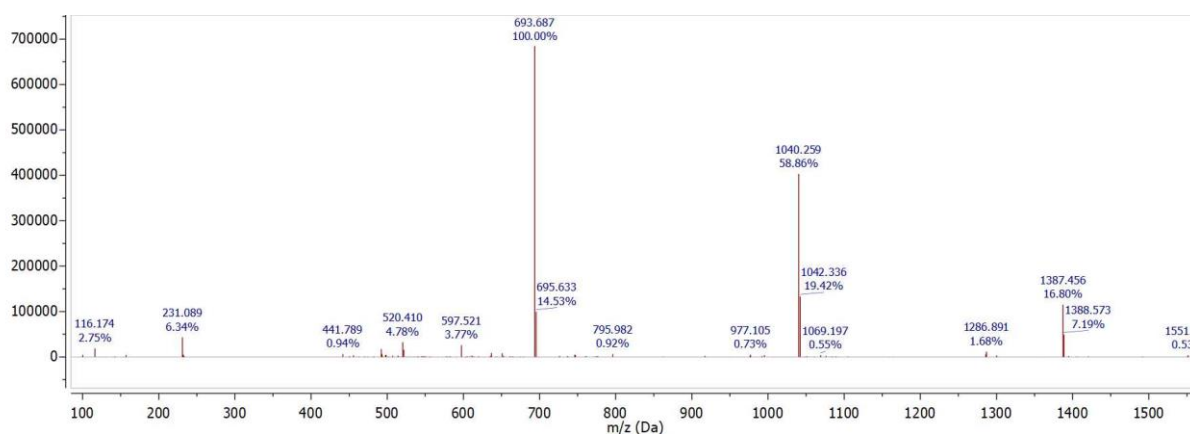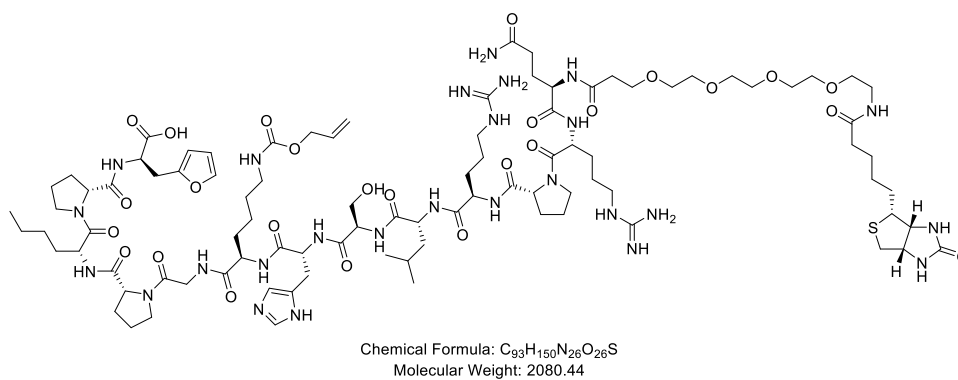
Figure 62:Experimental data on CC3.2
top: RP-HPLC, middle: ESI-MS and bottom: chemdraw structure.

81

10.2 Crosslinking experiments

10.2.1 General and reagents

All crosslinking experiments were performed on Campus Rommelaere of Ghent University in the labs of Professor Christophe Ampe and Professor Marleen Van Troys, guided by my thesis supervisor Laia Miret Casals. All cells were grown in Dulbecco Modified Eagle's Medium with 10% Fetal Bovine Serum (FBS, Gibco, Ref 10270-106) and 100 µg/mL of penicillin/streptomycin (Ref 15140-122, Gibco, Invitrogen) at 37 °C in a humidified atmosphere of 5% CO_2/95% air to maintain the pH of the medium. The medium is renewed three times per week. Since the cells used for this thesis show high proliferating rates, the cultures were subcultured at least twice a week. For the washing steps Dulbecco's Phosphate Buffered Saline (DPBS) was used. For the peptide incubation experiments Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L D-Glucose and no pyruvate was used to prepare the peptide solutions. Peptides were added to this medium from a stock solution of 25 mM in DMSO to reach desired peptide concentration. It is very important that no FBS was added to the medium used for the incubation experiments. The cells used for the apelin peptide experiment belong to the MDA-MB-231 cell line, cells for the opioid peptide experiment (performed by Laia Miret Casals) belong to the SH-SY5Y cell line.

10.2.2 Procedures and instrumentation

10.2.2.1 Cell culturing and seeding

The cells were grown in culture flasks with filter cap type and a surface of 75 cm^2 in a continuous layer at 37 degrees Celsius in 30 mL liquid volume. Cells were split every 2-3 days by harvesting, rediluting them in new cell growth medium and transferring them to new culture flasks. One almost confluent flask was diluted and divided over three new flasks. When cells would continue to grow for longer they would reach a confluent state, impeding further cell division and resulting in death over time. Cells were split inside a laminar flow fume hood. First, cell culture medium and trypsin are pre-heated in a 37°C bath. The cell culture medium is removed from the culture flask and discarded using a vacuum pump and Pasteur pipette. The cells were washed with 15 mL of PBS containing no $CaCl_2$ and no $MgCl_2$, gently added to the side of the cell flask to avoid cell layer disturbance. The wash solution was removed and discareded using a vacuum pump and Pasteur pipette. Next, 2 mL of trypsine/EDTA solution was added, covering the entire cell layer. Trypsine is a serine protease that disrupts the

proteins attaching the cells to the culture flask surface. The cell culture flask was shaken vigurously and incubated at 37 degrees Celsius for 3 minutes. 15 mL pre-warmed cell culture medium containing FBS and antibiotics was added and dispersed by pipetting over the cell layer surface several times to homogenize. The cells were counted in a Neubauer chamber. The mixture was divided over 3 new cell culture flasks, every flask was subsequently diluted .with another 25 mL of the growth medium with FBS and antibiotics. The cell culture flasks were shaken vigorously, checked under microscope to ensure no lumpy cell aggregates are present before putting them back in the 37 degrees incubator.

The same procedure was employed for seeding cells on 6-well (round) plates for the peptide incubation tests.

10.2.2.2 Cell counting

The cells were counted using a Neubauer chamber. The two semi-reflective rectangles are the counting chambers. Each chamber is engraved with a laser-etched grid of perpendicular lines. The device is carefully crafted so that the area covered by the lines is exactly 1 mm^2 and the depth is 0.1 mm. This allows counting the number of cells in a specific volume of solution, subsequently allowing calculation of the concentration of cells in the solution. After careful homogenization, 12 µL of cell suspension is applied between the coverslip (24 x 50 mm, Menzel-Glaser) and the Neubauer chamber (Figure 63). The cell suspension is applied on the edge of the coverslip and moves into the chamber by capillary action, completely filling the chamber with the sample. The number of cells in the chamber is determined by counting using a microscope (Nikon Eclipse TS100). The Neubauer chamber contains 9 squares (1 x 1 mm). The number of cells in each square is counted for at least 4 squared and the average number of cells is calculated. To calculate the number of cells in 1 mL of sample, this average value is multiplied by 10^4.
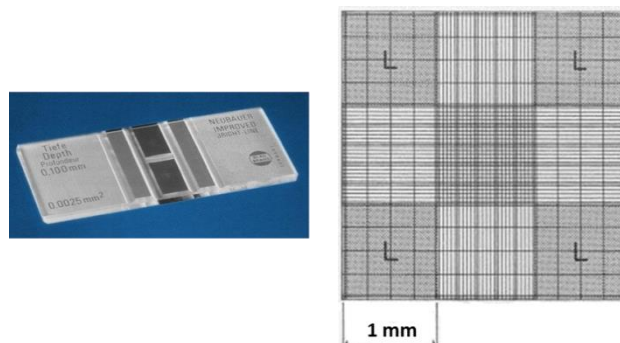
Figure 63: Neubauer chamber

10.2.2.3 Peptide incubation experiments

Experiments were performed on 6-well plates that contain about 825.000 cells. All manipulations were performed in a laminar flow fume hood. First, the growth medium was removed using a vacuum pump and Pasteur pipette. Next, the cells are rinsed with 1 mL of PBS containing no $CaCl_2$ and no $MgCl_2$, which was removed in the same way. 2 mL of the solution with desired concentration of peptide (or control solution containing 100 µM DMSO) was added (described in Section 10.2.1). Finally, the cells were incubated at 37 degrees for a desired period of time (30', 1 hour or 2 hours).

10.2.2.4 Cell harvesting and lysis

After the desired incubation time, the cells were washed with PBS and 100 µL of lysis buffer was added to the wells and the 6-well plates were immediately put (and kept) on ice to prevent enzymatic digestion of proteins. The lysis buffer used is specifically designed for solubilizing membrane proteins, it consists of: 7 M ureum, 2 M thioureum, 0.5% TRITON, 0.324 M DTT, phosphatase inhibitors (10 mM NaF, 1mM NaOVanadate) and protease inhibitors (1000x). The cells were scrubbed from the surface, transferred to an Eppendorf and cooled with ice. The cells are homogenized (10 strokes) using a syringe (1mL 25GA x 5/8in (0.5 x 16mm) BD Plastipak^{TM}). Next, the Eppendorfs containing the lysate solution were centrifuged at 4 °C , 0.7 x g for 10 minutes to pellet cellular debris. The supernatant (protein sample) was transferred to a new Eppendorf using a micropipette and put on ice. The Eppendorf containing the cell debris pellet was discarded. The protein samples were mixed with Loading Sample Buffer (LSB), to ensure that the proteins are properly denatured, and heated for 5 min at 95 °C. The LSB 4X is composed by Tris-HCl 0.4 M, pH 6.8, glycerol 69.6 %, SDS 8 % (p/v), bromophenol blue 0.05 mg/mL (allows sample visualization), and DTT 100 mM. DTT is used to break disulfide bridges between proteins. Once the samples are heated, the samples are centrifuged at maximum speed. Usually 40 µg of protein are loaded in each well in SDS-PAGE.

10.2.2.5 SDS-PAGE

The gels used for the SDS-PAGE were prepared manually, by combining 8 mL of 10% separation gel with 3 mL of 5% stacking gel. Composition of the gels can be found in Table 10, solutions in the table are in water.

10% Separation Gel		5% Stacking Gel	
Compound or solution	Percentage	Compound or solution	Percentage
H_2O	40	H_2O	68
30% acrylamide mix	33	30% acrylamide mix	17
1.5M Tris pH 8.8	25	1.5M Tris pH 6.8	13
10% SDS	1	10% SDS	1
10% $(NH_4)_2S_2O_8$	1	10% $(NH_4)_2S_2O_8$	1
TEMED	0.04	TEMED	0.1

Table 10: Composition of the used gels for SDS-PAGE gel preparation.

Before loading the gels, the wells were 'washed' using the running buffer by simply forcing some running buffer in the cavity to remove lumps of undesired polymer that might occupy the cavities. The gels were loaded with the samples of 30 µL in every lane. Some strategically chosen lanes are loaded with 4 µL of the protein marker (Precision Plus Protein™ All Blue Standards). The protein marker lane allows for easy recognition and orientation of the gel during experimental manipulations. For every gel, a current of 25 mA was applied for each gel. The gel was run until the blue color reached the end point of the gel (± 90 minutes).

10.2.2.6 Blotting

In order to make the proteins accessible to antibody detection they are moved from within the gel onto a membrane made of polyvinylidene difluoride (PVDF). In this thesis PVDF Immobilon™ membranes (Millipore) were used as they have very high protein binding capacities, paired with good handling characteristics, they are highly chemically inert, and effective for low-background staining. The method for transferring the proteins is called electroblotting and uses an electric current to force proteins from the gel into the PVDF utilizing the negative charge of the protein bound to SDS. The proteins move from within the gel onto the membrane while maintaining the organization they had within the gel. The transfer equipment used for this thesis was Mini Protean TransBlot Cell (BioRad), which employs a transfer tank apparatus filled with blotting buffer, in which the gel/membrane sandwich is placed. The blot sandwich consists of a sponge, three sheets of Whatman 3 MM filter paper, the gel, the PVDF membrane (cut to the size of the gel), three sheets of Whatman paper and a second sponge. All parts of this assembly (except the gel) are pre-wetted with blotting buffer.

10.2.2.7 Membrane blocking and antibody incubation

For every membrane, 10 mL of a 1:1 solution of PBS/blocking solution (Odyssey, Lyeor) was used. The membrane was incubated with this solution for 1 hour on a tilt shaker. Next, incubation with the primary antibody was performed on the tilt shaker overnight. For the apelin peptides experiment Apelin Receptor Antibody (5H CLC), Abfinity^{TM} Rabbit was used. For the opioid peptides experiment, Anti-μ-Opioid Receptor (OPRM1) (extracellular) Antibody (Alomone, AOR-011) was used. After this, the membrane gets washed three times with PBS containing 0.1% TWEEN, before final incubation with Goat Anti-Rabbit IgG H&L (HRP) (Abcam, ab205718) and IRDye 680RD Streptavidin (Li-COR) for 45 minutes. Final washing is performed: PBS containing 1% SDS (3x), regular PBS (2x) and a final washing step with water.

10.2.2.8 Scanning

The membranes were scanned using an Odyssey Infrared Imaging System at 685 nm and 800 nm to record fluorescence from the fluorescent antibodies.

10.2.3 Experimental data

10.2.3.1 Apelin peptides experiment: Western blot scan
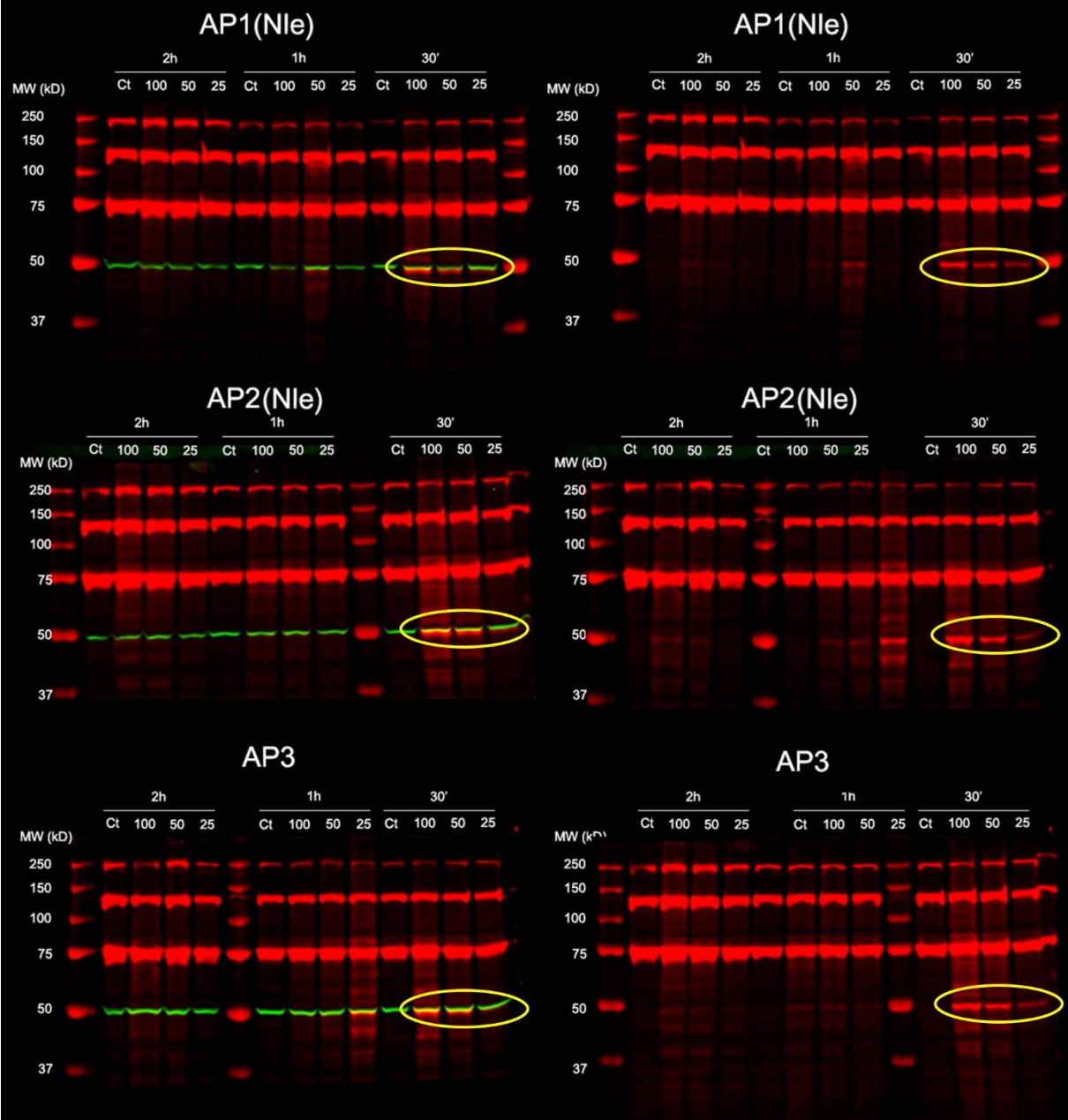
Figure 64: Western blot scans of apelin peptides experiments (left: red and green, right: red only).

10.2.3.2 Opioid peptides experiment: Western blot scan
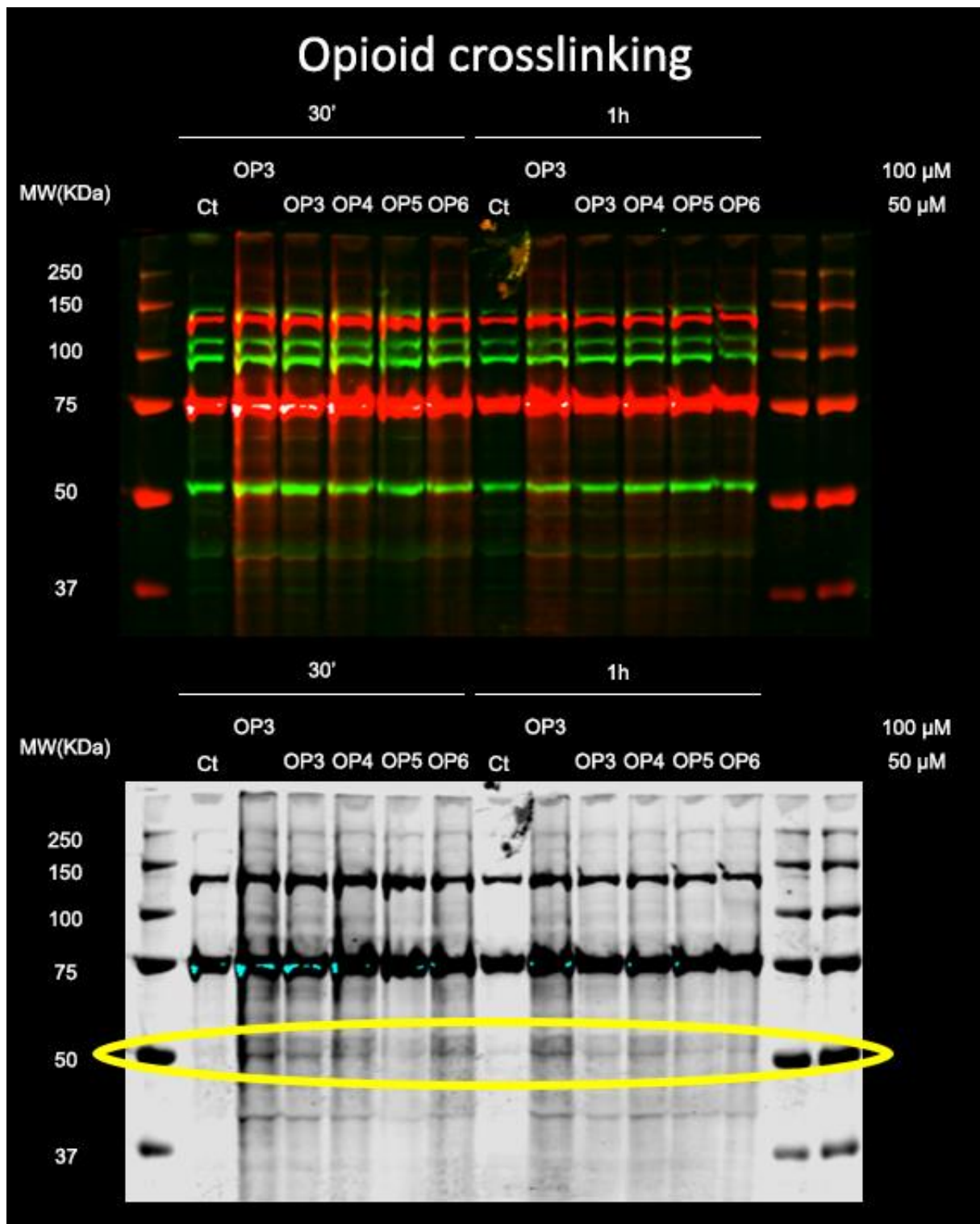
Figure 65: Western blot scans of opioid peptides experiment (top: red and green, bottom: red only, inversed and gray-scaled).

10.3 Visualization and simulation of ligand-receptor complexes

Schrödinger's Protein Preparation Wizard (PrepWizard) was used to prepare 5VBL[43] (downloaded from the Protein Data Bank (PDB)) structures for modeling and visualization. The PrepWizard adds missing hydrogen atoms, corrects metal ionization states, assigns bond orders, highlights residues with missing atoms, predicts protonation states, flips Asn/Gln and His residues to optimize the H-bond network, removes select crystallographic waters, optimizes the protein´s hydrogen bond network. Minimization of the structure could not be performed as no full license was available to me. 5VBL was imported as a biological unit and in both structures, all resolved crystal water molecules were maintained. Maestro Version 12.3.013, MMshare Version 4.9.013, Release 2020-12, Platform Windows-x64 Schrödinger was used to substitute the first, second or third residue of the ligand (counting from the C-terminus) by 2-furyl-L-alanine, as well as performing other modifications to obtain the peptide ligands AP1(Nle), AP2(Nle) and AP3.

The distance between the introduced furan and nucleophilic amino acids of the receptor in close proximity (less than 10 Å from the Cα of the furan-modified residues) were measured by rotating the side chains of the amino acids (including the 2-furyl-L-alanine) around selected dihedral angles. The number of possible dihedral angles depends on the residue. For example, lysine can be rotated around the various dihedral angles (Cα-Cβ bond, Cβ-Cγ bond, and Cγ–Cδ bond Cδ–Cϵ), such rotamers are chemically equivalent. Rotation was performed until a minimal distance value was achieved.

10.4 Molecular simulation of apelin 13 interacting with apelin receptor

10.4.1 CHARMM-GUI-solution builder workflow

In a modified pdb file, I discarded the SEQADV part, since I did not fully understand what was stated here and the numbering was very odd. On top of that, this part resulted in 'missing residues' on CHARMM-GUI. The CHARMM-GUI solution builder was executed on information coming from the SEQRES part only. This approach does probably imply the inclusion of the expression tags in the receptor for simulations done further down the road. The SEQADV part presumably eliminates these parts from the receptor, but I couldn't wrap my head around the numbering. The choice for discarding the SEQADV was made in order to allow me to keep an overview on what was happening inside the CHARMM-GUI solution builder, as well as making the mutations to be introduced and their corresponding numbering more straightforward (numbering taken straight from the SEQRES part in notepad)

- Page 1: upload modified pdb (no SEQADV part)

- Page 2: Chain A from **5** to 17, Chain B from **1** to **407** (for some reason, the standard values here were set to 19 -1054, no clue where these come from)

- Page 3: Renaming engineered residues:

    - HRG to ARG

    - ALC to LEU

    - OIC to PRO

    - NLE to MET

    - 200 to PHE

Still some residues were 'missing', I left 'model missing residues' checked, but didn't check any sub-options with specified areas (so I just left the standard options) Mutations (PROA = Apelin, PROB = Receptor):

    - PROA, 10 GLU to SER

    - PROA, 13 LYS to GLY

When trying to do the PROB mutations, it became clear that deleting the SEQADV didn't influence anything, as the dropdown box with options for mutations in the receptor didn't utilize the SEQRES numbering. Upon seeing this, the entire CHARMM-GUI procedure was repeated with the SEQADV left inside the pdb, apparently we're going to play by the rules of the authors of the paper after all…

- Page 1: upload original 5vbl pdb file

- Page 2: Chain A from **5** to 17, Chain B from 19 to **330** (When leaving the standard values -19 to 1054-, residue numbering jumps from 330 to 1000 in order to continue to 1054. As I don't know what these above 1000 residues are, I'll just ignore them, I can't seem to locate them in the pdb)

- Page 3: Renaming engineered residues:

  - HRG to ARG

  - ALC to LEU

  - OIC to PRO

  - NLE to MET

  - 200 to PHE

Still some residues of PROB (= receptor) were 'missing', I left 'model missing residues' checked, but didn't check any sub-options with specified areas (so I just left the standard options)

Mutations (PROA = Apelin, PROB = Receptor):

  - PROA, 10 GLU to SER

  - PROA, 13 LYS to GLY

Using the numbering used in the paper:

  - PROB, 117 ALA to VAL

  - PROB, 177 ASN to THR

  - PROB, 261 LYS to TRP

       ○   PROB, 325 LEU to CYS

       ○   PROB, 326 MET to CYS

For the remainder of the page, standard settings were used

- Page 4: include ions was unchecked, based on this being the case during the hands on sessions

- Page 5: Default

- Page 6: Force field was changed to AMBER

- Page 7: Download the tgz-file

10.4.2 OpenMM simulation source code (Python)

This source code includes more analyses compared to what was discussed in this thesis. The code was executed in two separate Jupyter notebooks.

10.4.2.1 Simulation part

```
#import necessities
from sys import stdout
from simtk.openmm.app import *
from simtk.openmm import *
from simtk.unit import *
import numpy as np
import nglview
import mdtraj
import pandas
import matplotlib.pyplot as plt
#Visualization of the apelin 13 generated by CHARMM-GUI, sadly the intended mutations
(back to natural) did not come through.
viewApelin = nglview.show_file('charmm-gui-amber/5vbl_proa.pdb')
viewApelin.clear_representations()
viewApelin.add_representation(repr_type='ball+stick')
viewApelin
#Visualization of the APJ-receptor, which sadly still contains all of the mutations performed
by Ma et al.
viewReceptor = nglview.show_file('charmm-gui-amber/5vbl_prob.pdb')
viewReceptor
glview.show_file('charmm-gui-amber/amber/step3_charmm2amber.complete.pdb')
viewEurekaPDB = nglview.show_file('charmm-gui-
amber/amber/step3_charmm2amber.complete.pdb')
viewEurekaPDB.clear_representations()
viewEurekaPDB.add_representation(repr_type='cartoon')
```

```
viewEurekaPDB
eurekapdb = PDBFile('charmm-gui-amber/amber/step3_charmm2amber.complete.pdb')
forcefield = ForceField('amber14-all.xml', 'amber14/tip3pfb.xml')
system = forcefield.createSystem(eurekapdb.topology, nonbondedMethod=PME,
constraints=HBonds)
temperature = 300 * kelvin
pressure = 1 * bar
integrator = LangevinIntegrator(temperature, 1/picosecond, 2*femtoseconds)
system.addForce(MonteCarloBarostat(pressure, temperature))
simulation = Simulation(eurekapdb.topology, system, integrator)
simulation.context.setPositions(eurekapdb.positions)
state0 = simulation.context.getState(getEnergy=True)
print(state0.getPotentialEnergy())
simulation.minimizeEnergy (maxIterations=100)
state1 = simulation.context.getState(getEnergy=True)
print(state1.getPotentialEnergy())
#without this line, the temperature rose slowly starting from 0 Kelvin, this shortens the
equilibration phase.
simulation.context.setVelocitiesToTemperature(300, 1)
simulation.reporters = []
simulation.reporters.append(DCDReporter('traj.dcd', 100))
simulation.reporters.append(StateDataReporter(stdout, 100, step=True, temperature=True,
elapsedTime=True))
simulation.reporters.append(StateDataReporter("scalars.csv", 100, time=True,
potentialEnergy=True, totalEnergy=True, temperature=True))
simulation.step(50000)
del simulation
traj = mdtraj.load('traj.dcd', top='charmm-gui-
amber/amber/step3_charmm2amber.complete.pdb')
view = nglview.show_mdtraj(traj)
view.clear_representations()
view.add_representation(repr_type='cartoon')
view.add_unitcell()
view
graph = pandas.read_csv("scalars.csv")
ax = plt.gca()
graph.plot(kind='line', x='#"Time (ps)"', y='Potential Energy (kJ/mole)', ax=ax)
graph.plot(kind='line', x='#"Time (ps)"', y='Total Energy (kJ/mole)', ax=ax)
graph.plot(kind='line', x='#"Time (ps)"', y='Temperature (K)')
eurekapdb2 = PDBFile('charmm-gui-amber/amber/step3_charmm2amber.complete.pdb')
forcefield2 = ForceField('amber14-all.xml', 'amber14/tip3pfb.xml')
system2 = forcefield2.createSystem(eurekapdb2.topology, nonbondedMethod=PME,
constraints=HBonds)
temperature2 = 900 * kelvin
pressure2 = 1 * bar
integrator2 = LangevinIntegrator(temperature2, 1/picosecond, 2*femtoseconds)
system2.addForce(MonteCarloBarostat(pressure2, temperature2))
```

```
simulation2 = Simulation(eurekapdb2.topology, system2, integrator2)
simulation2.context.setPositions(eurekapdb2.positions)
state02 = simulation2.context.getState(getEnergy=True)
print(state02.getPotentialEnergy())
simulation2.minimizeEnergy (maxIterations=100)
state12 = simulation2.context.getState(getEnergy=True)
print(state12.getPotentialEnergy())
#without this line, the temperature rose slowly starting from 0 Kelvin, this shortens the
equilibration phase.
simulation2.context.setVelocitiesToTemperature(500, 1)
simulation2.reporters = []
simulation2.reporters.append(DCDReporter('traj2.dcd', 10))
simulation2.reporters.append(StateDataReporter(stdout, 10, step=True, temperature=True,
elapsedTime=True))
simulation2.reporters.append(StateDataReporter("scalars2.csv", 10, time=True,
potentialEnergy=True, totalEnergy=True, temperature=True))
simulation2.step(2000)
del simulation2
traj2 = mdtraj.load('traj2.dcd', top='charmm-gui-
amber/amber/step3_charmm2amber.complete.pdb')
view2 = nglview.show_mdtraj(traj2)
view2.clear_representations()
view2.add_representation(repr_type='cartoon')
view2.add_unitcell()
view2
```

10.4.2.2 Analysis part

```
#import necessities
from sys import stdout
from simtk.openmm.app import *
from simtk.openmm import *
from simtk.unit import *
import numpy as np
import nglview
import mdtraj
import pandas
import matplotlib.pyplot as plt
traj = mdtraj.load ('traj.dcd', top = 'charmm-gui-
amber/amber/step3_charmm2amber.complete.pdb')
print(traj.topology)
#only select residues belonging to the receptor
traj.restrict_atoms(traj.topology.select("chainid 1"))
print(traj.topology)
#only select alpha carbons
traj.restrict_atoms(traj.topology.select("name CA"))
print(traj.topology)
traj.superpose(traj,0)
```
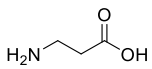
```python
rmsd = mdtraj.rmsd(traj,traj,0)
scalars = pandas.read_csv("scalars.csv")
plt.figure("RMSD")
plt.plot(scalars['#"Time (ps)"'],rmsd)
plt.xlabel("Time [ps]")
plt.ylabel("RMSD [nm]")
plt.show()
#reshape atomic positions array
print(traj.xyz.shape)
natom = traj.xyz.shape[1]
xyz = traj.xyz.reshape((-1,natom*3))
print (xyz.shape)
#calculate covariance matrix
covar = np.cov(xyz[50:],rowvar=False)
print (covar.shape)
#show matrix elements
plt.matshow(covar)
print(covar.min())
print(covar.max())
evals, evecs = np.linalg.eigh(covar)
plt.figure("eigh")
plt.semilogy(evals)
plt.xlabel("Eigenvalue index")
plt.ylabel("Eigenvalue [nm^2]")
plt.show()
plt.figure("projection")
av = xyz.mean(axis=0)
for i in range(5):
    proj = np.dot(xyz - av.reshape(1, -1), evecs[:, -i - 1])
    plt.plot(scalars['#"Time (ps)"'], proj, label="Proj. {}".format(i+1))
plt.xlabel("Time [ps]")
plt.ylabel("Displacement [nm]")
plt.legend(loc=0)
plt.show()
traj = mdtraj.load ('traj.dcd', top = 'charmm-gui-
amber/amber/step3_charmm2amber.complete.pdb')
print(traj.topology)
#only preserve alpha carbons
traj.restrict_atoms(traj.topology.select("name CA"))
print(traj.topology)
#only preserve the selected residues
traj.restrict_atoms(traj.topology.select(("residue <= 13 or residue == 20 or residue == 250 or
residue == 278")))
#distance from apelin 13 alpha carbons to Lys 250 alpha carbon
distances250 =
mdtraj.compute_distances(traj,[[0,14],[1,14],[2,14],[3,14],[4,14],[5,14],[6,14]\
                        ,[7,14],[8,14],[9,14],[10,14],[11,14],[12,14]])
```

```python
#plot for Lys250
plt.figure("250")
for counter, col in enumerate(distances250.T):
    plt.plot(scalars['#"Time (ps)"'], col, label=str(counter+1))
plt.xlabel("Time [ps]")
plt.ylabel("Distances [nm]")
plt.legend(loc=0)
plt.show()
```

11 Appendix A: peptide sequences

Specific abbreviations:

AP	=	Apelin		Fua	=	3-(2-furyl)alanine
Nle	=	Norleucine		SAP	=	Scrambled Apelin
OP	=	Opioid		D-Ala	=	D-alanine
Dmt	=	2,6-Dimethyl-L-tyrosine		β-Ala	=	
Bio	=	Glu(biotinyl-PEG)				

β-Ala structure: H_2N–CH$_2$–CH$_2$–COOH (drawn as chemical structure)

11.1 Apelin receptor ligands:

AP1(Met):   biotin-PEG4-GluArgProArgLeuSerHisLysGlyProMetProFua-OH

AP2(Met):   biotin-PEG4-GluArgProArgLeuSerHisLysGlyProMetFuaPhe-OH

AP1(Nle):   biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProFua-OH

AP2(Nle):   biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleFuaPhe-OH

AP3:   biotin-PEG4-GluArgProArgLeuSerHisLysGlyProFuaProPhe-OH

SAP1(Nle):   Biotin-PEG4-GlyProLysLeuNleArgProGluHisArgProSerFua-OH

SAP2(Nle):   Biotin-PEG4-GlyProLysLeuNleArgProGluHisArgProFuaSer-OH

SAP3(Nle):   Biotin-PEG4-GlyProLysLeuNleArgProGluHisArgFuaProSer-OH

Native AP(Nle): biotin-PEG4-GluArgProArgLeuSerHisLysGlyProNleProPhe-OH

11.2 μ-opioid receptor ligands:

OP1:   DmtD-AlaFuaPheßAlaGluBio-NH2

OP2:   DmtD-AlaPheFuaßAlaGluBio-NH2

OP3:   DmtD-AlaPheFuaßAlaBio-NH2

OP4:   DmtD-AlaFuaPheßAlaBio-NH2

OP5:   DmtD-AlaFuaßAlaBio-NH2

OP6:   DmtD-AlaPheFuaBio-NH2

12 References

[1]   S. K. Garg, A. H. Rewers, and H. K. Akturk, "Ever-Increasing Insulin-Requiring Patients Globally," *Diabetes Technol. Ther.*, vol. 20, pp. S21–S24, 2018, doi: 10.1089/dia.2018.0101.

[2]   D. Al Shaer, O. Al Musaimi, F. Albericio, and B. G. de la Torre, "2019 FDA TIDES (Peptides and oligonucleotides) harvest," *Pharmaceuticals*, vol. 13, no. 3, 2020, doi: 10.3390/ph13030040.

[3]   D. Agyei, I. Ahmed, Z. Akram, H. M. N. Iqbal, and M. K. Danquah, "Protein and Peptide Biopharmaceuticals: An Overview," *Protein Pept. Lett.*, vol. 24, no. 2, pp. 94–101, 2017, doi: 10.2174/0929866523666161222150444.

[4]   D. Goyal, S. Shuaib, S. Mann, and B. Goyal, "Rationally Designed Peptides and Peptidomimetics as Inhibitors of Amyloid-β (Aβ) Aggregation: Potential Therapeutics of Alzheimer's Disease," *ACS Comb. Sci.*, vol. 19, no. 2, pp. 55–80, 2017, doi: 10.1021/acscombsci.6b00116.

[5]   D. Van Lysebetten *et al.*, "A Thiolactone Strategy for Straightforward Synthesis of Disulfide-Linked Side-Chain-to-Tail Cyclic Peptides Featuring an N-Terminal Modification Handle," *ChemBioChem*, vol. 19, no. 6, pp. 641–646, 2018, doi: 10.1002/cbic.201700323.

[6]   S. B. Gunnoo and A. Madder, "Bioconjugation-using selective chemistry to enhance the properties of proteins and peptides as therapeutics and carriers," *Org. Biomol. Chem.*, vol. 14, no. 34, pp. 8002–8013, 2016, doi: 10.1039/c6ob00808a.

[7]   X. Xie, L. Gao, A. Y. Shull, and Y. Teng, "Stapled peptides: providing the best of both worlds in drug development," *Future Med. Chem.*, vol. 8, no. 16, pp. 1969–1980, Oct. 2016, doi: 10.4155/fmc-2016-0102.

[8]   L. Di, "Strategic Approaches to Optimizing Peptide ADME Properties," *AAPS J.*, vol. 17, no. 1, pp. 134–143, 2015, doi: 10.1208/s12248-014-9687-3.

[9]   J. L. Lau and M. K. Dunn, "Therapeutic peptides: Historical perspectives, current development trends, and future directions," *Bioorganic Med. Chem.*, vol. 26, no. 10, pp.

2700–2707, 2018, doi: 10.1016/j.bmc.2017.06.052.

[10] L. Atangcho, T. Navaratna, and G. M. Thurber, "Hitting Undruggable Targets: Viewing Stabilized Peptide Development through the Lens of Quantitative Systems Pharmacology," *Trends Biochem. Sci.*, vol. 44, no. 3, pp. 241–257, 2019, doi: 10.1016/j.tibs.2018.11.008.

[11] Y. M. Koen *et al.*, "Liver Protein Targets of Hepatotoxic 4 - Bromophenol Metabolites," 2012.

[12] J. Singh, R. C. Petter, T. A. Baillie, and A. Whitty, "The resurgence of covalent drugs," vol. 10, no. April, 2011, doi: 10.1038/nrd3410.

[13] H. M. Kumalo, S. Bhakat, and M. E. S. Soliman, "Theory and Applications of Covalent Docking in Drug Discovery: Merits and Pitfalls," no. i, pp. 1984–2000, 2015, doi: 10.3390/molecules20021984.

[14] J. M. Pereillo, M. F. Uzabiaga, J. Combalbert, C. Picard, M. Pascal, and J. M. Herbert, "Identification and Biological Activity of the Active Metabolite of Clopidogrel," no. 14, pp. 891–896, 2000.

[15] A. A. Adeniyi, R. Muthusamy, M. E. S. Soliman, A. A. Adeniyi, R. Muthusamy, and M. E. S. Soliman, "New drug design with covalent modifiers New drug design with covalent modifiers," vol. 0441, no. January, 2016, doi: 10.1517/17460441.2016.1115478.

[16] T. A. Baillie, "Targeted Covalent Inhibitors for Drug Design Angewandte Reviews," pp. 2–17, 2016, doi: 10.1002/anie.201601091.

[17] K. A. T. Makepeace *et al.*, "Improving identification of in-organello protein-protein interactions using an affinityenrichable, isotopically coded, and mass spectrometry-cleavable chemical crosslinker," *Mol. Cell. Proteomics*, vol. 19, no. 4, pp. 624–639, 2020, doi: 10.1074/mcp.RA119.001839.

[18] M. Schneider, A. Belsom, and J. Rappsilber, "Protein Tertiary Structure by Crosslinking/Mass Spectrometry," *Trends Biochem. Sci.*, vol. 43, no. 3, pp. 157–169, 2018, doi: 10.1016/j.tibs.2017.12.006.

[19] L. Seidel, B. Zarzycka, V. Katritch, and I. Coin, "Exploring Pairwise Chemical Crosslinking

To Study Peptide–Receptor Interactions," *ChemBioChem*, vol. 20, no. 5, pp. 683–692, 2019, doi: 10.1002/cbic.201800582.

[20] A. Sinz, C. Arlt, D. Chorev, and M. Sharon, "Chemical cross-linking and native mass spectrometry: A fruitful combination for structural biology," *Protein Sci.*, vol. 24, no. 8, pp. 1193–1209, 2015, doi: 10.1002/pro.2696.

[21] A. Sinz, "Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry," *Anal. Bioanal. Chem.*, vol. 397, no. 8, pp. 3433–3440, 2010, doi: 10.1007/s00216-009-3405-5.

[22] A. Leitner, M. Faini, F. Stengel, and R. Aebersold, "Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines," *Trends Biochem. Sci.*, vol. 41, no. 1, pp. 20–32, 2016, doi: 10.1016/j.tibs.2015.10.008.

[23] M. Zorn, C. H. Ihling, R. Golbik, R. G. Sawers, and A. Sinz, "Mapping cell envelope and periplasm protein interactions of Escherichia coli respiratory formate dehydrogenases by chemical cross-linking and mass spectrometry," *J. Proteome Res.*, vol. 13, no. 12, pp. 5524–5535, 2014, doi: 10.1021/pr5004906.

[24] A. Alegria-schaffer, *General Protein – Protein Cross-linking*, 1st ed., vol. 539. Elsevier Inc., 2014.

[25] P. Kleiner, W. Heydenreuter, M. Stahl, V. S. Korotkov, and S. A. Sieber, "A Whole Proteome Inventory of Background Photocrosslinker Binding Communications Angewandte," pp. 1–7, 2016, doi: 10.1002/anie.201605993.

[26] S. Halila, T. Velasco, P. De Clercq, and A. Madder, "Fine-tuning furan toxicity : fast and quantitative DNA interchain cross- link formation upon selective oxidation of a furan containing oligonucleotide," no. i, pp. 936–938, 2005, doi: 10.1039/b415092a.

[27] A. M. Jawalekar, O. De Beeck, L. Van Delft, and A. Madder, "ChemComm Synthesis and incorporation of a furan-modified adenosine building block for DNA interstrand crosslinking w," pp. 2796–2798, 2011, doi: 10.1039/c0cc04667a.

[28] M. O. De Beeck and A. Madder, "Unprecedented C-selective interstrand cross-linking

through in situ oxidation of furan-modified oligodeoxynucleotides," *J. Am. Chem. Soc.*, vol. 133, no. 4, pp. 796–807, 2011, doi: 10.1021/ja1048169.

[29]　M. Op De Beeck and A. Madder, "Sequence specific DNA cross-linking triggered by visible light," *J. Am. Chem. Soc.*, vol. 134, no. 26, pp. 10737–10740, 2012, doi: 10.1021/ja301901p.

[30]　L. L. G. Carrette, E. Gyssels, N. De Laet, and A. Madder, "Furan oxidation based cross-linking: a new approach for the study and targeting of nucleic acid and protein interactions," pp. 1539–1554, 2016, doi: 10.1039/c5cc08766j.

[31]　L. L. G. Carrette, E. Gyssels, J. Loncke, and A. Madder, "A mildly inducible and selective cross-link methodology for RNA duplexes," *Org. Biomol. Chem.*, vol. 12, no. 6, pp. 931–935, 2014, doi: 10.1039/c3ob42374c.

[32]　A. Manicardi, E. Gyssels, R. Corradini, and A. Madder, "Furan-PNA : a mildly inducible irreversible interstrand crosslinking system targeting single and double stranded DNA Electronic Supporting Information - Table of contents."

[33]　C. F. B. Pnas, J. Elskens, and A. Manicardi, "Synthesis and Improved Cross-Linking Properties of C5-Modified Furan Bearing PNAs," pp. 1–20, doi: 10.3390/molecules22112010.

[34]　E. Antonatou, Y. Verleysen, and A. Madder, "Singlet oxygen-mediated one-pot chemoselective peptide–peptide ligation," *Biomol. Chem.*, pp. 8140–8144, 2017, doi: 10.1039/c7ob02245j.

[35]　E. Antonatou, K. Hoogewijs, D. Kalaitzakis, A. Baudot, G. Vassilikogiannakis, and A. Madder, "Singlet Oxygen-Induced Furan Oxidation for Site-Specific and Chemoselective Peptide Ligation," *Chem. - A Eur. J.*, vol. 22, no. 25, pp. 8457–8461, 2016, doi: 10.1002/chem.201601113.

[36]　W. Vannecke, C. Ampe, M. Van Troys, M. Beltramo, and A. Madder, "Cross-Linking Furan-Modified Kisspeptin-10 to the KISS Receptor," *ACS Chem. Biol.*, vol. 12, no. 8, pp. 2191–2200, 2017, doi: 10.1021/acschembio.7b00396.

[37]　W. Vannecke, C. Ampe, M. Van Troys, M. Beltramo, and A. Madder, "Cross-Linking

Furan-Modified Kisspeptin-10 to the KISS Receiver," *ACS Chem. Biol.*, vol. 12, no. 8, pp. 2191–2200, 2017, doi: 10.1021/acschembio.7b00396.

[38] D. N. Langelaan, T. Reddy, A. W. Banks, G. Dellaire, D. J. Dupré, and J. K. Rainey, "Structural features of the apelin receptor N-terminal tail and first transmembrane segment implicated in ligand binding and receptor trafficking," *Biochim. Biophys. Acta - Biomembr.*, vol. 1828, no. 6, pp. 1471–1483, 2013, doi: 10.1016/j.bbamem.2013.02.005.

[39] X. Hu *et al.*, "Structural Insights into the Activation of Human Relaxin Family Peptide Receptor 1 by Small-Molecule Agonists," *Biochemistry*, vol. 55, no. 12, pp. 1772–1783, 2016, doi: 10.1021/acs.biochem.5b01195.

[40] A. S. Hauser *et al.*, "Pharmacogenomics of GPCR Drug Targets," *Cell*, vol. 172, no. 1–2, pp. 41-54.e19, 2018, doi: 10.1016/j.cell.2017.11.033.

[41] C. Holohan, S. Van Schaeybroeck, D. B. Longley, and P. G. Johnston, "Cancer drug resistance : an evolving paradigm," *Nat. Publ. Gr.*, vol. 13, pp. 714–726, 2013, doi: 10.1038/nrc3599.

[42] M. C. Fisher, N. J. Hawkins, D. Sanglard, and S. J. Gurr, "Worldwide emergence of resistance to antifungal drugs challenges humanhealth and food security," vol. 742, no. May, pp. 739–742, 2018.

[43] Y. Ma *et al.*, "Structural Basis for Apelin Control of the Human Apelin Receptor," *Structure*, vol. 25, no. 6, pp. 858-866.e4, 2017, doi: 10.1016/j.str.2017.04.008.

[44] Y. Kawamata *et al.*, "Molecular properties of apelin : tissue distribution and receptor binding," vol. 1538, pp. 162–171, 2001.

[45] S. C. Chng, L. Ho, J. Tian, and B. Reversade, "ELABELA: A hormone essential for heart development signals via the apelin receptor," *Dev. Cell*, vol. 27, no. 6, pp. 672–680, 2013, doi: 10.1016/j.devcel.2013.11.002.

[46] M. J. Kleinz and A. P. Davenport, "Emerging roles of apelin in biology and medicine," vol. 107, pp. 198–211, 2005, doi: 10.1016/j.pharmthera.2005.04.001.

[47] A. M. O'Carroll, S. J. Lolait, L. E. Harris, and G. R. Pope, "The apelin receptor APJ: Journey

from an orphan to a multifaceted regulator of homeostasis," *J. Endocrinol.*, vol. 219, no. 1, 2013, doi: 10.1530/JOE-13-0227.

[48] A. D. Corbett, G. Henderson, A. T. McKnight, and S. J. Paterson, "75 Years of opioid research: The exciting but vain quest for the Holy Grail," *Br. J. Pharmacol.*, vol. 147, no. SUPPL. 1, pp. 153–162, 2006, doi: 10.1038/sj.bjp.0706435.

[49] R. Chou *et al.*, "Clinical Guidelines for the Use of Chronic Opioid Therapy in Chronic Noncancer Pain," *J. Pain*, vol. 10, no. 2, pp. 113-130.e22, 2009, doi: 10.1016/j.jpain.2008.10.008.

[50] M. Noble *et al.*, "Long-term opioid management for chronic noncancer pain (Review)," no. 1, 2010.

[51] W. Huang *et al.*, "Structural insights into μ-opioid receptor activation," *Nature*, vol. 524, no. 7565, pp. 315–321, 2015, doi: 10.1038/nature14886.

[52] S. S. G. Ferguson, "Evolving Concepts in G Protein-Coupled Receptor Endocytosis : The Role in Receptor Desensitization and Signaling," vol. 53, no. 1, pp. 1–24, 2001.

[53] B. F. O'Dowd *et al.*, "A human gene that shows identity with the gene encoding the angiotensin receptor is located on chromosome 11," *Gene*, vol. 136, no. 1–2, pp. 355–360, 1993, doi: 10.1016/0378-1119(93)90495-O.

[54] S. M. K. McKinnie *et al.*, "The Metalloprotease Neprilysin Degrades and Inactivates Apelin Peptides," *ChemBioChem*, pp. 1495–1498, 2016, doi: 10.1002/cbic.201600244.

[55] C. A. Flanagan, *GPCR-radioligand binding assays*, vol. 132. Elsevier Ltd, 2016.

[56] S. D. Katugampola, J. J. Maguire, S. R. Matthewson, and A. P. Davenport, "[125I]-(Pyr1)Apelin-13 is a novel radioligand for localizing the APJ orphan receptor in human and rat tissues with evidence for a vasoconstrictor role in man," *Br. J. Pharmacol.*, vol. 132, no. 6, pp. 1255–1260, 2001, doi: 10.1038/sj.bjp.0703939.

[57] P. P. Roller *et al.*, "Norleucine as a replacement for methionine in phosphatase-resistant linear and cyclic peptides which bind to p85 SH2 domains," *Bioorganic Med. Chem. Lett.*, vol. 4, no. 15, pp. 1879–1882, 1994, doi: 10.1016/S0960-894X(01)80389-9.

[58] B. Janota, U. Karczmarczyk, E. Laszuk, P. Garnuszek, and R. Mikołajczak, "Oxidation of

methionine - Is it limiting the diagnostic properties of 99mTc-labeled exendin-4, a glucagon-like peptide-1 receptor agonist?," *Nucl. Med. Rev.*, vol. 19, no. 2, pp. 104–110, 2016, doi: 10.5603/NMR.2016.0021.

[59] S. Roosenburg *et al.*, "Stabilized 111In-labeled sCCK8 analogues for targeting CCK2-receptor positive tumors: Synthesis and evaluation," *Bioconjug. Chem.*, vol. 21, no. 4, pp. 663–670, 2010, doi: 10.1021/bc900465y.

[60] A. Manglik *et al.*, "Crystal structure of the μ-opioid receptor bound to a morphinan antagonist," *Nature*, vol. 485, no. 7398, pp. 321–326, 2012, doi: 10.1038/nature10954.

[61] S. Ballet *et al.*, "Design of novel neurokinin 1 receptor antagonists based on conformationally constrained aromatic amino acids and discovery of a potent chimeric opioid agonist-neurokinin 1 receptor antagonist," *J. Med. Chem.*, vol. 54, no. 7, pp. 2467–2476, 2011, doi: 10.1021/jm1016285.

[62] T. Willemse *et al.*, "Chemical space screening around Phe3 in opioid peptides: Modulating μ versus δ agonism by Suzuki-Miyaura cross-couplings," *Bioorganic Med. Chem. Lett.*, vol. 28, no. 13, pp. 2320–2323, 2018, doi: 10.1016/j.bmcl.2018.05.015.

[63] R. J. Bodnar, "Endogenous Opiates and Behavior: 2016," *Peptides*, vol. 101, no. November 2017, pp. 167–212, 2018, doi: 10.1016/j.peptides.2018.01.011.

[64] P. Melchiorri and L. Negri, "The dermorphin peptide family," *Gen. Pharmacol.*, vol. 27, no. 7, pp. 1099–1107, 1996, doi: 10.1016/0306-3623(95)02149-3.

[65] M. BROCCARDO *et al.*, "Pharmacological Data on Dermorphins, a New Class of Potent Opioid Peptides From Amphibian Skin," *Br. J. Pharmacol.*, vol. 73, no. 3, pp. 625–631, 1981, doi: 10.1111/j.1476-5381.1981.tb16797.x.

[66] S. D. Bryant, Y. Jinsmaa, S. Salvadori, Y. Okada, and L. H. Lazarus, "Dmt and opioid peptides: A potent alliance," *Biopolym. - Pept. Sci. Sect.*, vol. 71, no. 2, pp. 86–102, 2003, doi: 10.1002/bip.10399.

[67] R. B. Merrifield, "Solid Phase Peptide Synthesis. I. The Synthesis of a Tetrapeptide," *J. Am. Chem. Soc.*, vol. 85, no. 14, pp. 2149–2154, 1963, doi: 10.1021/ja00897a025.

[68] D. M. M. Jaradat, "Thirteen decades of peptide synthesis: key developments in solid

phase peptide synthesis and amide bond formation utilized in peptide ligation," *Amino Acids*, vol. 50, no. 1, pp. 39–68, 2018, doi: 10.1007/s00726-017-2516-0.

[69] K. Hoogewijs, A. Deceuninck, and A. Madder, "Aromatic capping surprisingly stabilizes furan moieties in peptides against acidic degradation," *Org. Biomol. Chem.*, vol. 10, no. 20, pp. 3999–4002, 2012, doi: 10.1039/c2ob25548k.

[70] J. M. Niers, J. W. Chen, R. Weissleder, and B. A. Tannous, "Enhanced in vivo imaging of metabolically biotinylated cell surface reporters," *Anal. Chem.*, vol. 83, no. 3, pp. 994–999, 2011, doi: 10.1021/ac102758m.

[71] E. O'Flaherty *et al.*, "BS16 A novel fluorescent apelin ligand tracks apelin receptor internalisation," *Heart*, vol. 105, no. Suppl 6, p. A151 LP-A151, May 2019, doi: 10.1136/heartjnl-2019-BCS.179.

[72] K. Gach, J. Szemraj, O. Stasikowska-Kanicka, M. Danilewicz, and A. Janecka, "Opioid-receptor gene expression and localization in cancer cells," *Cent. Eur. J. Biol.*, vol. 6, no. 1, pp. 10–15, 2011, doi: 10.2478/s11535-010-0097-y.

[73] P. R. Moody *et al.*, "Receptor Crosslinking: A General Method to Trigger Internalization and Lysosomal Targeting of Therapeutic Receptor:Ligand Complexes," *Mol. Ther.*, vol. 23, no. 12, pp. 1888–1898, 2015, doi: 10.1038/mt.2015.178.

Cross-Linking Furan-Modified Apelin-13 to the Apelin Receptor

E. Ongenae[a], L. M. Casals[a], M. Van Troys[b], C. Ampe[b], A. Madder[a]

[a] Organic and Biomimetic Chemistry Research Group, Department of Organic and Macromolecular Chemistry, Ghent University, Krijgslaan 281-S4, 9000 Ghent, Belgium
[b] Center for Medical Biotechnology, VIB-UGent, Technologiepark 71, 9052 Ghent, Belgium

Peptide-protein crosslinking is a well-established technique for studying peptide-protein interactions. Crosslinking often comes with the downside of unwanted toxicity of activating reagents (or UV-light) to trigger the crosslink formation, limiting applications. A novel crosslink strategy has been developed which is based on the incorporation of a furan moiety in the peptide of interest. The furan moiety is oxidized *in situ* to its reactive form by endogenously produced reactive oxygen species, thus bypassing the need for toxic activating reagents or radiation. This novel strategy was previously demonstrated on the Kisspeptin-10 – KISS receptor system in a proof of concept study. In this article, the applicability of the strategy is further expanded and demonstrated on the Apelin 13 – Apelin receptor interaction.

Keywords: Crosslinking, Furan chemistry, Apelin receptor (APJR)

Introduction

While peptide-protein interactions are dynamic by nature, chemically crosslinking peptides to proteins by means of forming a covalent bond allows to freeze these interactions in place. Chemical crosslinking has been heavily explored over the past decade in the fields of structural biology and proteomics[1], [2]. Typically, new information on the structure of the formed complex, binding partner identity, binding site topology and more is gathered by performing a crosslink experiment followed by enzymatic digestion and advanced mass spectrometric analysis. This strategy has successfully led to numerous new insights in the complex worlds of structural and chemical biology in the recent past and will continue to elucidate nature's biochemical cellular processes in the coming years[3].

The two most widely used classes of crosslinking reagents for the strategy described above are amine- or thiol-reactive crosslinkers (e.g. carbodiimides, reactive esters, isocyanates…) and photoreactive crosslinkers (e.g. azides, benzophenones…)[4]. While the first class relies on specific reactivity from certain amino acids (e.g. lysine, cysteine, …) present in the target protein, the latter can react with a larger variety of amino acids to form the covalent crosslink-bond upon activation of the photoreactive chemical moiety. The two can also be combined, where the introduced bifunctional molecule possesses for example an amide-sensitive group on one side of the molecule in combination with a photoreactive moiety on the other side. Alternatively, reactive esters can be used in the form of homobifunctional molecules, serving as glue to covalently bind many interacting

protein pairs at once (often called 'shotgun approach'). Photoreactive crosslinkers on the other hand can be incorporated by means of an unnatural amino acid in the peptide or protein sequence in order to enable a more selective crosslinking experiment. This unnatural amino acid can be introduced either during chemical peptide synthesis, or through genetic incorporation. In order to extract accurate information from crosslinking experiments, it is of high importance that the crosslinking event occurs in a natural cellular environment, as altered conditions like addition of certain chemical reagents or UV-light can have detrimental effects on cells and can even yield false positive results[5].

Furan-based crosslinking

Over the past ten years, the Organic and Biomimetic Chemistry Research group (OBCR) at the Ghent University has built up considerable expertise in furan-based modification of biomolecules and subsequent interstrand crosslinking of DNA[6]–[9], RNA[10] and PNA[11], [12] (peptide nucleic acid), peptide-peptide ligations[13], [14] or ligand-receptor crosslinking[15]. The reactivity of the otherwise stable furan moiety is based upon its oxidation to the corresponding keto-enal, which is receptive for nucleophilic attack by proximal amine or sulfhydryl groups present in the binding partner (Figure 1). Furan oxidation can be achieved by selective N-bromosuccinimide (NBS) oxidation[6] or by singlet oxygen produced by photosensitizers when irradiated with a visible light source[9]. Alternatively, oxidation can occur spontaneously in living cells by endogenously produced reactive oxygen species (ROS)[15]
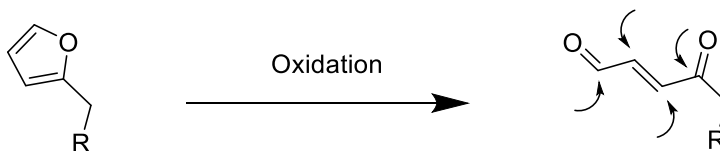
Figure 1: Furan oxidation to keto-enal, arrows indicate possible nucleophilic attack sites.

In the field of structural biology and proteomics, this possibility for endogenous activation through oxidation in cells has substantial benefits over the widely used alternatives. Experiments can be performed with complete omission of the formerly required addition of chemicals or administration of UV-light, which are known to have an influence on the cellular environment and thus interfere with the results of experiments[5]. In addition, this discovery even offers the possibility to further expand the applicability of furan-based crosslinking beyond fundamental research in structural biology and proteomics. The furan moiety is intrinsically stable prior to oxidation and is easily incorporable in peptide sequences through the commercially available Fmoc protected 3-(2-furyl)alanine, which is suitable for solid phase peptide synthesis (SPPS). These two attractive characteristics of furan-based crosslinking could potentially pave the road for novel pharmaceutical applications.

<u>Apelin 13 – Apelin receptor</u>

The work describe herein further expands our earlier proof of concept study[15] of the furan-based peptide-protein crosslinking triggered by endogenous reactive oxygen species to the Apelin 13 – Apelin receptor system. The Apelin receptor, also known as the APJR receptor, is a G protein-coupled receptor (GPCR) belonging to class A, also known as the rhodopsin-like GPCRs. The Apelin receptor is consistently expressed in human cells, most notably in the central nervous system, lung and mammary gland[16]. Several natural Apelin peptide ligands are known: Apelin 36, Apelin 17 and Apelin 13 being the shortest (the number index indicates the number of amino acids in the peptide ligand, for complete structures *vide infra* section **Error! Reference source not found.**). On top of this, Elabela, a second type of endogenous peptide ligand of the Apelin receptor has recently been discovered[17]. The Apelin receptor has been associated with cardiovascular regulations and fluid homeostasis. It also plays a role in the inhibition of HIV infection and has an influence on the glucose and energy metabolism, making it an interesting drug target. The different Apelin peptide ligands have non-identical potencies in regulating the above described functions, the longest ligand (Apelin 36) plays a more effective role in the HIV infection inhibition, while the shortest peptide (Apelin 13) has a more pronounced effect on the cardiovascular system[18], [19].

In preliminary studies, it was tested at which positions the 3-(2-furyl)alanine could be inserted in the Apelin 13 peptide, without losing the interaction with the Apelin receptor. In addition, it was found that a C-terminal carboxylic acid functionality allowed a better receptor interaction than a C-terminal amide. Based on these results, three different furan-modified Apelin 13 peptides were synthesized and tested: AP1(Nle), AP2(Nle) and AP3. Each of these peptides was coupled to biotin-PEG4-propionic acid at the N-terminus and contained the natural carboxylic acid functionality at the C-terminus. The number stands for the position of the 3-(2-furyl)alanine (counting from the C-terminus). To prevent oxidation of the naturally occurring methionine at position 3 (counting from the C-terminus), this residue was replaced by the isosteric norleucine, an often performed substitution which is known to preserve bio-activity[20]–[22].

Experimental

<u>Synthesis of furan-modified Apelin 13 analogues</u>

For the synthesis of the furan modified Apelin 13 peptides, SPPS was used. Fmoc-protected amino acids used for the peptide synthesis were obtained from Iris Biotech GmbH. Amino acids with sensitive side-chains were purchased side-chain protected: Fmoc-Lys(Boc)-OH, Fmoc-His(Trt)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Arg(Pbf)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Glu(tBu)-OH. All peptides were synthesized using the fully automated SYRO Multiple Peptide Synthesizer robot, equipped with a shaker (vortex) for the reactor block accommodating a maximum of 24 reactors (polypropylene syringe tubes with a polyethylene filter at the bottom). Usually, synthesis with double coupling steps was performed as following: a mixture of 5 equivalents (eq.) amino acid in DMF (0.5 M), 5 eq. HBTU in DMF (0.5 M) and 10 eq. DIPEA in NMP (2 M) is added to the resin, with subsequent reaction for 40 minutes at room temperature. The Fmoc-3-(2-furyl)-L-alanine was incorporated as a standard amino acid. 2-Chlorotrityl chloride polystyrene resin (1% DVB) was used, synthesis was performed at 0.12 mmol scale. The

peptides were cleaved using a cleavage cocktail consisting of 95% TFA, 2.5% TIS and 2.5% thioanisole, over a period of 2 hours. MTBE precipitation was performed prior to purification. Purification was performed using a preparative RP-HPLC machine equipped with a UV-detector measuring absorbance at 214 nm (EOS, maximum loading of 200 mg, loop size 20 mL). The gradient for the purifications was composed of MQ water containing 0.1% TFA (A) and ACN containing 0.1% TFA (B): 0% to 70% B in 35 minutes. Purity determination and validation were performed using RP-HPLC (Agilent 1100 series HPLC system equipped with a Phenomenex Luna 5u C18 column (250 mm x 4.6 mm, 5 µm)) and LCMS (Agilent 1100 with diode array detector (set to 214, 254, 280, 310, 360 nm), equipped with a Phenomenex Kinetex C18 100Å (150 mm x 4.6 mm, 5 µm, 35 °C) coupled to an Agilent ESI-single quadrupole MS detector type VL, mass detection was performed in positive mode).

Crosslinking experiments

The cells used for the crosslinking experiments belong to the MDA-MB-231 cell line. All cells were grown in Dulbecco Modified Eagle's Medium with 10% Fetal Bovine Serum (FBS, Gibco, Ref 10270-106) and 100 µg/mL of penicillin/streptomycin (Ref 15140-122, Gibco, Invitrogen) at 37 °C in a humidified atmosphere of 5% CO_2/95% air to maintain the pH of the medium. The medium was renewed three times per week. Since the cells used for this thesis show high proliferating rates, the cultures were subcultured at least twice a week. For the washing steps Dulbecco's Phosphate Buffered Saline (DPBS) was used. For the peptide incubation experiments Dulbecco's Modified Eagle Medium (DMEM) containing 4.5 g/L D-Glucose and no pyruvate was used to prepare the peptide solutions. Peptides were added to this medium from a stock solution of 25 mM in DMSO to reach desired peptide concentration. It is very important that no FBS was added to the medium used for the incubation experiments.

Western Blot

The gels used for the SDS-PAGE were prepared manually, by combining 8 mL of 10% separation gel with 3 mL of 5% stacking gel. The gels were loaded with 30 µL of the protein samples in every lane. Two lanes in each gel were loaded with 4 µL of protein marker: Precision Plus Protein™ All Blue Standards. Blotting was performed to a PVDF Immobilon membrane (Millipore), transfer equipment used was Mini Protean TransBlot Cell (BioRad). Each membrane was blocked using a 1/1 solution of PBS/blocking solution (Odyssey, Lyeor) for 1 hour. Incubation with the primary antibody (Apelin Receptor Antibody (5H CLC), Abfinity™ Rabbit) was performed on the tilt shaker overnight. After this, the membrane gets washed three times with PBS containing 0.1% TWEEN, before final incubation with Goat Anti-Rabbit IgG H&L (HRP) (Abcam, ab205718) and IRDye 680RD Streptavidin (Li-COR) for 45 minutes. Final washing is performed: PBS containing 1% SDS (3x), regular PBS (2x) and a final washing step with water. The membranes were scanned using an Odyssey Infrared Imaging System at 685 nm and 800 nm.

Results and discussion

Peptide synthesis

Commercially available Fmoc protected 3-(2-furyl)alanine (= Fua) was included seamlessly in the automated SPPS procedures. The three envisaged peptides were obtained with moderate yields, displayed in Table 1. When the furan is inserted at the first position counting from the C-terminus, the yield is substantially lower. This can be attributed to the furan in this position being more exposed and thus more sensitive to the acidic cleavage conditions, causing more degradation. The coupling of the biotin-PEG-propionic acid to each N-terminal amine group was successful. This was performed to allow easy visualization using fluorescent streptavidine in the following western blot experiments.

TABLE I. Furan-modified Apelin 13 peptide synthesis

Peptide	Sequence (Fua = X, X' = Nle)	Yield
AP1(Nle)	biotin-PEG4-QRPRLSHKGPX'PX-OH	7.6%
AP2(Nle)	biotin-PEG4-QRPRLSHKGPX'XF-OH	28.1%
AP3	biotin-PEG4-QRPRLSHKGPXPF-OH	15.4%

Crosslinking experiments

The crosslinking experiments encompass two distinct phases. The first phase covers the treatment of the living cells with the synthesized furan-modified peptides at different concentrations during different incubation periods. During this phase, the endogenously activated crosslinking event takes place. The second phase consists of a Western blot experiment in order to examine if crosslinking of the furan-modified peptide to the GPCRs has occurred. In order to perform the Western blot, the cells need to be lysed, which is achieved by addition of lysis buffer to the cells. Western blotting covers a series of sample processing steps, starting with gel electrophoresis to separate the different lysate components by size. This is followed by a blotting step (transfer of the proteins from the gel to a membrane) and finally the separated compounds are visualized by means of dual immunodetection.

   Living cells treatment. The human cells used to perform the crosslinking experiments originate from the MDA-MB-231 cell line, a late-stage breast cancer cell line. The cells were seeded on six-well plates (825.000 cells per well) before treatment with the furan-modified Apelin peptides the next day, when the cells were about 80% confluent. The cells were treated with three different furan-modified Apelin peptides (AP1(Nle), AP2(Nle) and AP3) at different concentrations (25, 50 or 100µM) during different periods of time (30 minutes, 1 hour, 2 hours). A control sample was included in the crosslinking experiment: cells were treated with cell medium containing 100µM DMSO (without addition of peptide) at the same conditions.

After incubation of the cells with the furan-modified peptides for a desired period of time, the cells were lysed to obtain a lysate solution that contains all proteins, including the crosslinked ligand-receptor complex. Upon addition of the lysis buffer, which contains a protease-inhibitor, the cell lysate is meticulously kept on ice to further reduce protease activity. After careful homogenization of the cell lysate solution by forcing it through a thin needle for ten times and subsequent centrifugation, the obtained supernatant no

longer contains any cell debris and is a clear solution. To this solution, Laemmli sample buffer 5X is added which denatures the proteins and gives them a uniform negative charge distribution. Bromophenol blue present in the Laemmli sample buffer serves as a tracking dye. Before proceeding to the next step, the samples were heat-treated at 95 degrees Celsius for 5 minutes.

Western blot. A dual immunodetection approach is used, where two different antibodies having fluorescent properties at different wavelengths are utilized. For the performed experiment with the Apelin peptides, the aim was to visualize the Apelin receptor using a primary Apelin receptor antibody, amplified by a secondary antibody which features green fluorescence. On top of that, biotin was visualized using red fluorescent streptavidin (Figure 2). This dual immunodetection should allow for easy identification of the crosslinked ligand-protein complex. The Apelin receptor itself should be clearly visible as a green fluorescent signal. If crosslinking occurred, a red signal should be observable in close proximity on the membrane. From the crosslinking experiment performed with three Apelin peptides, promising results were obtained. The results are based upon the scan of the three membranes obtained after blotting and antibody incubation. Each membrane contains experiments with only one single type of Apelin peptide. The three membrane scans are displayed in Figure 3.

It can be noted that every lane contains at least three intense red signals (±250 kDa, ±130 kDa and ±75 kDa). Human cells are known to contain multiple endogenously biotinylated proteins, which explains these red signals: coA

Figure 2: Western Blot for three different administered apelin peptides at different concentrations (100, 50, or 25 µM) and different incubation times (2 hours, 1 hour or 30 minutes). Observed crosslinking signal is indicated with a yellow ellipsoid.

carboxylase at 220 kDa, pyruvate carboxylase at 130 kDa, 3-methylcrotonyl coA carboxylase at 75 kDa and propionyl coA carboxylase at 72 kDa[23]. These last two give rise to a single signal due to overlap, as the resolution of the performed SDS-PAGE is limited. Around 50 kDa, one or two different signals can be observed: a green signal and a red signal. As the Apelin receptor has a mass of 43 kDa and the furan-modified Apelin peptides have a mass around 2 kDa, the crosslinked Apelin ligand-receptor complex
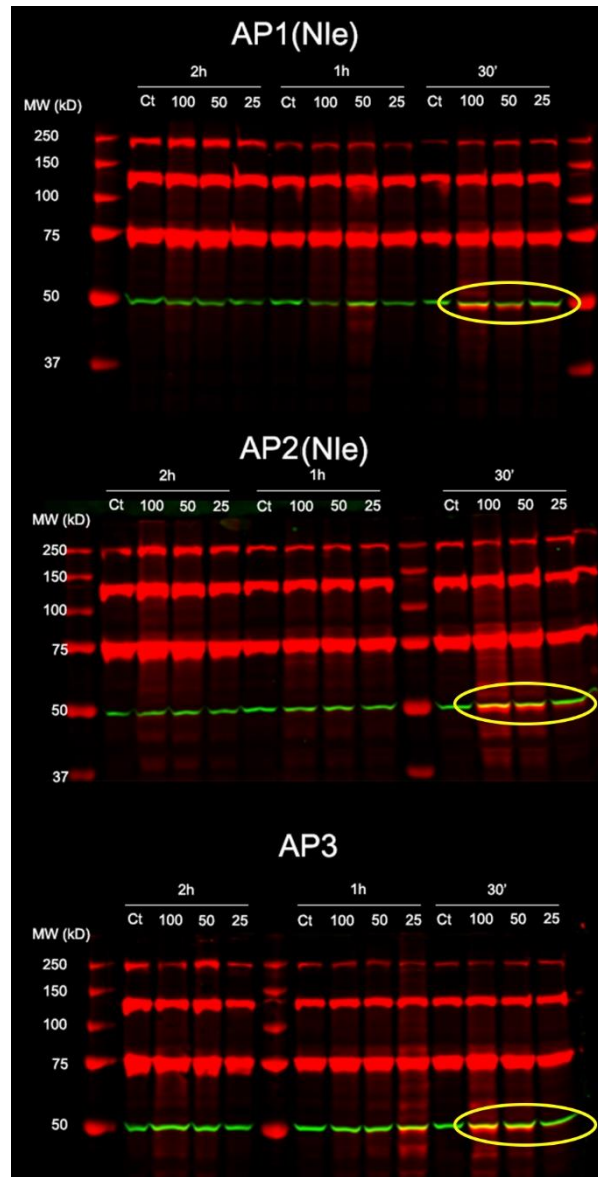
should have a mass of 45 kDa. This means that the signals around 50 kDa in the western blot carry the most valuable information of the crosslinking experiment. As the Apelin receptor is not reported to be biotinylated, these red signals should originate from the crosslinked ligand-receptor complex, as a biotin moiety was attached to the furan-modified Apelin peptides for easy detection in western blot experiments. No red signal around 50 kDa is observed in the control experiment, where no furan-modified peptide was added to the cells. This also strengthens the conclusion of the red signal around 50 kDa representing the Apelin receptor crosslinked to the synthesized furan-modified Apelin peptide.

As a conclusion, the spontaneous oxidation of the furan moiety and subsequent crosslink reaction between the furan-modified Apelin 13 analogue and the Apelin receptor in living cells appears to work. Using the described methods, it is possible to crosslink the synthesized furan-modified peptide ligand to its receptor selectively and subsequently visualize the crosslinked product via western blotting. The lack of crosslinked product signal in the western blot when the cells were incubated with the synthesized peptides for one hour or longer can potentially be attributed to internalization of the receptor over time after the occurrence of the crosslinking event. When an incubation time of 30 minutes is utilized, the crosslinking signal is most pronounced. Further investigation of all factors influencing the crosslink event are ongoing in order to fully understand and further develop the furan crosslinking technology on living cells.

Summary

A series of furan-modified Apelin 13 analogues was synthesized and crosslinking of the synthesized furan-modified peptides to their corresponding GPCR triggered after endogenous furan oxidation, was observed in experiments on living human breast-cancer cells (MDA-MB-231). Upon treatment of the cells with the furan-modified peptides for 30 minutes at 37 °C in medium containing no FBS, the crosslinked ligand-receptor complex was detected by Western blot. The applicability of the spontaneous oxidation and subsequent covalent crosslinking of a furan-containing peptide to its receptor when added to a cell culture, has thus successfully been expanded to the Apelin 13 – Apelin receptor system.

References

[1] L. Seidel, B. Zarzycka, V. Katritch, and I. Coin, "Exploring Pairwise Chemical Crosslinking To Study Peptide–Receptor Interactions," *ChemBioChem*, vol. 20, no. 5, pp. 683–692, 2019, doi: 10.1002/cbic.201800582.
[2] M. Schneider, A. Belsom, and J. Rappsilber, "Protein Tertiary Structure by Crosslinking/Mass Spectrometry," *Trends Biochem. Sci.*, vol. 43, no. 3, pp. 157–169, 2018, doi: 10.1016/j.tibs.2017.12.006.
[3] A. Sinz, "Investigation of protein-protein interactions in living cells by chemical crosslinking and mass spectrometry," *Anal. Bioanal. Chem.*, vol. 397, no. 8, pp. 3433–3440, 2010, doi: 10.1007/s00216-009-3405-5.
[4] A. Leitner, M. Faini, F. Stengel, and R. Aebersold, "Crosslinking and Mass Spectrometry: An Integrated Technology to Understand the Structure and Function of Molecular Machines," *Trends Biochem. Sci.*, vol. 41, no. 1, pp. 20–32, 2016, doi: 10.1016/j.tibs.2015.10.008.

[5]   P. Kleiner, W. Heydenreuter, M. Stahl, V. S. Korotkov, and S. A. Sieber, "A Whole Proteome Inventory of Background Photocrosslinker Binding Communications Angewandte," pp. 1–7, 2016, doi: 10.1002/anie.201605993.

[6]   S. Halila, T. Velasco, P. De Clercq, and A. Madder, "Fine-tuning furan toxicity : fast and quantitative DNA interchain cross- link formation upon selective oxidation of a furan containing oligonucleotide," no. i, pp. 936–938, 2005, doi: 10.1039/b415092a.

[7]   A. M. Jawalekar, O. De Beeck, L. Van Delft, and A. Madder, "ChemComm Synthesis and incorporation of a furan-modified adenosine building block for DNA interstrand crosslinking w," pp. 2796–2798, 2011, doi: 10.1039/c0cc04667a.

[8]   M. O. De Beeck and A. Madder, "Unprecedented C-selective interstrand cross-linking through in situ oxidation of furan-modified oligodeoxynucleotides," *J. Am. Chem. Soc.*, vol. 133, no. 4, pp. 796–807, 2011, doi: 10.1021/ja1048169.

[9]   M. Op De Beeck and A. Madder, "Sequence specific DNA cross-linking triggered by visible light," *J. Am. Chem. Soc.*, vol. 134, no. 26, pp. 10737–10740, 2012, doi: 10.1021/ja301901p.

[10]   L. L. G. Carrette, E. Gyssels, J. Loncke, and A. Madder, "A mildly inducible and selective cross-link methodology for RNA duplexes," *Org. Biomol. Chem.*, vol. 12, no. 6, pp. 931–935, 2014, doi: 10.1039/c3ob42374c.

[11]   A. Manicardi, E. Gyssels, R. Corradini, and A. Madder, "Furan-PNA : a mildly inducible irreversible interstrand crosslinking system targeting single and double stranded DNA Electronic Supporting Information - Table of contents."

[12]   C. F. B. Pnas, J. Elskens, and A. Manicardi, "Synthesis and Improved Cross-Linking Properties of C5-Modified Furan Bearing PNAs," pp. 1–20, doi: 10.3390/molecules22112010.

[13]   E. Antonatou, K. Hoogewijs, D. Kalaitzakis, A. Baudot, G. Vassilikogiannakis, and A. Madder, "Singlet Oxygen-Induced Furan Oxidation for Site-Specific and Chemoselective Peptide Ligation," *Chem. - A Eur. J.*, vol. 22, no. 25, pp. 8457–8461, 2016, doi: 10.1002/chem.201601113.

[14]   E. Antonatou, Y. Verleysen, and A. Madder, "Singlet oxygen-mediated one-pot chemoselective peptide–peptide ligation," *Biomol. Chem.*, pp. 8140–8144, 2017, doi: 10.1039/c7ob02245j.

[15]   W. Vannecke, C. Ampe, M. Van Troys, M. Beltramo, and A. Madder, "Cross-Linking Furan-Modified Kisspeptin-10 to the KISS Receptor," *ACS Chem. Biol.*, vol. 12, no. 8, pp. 2191–2200, 2017, doi: 10.1021/acschembio.7b00396.

[16]   Y. Kawamata *et al.*, "Molecular properties of apelin : tissue distribution and receptor binding," vol. 1538, pp. 162–171, 2001.

[17]   S. C. Chng, L. Ho, J. Tian, and B. Reversade, "ELABELA: A hormone essential for heart development signals via the apelin receptor," *Dev. Cell*, vol. 27, no. 6, pp. 672–680, 2013, doi: 10.1016/j.devcel.2013.11.002.

[18]   M. J. Kleinz and A. P. Davenport, "Emerging roles of apelin in biology and medicine," vol. 107, pp. 198–211, 2005, doi: 10.1016/j.pharmthera.2005.04.001.

[19]   A. M. O'Carroll, S. J. Lolait, L. E. Harris, and G. R. Pope, "The apelin receptor APJ: Journey from an orphan to a multifaceted regulator of homeostasis," *J. Endocrinol.*, vol. 219, no. 1, 2013, doi: 10.1530/JOE-13-0227.

[20]   P. P. Roller *et al.*, "Norleucine as a replacement for methionine in phosphatase-resistant linear and cyclic peptides which bind to p85 SH2 domains," *Bioorganic Med. Chem. Lett.*, vol. 4, no. 15, pp. 1879–1882, 1994, doi: 10.1016/S0960-894X(01)80389-9.

[21] B. Janota, U. Karczmarczyk, E. Laszuk, P. Garnuszek, and R. Mikołajczak, "Oxidation of methionine - Is it limiting the diagnostic properties of 99mTc-labeled exendin-4, a glucagon-like peptide-1 receptor agonist?," *Nucl. Med. Rev.*, vol. 19, no. 2, pp. 104–110, 2016, doi: 10.5603/NMR.2016.0021.

[22] S. Roosenburg *et al.*, "Stabilized 111In-labeled sCCK8 analogues for targeting CCK2-receptor positive tumors: Synthesis and evaluation," *Bioconjug. Chem.*, vol. 21, no. 4, pp. 663–670, 2010, doi: 10.1021/bc900465y.

[23] J. M. Niers, J. W. Chen, R. Weissleder, and B. A. Tannous, "Enhanced in vivo imaging of metabolically biotinylated cell surface reporters," *Anal. Chem.*, vol. 83, no. 3, pp. 994–999, 2011, doi: 10.1021/ac102758m.