UNIVERSITAT ROVIRA i VIRGILI

Facultat de Química Tarragona

THE UNIVERSITY OF ARIZONA

COLLEGE OF SCIENCE
COLLEGE OF MEDICINE TUCSON
Chemistry & Biochemistry

BACHELOR'S THESIS

CONFORMATIONAL VARIATION AND SELECTIVITY OF GLYCOSYLATED PACAP$_{23}$ REPLACING THE FOURTH RESIDUE BY BETA-TURN OR ALPHA-HELIX INDUCERS

Student: Thiago Gilmartin
Academic Tutor: Mª Elena Fernández Gutiérrez
Professional Tutor: Robin Polt
Bachelor's Degree in Chemistry
Year: 2020-2021

TABLE OF CONTENTS

1. SUMMARY

In this project I will discuss the variation in the conformation and affinity of the glycosylated neurohormone $PACAP_{23}$ in function of the substitution of the fourth amino acid of the peptide chain by Glycine, Alanine, or Sarcosine; and their affinity towards the G protein-coupled receptors PAC_1, $VPAC_1$, and $VPAC_2$. This substitution may cause the variation of the three-dimensional structure of the peptide, producing ß-turn or α-helix conformations. It has been demonstrated through several studies of $PACAP_{27}$ and $PACAP_{38}$ that ß-turns favors the affinity to PAC_1 receptors, while α-helix conformations will rather interact with $VPAC_1$ and $VPAC_2$ receptors. This study has been centered on PAC_1 affinity and activation, due to its neuroprotective and neuro-repairing activity.

The production and purification of three different glycosylated $PACAP_{23}$ batches have been done using Solid-Phase Peptide Synthesis (SPPS) and preparative HPLC purification, at Dr. Robin Polt's laboratory group at the University of Arizona. Their characterization has been determined by analytical HPLC and MS.

En aquest projecte discutiré la variació de la conformació i afinitat de la neurohormona glicosilada $PACAP_{23}$ en funció de la substitució del quart aminoàcid de la cadena peptídica per Glicina, Alanina o Sarcosina; i la seva afinitat amb els receptors acoblats a proteïnes G: PAC_1, $VPAC_1$, i $VPAC_2$. Aquesta substitució pot causar la variació de l'estructura tridimensional de la proteïna, produint conformacions de gir beta o d'hèlix alfa. S'ha demostrat a través de diversos estudis de $PACAP_{27}$ i $PACAP_{38}$ que els girs beta afavoreixen l'afinitat als receptors PAC_1, mentre que les conformacions d'hèlixs alfa més aviat interaccionen amb els receptors $VPAC_1$ i $VPAC_2$. Aquest estudi s'ha centrat en l'activació i l'afinitat per PAC_1, a causa de la seva activitat neuroprotectora i neuro-reparadora.

La producció i purificació dels tres lots de $PACAP_{23}$ glicosilats diferents s'han fet utilitzant la síntesi de pèptids de fase sòlida (SPPS) i la purificació preparativa HPLC, en el grup de laboratori del Dr. Robin Polt a la Universitat d'Arizona. La seva caracterització ha sigut feta a través de HPLC i MS.

2. OBJECTIVE

The main objective of this thesis is to compare the variation of the conformation and affinity of the glycosylated neurohormone $PACAP_{23}$ by substituting the fourth amino acid of the chain by three different residues: Glycine, Alanine and Sarcosine. To do so, three different $PACAP_{23}$ batches have been synthesized by using a technique called Solid Phase Peptide Synthesis, which simplifies the synthetic procedure automatizing and monitoring a great part of the process, together with purification systems like Preparative HPLC, and analytical methods such as HPLC and MS. In addition, being able to discuss results cooperatively in the research group, being able to access and operate in large facilities at the University of Arizona such as the NMR room for sample loading, and delving into complex bibliography, are also some of the objectives that I have been able to carry out.

Previous studies have been done with $PACAP_{38}$ and $PACAP_{27}$, however, the use of $PACAP_{23}$ would carry a huge simplification of the manufacture of the drug, being able to dispense with several amino acids of the chain.

3. THEORETICAL BACKGROUND

The development of neurodegenerative disorders as one gets older becomes more common, giving rise to the appearance of diseases such as Parkinson's Disease or Alzheimer's Disease. For this reason, the research of a therapy that treats not only the symptoms but also presents the capacity of slowing down the rate of neurological decline, or even prolongs the time to the appearance of clinical these symptoms, would produce a massive change in life quality of any patient.

3.1. PACAP: A NOVEL NEUROPROTECTIVE DRUG

The *pituitary adenylate cyclase-activating polypeptide*, also known as PACAP, is a 38-amino acid peptide originally isolated from the ovine hypothalamus, and encoded by the gene ADCYAP1, that presents the capacity of stimulating the production of cyclic adenosine monophosphate (cAMP) in pituitary cells from rats [1][2]. Furthermore, in the central nervous system (CNS), this protein acts as a neurotransmitter, neurohormone, also as a neurotrophic factor and a neuroprotective agent due to its strong antiapoptotic activity. Because of its behavior in neurodevelopment and neuroprotection, PACAP is considered a powerful drug candidate for the treatment of neurodegenerative diseases like Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis (ALS), and Huntington's disease; slowing down the neurological decline. Moreover, this neurohormone has demonstrated to present a great number of biological effects such as vasodilatation, bronchodilation, and control of tumor cell proliferation [3].

Although the first isolated PACAP had been its 38 residues configuration, a shorter structure can also be found in the form of 27 amino acids. This can be explained due to a possible amidation that produces an internal cleavage site in residues Gly^{28}-Lys^{29}-Arg^{30} [2]. This C-terminally truncated 27-amino acid peptide is also biologically active, even though it is present in lower concentrations in the central nervous system [4].

Both $PACAP_{38}$ and $PACAP_{27}$ are members of the vasoactive intestinal peptide-like peptide family, belonging to the same group of peptide hormones as the vasoactive intestinal peptide (VIP), glucagon, glucagon-like peptides such as GLP-1 and GLP-2, secretin, gastric inhibitory peptide, and growth hormone-releasing hormone (GHRH). In the case of $PACAP_{27}$, it shares 68% of the

sequence with VIP, having a highly similar structure. The structural variation in the amino acid chain of this superfamily of peptide hormones can be observed in Figure 2 [1].

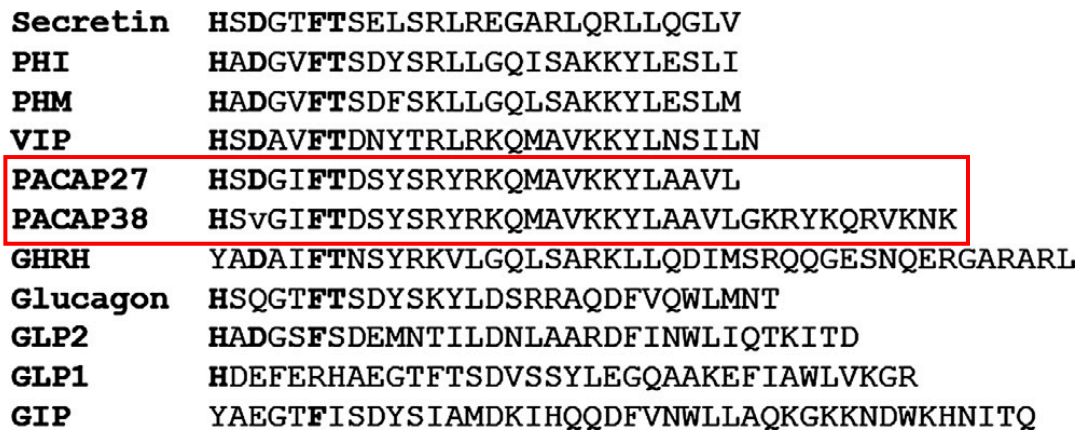


Figure 2. Sequence of amino acids of PACAP/VIP neuropeptide family.

It has been seen that the variation of the amino acids' composition among this family of neurohormones is different if we look at the chain from the N-terminal or the C-terminal domain. On one hand, it has been observed that the N-terminal domain of most of these peptides is conserved, changing only a few numbers of residues but maintaining the main amino acid skeleton. This might explain why this domain is extremely crucial, chemically, and structurally, for the biological activity of these neurohormones. On the other hand, the C-terminal domain changes significantly. For instance, it has been demonstrated that the segments from 28 to 38 of PACAP are not essential for the biological activity of this peptide, giving a reasonable hypothesis for the existence of $PACAP_{27}$ [5][6]. Moreover, speaking in terms of $PACAP_{38}$ and $PACAP_{27}$ comparison, the 38-residues form is prone to be degraded rapidly by the enzyme dipeptidyl peptidase IV (DPP IV), which releases dipeptides from the N-terminus, decreasing the efficiency of the PACAP of being attached to the receptor. The activity of this enzyme presents an enormous drawback in the use of $PACAP_{38}$ for the development of pharmacological agents because N-terminally truncated peptides are released with a receptor antagonist behavior. Although it is far from the bonds to be degraded, this enzyme interacts deeply with the 28 to 38 segments of the C-terminal, being these segments critical for the binding of the enzyme to the protein. For this reason, it mostly affects $PACAP_{38}$ and not $PACAP_{27}$, as the lack of these eleven amino acids in the 27-residues form decreases the probabilities of being degraded, increasing so the efficiency towards the receptors [7].

In addition to these two analogs, the synthesis of an even shorter version of this neurohormone has been a point of study for years, implying a significant simplification in the synthesis of therapeutic drugs. For this reason, several investigations are focusing on the study and development of $PACAP_{23}$. It has been demonstrated that $PACAP_{23}$ presents a potent neuroprotective effect in Parkinson's disease models in addition to present high affinity and selectivity towards PACAP receptors. Although the 23-mer showed that this protecting action over neurons is not identical to those mechanisms produced by the native PACAP isoforms, it is similar enough to induce neuron cell survival [8].

3.2. RECEPTOR SELECTIVITY: PAC_1, $VPAC_1$, AND $VPAC_2$

G protein-coupled receptors (GPCRs) are the most important category of membrane proteins within the human genome, being the largest class of receptors. This type of receptor presents seven-transmembrane helical domains, in which three of them are extracellular and three of them are located in the intracellular domain. What is more, all of them have a similar design, with the N-terminus placed in the extracellular domain, and the C-terminus in the intracellular domain. There are more than eight hundred of this kind of receptors in the human organism, and they can be categorized into five different classes: rhodopsin-like family (class A), secretin receptor family (also called class B), the glutamate family (class C), adhesion family, and frizzled family. [9]

PACAP affinity and selectivity towards receptors has been a subject of study for years, leading to the discovery of three different types of class B G-protein coupled receptors (B GPCR) called PAC_1, $VPAC_1$, and $VPAC_2$. So, focusing on the secretin family of peptides, it is important to note that the N-terminal domain is able to participate in the binding of the ligand towards the receptor, while the C-terminal domain, as it is common for all GPCRs, is placed inside the cell. Although the secretin family of receptors is not a very large class, made up of only twenty receptors that are activated by neurohormones, the functions they carry out are crucial for homeostasis, as they are responsible for the control and regulation of vital operations in the organism, such as brain, heart, and lungs functionality.

The three receptors mentioned above have a high affinity towards PACAP, however, PAC_1 is the one that presents the highest affinity for this peptide hormone, when $VPAC_1$ and $VPAC_2$ bind to both PACAP and VIP with great

similarity [1]. It is important to mention that the three of them produce different biological responses. In the case of PAC_1, it is responsible for the neuroprotective and neuro-repairing response, whereas $VPAC_1$ and $VPAC_2$ are in control of the anti-inflammatory response.

These B GPCRs are distributed along the body within the peripheral organs and the central nervous system, giving as a response of PACAP activation the release of hormones from the pituitary, gonads, hypothalamus, pancreas, and adrenal glands. As it has been said previously, the activation of PAC_1, $VPAC_1$, and $VPAC_2$ receptors in the central nervous system is closely related to neuronal differentiation and neuroprotection, being the reason why PACAP is a prominent therapy to treat central and peripheral neurodegenerative disorders like Parkinson's and Alzheimer's diseases, and ALS respectively. Moreover, in the case of PAC_1 receptors, its activation with exogenous PACAP has shown significant results in the treatment of traumatic brain injuries in rats [10]. What is more, these receptors are distributed not only along the CNS, but also through the respiratory, cardiovascular, hepatic, and renal tissue, together with the gastrointestinal tract and gonads. However, activation of $VPAC_1$ and $VPAC_2$ causes other effects like vasodilatation or water retention.

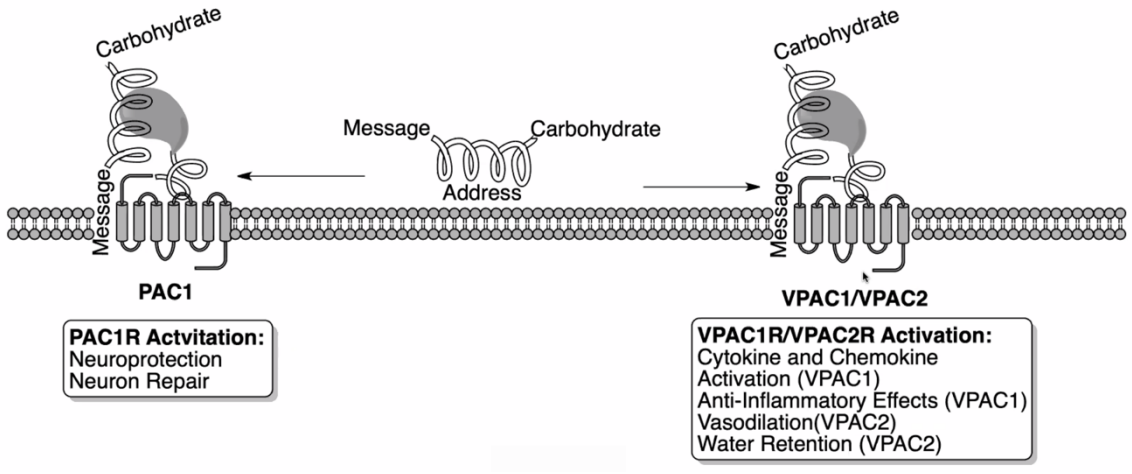


Figure 2. Comparison of PAC_1 and $VPAC_1$/$VPAC_2$ receptors activation.

Concerning the structural information of these family of receptors, although there is not a completed crystal structure that defines the full length of the receptors, there is in fact, structural information in terms of NMR and crystal structure for the N-terminal extracellular domain for some of the members of this class, including PAC_1. This information is considered very significant, as the N-

terminal part located at the extracellular domain of the receptor is the one that participates in ligand binding and receptor activation [11].

It is important to note that all the peptides related to B GPCRs' activation form unstructured random coils when they are in an aqueous system. Nevertheless, when these peptides associate with the membrane, the C-terminus is induced to form an α-helix, while the N-terminus remains disordered. The reason for this relies upon reducing their entropy when they bind to the receptor. This helix formation is very usual in many types of peptides, especially when they bind to a membrane because, in the function of the composition of the amino acid chain, they are able to arrange in an amphipathic helix. The amphipathicity of the chain permits to project the hydrophilic part of the chain in one side of the helix and the hydrophobic face towards the opposite direction. This allows the interaction between this hydrophobic face of the peptide with the lipid membrane of the cell while the hydrophilic part is in contact with the aqueous environment [12].

The way PACAP activates its receptors consists of a two-step process, in which in the first one, the C-terminal part of PACAP interacts and binds to the cell membrane at the extracellular domain. After this interaction, the N-terminal part of the hormone can be placed near the transmembrane regions of the corresponding receptor to start its activation. It has been demonstrated that in the case of PACAP, the peptide chain is oriented parallel to the cell membrane, while other antagonists of the same peptide family, like GLP-1, are oriented in a perpendicular way [3]. This specific behavior between PACAP and receptor shows how important the N-terminal domain is for activation. In fact, this would explain why most of the first ten amino acids that compose the N-terminal part of the chain of the vasoactive intestinal peptide-like peptide family have been conserved. For example, positions six, seven, and ten are equal or homologous in the vast majority of the peptides, being the sixth and tenth positions hydrophobic, and residues with polar radicals in position seven. Furthermore, the first five amino acids at the N-terminal domain are known to be crucial for the activation, whereas residues from positions six to ten are responsible to stabilize the α-helix of the chain. This stabilization takes place through a process called N-capping, in which the hydrophobic residues provide hydrogen bonding to prevent the helix to unwind.

3.3. THE IMPORTANCE OF THE FOURTH AMINO ACID

As it has been said before, PACAP has an affinity to three different receptors, PAC_1, $VPAC_1$, and $VPAC_2$, inducing different responses when binding to one or another. For this reason, ways are being developed to direct the selectivity of PACAP towards a specific receptor, without having affinity (or having a very low affinity) for the rest of the receptors. To do so, investigations have been carried out to determine unique aspects in the interaction between PACAP-receptor, discovering an exclusive pattern found only between PACAP and PAC_1 receptor [13]. This interaction is done demonstrating an amphipathic character, in which once PACAP binds to PAC_1, a ß-coil with residues D3, G4, I5, and T7 is formed creating a patch of hydrophobic residues crucial for the receptor binding [14]. The amino acids' classifications and abbreviations can be found in Annex 1.

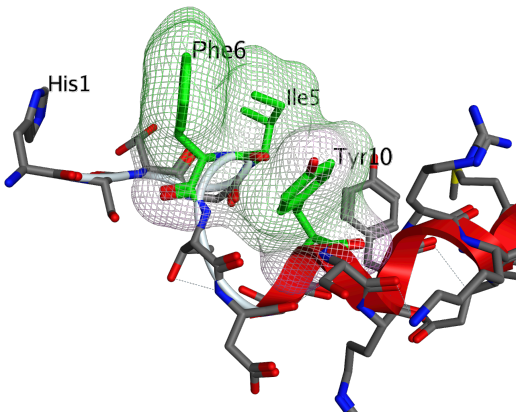


Figure 3. Hydrophobic pocket formed in receptor bound PACAP by I5, F6, Y10 [3].

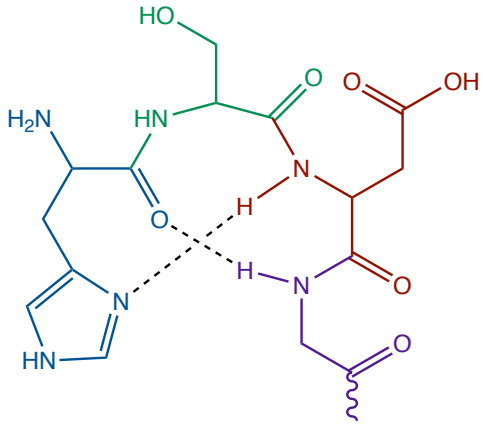


Figure 4. Hydrogen bonds that induce ß-turn in the N-terminal domain of PACAP [15].

It has been observed that the G4 residue (Glycine located at position four), is needed for the stability of the ß-turn and, in addition to this, it is essential to keep together a hydrophobic cluster formed by the side chains of residues I5, F6, and Y10 (Figure 3) [3]. This clustering behavior is only seen when the protein is bonded to the receptor. The ß-turn is induced and stabilized by the hydrogen bond interactions between the carbonyl group of the H1 residue with the amino group of G4, and between the imidazole of H1 with the amino group of D3 (Figure 4) [15][16].

In fact, the change of the G4 residue by another amino acid produces a conformational modification that leads to a lower biological activity towards the PAC_1 receptor. The reason why a neurohormone like VIP has a low affinity to

PAC_1 receptor might be the lack of this G4 residue, as VIP contains an Alanine in the fourth position instead. Thus, VIP is not able to stabilize a ß-turn, as Alanine is not a ß-turn inducer, but an α-helix stabilizer, forbidding the formation of the hydrophobic cluster produced when the peptide is bonded to the receptor. Moreover, it has been demonstrated that the replacement of Glycine by Alanine not only reduces PAC_1 affinity but increases the binding towards $VPAC_1$ [15].

This fact has been taken as the starting point for the development of selective PACAP drug candidates, able to discriminate between PAC_1 in front of $VPAC_1$ and $VPAC_2$ receptors, to rather enhance the neuroprotective and neuro-repairing response than an anti-inflammatory one. Also, other side effects corresponding $VPAC_1$ and $VPAC_2$ activation are to be avoided, like vasodilatation.

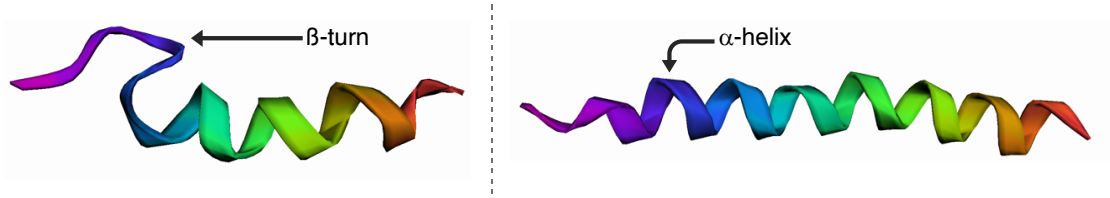


Figure 5. Representation of the ß-turn (left) and α-helix (right) of the N-terminal domain of PACAP when bonded to the receptor [17][18].

What is more, it has been also shown that N-methylated amino acids, especially for Glycine, induce even more the formation of a ß-turn [19]. N-methylated Glycine is also known as Sarcosine. Several studies have demonstrated the facility of Sarcosine to stimulate the formation of ß-sheet conformation regardless of the amino acids present in the turn.

Because of this, three different glycosylated $PACAP_{23}$ have been synthesized by substituting the amino acid placed at the fourth position of the chain. In the first 23-mer, a Glycine was used, in the second one an Alanine, and in the last one a Sarcosine. The three of them were synthesized in Dr. Polt's Laboratory Group at the University of Arizona.

3.4. GLYCOSYLATION OF PACAP

The stability and ability of PACAP to cross the blood-brain barrier (BBB) is a matter of great importance because it is a crucial step that determines the capacity of the neurohormone to penetrate the CNS to produce its neuroprotective effects. This BBB is an interface composed of a thin endothelial layer that regulates the transport of substances between the peripheral circulation

and the CNS. The BBB restricts the entry of molecules to the CNS, prohibiting the crossing of more than 95% of the drug candidates developed to achieve the CNS. There are mainly two ways for biomolecules to cross the BBB, the first one is using a specific transport mechanism, and the second one, through passive diffusion across cell membranes, although the latest one is generally restricted to small uncharged lipophilic molecules [2].

Concerning PACAP, it has been seen that $PACAP_{27}$ can pass through the BBB by passive diffusion, while $PACAP_{38}$ crosses the barrier through a specific peptide transporter. In fact, $PACAP_{38}$ is the only one of its peptide family that is able to cross the BBB by a specific transport system, providing PACAP a unique advantage for neuroprotective treatments. However, even though its high capacity to cross the barrier, it has been observed that after five minutes of an intravenous injection of radiolabeled $PACAP_{27}$ and $PACAP_{38}$ in rodents, the amount of each peptide that reached the CNS was lower than 0.1%. Although this percentage was enough to produce a neuro-repairing response, it is believed that increasing the BBB penetration rate will enhance the neuroprotective activity of the neurohormone.

Several studies suggest that glycosylation of peptides promotes BBB penetration through a process called transcytosis. This process consists of a transcellular transport in which various macromolecules are transported across the interior of a cell by being captured in vesicles on one side of the barrier and ejected on the other side (Figure 6) [20].

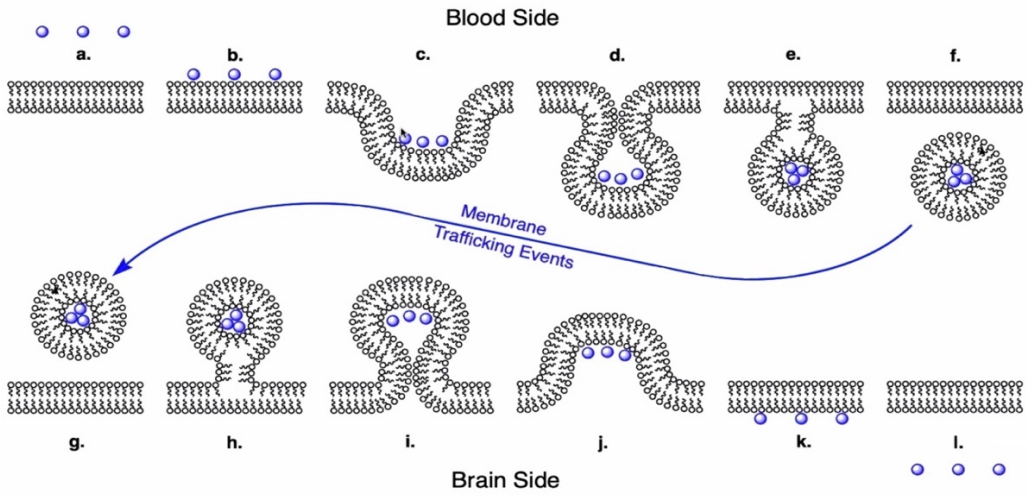


Figure 6. Blood-brain barrier penetration via Transcytosis [20].

In addition to this enhancement in BBB penetration, it has been postulated that a sugar portion bonded to the peptide chain in a specific position, helps the

neurohormone to find a receptor. This point has been hypothesized to happen because of the way in which the helix formed by glycosylated peptide integrates with the cell membrane, where the sugar portion acts to modulate the residence time on the membrane surfaces together with the "on" and "off" rates of the peptide. In other words, the fact of having a sugar portion increases the probabilities of the peptide encountering the receptor because of the amphipathic character that produces in the glycopeptide. Going into details, if the glycopeptide finds a receptor, the helix binds with the extracellular domain and activates the receptor. After that, because of the enhanced solubility property due to the presence of the sugar, the glycopeptide dissociates from the receptor and the membrane, traveling through the aqueous extracellular fluid ready to find another part of the membrane to activate another receptor. The process of lifting the peptide helix out of the membrane and solubilizing it in the extracellular fluid to then attach to another part of the membrane is called "membrane hopping". The direct consequence of this mechanism is not only the increase of the probabilities of the glycopeptide to encounter and activate a receptor but the possibility of activating multiple receptors with only one glycopeptide molecule [21].

What is more, it has been observed by several studies done on different glycosylated endogenous neuropeptides that lactosides presented better characteristics than glucosides. While glucosides showed a low affinity for receptors and a weak BBB penetration, lactosides presented a stronger peptide-receptor affinity and better BBB penetration. This can be explained by the higher hydrophilicity produced by lactose. For this reason, a ß-lactose was chosen for the glycosylation of the three different PACAP$_{23}$ analogs produced in this experimental study.

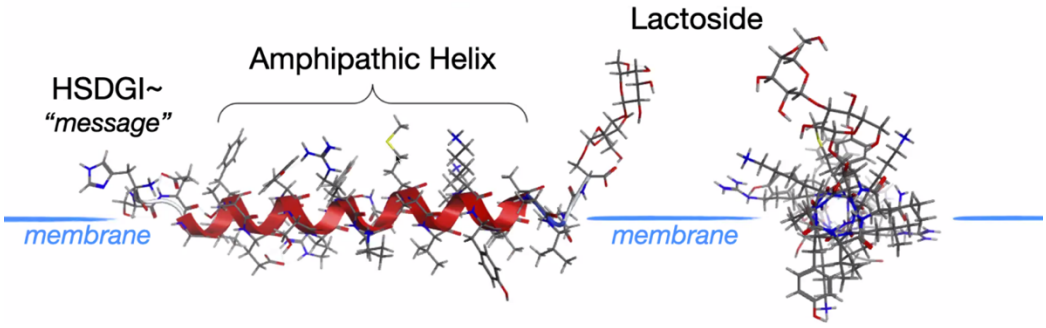


Figure 7. 3D representation of the PACAP lactoside chain attached to the cell membrane.

3.5. SOLID PHASE PEPTIDE SYNTHESIS

The syntheses of the three PACAP$_{23}$ analogs were done using a technique called Solid Phase Peptide Synthesis.

The idea of SPPS is based on a successive addition of protected amino acid derivatives to a growing peptide chain immobilized on a solid phase, which usually is a resin. This process includes the deprotection and washing steps to remove the unreacted groups and the side products formed during the reactions.

Focusing on the resin, it is a polymeric solid support that is linked permanently to a linker that enables a provisional anchoring for the binding of the first amino acid. For a proper operation, the resin must be swollen, as the peptide synthesis happens mostly inside the swollen network of the resin beads. This process takes place when the solvent molecules occupy positions between polymer chains, increasing the volume of the spaces between chains. As the kinetics of the reaction is diffusion controlled, a resin that swells more will have a higher diffusion rate of substrates within the polymeric network, giving as a result a faster reaction and better conversions. For this reason, the swelling concept is of great importance for SPPS [22].

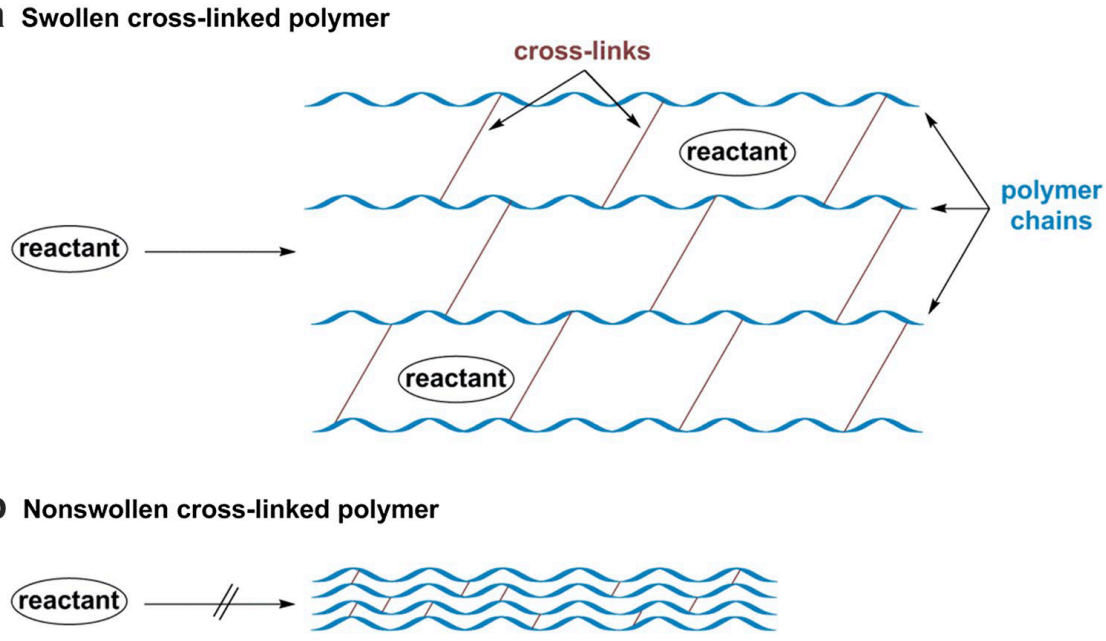


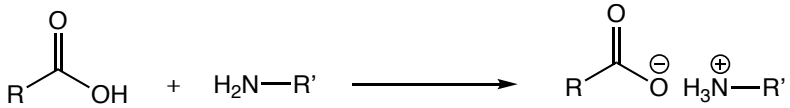
Figure 8. Swelling effect on SPPS [22].

Another important concept involving the resin is its deprotection. In the case of the resin used in this study, it presented a Fmoc protection in the amino group of

the resin, in which the first amino acid of the chain binds, so it had to be deprotected using a strong base.

With respect to the synthesis itself, it is usually carried out starting from the C-terminus towards the N-terminus, although it can also be done the other way around. To do so, the first N-protected amino acid (usually Fmoc/t-Bu or Boc/Bzl protected) must be immobilized by bonding the carboxylic group with the resin giving as a result an amide, ester, thioester, O-substituted oxime, or hydrazide in function of the resin used. Once the first amino acid is attached, in order to avoid the following residue to react with unreactive sites of the resin, a capping treatment must be done. After that, the building of the peptide chain can start. To do so, the amino group of the last amino acid added to the chain has to react with a new one forming a peptide bond. The main problem in this crucial step is the possible acid/base reaction between the carboxylic group of the incoming amino acid and the amino group of the last amino acid of the chain to give an ammonium salt as a result. In order to avoid this, a substance called coupling reagent is added to activate the carboxylic group of the new amino acid, increasing the leaving group ability of the OH. This activation allows the nucleophilic attack of the amino group of the chain towards the incoming amino acid (Figure 9).

a) Acid-Base Reaction



b) Amide Formation by Nucleophilic Acyl Substitution Reaction

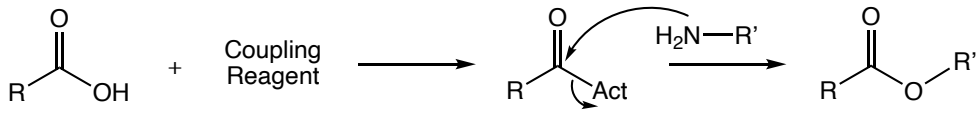


Figure 9. Comparison of the acid-base reaction and the amide formation reaction.

The development of a technique like this has allowed the synthesis of high pure peptides chains on a large scale, producing an important enhancement in the development of powerful therapeutic agents and vaccines.

4. EXPERIMENTAL PART

4.1. SYNTHESIS OF PACAP23

Three different glycosylated PACAP$_{23}$ compounds were synthesized by Christopher R. Apostol and Thiago Gilmartin in Dr. Robin Polt's Laboratory Group at the University of Arizona, labeled as CRA_TG_3025, CRA_TG_3026, and CRA_TG_3027. The first one corresponds to the 23-mer containing a Glycine in the fourth position, the second one an Alanine, and the third one a Sarcosine. The amino acid sequences of the three compounds can be seen in Table 1, and their complete structural formula in Figures 11, 12, and 13, respectively.

Table 1. Amino acid sequences for the three glycosylated PACAP$_{23}$ analogs.

Compound ID	Structures
CRA_TG_3025	HSDGIFTDSY$_{10}$SRYRKQÑAVK$_{20}$KYL-Ser(Lac)-CONH$_2$
CRA_TG_3026	HSDAIFTDSY$_{10}$SRYRKQÑAVK$_{20}$KYL-Ser(Lac)-CONH$_2$
CRA_TG_3027	HSD-Sar-IFTDSY$_{10}$SRYRKQÑAVK$_{20}$KYL-Ser(Lac)- CONH$_2$

The three different glycosylated PACAP$_{23}$ compounds were synthesized using the Solid Phase Peptide Synthesis technique on a Prelude® automated peptide synthesizer. The procedure consisted of both automated and semi-manual ways, wherein the latest, reagents were loaded into the reaction vessels with the help of a syringe. The stirring and agitation of the resin during the process were done using a steady flow of argon. The washing steps were done for 2 minutes each with dimethylformamide (DMF) and dichloromethane (DCM). Previously to these syntheses, the Fmoc protected glycosylated Serine starting material had to be done.

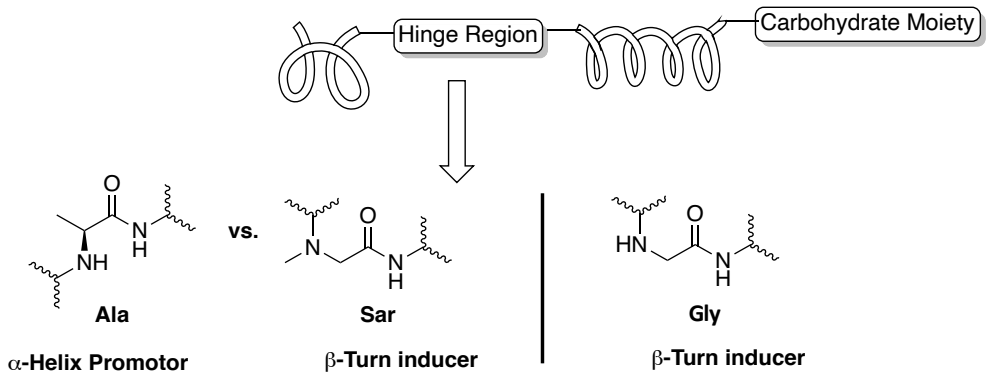
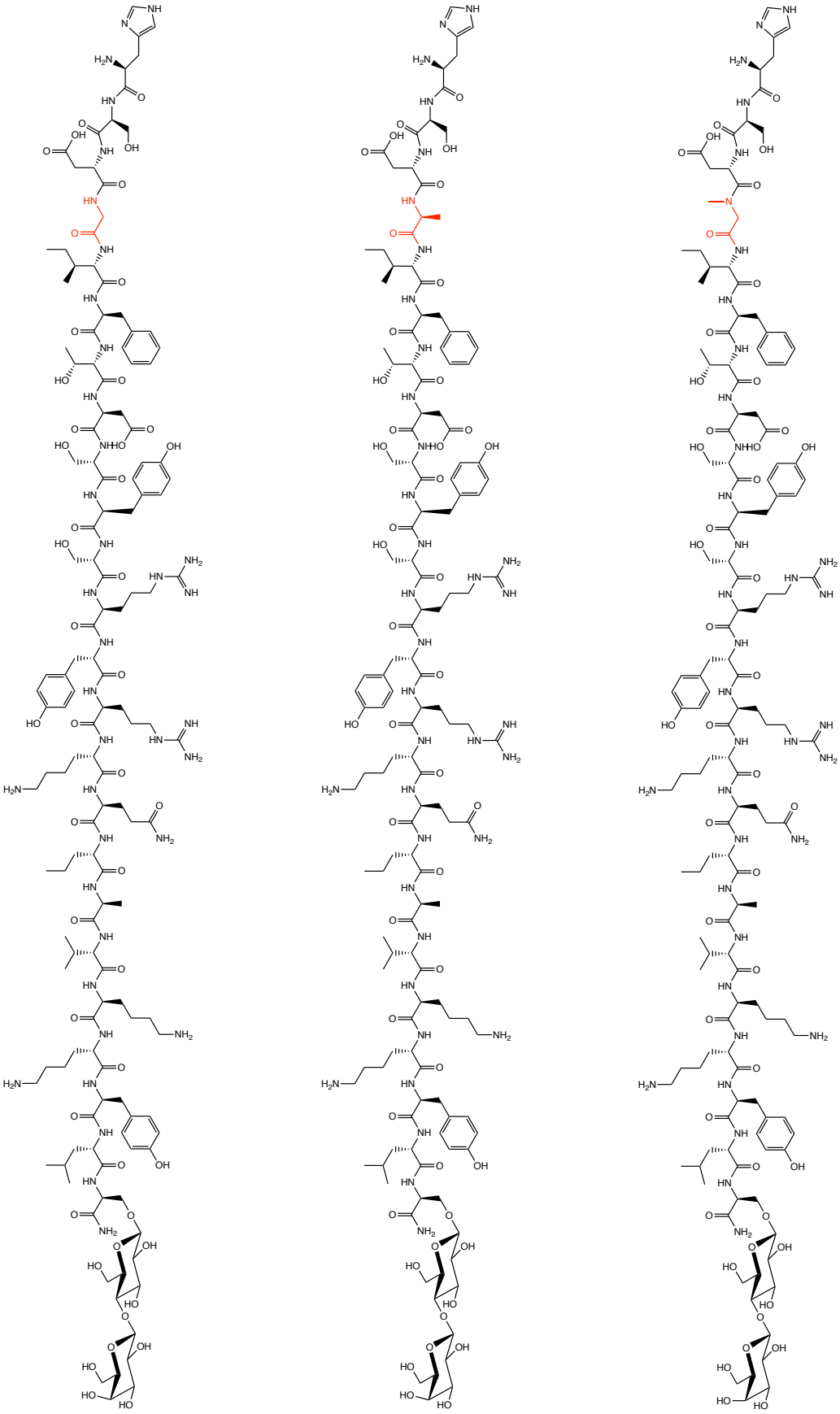


Figure 10. Schematic representation of the synthesis objective.

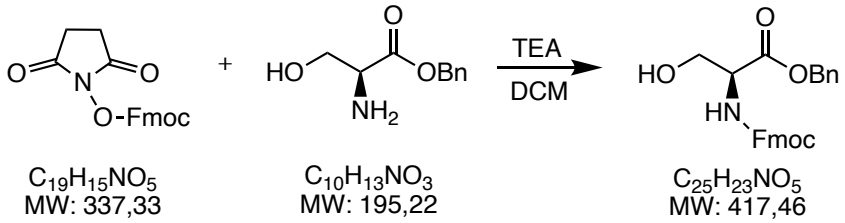


Figures 11, 12, and 13. CRA_TG_3025 (Gly^4 in red, left), CRA_TG_3026 (Ala^4 in red, middle), and CRA_TG_3027 (Sar^4 in red, right) respectively.

4.1.1. SYNTHESIS OF Fmoc-Ser(OBn)

The first step in this long experimental procedure was the synthesis of the Fmoc protected Serine benzyl ester, the first building block in the three PACAP$_{23}$ analogs, in which the lactose will be bonded afterward.

For the synthesis of Fmoc-Ser(OBn), 1 equivalent of Fmoc N-hydroxy-succinimide ester protecting group, also known as Fmoc; together with 1.1 equivalents of L-serine benzyl ester hydrochloride, were dissolved in dichloromethane in a round-bottomed flask.



$C_{19}H_{15}NO_5$
MW: 337,33

$C_{10}H_{13}NO_3$
MW: 195,22

$C_{25}H_{23}NO_5$
MW: 417,46

Figure 14. Fmoc-Ser(OBn) synthesis reaction.

Despite it was left stirring for some minutes, starting materials did not fully dissolve, forming a white solution. Then, 1.5 equivalents of triethylamine (TEA) were added dropwise to the reaction flask, making the solution completely clear. The mixture was left stirring overnight.

The following day, a TLC (mobile phase 1:1 ethyl acetate/hexanes) of the crude and the starting materials was done to make sure that everything has reacted. Only one spot was seen for the product, therefore, everything had reacted. After that, the crude solution was washed three times with hydrochloric acid 0.1 M, collecting the organic layer to wash it twice with water. The organic fraction was dried by adding anhydrous magnesium sulfate and collected in a round-bottomed flask to evaporate in a rotavapor. Once all the DCM was evaporated, a white solid appeared. This solid was dissolved in a mixture of composition 1:1 DCM/hexanes for its recrystallization. To ease this process, an excess of hexanes was added until the solution became a bit cloudy. Then, it was left overnight to recrystallize.

The next morning, a great amount of solid was observed, which was filtered under vacuum to get white foamy crystals. Once they were dried, they were weighted, obtaining a yield of 86 %.

To proceed with the synthesis, the purity of this compound had to be checked. To do so, different analytical techniques were used. Firstly, another TLC of the

compound was done in 1:1 ethyl acetate/hexanes. Apparently, there was no L-serine benzyl ester from the starting material in the product. Secondly, a proton NMR was done (dissolving 10 mg of the sample in 0.7 mL of deuterated chloroform). To compare the results, a different spectrum from Dr. Polt's laboratory research group was taken as a reference, observing that all the peaks in the product spectrum matched the peaks of the reference one. Finally, to ensure purity, a HPLC was done (dissolving 1 mg of the sample in 1 mL of methanol) demonstrating a purity of 99.8 %.

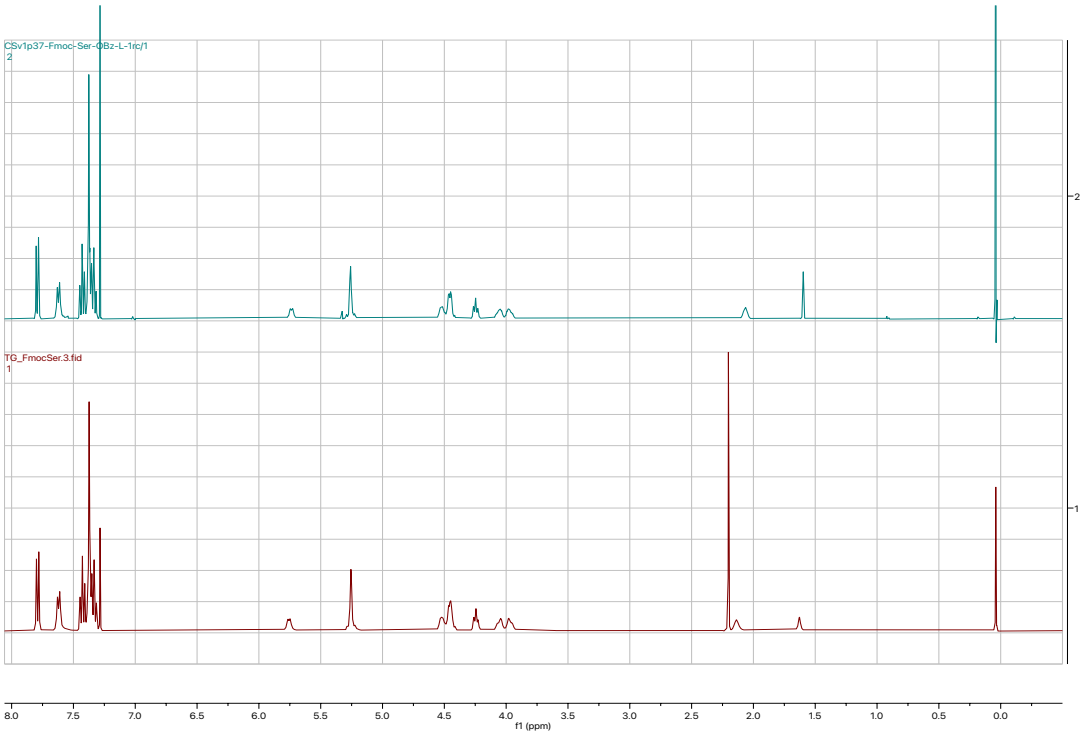
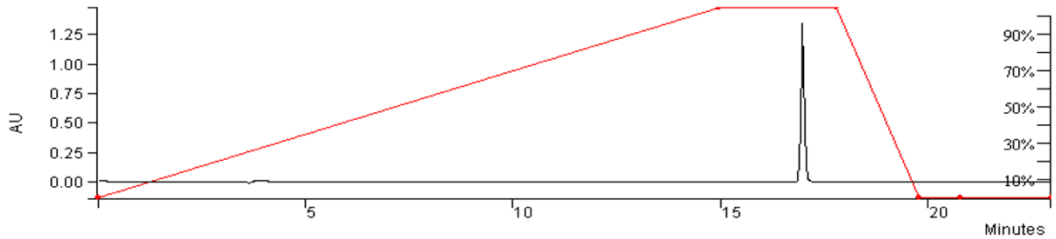


Figure 15. H^{1}NMR Fmoc Serine Benzyl Ester, comparison between my results (below) and spectrum provided by the laboratory (above).



Peak No	Peak Name	Result ()	Ret. Time (min)	Time Offset (min)	Area (counts)	Rel Ret Time	Sep. Code	Width 1/2 (sec)	Status Codes	Group
1		0.0619	15.507	0.000	24370	0.00	BB	7.8		0
2		99.8410	16.973	0.000	39279512	0.00	BB	5.3		0
3		0.0971	18.939	0.000	38187	0.00	BB	7.2		0
	Totals	**100.0000**		**0.000**	**39342068**					

Figure 16. HPLC of the Fmoc Serine Benzyl Ester reaction.

4.1.2. GLYCOSYLATION OF Fmoc-Ser(OBn) WITH ß-PERACETYL LACTOSE

For the reaction, 1 equivalent of Fmoc-Ser(OBn) and 1.25 equivalents of ß-peracetyl lactose were weighed and placed in a round-bottomed flask and were dissolved in a minimum amount of carbon tetrachloride. The mixture was left stirring under reflux for two hours before adding the $InBr_3$. After two hours, 0.1 equivalents of $InBr_3$ were added, leaving the reaction under reflux overnight at around 95 ºC.

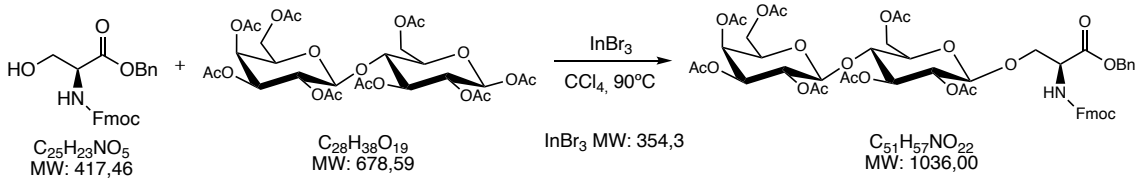
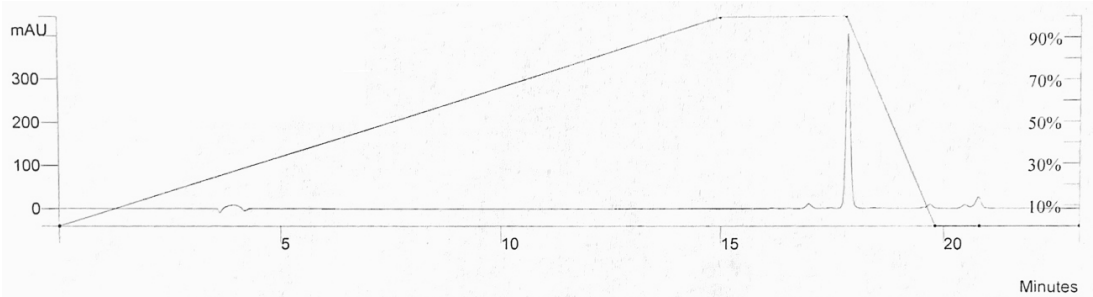


Figure 17. Fmoc-Ser(OBn) and ß-peracetyl lactose glycosylation reaction.

The following day, the reaction was monitored by HPLC to check the reaction status. After that, the solvent was evaporated by using a rotavapor, and co-evaporated with ethyl acetate three times, and diethyl ether also three times. Then, the compound was redissolved in ethyl acetate and was transferred to a separatory funnel to rinse twice with water and twice with brine. The organic layer was then dried using anhydrous magnesium sulfate, and evaporated in the rotavapor, obtaining a white solid.

Afterward, to isolate the desired compound from the solution, a purification using a Biotage auto-column was done. To do so, a TLC with 6:4 hexanes/ethyl acetate was done to determine the order in which each compound will be eluted. The product should stain a different color than ß-peracetyl lactose. The auto-column was filled with silica and was packed properly. Then, it was loaded and flushed with 100% hexanes for at least three minutes. After that, about six grams of the product were dissolved in DCM and loaded into the column (six grams must not be exceeded to guarantee an effective purification), because of this, purification had to be done several times. Once the sample was loaded, the eluent was switched to 6:4 hexanes/ethyl acetate, and fractions were tested by TLC to localize the desired compound. It was observed that the product was the last one to be eluted, so the eluent was changed to 100% ethyl acetate to collect all of it. The fractions containing the product were combined and evaporated to obtain a white solid. Once the three auto-columns were done, pure Fmoc-Ser(OBn) ß-peracetyl lactose was obtained, resulting in a yield of 71%.



Peak No	Peak Name	Result ()	Ret. Time (min)	Time Offset (min)	Area (counts)	Rel Ret Time	Sep. Code	Width 1/2 (sec)	Status Codes	Group
1		2.5605	16.984	0.000	425386	0.00	BB	7.2		0
2		84.8541	17.861	0.000	14097316	0.00	BB	6.2		0
3		2.0368	19.696	0.000	338381	0.00	BB	6.4		0
4		2.3282	20.480	0.000	386806	0.00	BV	8.6		0
5		8.2204	20.776	0.000	1365705	0.00	VB	8.0		0
	Totals	100.0000		0.000	16613594					

Figure 18. HPLC of Fmoc-Ser(OBn) ß-peracetyl lactose reaction.

4.1.3. Fmoc-Ser(OBn) ß-PERACETYL LACTOSE HYDROGENATION

The hydrogenation of the benzyl ester group in Fmoc-Ser(OBn) ß-peracetyl lactose was performed by using a Schlenk line and working in an air-free atmosphere with Palladium/Charcoal catalysts and Hydrogen gas.

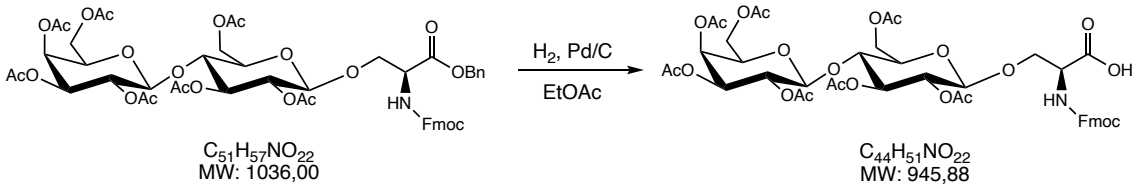


Figure 19. Hydrogenation reaction.

Pd/C was added to a round-bottomed flask and was degassed and flushed with Argon three times. The serine lactoside was dissolved in another flask with ethyl acetate and some drops of acetic acid, then it was degassed for 10 minutes. The Pd/C was transferred to a three-necks round-bottomed flask and connected to the Schlenk line with Argon and vacuum. Then the solution containing the sugar was added to the flask containing the palladium through cannulation, to avoid contact with air. Then H_2 was added through one of the necks of the flask. The reaction was left stirring overnight, and the next morning more H_2 was added. The reaction was monitored by TLC (6:4 EtOAc:Hex).

Once the reaction finished, the solution was filtered under vacuum and the collected solution was evaporated to obtain a white solid. This solid was redissolved in a solution 1:1 ACN and water for its lyophilization, obtaining a white and foamy white solid.

After these syntheses, the Solid Phase Peptide Synthesis could start.

4.1.4. RINK AMIDE RESIN PREPARATION

For the solid phase used in the synthesis 0.25 mmol of Rink Amide-MBHA resin (0.6 g) were used (Figure 20). The resin was placed in a 45 mL reaction vessel and was left for swelling in DMF for 1 hour. The active site of the resin presented a Fmoc protected amino group. The deprotection was done by the addition of a solution containing 2%DBU-2%piperidine in DMF (6 mL), mixing for 4 minutes. Then, the mixture was drained, and the resin was washed once with 6 mL of DMF. The Fmoc deprotection was repeated for an additional 8 minutes followed by 6 DMF washes (6 mL, for 2 minutes). The reason why many DMF washes were done is that DBU cannot perform an effective neutralization of the dibenzofulvene (DBF) released during the reaction, so DMF washes away the DBF produced, avoiding any side reaction (Figure 21). [23]

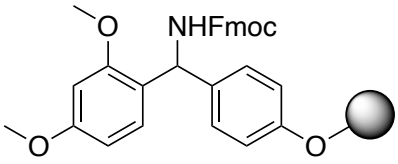


Figure 20. Active site of the Rink Amide-MBHA resin.

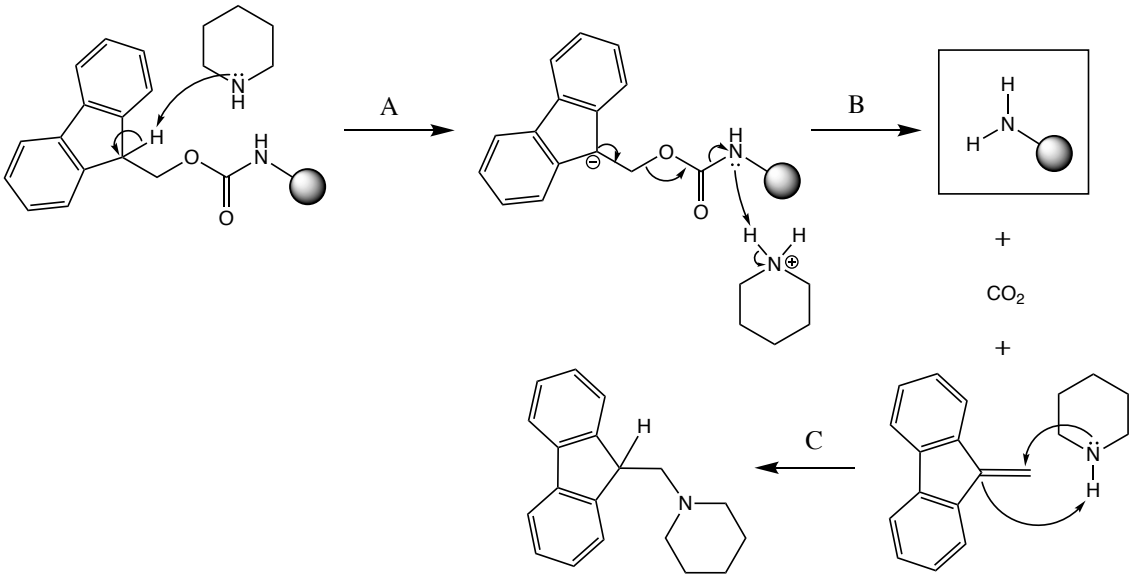


Figure 21. Fmoc deprotection mechanism of the resin through an elimination reaction, A. Acid-base reaction between the highly acidic proton of the fluorene group and piperidine (or DBU); B. The anion moves to the adjacent atom releasing CO_2 and dibenzofulvene (DBF) deprotecting the amino group of the resin; C. Piperidine (or DBU) neutralizes the reactive DBF.

4.1.5. GLYCOSYL AMINO ACID LOADING

The first amino acid to be loaded was the glycosylated one, in specific, a ß-peracetyl lactose Serine Fmoc protected synthesized previously (Figure 22). To do so, 0.2 mmol, (0.8 eq.) o Fmoc-Ser(Lac (OAc)$_7$)-OH) and 0.2 mmol (0.8 eq.) 6-Cl-HOBt were placed into a vial and were dissolved in 4 mL of N-methyl-2-pyrrolidone (NMP). Then, 0.2 mmol (0.8 eq.) of N,N′-diisopropylcarbodiimide (DIC) was added into the solution. Then, the solution was vortexed for 1 minute to mix it and if it was necessary, was sonicated for another minute before adding it to the resin.

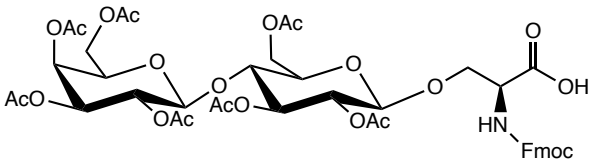


Figure 22. ß-peracetyl lactose Serine Fmoc protected amino acid block.

After doing that, the vessel of the reaction was left mixing with an argon stream for 16 hours. Once the mixture step was done, the mixture was diluted in 10 mL of DMF and was drained instantly. The next step was washing 6 times the resin with 6 mL of DMF and then 6 times with 6 mL of DCM.

Then, the unreacted NH_2 sites of the resin were capped with 8 mL of a DCM solution containing 10% N,N-diisopropylethylamine and 10% Ac_2O. This step was left to react for 1 hour, and when it was done, it was washed 6 times with 6 mL of DCM and 4 times with 6 mL of DMF. After the washing steps, the resin is ready for the automated steps.

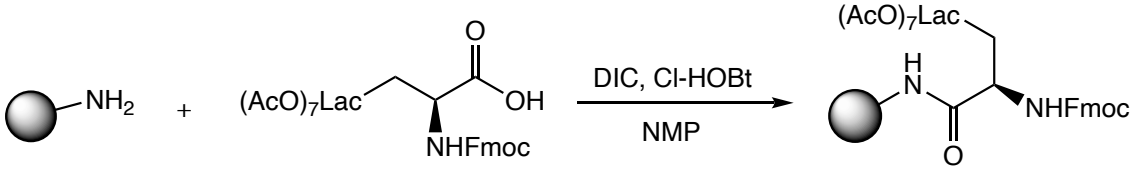


Figure 23. Reaction of the resin loading with the glycosyl amino acid.

4.1.6. Tyr^{10}-Leu^{23} AUTOMATED SYNTHESIS

The amino acid chain from Leucine^{23} to Tyrosine^{10} was build using the automated SPPS feature on a Prelude® automated peptide synthesizer. The same steps were applied for each of the thirteen amino acids that make up the peptide chain. These steps consisted of first the removal of the Fmoc group, and then the coupling reaction with the Fmoc protected amino acid. To do so, the deprotection of the Fmoc group was done as explained in the previews section. The amino acid coupling was done by adding 2 equivalents of the Fmoc-amino acid, 2 equivalents of HBTU, and 10 equivalents of N-methylmorpholine to the reaction vessel. After adding it to the resin, the mixture was mixed for 30 minutes and then washed once with 6 mL of DMF. The coupling reaction was repeated a second time for 30 minutes, and the resin was then washed 6 times with DMF (6mL). These steps were repeated thirteen times, once per amino acid coupled to the peptide chain.

Figure 24. Prelude® Automated Peptide Synthesizer.

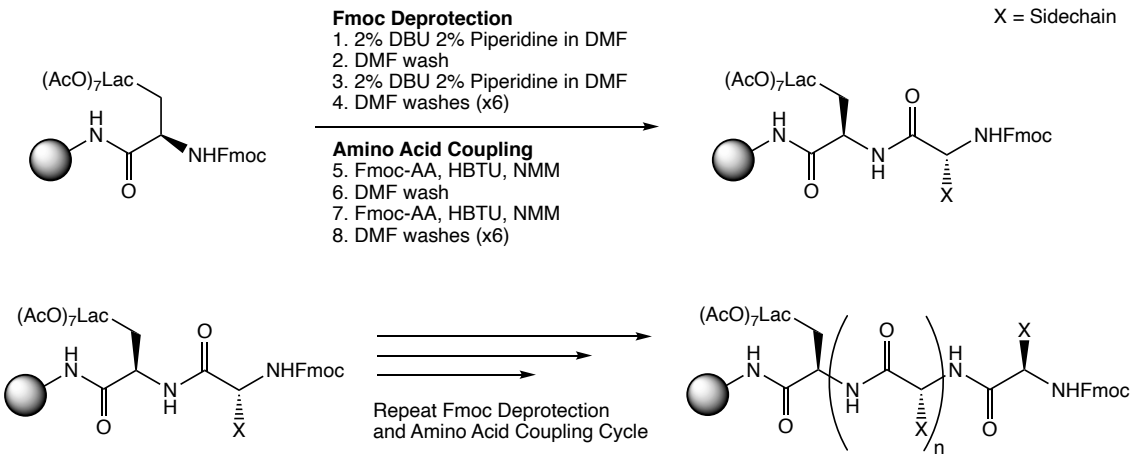


Figure 25. General scheme of automated peptide synthesis for PACAP$_{23}$.

4.1.7. Asp^{8}-Ser^{9} DIPEPTIDE ADDITION

The reason why the automated part was done only until Tyr^{10} is because the addition of single Aspartic acid to the peptide chain can cause the formation of Aspartimide as a side reaction, which leads to low yields in addition to costly purification or even inaccessible peptide sequences. Aspartimide formation implies the cyclization of the Aspartic acid side chain with the α-amino nitrogen of the proceeding residue, being Serine, Glycine, Threonine, and Cysteine the most problematic ones. Under basic conditions, like in the presence of piperidine, ring-opening happens, leading to racemization of the peptide chain, as two different isomers can be formed (Figure 26) [24]. To avoid this situation, Aspartic acid is coupled in the form of pseudo-proline protected building dipeptides blocks.

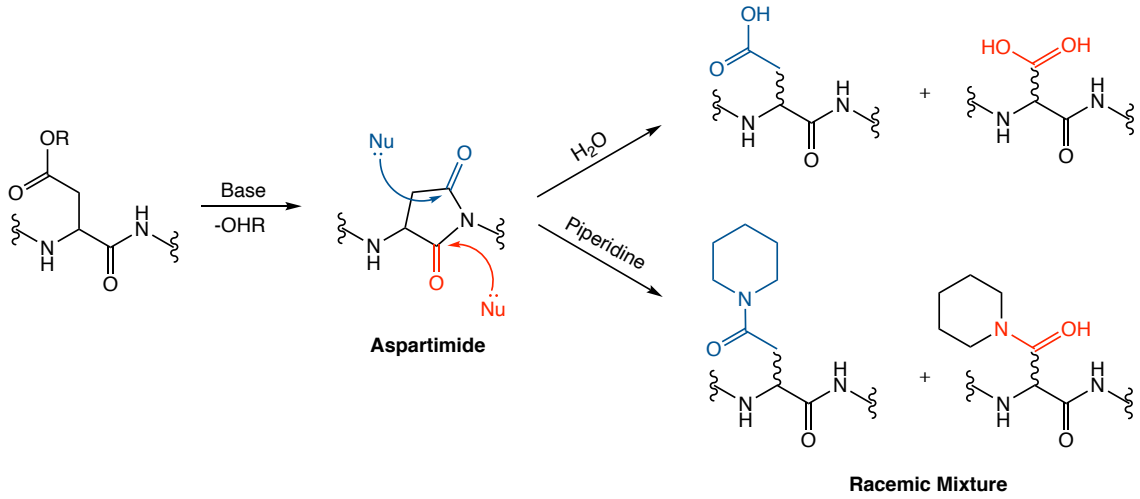


Figure 26. Aspartimide formation reaction and racemization produced by ring-opening.

So, the following step in the synthesis consisted of adding 0.3 mmol (1.2 equiv.) of Fmoc-DS-OH (Serine-Aspartic acid dipeptide), together with 0.3 mmol of 6-Cl-HOBt (1.2 equiv.), and 0.3 mmol of DIC (1.2 equiv.), in 4 mL of NMP. This mixture was vortexed and sonicated for 1 minute and then was added manually to the resin with a syringe. The reason why only 1.2 equivalents were used was that the reaction was left for 16 hours. Once the reaction finished, the mixture was diluted with 10 mL of DMF and was drained instantly. Finally, the resin was washed with 6 mL of DMF 6 times.

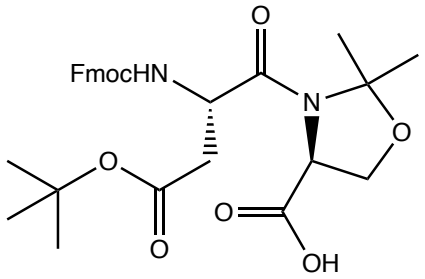


Figure 27. Arg^{8}-Ser^{9} dipeptide block, Fmoc-Asp(OtBu)-Ser[Psi$^{(Me,Me)}$Pro]-OH

4.1.8. Ile^5-Thr^7 AUTOMATED SYNTHESIS

The following three amino acid sequence Ile^5-Phe^6-Thr^7 was coupled using the automated feature on the Prelude® automated peptide synthesizer. The steps to follow were similar to the Tyr^{10}-Leu^{23} automated synthesis explained above, but in this case, 4 equivalents of Fmoc-amino acid, 4 equivalents of HBTU, and 16 equivalents of NMM in 10 mL of DMF were used instead. After adding the reagents, the reaction vessel was mixed for half an hour and washed once with 10 mL of DMF. Then, the coupling reaction was repeated a second time for another half an hour. Once the reaction finished, the resin was washed 6 times with DMF. The final step was Isoleucine^5 deprotection, preparing the chain for the next manual step.

4.1.9. Asp^3-Gly^4 DIPEPTIDE ADDITION TO CRA_TG_3025

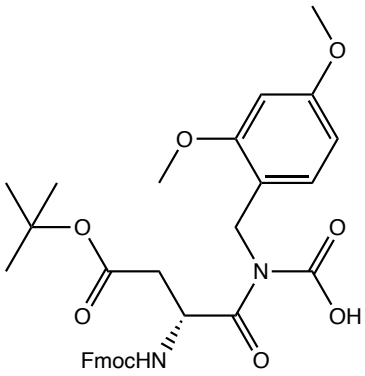


Figure 28. Asp^3-Gly^4 dipeptide block, Fmoc-Asp(OtBu)-Gly(DMB)-OH

For the same reason as in Asp^8-Ser^9 synthetic step, the addition of Asp^3-Gly^4 was done in the form of a dipeptide to avoid Aspartimide formation. To do so, the Fmoc group was firstly removed as described in previews steps, and then 0.5 mmol (2.0 equiv.) of Fmoc-DG-OH (Figure 28) and 0.5 mmol (2.0 equiv.) of 6-Cl-HOBt were dissolved in a vial with 4 mL of NMP. Then, 0.5 mmol (2.0 equiv.) of DIC were added to the mixture. The solution was vortexed and was sonicated for 1 minute. Once everything was dissolved, the mixture was added to the resin and mixed for 6 hours under argon flow. Afterward, the mixture was diluted with 10 mL of DMF, and drained immediately. Finally, the resin was washed with 6 mL of DMF 6 times.

4.1.10. ADDITION OF Fmoc-Ala-OH TO CRA_TG_3026, AND Fmoc-Sar-OH TO CRA_TG_3027

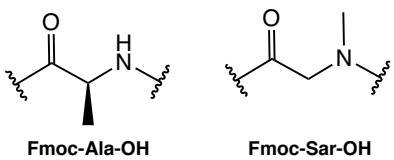


Figure 29. Fmoc-Ala-OH and Fmoc-Sar-OH

In the case of CRA_TG_3026 and CRA_TG_3027, a dipeptide was not used, so Fmoc-Ala-OH and Fmoc-Sar-OH were used respectively. It is believed that Aspartimide is not prone to be formed due to the presence of the methyl group in both Alanine (in the α-carbon) and Sarcosine (N-methylation), due to steric hindrance.

So, the synthetic step began with the Fmoc group deprotection. Then, 1.0 mmol (4 equiv.) of amino acid and 1.0 mmol (4 equiv.) of 6-Cl-HOBt were added to a vial and dissolved in 4 mL of NMP. 1.0 mmol (4 equiv.) of DIC was then added to the solution. The mixture was vortexed for 1 minute and then added to the resin. The reaction mixture was mixed for 2 hours. The mixture was diluted with 10mL DMF and drained immediately. Then the resin was washed 6 times with 6 mL of DMF.

4.1.11. MANUAL LOADING OF REMAINING AMINO ACIDS

The Asp^3 was then added manually to CRA_TG_3026 and CRA_TG_3027 and, after that, only two amino acids were left to add to each PACAP$_{23}$ analog, His^1, and Ser^2. These two last residues were also added manually. To do so, as it was done in each step, the Fmoc group was removed. After that, 1.0 mmol (4 equiv.) of the corresponding amino acid and 1.0 mmol (4 equiv.) of 6-Cl-HOBt were added to a vial and dissolved in 4 mL of NMP. Then, 1.0 mmol (4 equiv.) of DIC was added to the mixture. The solution was vortexed for 1 minute and then added to the resin. The reaction mixture was mixed for 1 hour. The mixture was diluted with 10 mL of DMF and drained instantly. Then the resin was washed 6 times with 6 mL of DMF. Finally, the Fmoc group of His^1 was removed as explained previously.

4.1.12. ACETYL CLEAVAGE

After the coupling of the last amino acid to the peptide chain, the acetyl groups present in the ß-lactose must be cleaved and substituted by hydroxyl groups.

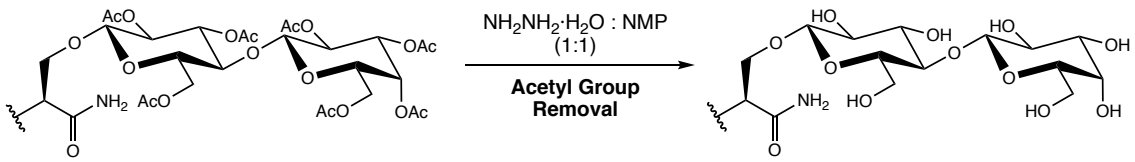


Figure 30. Acetyl group removal from the ß-peracetyl lactose bonded to the Serine.

To do so, a solution containing 50% of hydrazine ($NH_2NH_2{\cdot}H_2O$) in NMP was made, and 10 mL of it were added to each reaction vessel to react with the resin. Then, the solution was mixed with argon overnight for 16 hours. The next day, resins were drained, another 10 mL of the cleavage solution were added to each reaction vessel and were left to react mixing for an additional 2 hours. After this time, resins were drained again and were washed first 8 times with 10 mL of DMF,

and then 8 times with 10 mL of DCM. Once all the washing steps were done, each resin was left to dry under vacuum for 3 hours.

4.1.13. CLEAVAGE FROM THE RESIN AND GLOBAL SIDE CHAIN DEPROTECTION

After the elimination of the acetyl groups from the lactose, the dried resin was treated with an acidic cleavage cocktail containing TFA, DCM, H_2O, triethylsilane, and anisole in the following composition: 90:10:2:3:0.5. This process was done to cleave the peptide chain from the resin, and to deprotect any side chain of any amino acid constituting the peptide.

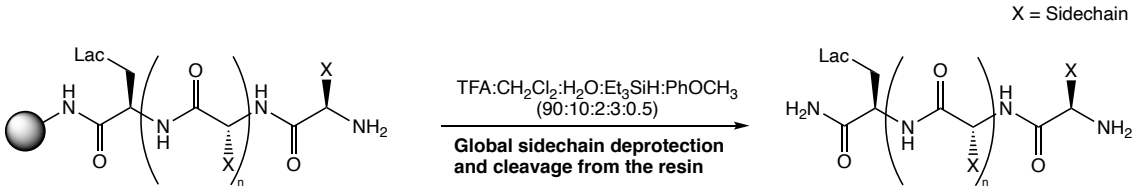


Figure 31. Global sidechain deprotection and cleavage from the resin.

So, the resin was mixed for 1 hour, and the solution was collected in a 45 mL centrifuge tube. This cleavage step was repeated two more times for 10 minutes periods. After repeating the cleavage, the combined fractions were evaporated gradually over a stream of argon until the peptide began to crash out (when there were only 5 mL of solution left in the centrifuge tube). Then, about 40 mL of cold diethyl ether were added for the peptide to precipitate (solution became milky). To isolate the solid, the tubes were centrifuged for 10 minutes at 5 G. After that, it could be observed that a white solid precipitated at the bottom of the tubes. The ether was decanted off, and another 40 mL of cold diethyl ether were added to the crude peptide to centrifuge once more. This process was repeated a third time. After decanting the ether layer, the crude peptide was dried under vacuum overnight.

4.2. HPLC PURIFICATION AND CHARACTERIZATION OF CRUDE PEPTIDES

The last step consisted of an HPLC purification of the crude samples on a Gilson system with a UV detector set at 280 nm, using a Vydak C18 preparative reversed-phase column (250 mm x 50 mm) with a gradient of 5-80% CH_3CN vs 0.1% CF_3COOH in H_2O over 60 minutes to give the glycopeptides in pure form (Figure 32)

Previews to the purification and to monitor it, an analytical HPLC was done (in an Inspire C18 5µm 250 mm x 4.6 mm column) on a Varian LC with a diode array detector system (at 280 nm) employing the same gradient over a period of 15 minutes.

The pure fractions obtained from preparative HPLC purification were frozen at a temperature of -80°C and then lyophilized to afford the pure peptides as white and fluffy solids. The purity of the compounds was checked by analytical HPLC using the same technique as before but running it over a period of 35 minutes to ensure a better separation in the case other peaks appeared. The pure peptides were then characterized using mass spectrometry (ESI-MS).

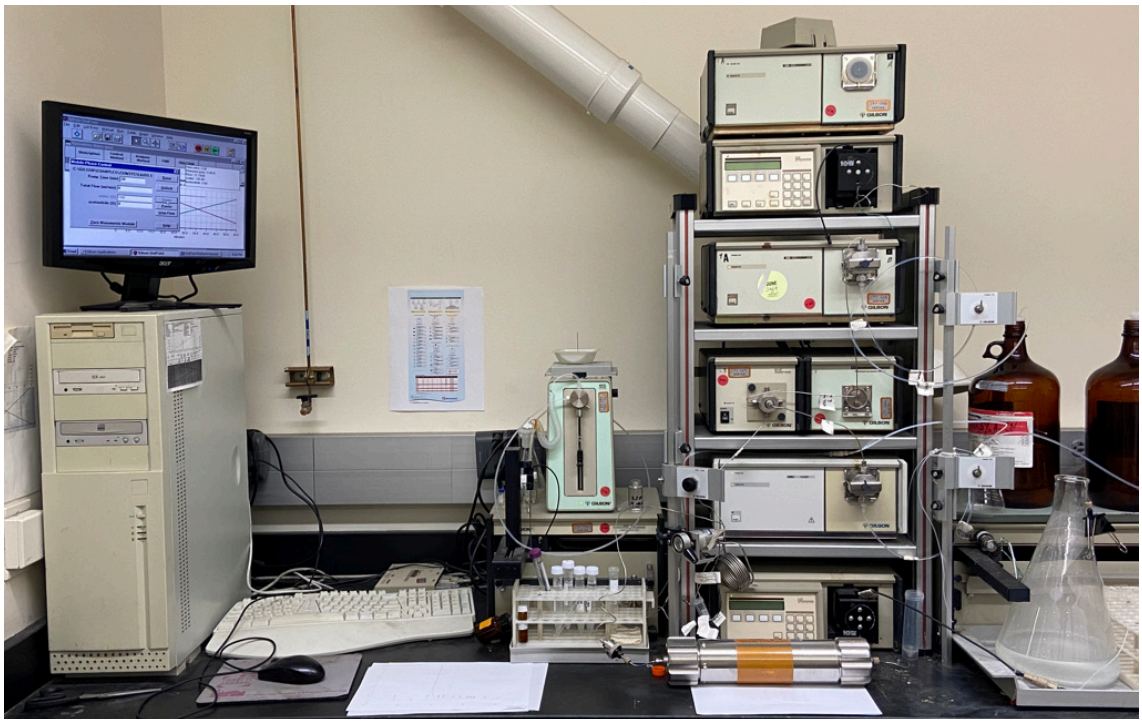


Figure 32. HPLC Preparative Reversed-Phase Column system used in Dr. Robin Polt's Laboratory Group.

5. RESULTS AND DISCUSSION

The three glycosylated $PACAP_{23}$ analogs were characterized through two different techniques: High-Performance Liquid Chromatography and Mass Spectrometry.

The amounts obtained were similar for the Glycine and Sarcosine analogs, but lower for the Alanine one. All three compounds were white solids.

The yields found were high enough for this kind of synthesis, with a yield of 36.0 mg for CRA_TG_3025, 19.7 mg for CRA_TG_3026, and 31.2 mg for CRA_TG_3027. The Glycine and Sarcosine analogs were easier to purify through Preparative HPLC than the Alanine one, for this reason, their yields are higher. This fact might be attributed to a possible amount of Aspartimide formed during the Alanine analog synthesis.

Table 2. Characterization of synthesized $PACAP_{23}$ Glycopeptides.

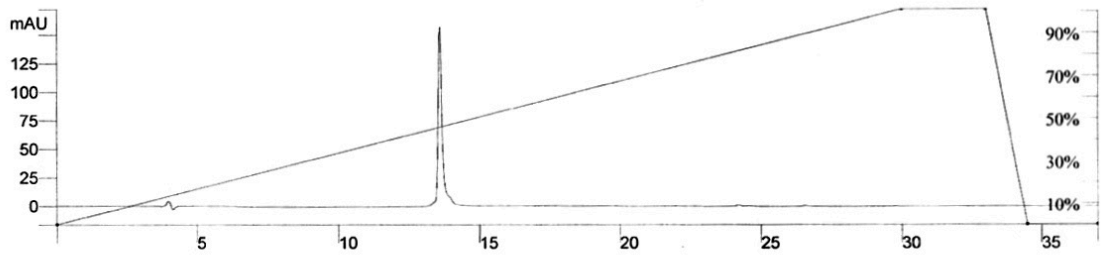
Compound	Molecular Formula	Exact Mass	Mass Found (m/z values)	HPLC Retention Time (min)	Yield (mg)
CRA_TG_3025	$C_{140}H_{219}N_{37}O_{47}$	3170.59	$793.93(M+4H)^{4+}$	13.587	36.0
CRA_TG_3026	$C_{141}H_{221}N_{37}O_{47}$	3184.60	$638.09(M+5H)^{5+}$	13.477	19.7
CRA_TG_3027	$C_{141}H_{221}N_{37}O_{47}$	3184.60	$638.01(M+5H)^{5+}$	13.424	31.2

5.1. HIGH PERFORMANCE LIQUID CHROMATOGRAPHY

The purity of the three 23-mers was determined by analytical HPLC in an Inspire C18 5µm 250 mm x 4.6 mm column, on a Varian Liquid Chromatography with a diode array detector system at 280 nm, employing a gradient of 5-80% CH_3CN vs 0.1% CF_3COOH in H_2O over a period of 35 minutes.

5.1.1. CRA_TG_3025 HPLC

The glycosylated $PACAP_{23}$ containing a Glycine in the fourth position, named CRA_TG_3025, was the first one to purify and characterize. This analysis showed high purity and a narrow peak at a retention time placed at 13.587 minutes, as expected (Figure 33).

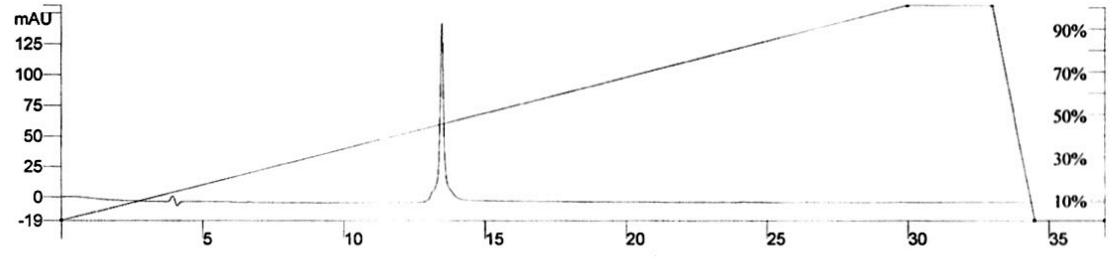


Peak No	Peak Name	Result ()	Ret. Time (min)	Time Offset (min)	Area (counts)	Rel Ret Time	Sep. Code	Width 1/2 (sec)	Status Codes	Group
1		100.0000	13.587	0.000	8293504	0.00	BB	7.8		0
	Totals	100.0000		0.000	8293504					

Figure 33. HPLC of CRA_TG_3025 with a gradient of 5-80% CH_3CN vs 0.1% CF_3COOH in H_2O over 35 minutes.

5.1.2. CRA_TG_3026 HPLC

The glycosylated PACAP$_{23}$ containing an Alanine in the fourth position, named CRA_TG_3026 was the most difficult to isolate and purify, for this reason, two shoulders can be observed on the right and on the left of the main peak. The retention time was 13.477 minutes, as expected (Figure 34). Although the presence of two shoulders, the purity was high enough for the *in vitro* and *in vivo* experimental studies. These impurities might correspond to Aspartimide formation, as no Aspartic Acid-Alanine dipeptides were used for this synthesis (page 18). This fact could have produced minor racemization of the product, decreasing gently the purity. When purifying the compound by Preparative HPLC, overlapped weak peaks before and after the main peak were observed, corresponding to those impurities. This has been hypothesized because the retention times of these peaks were similar to the one of the product of interest.

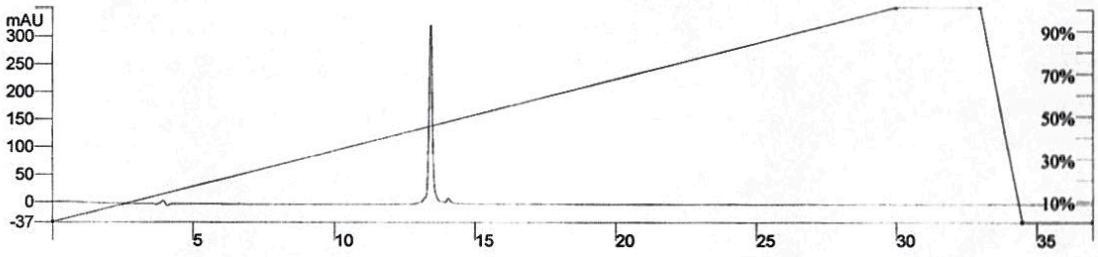


Peak No	Peak Name	Result ()	Ret. Time (min)	Time Offset (min)	Area (counts)	Rel Ret Time	Sep. Code	Width 1/2 (sec)	Status Codes	Group
1		100.0000	13.477	0.000	8895272	0.00	BB	8.0		0
	Totals	100.0000		0.000	8895272					

Figure 34. HPLC of CRA_TG_3026 with a gradient of 5-80% CH_3CN vs 0.1% CF_3COOH in H_2O over 35 minutes.

5.1.3. CRA_TG_3027 HPLC

The glycosylated PACAP$_{23}$ containing a Sarcosine in the fourth position, named CRA_TG_3026 was easy to purify, and therefore, it showed high purity and a narrow peak at 13.424 minutes (Figure 35). Although a very short peak can be observed at 14.045 minutes, it was considered insignificant, as the purity was high enough.



Peak No	Peak Name	Result ()	Ret. Time (min)	Time Offset (min)	Area (counts)	Rel Ret Time	Sep. Code	Width 1/2 (sec)	Status Codes	Group
1		97.8103	13.424	0.000	15267866	0.00	BB	7.2		0
2		2.1897	14.045	0.000	341804	0.00	TS	0.0		0
	Totals	100.0000		0.000	15609670					

Figure 35. HPLC of CRA_TG_3027 with a gradient of 5-80% CH_3CN vs 0.1% CF_3COOH in H_2O over 35 minutes.

5.2. MASS SPECTROMETRY

For these analyses, the samples were dissolved in 200 μL of 50% ACN, and two drops of 0.3% formic acid were added. Then they were sent to the Mass Spectrometry Laboratory at the University of Arizona for Low-Resolution MS analysis. The purpose of the formic acid added was to ease the protonation of the molecules during the analysis. Once the compound was ionized by Electrospray Ionization (ESI), it was charged and promoted to the gas phase. As PACAP is a large peptide composed of a lot of basic residues, the signals obtained in the MS do not belong to a single charge, but to multiple charges. This fact is very common in large proteins. For this reason, 3+, 4+, 5+ and 6+ charges can be observed.

The way of characterizing the different peaks consisted of taking the exact mass of each peptide, sum the number of protons and divide it by the charge produced by those protons (m/z). For example, for a 5+ charge, 5 grams have to be summed to the exact mass and divided by 5, then find this m/z in the graphic.

The general Mass Spectra and the isotope distributions were obtained, however, the characterization of the three 23-mers could be done by looking at the general one alone. The complete spectra can be found in Annex 2, 3, and 4.

5.2.1. CRA_TG_3025 MS

The CRA_TG_3025 characterization was done by identifying the peaks corresponding to $(M+3H)^{3+}$, $(M+4H)^{4+}$, $(M+5H)^{5+}$, and $(M+6H)^{6+}$, placed at m/z of 1058.24, 793.93, 635.35, and 529.62 respectively (Table 3, Figure 36).

Table 3. Comparison of the m/z ratio calculated and found in the graphic for CRA_TG_3025.

Exact Mass: 3170.59	m/z Calculated	m/z Found
$(M+3H)^{3+}$	1057.86	1058.24
$(M+4H)^{4+}$	793.64	793.93
$(M+5H)^{5+}$	635.12	635.35
$(M+6H)^{6+}$	529.43	529.62

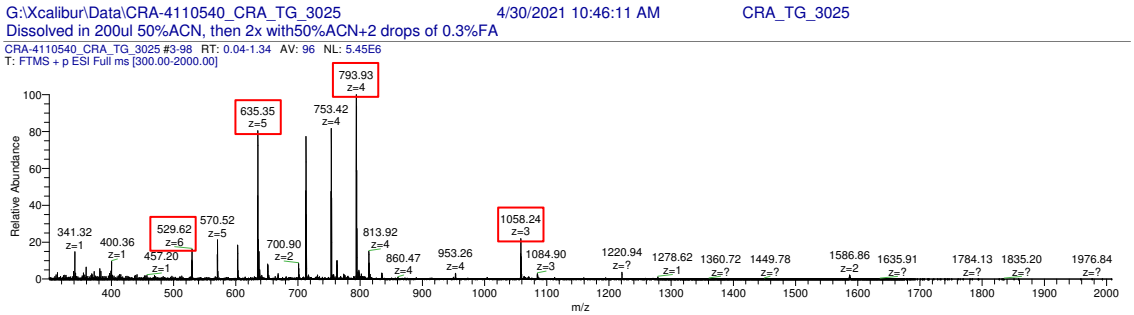


Figure 36. MS spectra of CRA_TG_3025.

5.2.2. CRA_TG_3026 MS

The characterization of CRA_TG_3026 was performed by identifying the peaks corresponding to $(M+3H)^{3+}$, $(M+4H)^{4+}$, $(M+5H)^{5+}$, and $(M+6H)^{6+}$, placed at m/z of 1063.10, 797.26, 638.09, and 532.09 respectively (Table 4, Figure 37).

Table 4. Comparison of the m/z ratio calculated and found in the graphic for CRA_TG_3026.

Exact Mass: 3184.60	m/z Calculated	m/z Found
$(M+3H)^{3+}$	1062.53	1063.10
$(M+4H)^{4+}$	797.15	797.26

$(M+5H)^{5+}$	637.92	638.09
$(M+6H)^{6+}$	531.77	532.09

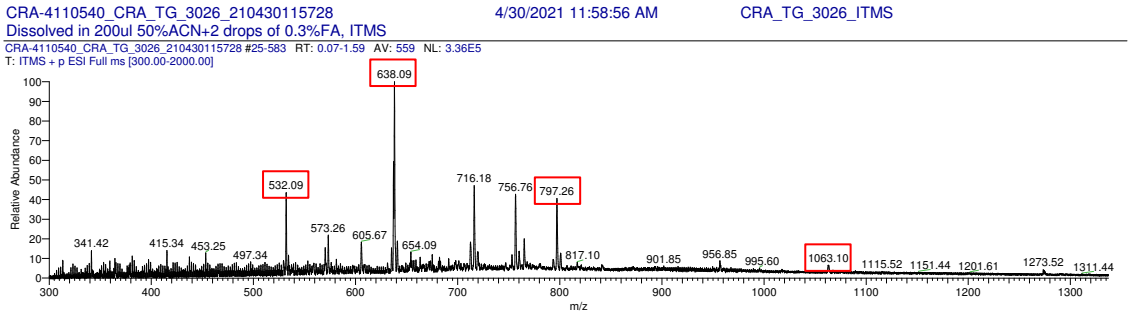


Figure 37. MS spectra of CRA_TG_3026.

5.2.3. CRA_TG_3027 MS

The exact mass of CRA_TG_3027 is the same as CRA_TG_3026, as the only thing that changes is the position of a methyl group. Its characterization was done by identifying the peaks corresponding to $(M+3H)^{3+}$, $(M+4H)^{4+}$, $(M+5H)^{5+}$, and $(M+6H)^{6+}$, placed at m/z of 1063.10, 797.18, 638.01, and 532.01 respectively (Table 5, Figure 38).

Table 5. Comparison of the m/z ratio calculated and found in the graphic for CRA_TG_3027.

Exact Mass: 3184.60	**m/z Calculated**	**m/z Found**
$(M+3H)^{3+}$	1062.53	1063.10
$(M+4H)^{4+}$	797.15	797.18
$(M+5H)^{5+}$	637.92	638.01
$(M+6H)^{6+}$	531.77	532.01

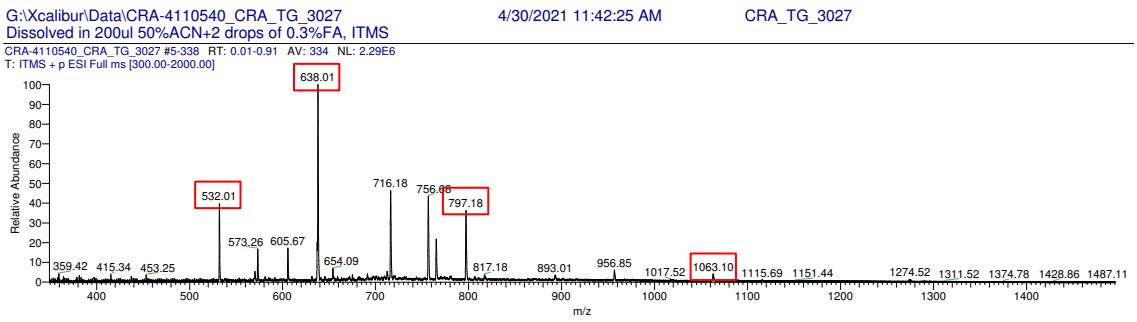


Figure 38. MS spectra of CRA_TG_3027.

5.3. *IN VITRO* RESULTS AND FUTURE EXPERIMENTS

Previews experiments have been done in Dr. Polt's Laboratory Group regarding $PACAP_{27}$ selectivity and affinity, including their syntheses, purification, characterization, and testing both *in vitro* and *in vivo* experimental techniques. Such studies provided promising results regarding PACAP with Gly^4 and Sar^4 selectivity towards PAC_1 receptor, and Ala^4 affinity to $VPAC_1$ and $VPAC_2$. For this reason, the development of a shorter analog of $PACAP_{27}$ was such an important issue, being $PACAP_{23}$ an easier to synthesize drug candidate.

The *in vitro* studies for these three 23-mers were sent for their analysis in Professor John Streicher's laboratory by Parthasaradhi Tanguturi. These facilities are located at the University of Arizona College of Medicine, in specific at the Department of Pharmacology, and will study the functional activity of the glycopeptides through cAMP stimulation, and their stability. The results are obtained in terms of the half-maximal effective concentration (EC_{50}), which is the concentration required to obtain a 50% of the desired effect. In this case, a lower EC_{50} indicates a lower concentration of the peptide needed to stimulate the receptor. The experimental part done by Professor John Streicher's laboratory group can be seen in Annex 5.

These *in vitro* studies have been carried out during June of 2021, and after their completion, the *in vivo* studies in mice will start. A total of six different laboratories will be involved in the analyses of these glycopeptides.

After the publication of the provisional patent of the structures and their performance in vitro, the results obtained could be shared. It has been observed that all $PACAP_{23}$ analogs showed strong efficacy at the primary target, PAC_1. Moreover, all compounds also showed reasonable potency, indicating that the modifications made do not compromise the ability of the ligands to engage the target. Most compounds had low EC_{50}, in terms of nM, being just a bit high for CRA_TG_3027.

With respect to the other two targets, a consistent loss of potency was seen with $VPAC_1$, increasing even more at $VPAC_2$, being consistent with past experiences screening these targets. The full efficacy was maintained. It can be observed that among these three compounds, the most selective ligand towards PAC_1 is CRA_TG_3025, with ~12-fold selectivity versus $VPAC_1$, and ~1080-fold selectivity versus $VPAC_2$. Also, it can be observed that the selectivity of

CRA_TG_3026 to PAC_1 is only ~5 times bigger than the selectivity to $VPAC_1$, and only ~46 bigger than the selectivity towards $VPAC_2$. With this, it can be seen that the selectivity of a ß-turn inducer towards PAC_1 is higher than the one of an α-helix inducer. Although the concentration of CRA_TG_3027 needed to activate PAC_1 was the highest, the selectivity towards PAC_1 versus $VPAC_1$ is almost triple. Lastly, to comment that the results for $VPAC_2$ binding affinity related to CRA_TG_3027 couldn't be calculated.

Table 6. *In vitro* results for binding affinity of $PACAP_{27}$ and the three $PACAP_{23}$ analogs.

Compound	**PAC_1**		**$VPAC_1$**		**$VPAC_2$**	
	EC_{50} (nM)	E_{MAX} (%)	EC_{50} (nM)	E_{MAX} (%)	EC_{50} (nM)	E_{MAX} (%)
$PACAP_{27}$	20, 9.0	100, 100	30, 6.5	100, 100	71, 29	100, 100
CRA_TG_3025	1.8	174	21	118	1943	169
CRA_TG_3026	7	181	38	124	321	139
CRA_TG_3027	67	134	195	107	NC	NC

6. CONCLUSIONS

After the analytical characterization of the three different glycosylated $PACAP_{23}$, it has been demonstrated that the syntheses and purification through SPPS and Preparative HPLC respectively, are viable, presenting high yield for this kind of synthesis technique, together with a purity high enough for *in vitro* and *in vivo* studies, which are the following steps in this research. These future studies will determine the affinity of each of the three drug candidates for each of the three B GPCR receptors, providing a significant advance to this branch of the science.

Once again, the use of automated SPPS for the synthesis of PACAP has resulted feasible, even when presenting complicated sequences such as those with Aspartic Acid. Moreover, the purification technique permitted a proper separation from impurities. Both analytical techniques used to characterize the compounds presented clear results, which in comparison to previous studies done to $PACAP_{27}$, it could be observed that the HPLC peaks obtained for $PACAP_{23}$ were narrower.

The results obtained in the *in vitro* experiments demonstrated a clear selectivity of ß-inducers towards PAC_1 in front of an α-helix inducer, which presents a lower selectivity due to a higher affinity towards $VPAC_1$ and in specific, $VPAC_2$. These results demonstrate the ability to make modified PACAP ligands that will preferentially engage the primary receptor target while developing selectivity versus the two most closely related off-target receptors, all while maintaining efficacy with expected pharmacokinetic benefits.

The fact of working in this field has meant both intellectual and personal development for me by being able to contribute to this noble and innovative initiative that aims to improve the quality of life of people affected by diseases such as Alzheimer's. In particular, having lived this disease so closely through a family member has made me feel proud and excited about the search for a possible cure.

This experience has given me the opportunity of joining the professional and specialized research group at the University of Arizona, being a part of the team, and having even the privilege of appearing as a collaborator and author in a scientific article that is meant to be published in the following months by the MDPI. Attending results discussion meetings every week, synthesizing starting materials like Fmoc protected Serine Disaccharides useful for laboratory

partners, having access to highly qualified facilities, and learning first-hand about this branch of the chemistry has given me a new perspective about the pathway I want to follow in my academic and professional growth.

Acknowledgements to Dr. Polt's Laboratory Group members at the University of Arizona including Dr. Robin Polt, Dr. Lajos Szabò, Christopher Apostol, Hannah Goodman, Alex Marciniak, Wafaa Alabsi, Nick Christie, John Ely, Alejandra Katz, and Natalie Wallace; collaborators in Professor John Streicher's laboratory group; together with my academic tutor at the Universitat Rovira i Virgili, Maria Elena Fernández Gutiérrez.

7. BIBLIOGRAPHY

(1) Shen, S.; Gehlert, D. R.; Collier, D. A. PACAP and PAC1 Receptor in Brain Development and Behavior. *Neuropeptides* **2013**, *47* (6), 421–430. https://doi.org/10.1016/j.npep.2013.10.005.

(2) Bourgault, S.; Chatenet, D.; Wurtz, O.; D. Doan, N.; Leprince, J.; Vaudry, H.; Fournier, A.; Vaudry, D. Strategies to Convert PACAP from a Hypophysiotropic Neurohormone Into a Neuroprotective Drug. *Curr. Pharm. Des.* **2011**, *17* (10), 1002–1024. https://doi.org/10.2174/138161211795589337.

(3) Anglin, B. L. Design and Synthesis of PACAP Based Glycopeptide Analogs; Effects of Glycosylation on Activity and Blood-Brain Barrier Penetration. *ProQuest Diss. Theses* **2014**, 22-23.

(4) Miyata, A.; Jiang, L.; Dahl, R. D.; Kitada, C.; Kubo, K.; Fujino, M.; Minamino, N.; Arimura, A. Isolation of a Neuropeptide Corresponding to the N-Terminal 27 Residues of the Pituitary Adenylate Cyclase Activating Polypeptide with 38 Residues (PACAP38). *Biochem. Biophys. Res. Commun.* **1990**, *170* (2), 643–648. https://doi.org/https://doi.org/10.1016/0006-291X(90)92140-U.

(5) Chartrel, N.; Tonon, M. C.; Vaudry, H.; Conlon, J. M. Primary Structure of Frog Pituitary Adenylate Cyclase-Activating Polypeptide (PACAP) and Effects of Ovine PACAP on Frog Pituitary. *Endocrinology* **1991**, *129* (6), 3367–3371. https://doi.org/10.1210/endo-129-6-3367.

(6) Vaudry, D.; Falluel-Morel, A.; Bourgault, S.; Basille, M.; Burel, D.; Wurtz, O.; Fournier, A.; Chow, B. K. C.; Hashimoto, H.; Galas, L.; Vaudry, H. Pituitary Adenylate Cyclase-Activating Polypeptide and Its Receptors: 20 Years after the Discovery. *Pharmacol. Rev.* **2009**, *61* (3), 283–357. https://doi.org/10.1124/pr.109.001370.

(7) Zhu, L.; Tamvakopoulos, C.; Xie, D.; Dragovic, J.; Shen, X.; Fenyk-Melody, J. E.; Schmidt, K.; Bagchi, A.; Griffin, P. R.; Thornberry, N. A.; Roy, R. S. The Role of Dipeptidyl Peptidase IV in the Cleavage of Glucagon Family Peptides: In Vivo Metabolism of Pituitary Adenylate Cyclase-Activating Polypeptide-(1-38). *J. Biol. Chem.* **2003**, *278* (25), 22418–22423. https://doi.org/10.1074/jbc.M212355200.

(8) Lamine, A.; Poujol de Molliens, M.; Létourneau, M.; Hébert, T. E.; Vaudry, D.; Fournier, A.; Chatenet, D. The Amidated PACAP1–23 Fragment Is a Potent Reduced-Size Neuroprotective Agent. *Biochim. Biophys. Acta - Gen. Subj.* **2019**, *1863* (11), 129410. https://doi.org/10.1016/j.bbagen.2019.08.003.

(9) Alexander, S. P. H.; Benson, H. E.; Faccenda, E.; Pawson, A. J.; Sharman, J. L.; Mcgrath, J. C.; Catterall, W. A.; Spedding, M.; Peters, J. A.; Harmar, A. J.; Collaborators, C.; Alexander, S. P. H. The Concise Guide to PHARMACOLOGY.

Overview. Br. J. Pharmacol. **2013**, *14*, 1449–1458. https://doi.org/10.1111/bph.12444/full.

(10) Chen, Y.; Samal, B.; Hamelink, C. R.; Xiang, C. C.; Chen, Y.; Chen, M.; Vaudry, D.; Brownstein, M. J.; Hallenbeck, J. M.; Eiden, L. E. Neuroprotection by Endogenous and Exogenous PACAP Following Stroke. *Regul. Pept.* **2006**, *137* (1–2), 4–19. https://doi.org/10.1016/j.regpep.2006.06.016.

(11) Martin, B.; Lopez De Maturana, R.; Brenneman, R.; Walent, T.; Mattson, M. P.; Maudsley, S. Class II G Protein-Coupled Receptors and Their Ligands in Neuronal Function and Protection. NeuroMolecular Med. 2005, 7 (1–2), 3–36. https://doi.org/10.1385/nmm:7:1-2:003.

(12) Dhanasekaran, M.; Palian, M. M.; Alves, I.; Yeomans, L.; Keyari, C. M.; Davis, P.; Bilsky, E. J.; Egleton, R. D.; Yamamura, H. I.; Jacobsen, N. E.; Tollin, G.; Hruby, V. J.; Porreca, F.; Polt, R. Glycopeptides Related to β-Endorphin Adopt Helical Amphipathic Conformations in the Presence of Lipid Bilayers. *J. Am. Chem. Soc.* **2005**, *127* (15), 5435–5448. https://doi.org/10.1021/ja0432158.

(13) Stern, J.; Son, S.; Biancardi, V.; Zheng, H.; Sharma, N.; K. P. P. Myocardial Extraction from Newborn Rats. HHS Public Access. *Physiol. Behav.* **2016**, *176* (1), 139–148. https://doi.org/10.1016/j.peptides.2015.01.009.A.

(14) Inooka, H.; Ohtaki, T.; Kitahara, O.; Ikegami, T.; Endo, S.; Kitada, C.; Ogi, K.; Onda, H.; Fujino, M.; Shirakawa, M. Conformation of a Peptide Ligand Bound to Its G-Protein Coupled Receptor. *Nat. Struct. Biol.* **2001**, *8* (2), 161–165. https://doi.org/10.1038/84159.

(15) Doan, N. D.; Bourgault, S.; Dejda, A.; Létourneau, M.; Detheux, M.; Vaudry, D.; Vaudry, H.; Chatenet, D.; Fournier, A. Design and in Vitro Characterization of PAC1/VPAC1-Selective Agonists with Potent Neuroprotective Effects. *Biochem. Pharmacol.* **2011**, *81* (4), 552–561. https://doi.org/10.1016/j.bcp.2010.11.015.

(16) Jwad, R.; Weissberger, D.; Hunter, L. Strategies for Fine-Tuning the Conformations of Cyclic Peptides. *Chem. Rev.* **2020**, *120* (17), 9743–9789. https://doi.org/10.1021/acs.chemrev.0c00013.

(17) Inooka, H.; Ohtaki, T.; Kitahara, O.; Ikegami, T.; Endo, S.; Kitada, C.; Ogi, K.; Onda, H.; Fujino, M.; Shirakawa, M. Biological Magnetic Resonance Data Bank. BMRB Entry 4916 https://bmrb.io/data_library/summary/index.php?bmrbId=4916 (accessed May 5, 2021).

(18) Yoshitaka, U.; Takeshi, T.; Natsuko, G.; Takahisa, I.; Hiroaki, H.; Biological Magnetic Resonance Data Bank. BMRB Entry 11419 https://bmrb.io/data_library/summary/index.php?bmrbId=11419 (accessed May 5, 2021).

(19) Ghosh, D.; Lahiri, P.; Verma, H.; Mukherjee, S.; Chatterjee, J. Engineering β-Sheets Employing: N -Methylated Heterochiral Amino Acids. *Chem. Sci.* **2016**, *7* (8), 5212–5218. https://doi.org/10.1039/c6sc00518g.

(20) Egleton, R. D.; Bilsky, E. J.; Tollin, G.; Dhanasekaran, M.; Lowery, J.; Alves, I.; Davis, P.; Porreca, F.; Yamamura, H. I.; Yeomans, L.; Keyari, C. M.; Polt, R. Biousian Glycopeptides Penetrate the Blood-Brain Barrier. *Tetrahedron Asymmetry* **2005**, *16* (1), 65–75. https://doi.org/10.1016/j.tetasy.2004.11.038.

(21) Li, Y.; Louis, L. S.; Knapp, B. I.; Muthu, D.; Anglin, B.; Giuvelis, D.; Bidlack, J. M.; Bilsky, E. J.; Polt, R. Can Amphipathic Helices In Fl Uence the CNS Antinociceptive Activity of Glycopeptides Related to β - Endorphin? **2014**.

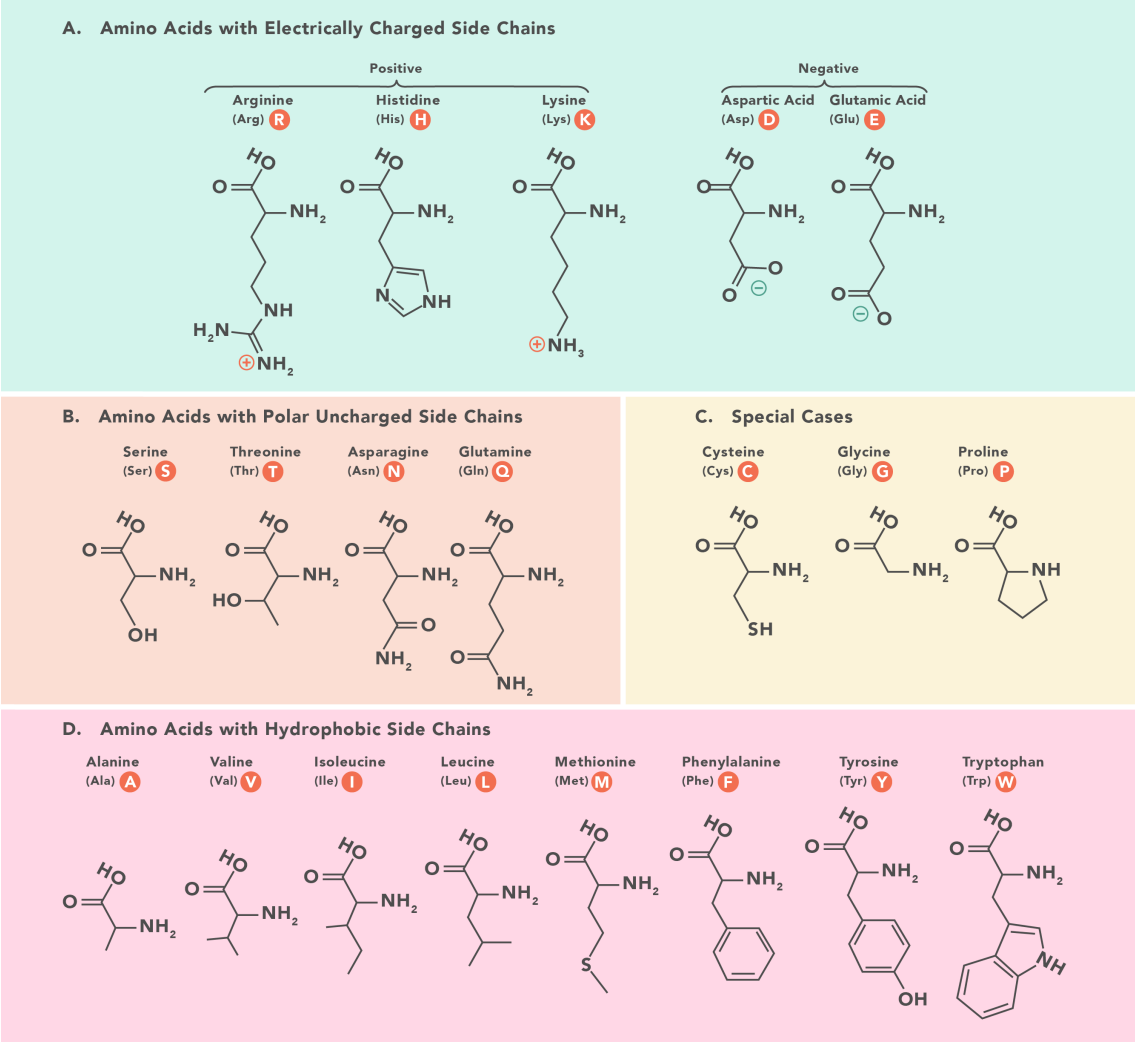
(22) Jaradat, D. M. M. Thirteen Decades of Peptide Synthesis: Key Developments in Solid Phase Peptide Synthesis and Amide Bond Formation Utilized in Peptide Ligation. *Amino Acids* **2018**, *50* (1), 39–68. https://doi.org/10.1007/s00726-017-2516-0.

(23) Peptide Chemistry Portal. Fmoc Deprotection in Peptide Synthesis https://peptidechemistryportal.com/fmoc-deprotection-in-peptide-synthesis/#:~:text=Fmoc (9-fluorenylmethoxycarbony-) group, (besides amino acids coupling). (accessed May 12, 2021).

(24) Neumann, K.; Farnung, J.; Baldauf, S.; Bode, J. W. Prevention of Aspartimide Formation during Peptide Synthesis Using Cyanosulfurylides as Carboxylic Acid-Protecting Groups. *Nat. Commun.* **2020**, *11* (1), 1–10. https://doi.org/10.1038/s41467-020-14755-6.

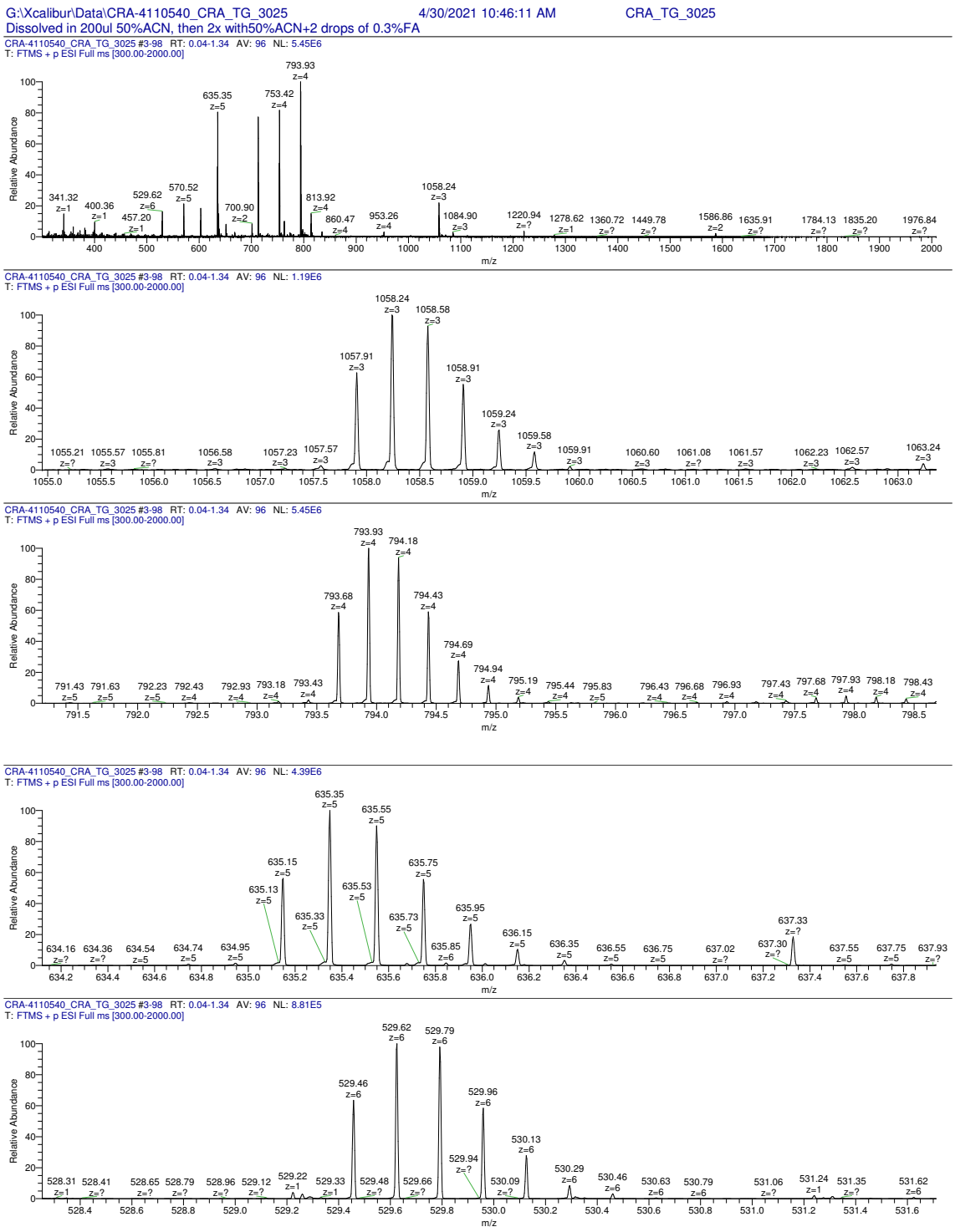
8. ANNEXES

Annex 1: Amino acid classification, structures, and abbreviations.

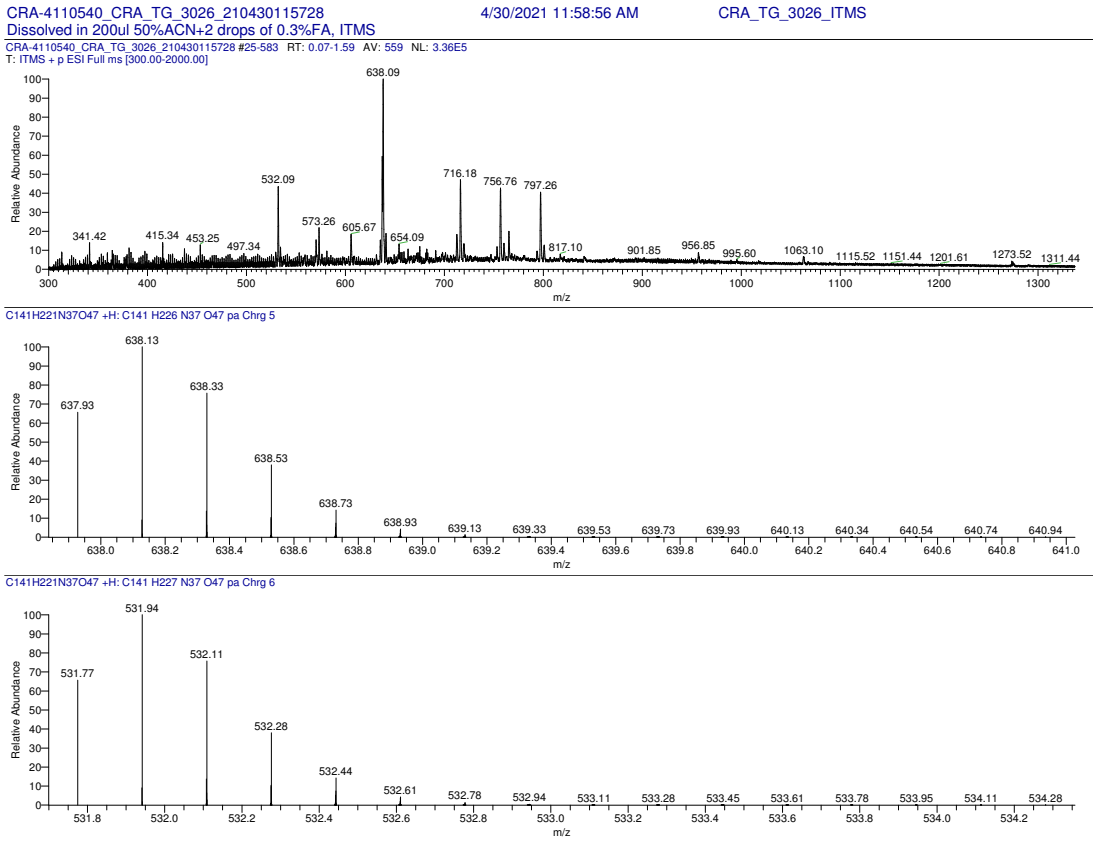


Networks, T. Amino Acids - The Building Blocks of Proteins https://www.technologynetworks.com/applied-sciences/articles/essential-amino-acids-chart-abbreviations-and-structure-324357 (accessed May 26, 2021).

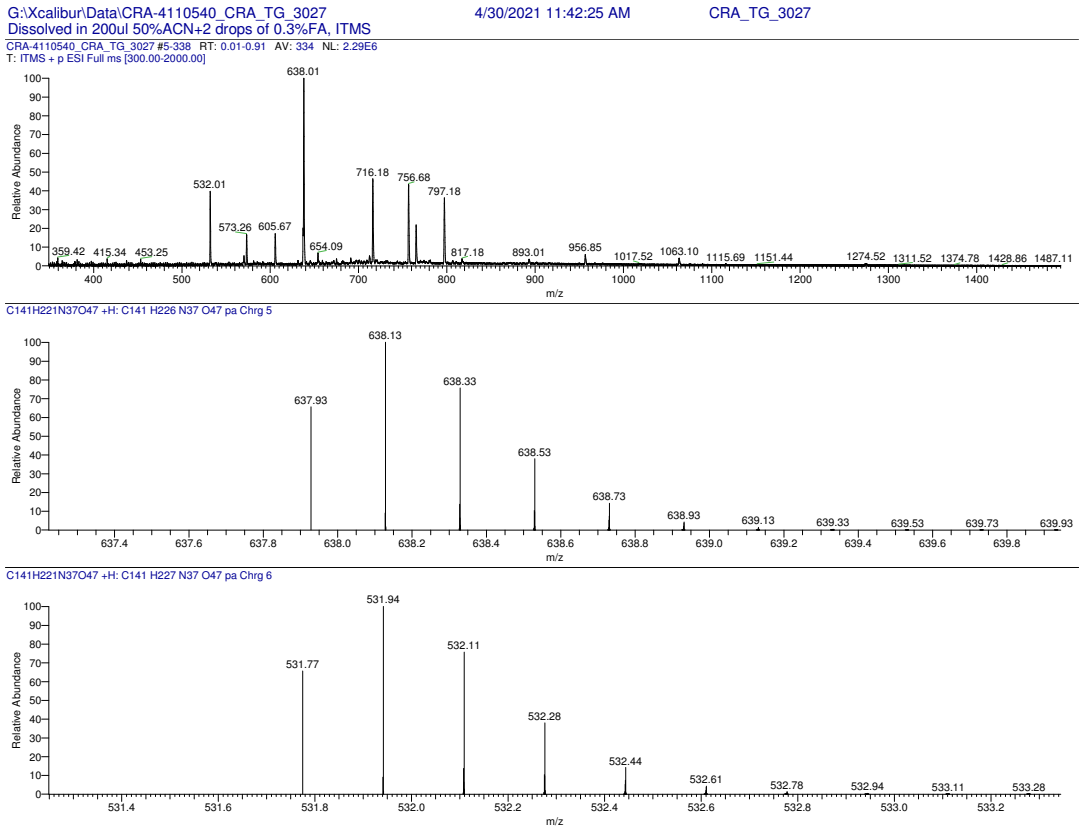
Annex 2: Isotopic Mass Spectra of CRA_TG_3025



Annex 3: Isotopic Mass Spectra of CRA_TG_3026



Annex 4: Isotopic Mass Spectra of CRA_TG_3027



Annex 5. In vitro experiments done by Professor John Streicher's laboratory group at the University of Arizona.

- Cell Culture: CHO cells stably expressing cloned PAC_1, $VPAC_1$, and $VPAC_2$ were produced by electroporation with human PAC_1/$VPAC_1$/$VPAC_2$ N-3xHA tag cDNA constructs (GeneCopoeia). Cells were grown on 10cm dishes in DMEM/F-12 50/50 mix w/ L-glutamine & 15mM HEPES (Corning) containing 10% heat inactivated fetal bovine serum, 100 units/mL penicillin, 100 µg/mL streptomycin, and 500 µg/mL G418 under 5% CO_2 at 37°C. The cells were enriched into high expressing populations using flow cytometry, selecting the top ~2% of expressing cells.

- cAMP Accumulation Assay: At ~80% confluence, cells were plated into 96-well plates (20,000 cells/well) and grown in the same medium and conditions as described above for 24 hrs. The cells were then serum starved for 4 hr. After a 20 min incubation at 37°C with 500 µM 3-Isobutyl-1-methylxanthine (IBMX), serum free medium containing 500 µM IBMX and the appropriate agonists were added and then incubated for 10min at 37°C. The reaction was terminated by removing the medium and adding 60µL of ice-cold assay buffer (50mM Tris-HCl pH 7.4, 100mM NaCl, 5mM ethylenediaminetetraacetic acid [EDTA]). Plates were sealed with boiling mats and then boiled at 95°C for 10 min. Plates were then centrifuged at 4000 rpm, 4°C, for 10 min to remove debris. 50µL of lysate was transferred to a 96-well plate. Lysate was incubated with ~1 pmol 3H-cAMP (PerkinElmer), and 7 µg protein kinase A (Sigma Aldrich) with 0.05% Bovine Serum Albumin (BSA). The assay was incubated at room temperature for 1 hr. The reactions were then harvested onto GF/B filter plates (PerkinElmer) via rapid filtration by a 96-well plate Cell Harvester (Brandel) and washed 3 times with ice-cold water. Filter plates were dried, 40µL of Microscint-PS scintillation cocktail was added to each well, and then counted in a TopCount or Microbeta2 (PerkinElmer) microplate scintillation counter.