


2015

# Investigation into the Biological Importance and Function of Proinsulin C-Peptide

Christina L. Newsome  
newsome34@marshall.edu

Follow this and additional works at: <http://mds.marshall.edu/etd>

 Part of the [Biochemistry Commons](#), and the [Inorganic Chemistry Commons](#)

---

## Recommended Citation

Newsome, Christina L., "Investigation into the Biological Importance and Function of Proinsulin C-Peptide" (2015). *Theses, Dissertations and Capstones*. 958.  
<http://mds.marshall.edu/etd/958>

This Thesis is brought to you for free and open access by Marshall Digital Scholar. It has been accepted for inclusion in Theses, Dissertations and Capstones by an authorized administrator of Marshall Digital Scholar. For more information, please contact [zhangj@marshall.edu](mailto:zhangj@marshall.edu), [martj@marshall.edu](mailto:martj@marshall.edu).

# INVESTIGATION INTO THE BIOLOGICAL IMPORTANCE AND FUNCTION OF PROINSULIN C-PEPTIDE

A thesis submitted to  
the Graduate College of  
Marshall University

In partial fulfillment of  
the requirements for the degree of  
Master of Science

In  
Chemistry  
By  
Christina L. Newsome

Approved by  
Dr. Leslie Frost, Committee Chairperson  
Dr. John Hubbard  
Dr. Bin Wang

Marshall University  
December 2015

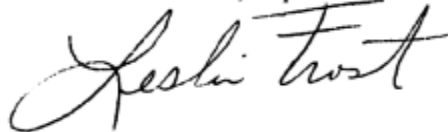
APPROVAL OF THESIS

We, the faculty supervising the work of Christina L. Newsome, affirm that the thesis, *Investigation into the Biological Importance and Function of C-peptide*, meets the high academic standards for original scholarship and creative work established by the Master of Science in Chemistry and the College of Science. This work also conforms to the editorial standards of our discipline and the Graduate College of Marshall University. With our signatures, we approve the manuscript for publication.

Dr. Leslie Frost, Department of Chemistry

Committee Chairperson

Date

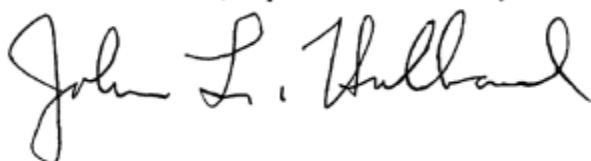


12/7/15

Dr. John Hubbard, Department of Chemistry

Committee Member

Date



12/7/15

Dr. Bin Wang, Department of Chemistry

Committee Member

Date



12/7/15

## **ACKNOWLEDGMENTS**

This thesis would have never been completed without the encouragement and devotion of my loving husband and wonderful parents. I wish to express sincere appreciation to Dr. Frost for all of her help and guidance. In addition, I would like to thank the faculty of the Department of Chemistry for their wonderful support.

## **DEDICATION**

I would like to dedicate this work to my son, Cooper, who is his mommy's joy and inspiration. Mommy loves you to the moon Cooper Logan.

## TABLE OF CONTENTS

Signature Page	ii
Acknowledgements	iii
Dedication	iv
List of Tables	vii
List of Figures	viii
Abstract	xi
Introduction	1
Membrane Proteins	7
G-protein Coupled Receptors	9
Instrumentation	17
Mass Spectrometry	17
MALDI-TOF	18
Peptide Mass Fingerprinting and Mascot	22
Radioimmunoassay	25
Experimental	31
Preparation of Pre-clear Beads	32
Preparation of C-Peptide Affinity Beads	32
Tissue Lysis	33
NP-40 Lysates	33
Cytoplasmic and Nuclear Tissue Lysates	34
Affinity Column Protocol	38
SDS-PAGE	35
Preparation of Tryptic Digests of Protein Samples from SDS-PAGE Bands	37

Analysis of Tryptic Digests of Proteins by MALDI-TOF Mass Spectrometry	37
Results and Discussion	38
Bovine Kidney Trial 1	41
Bovine Kidney Trial 2	42
Bovine Heart	49
Bovine Thymus	54
NP-40 Protein Lysates	55
Conclusions	59
Future	62
Literature References	63
Appendix A	67
Appendix B	68

## LISTS OF TABLES

- |   |   |    |
|---|---|----|
| 1 | Molecular weights of tryptic peptides obtained from the mass spectra of the four protein bands excised from the gel shown in Figure 32. | 57 |
| 2 | Table of proteins identified in this study indicating their cellular location and function.   | 60 |



## LIST OF FIGURES

1	Synthesis of insulin.	1
2	Amino acid sequences of proinsulin C-peptides from different species.	2
3	Chemical structure for proinsulin C-peptide.	3
4	Summary of clinical, in vivo animal and in vitro cellular effects of C-peptide.	4
5	Preliminary model of C-peptide signaling pathways.	6
6	Schematic diagram of the general structure of G-protein linked receptors.	10
7	Primary structure of the hamster $\beta$ AR showing the proposed topology of the seven transmembrane helices.	11
8	Structural topology of the receptors within each family: Family 1, Family 2, and Family 3	14
9	Schematic membrane topology and key structural features of class I and II GPCRs.	16
10	A basic schematic of a mass spectrometer.	18
11	A schematic of the principle of MADLI.	19
12	A more detailed schematic of the principles of MALDI TOF.	20
13	Schematic for TOF Reflectron.	21
14	Conversion of a TOF spectrum to a mass spectrum. TOF measures a particle's m/z.	22
15	Protease Specificity.	24
16	Schematic for a RIA.	27
17	Preparation of NP-40 protein lysates.	33
18	Schematic representation of affinity column protocol.	36

19	Basic experimental protocol for isolating proteins from bovine kidney tissue using C-peptide affinity beads.	39
20	Overview of the Peptide Mass Fingerprint technique for identifying proteins by mass spectrometry.	40
21	SDS-PAGE gel of proteins isolated from bovine kidney tissue in trial 1.	42
22	SDS-PAGE gel of proteins isolated from bovine kidney tissue in trial 2.	43
23	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at 45 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 22.	45
24	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at 51 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 22.	46
25	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~57 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 21.	48
26	Ketone bodies synthesis pathway.	49
27	SDS-PAGE gel of proteins isolated from bovine heart tissue.	50
28	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~150 kDa in lane 4 of the SDS-PAGE gel for bovine heart show in Figure 27.	51
29	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~250 kDa in lane 4 of the SDS-PAGE gel for bovine heart shown in Figure 27.	52

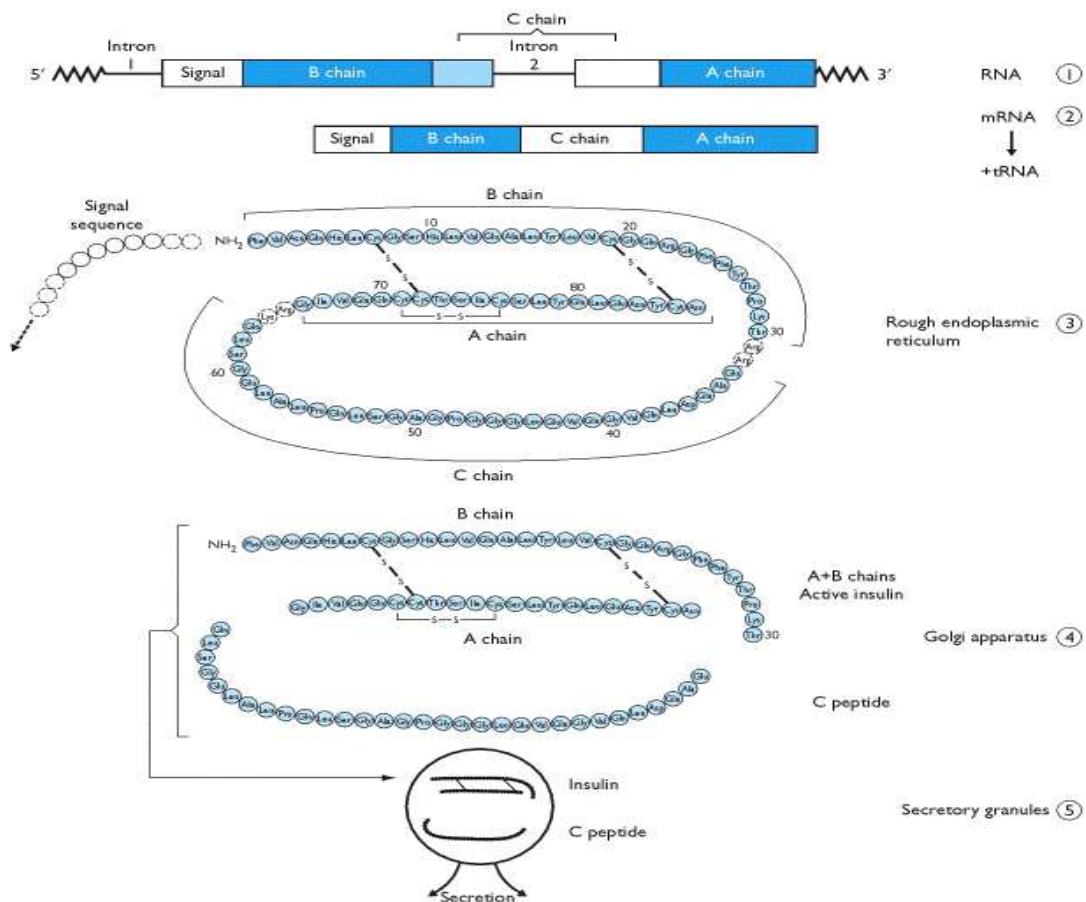
30	MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~80 kDa in lane 4 of the SDS-PAGE gel for bovine heart shown in Figure 27.	53
31	SDS-PAGE gel of proteins isolated from bovine thymus tissue.	54
32	Basic experimental protocol for isolating proteins extracted from bovine thymus using NP-40 lysis buffer with the C-peptide affinity beads.	56
33	SDS-PAGE gel of proteins isolated from bovine thymus using NP-40 lysis buffer.	57
34	SDS-PAGE separation of proteins eluted from the C-peptide beads using human cheek lysate, rat liver lysate and bovine thymus lysate.	58
35	SDS-PAGE separation of proteins eluted from the C-peptide beads using rat liver lysate and bovine thymus lysate.	59

## ABSTRACT

The C-peptide of insulin was thought to be biologically inactive, but recent studies have shown that the C-peptide causes multiple molecular and physiological effects. Evidence has shown that C-peptide binds to a cell surface receptor, probably a G-protein coupled receptor, and that the COOH-terminal pentapeptide is essential for binding and constitutes an active site. For a further understanding of the detailed nature of the physiological effects of C-peptide, the receptor structure needs to be determined. We designed an affinity column using C-peptide to try and gain a better understanding of the biological effects by examining what proteins the affinity column with attached C-peptide would isolate from bovine tissue samples. Since the C-peptide was shown to be internalized in the cytosol and nucleus of kidney cells, we started with cytoplasmic and nuclear tissue lysates obtained from bovine kidney tissue. The isolated proteins were eluted from the beads, and separated by reducing SDS-PAGE. Protein bands of interest were then excised from the gel, digested with trypsin, and analyzed via MALDI-TOF mass spectrometry. We were able to identify a couple of proteins using bovine heart tissue lysates isolated with C-peptide affinity beads. Fatty acid synthase and fatty acyl-CoA ligase were identified. The isolation of both fatty acid synthase and long chain fatty acyl-CoA ligase indicates that C-peptide may play a role in stimulating the production of fatty acids from excess glucose and converting those fatty acids to triacylglycerides for storage inside muscle cells. Our results indicate that C-peptide may be involved in modulating lipid metabolism within cells and may play a role in determining the fate of the excess glucose that enters cells by stimulating the production of fatty acids and conversion of those fatty acids to triacylglycerides for short-term intracellular storage instead of sending the fatty acids to the adipose tissue for long-term storage.

## INTRODUCTION

Insulin is synthesized by pancreatic  $\beta$  cells in a proinsulin form which is then cleaved at two positions to produce the mature insulin and the C-peptide (Figure 1). C-peptide originates from the mid portion of proinsulin, corresponding to the A and B chains of insulin. Both mature insulin and the C-peptide are packaged into secretory granules for storage and released in equimolar quantities into the bloodstream in response to elevated blood glucose levels.<sup>2</sup> It was first thought that C-peptide's main physiological role was in facilitating the folding of the proinsulin molecule in a manner that allowed for the appropriate formation of the disulfide bonds between the cysteine residues of the A- and B-chains of insulin.<sup>1</sup>



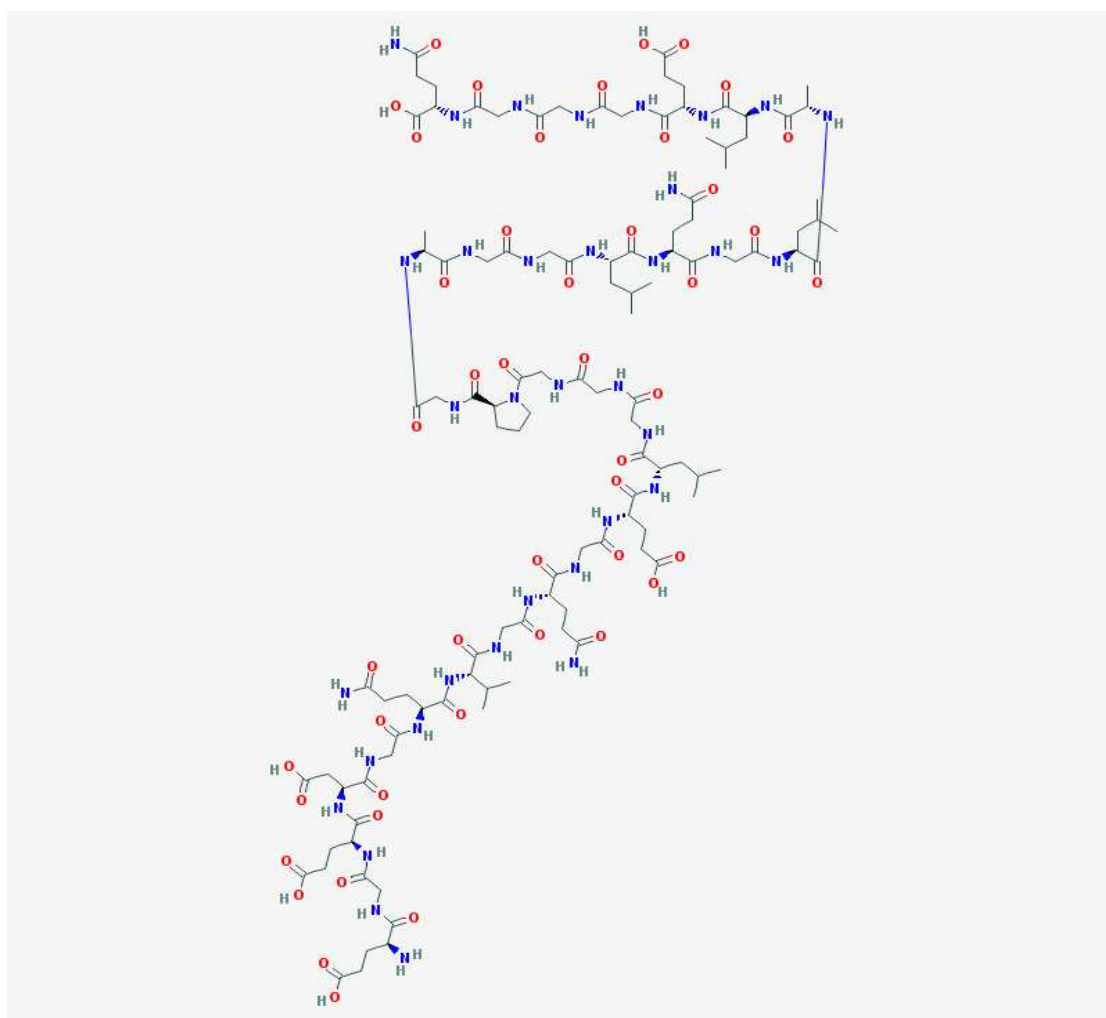
**Figure 1.** Synthesis of insulin.<sup>1</sup>

C-peptide is typically 31 amino acids in length for mammals and contains 4 to 5 acidic residues.<sup>2</sup> Very few species have just a single basic residue (Figure 2).<sup>3</sup> C-peptide is a highly variable structure and generally devoid of aromatic amino acids. C-peptide's sequence variability resembles relaxin, another hormone in the same protein family. Eight residues in C-peptide are highly conserved in mammals; 1 (Glu), 3 (Glu), 6 (Gln) 11 (Glu), 12 (Leu), 26 (Leu), and 31 (Gln).<sup>2</sup> From this conservation, one would conclude that acidic residues, especially (Glu) are important for C-peptide function overall, as well as at specific positions.<sup>3</sup>

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38		
HUMAN	Glu	Ala	Glu	Asp	Leu	Gln	Val	Gly	Gln	Val	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Ser	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Gly	Ser	Leu	Gln									
MONKEY	Glu	Ala	Glu	Asp	Pro	Gln	Val	Gly	Gln	Val	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Ser	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Gly	Ser	Leu	Gln									
HORSE	Glu	Ala	Glu	Asp	Pro	Gln	Val	Gly	Glu	Val	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Leu	Gly	Gly	Leu	Gln	Pro	Leu	Ala	Leu	Ala	Gly	Pro	Gln	Gln									
PIG	Glu	Ala	Glu	Asn	Pro	Gln	Ala	Gly	Ala	Val	Glu	Leu	Gly	Gly	Gly	Leu	Gly		Gly		Leu	Gln	Ala	Leu	Ala	Leu	Glu	Gly	Pro	Pro	Gln									
COW, LAMB	Glu	Val	Glu	Gly	Pro	Gln	Val	Gly	Ala	Leu	Glu	Leu	Ala	Gly	Gly	Pro	Gly	Ala	Gly	Gly	Leu																			
RABBIT	Glu	Val	Glu	Glu	Leu	Gln	Val	Gly	Gln	Ala	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Gly	Leu	Gln	Pro	Ser	Ala	Leu	Glu		Ala	Leu	Gln									
DOG	Glu	Val	Glu	Asp	Leu	Gln	Val	Arg	Asp	Val	Glu	Leu	Ala	Gly	Ala	Pro	Gly	Glu	Gly	Gly	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Gly	Ala	Leu	Gln									
RAT I	Glu	Val	Glu	Asp	Pro	Gln	Val	Pro	Gln	Leu	Glu	Leu	Gly	Gly	Gly	Pro	Glu	Ala	Gly	Asp	Leu	Gln	Thr	Leu	Ala	Leu	Glu	Val	Ala	Arg	Gln									
RAT II	Glu	Val	Glu	Asp	Pro	Gln	Val	Ala	Gln	Leu	Glu	Leu	Gly	Gly	Gly	Pro	Gly	Ala	Gly	Asp	Leu	Gln	Thr	Leu	Ala	Leu	Glu	Val	Ala	Arg	Gln									
GUINEA PIG	Glu	Leu	Glu	Asp	Pro	Gln	Val	Glu	Gln	Thr	Glu	Leu	Gly	Met	Gly	Leu	Gly	Ala	Gly	Gly	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Met	Ala	Leu	Gln									
CHINCHILLA	Glu	Leu	Glu	Asp	Pro	Gln	Val	Gly	Gln	Ala	Asp	Pro	Gly	Val	Val	Pro	Glu	Ala	Gly	Arg	Leu	Gln	Pro	Leu	Ala	Leu	Glu	Met	Thr	Leu	Gln									
DUCK	Asp	Val	Glu	Gln	Pro	Leu	Val	Asn	Gly	Pro		Leu	His	Gly	Glu	Val	Gly	Glu			Leu	Pro	Phe	Gln	His	Glu	Glu		Tyr	Gln										
CHICKEN	Asp	Val	Glu	Gln	Pro	Leu	Val	Ser	Ser	Pro		Leu	Arg	Gly	Glu	Ala	Gly	Val			Leu	Pro	Phe	Gln	Gln	Glu	Glu	Tyr	Glu	Lys	Val									
ANGLERFISH	Asp	Val	Asp	Gln	Leu	Leu	Gly	Phe	Leu	Pro	Pro	Lys	Ser	Gly	Gly	Ala	Ala	Ala	Ala	Gly	Ala	Asp	Asn	Glu	Val	Ala	Glu	Phe	Ala	Phe	Lys	Asp	Gln	Met	Glu	Met	Met	Val		
HAGFISH	Asp	Thr	Gly	Ala	Leu	Ala	Ala	Phe	Leu	Pro	Leu	Ala	Tyr	Ala	Glu	Asp	Asn	Glu	Ser	Gln	Asp	Asp	Glu	Ser	Ile	Gly	Ile	Asn	Glu	Val	Leu	Lys	Ser							
AMPHIOXUS	Ser	Val	Ser	Lys	Arg	Ala	Ile	Asp	Phe	Ile	Ser	Glu	Gln	Gln	Ala	Lys	Asp	Tyr	Met	Gly	Ala	Met	Pro	His	Ile															
MOLLUSC	Asn	Ala	Glu	Thr	Asp	Leu	Asp	Asp	Pro	Leu	Arg	Asn	Ile	Lys	Leu	Ser	Ser	Glu	Ser	Ala	Leu	Thr	Tyr	Leu	Thr															

Figure 2. Amino acid sequences of proinsulin C-peptides from different species.<sup>3</sup>

Many studies have failed to detect the ordered structures in C-peptide. More recent studies have indicated that C-peptide is actually not a random coil, but contains detectable ordered structure both free and attached to insulin and proinsulin. The presence of the C-peptide in proinsulin influences insulin core packing evidently through the formation of some kind of local stable structure at the C-A junction. The central region of C-peptide usually contains a glycine rich segment that would be expected to confer great flexibility, allowing it to bend back upon itself to form a U-shaped or hairpin structure when attached to the insulin chains in native proinsulin (Figure 3).<sup>3-4</sup>



**Figure 3.** Chemical structure for proinsulin C-peptide.<sup>4</sup>

The proinsulin C-peptide was originally thought to be biologically inactive, but the cellular and molecular mechanisms of C-peptide activity are just now being studied. The table below (Figure 4) illustrates the biological activity of proinsulin C-peptide.<sup>5</sup> Specific binding sites for C-peptide on the surface of human renal tubular cells, endothelial cells, and fibroblasts have been observed using fluorescence correlation spectroscopy.<sup>6</sup> The C-peptide has also been observed to bind to a membrane receptor coupled to a pertussis toxin sensitive G-protein.<sup>7-9</sup> Incubation of fibroblast cells with the C-peptide elicited a transient increase in the intracellular  $Ca^{+2}$  concentration and stimulated mitogen-activated protein kinase (MAPK)-dependent signaling pathways.<sup>7</sup> Specific cellular binding of the peptide, its intracellular signaling characteristics and end effects, including its action on eNOS,  $Na^{+}/K^{+}$ -ATPase and several transcription factors, have been established for many cell systems.<sup>5</sup>

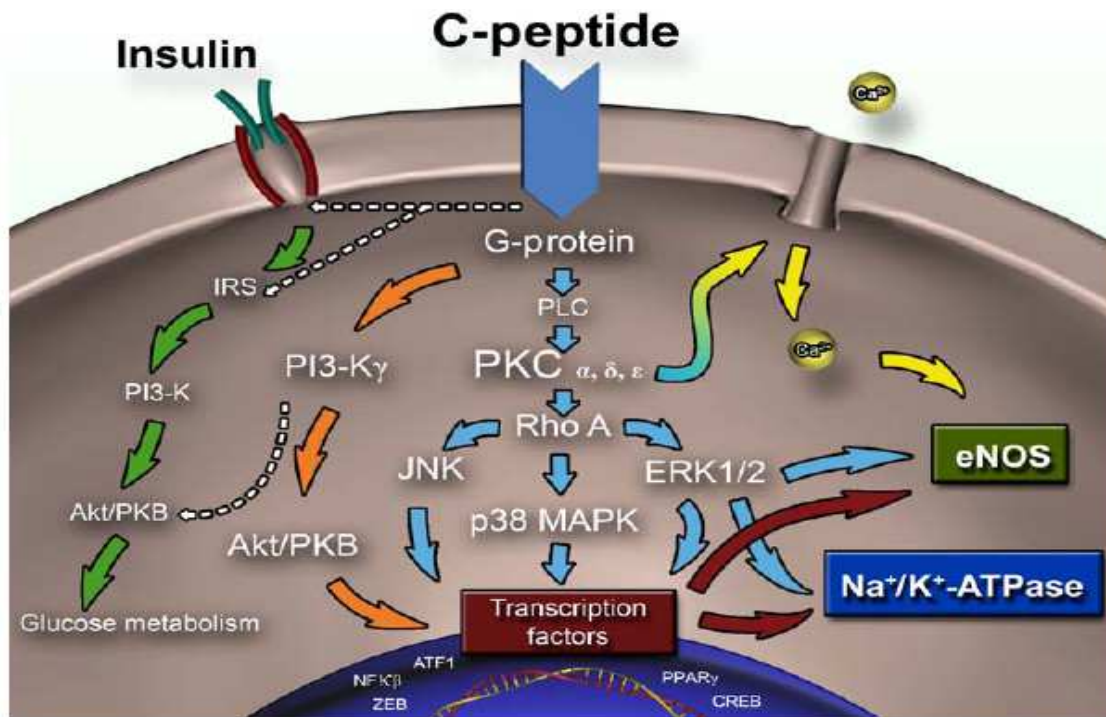
In vivo effects	In vitro effects
<p><b>Renal</b></p> <ul style="list-style-type: none"> <li>Functional reserve ↑</li> <li>Glomerular hyperfiltration ↓</li> <li>Urinary albumin excretion ↓</li> <li>Structural abnormalities ↓</li> </ul> <p><b>Nerve</b></p> <ul style="list-style-type: none"> <li>Conduction velocity ↑</li> <li>Vibration perception ↑</li> <li>Blood Flow ↑</li> <li><math>Na^{+}/K^{+}</math>-ATPase activity ↑</li> <li>Hyperalgesia ↓</li> <li>Structural abnormalities ↓</li> </ul> <p><b>Circulation</b></p> <ul style="list-style-type: none"> <li>Muscle blood flow ↑</li> <li>Skin blood flow ↑</li> <li>Myocardial blood flow and contraction rate ↑</li> <li>QT interval ↓</li> </ul>	<p><b>Membrane interaction</b></p> <ul style="list-style-type: none"> <li>Specific binding in nanomolar range</li> </ul> <p><b>Intracellular signaling</b></p> <ul style="list-style-type: none"> <li>G-protein involvement</li> <li>Intracellular <math>Ca^{2+}</math> ↑</li> <li>PKC, MAPK, and PI-3K ↑</li> <li>NFκB, PPARγ, Bcl2, c-Fos, ZEB ↑</li> </ul> <p><b>End effects</b></p> <ul style="list-style-type: none"> <li>eNOS activity and protein levels ↑</li> <li><math>Na^{+}/K^{+}</math>-ATPase activity and protein levels ↑</li> <li>Cell growth ↑</li> <li>Apoptosis ↓</li> <li>Insulinomimetic effects</li> <li>Anti-thrombotic effects</li> </ul> <p><b>Other</b></p> <ul style="list-style-type: none"> <li>Disaggregation of insulin hexamers</li> </ul>

**Figure 4.** Summary of clinical, in vivo animal and in vitro cellular effects of C-peptide.<sup>5</sup>



Figure 5, pictured below, illustrates an overview of the intracellular responses to C-peptide. Most actions can be blocked by the pre-incubation of cells with pertussis toxin, which indicates the involvement of a  $G_i/G_o$  linked protein in C-peptide signaling. This has been confirmed by direct studies of guanosine triphosphate  $\gamma S$  binding to a  $G_i$  protein after being exposed to C-peptide. When C-peptide is exposed to renal tubular and endothelial cells at physiological concentrations, there is an immediate elevation of intracellular  $Ca^{2+}$  concentrations. The activation of a number of components in the MAPK, in a concentration dependent manner, is consistently observed in all cell types exposed to C-peptide. In muscle cells, C-peptide is also found to mirror the effects of insulin. C-peptide prompts the release of NO in the endothelial cells in a concentration and time dependent manner. The effect occurs within a few minutes and is abolished in  $Ca^{2+}$  free medium in the presence of pertussis toxin or a NOS inhibitor. Increased expression of eNOS mRNA and protein has been demonstrated after endothelial cells were exposed to C-peptide. eNOS expression is enhanced by MAPK dependent transcriptional activation.<sup>5</sup>

$Na^+/K^+$ -ATPase C-peptide uses a stimulatory effect on  $Na^+/K^+$ -ATPase activity and protein expression in renal tubular cells. This effect is concentration dependent, could be blocked by pertussis toxin, and is dependent on  $Ca^{2+}$ . Secondary effects of the decreased red cell  $Na^+/K^+$ -ATPase activity are impaired deformability of the cells, but corrected after exposure to C-peptide. The evidence confirms support for a direct relationship between C-peptide levels and  $Na^+/K^+$ -ATPase activity in renal and nerve tissue and red blood cells under in vitro and in vivo conditions.<sup>5</sup>



**Figure 5.** Preliminary model of C-peptide signaling pathways. The influence of C-peptide on eNOS and Na<sup>+</sup>/K<sup>+</sup>-ATPase include both activation and induction. Dashed lines indicate insulinomimetic signaling demonstrated in muscle cells.<sup>5</sup>

C-peptide also causes activation and DNA binding of several transcription factors. C-peptide also elicits decreased surface expression of the cell adhesion molecules P-selectin and intercellular adhesion molecule 1 on vascular endothelium, inhibiting leucocyte-endothelium interactions. There has also been a reported stimulatory influence on cell proliferation for renal tubular and neuroblastoma cells. In neuroblastoma cells, there was activation of phosphoinositol 3-kinase and p38 MAPK. After C-peptide exposure, this resulted in enhanced expression and translocation of NFκB.<sup>5</sup>

It has been shown that the COOH-terminal pentapeptide from the C-peptide appears to mediate the effects of the C-peptide.<sup>6</sup> The binding of the pentapeptide occurs in the low nanomolar concentration range, and its interaction fits the classic ligand-receptor concept. For a

further understanding of the detailed nature of the physiological effects of C-peptide, the receptor structure will have to be determined. An experimental protocol was set up (see attached) in which the C-peptide was used to isolate any proteins capable of binding to C-peptide.

### **Membrane Proteins**

Cells are essential to all known life forms, with their main function being to sense and interact with their environment or other cells and produce physiological responses. These responses are vital to the proper functioning of cells. Cells need to process different signals inside and outside of the cell. Signaling proteins have progressed to sense signals directly or binding to other proteins. Different cells are programmed to sense different signals depending on the proteins expressed in their plasma membrane. These membrane proteins enable signal transmission, and signal modulation and diversification by undergoing conformational changes through protein-protein interactions.<sup>10</sup>

Cellular signaling processes determine how a specific cell in an organism behaves. Cellular signaling refers to highly evolved networks of signaling events that allow a cell to function. These networks or cascade of events are a series of biochemical processes in which each process is initiated by the appearance of a signal. It is then followed by sensing, processing and transmission as another signal for the next downstream process in the signaling cascade. The signal may be inside or outside the cell. Plasma membrane proteins usually sense and process the extracellular signals, while soluble proteins or membrane proteins process intracellular signals.<sup>10</sup>

There are a number of different types of signaling. In endocrine signaling, signal molecules or hormones are released by a cell and travel long distances causing an effect in a different part of the organism. The processing of light, taste, and smell are considered endocrine. Paracrine signaling is a short range version of endocrine signaling. In paracrine signaling, the

signal is produced by a cell that is sensed locally and processed by the neurons. In juxtacrine signaling, the signal is membrane bound to one cell and is sensed by a receptor that is on an adjacent cell. Autocrine signaling is when cells release a signal outside of the cell. The signal is sensed by a membrane protein and the same cell leading to self-stimulation, such as a breast cancer cell releasing transforming growth factor alpha that interacts with its epidermal growth factor. A signal generated inside the cell and sensed by another receptor outside of the cell is known as intracrine signaling. Electrical signaling is a specialized process that propagates an electrical potential along the length of the cell and occurs on a long spatial scale.<sup>10</sup>

There is an immense diversity of signals. Chemical signals range from very small, like oxygen, to peptides and large proteins. Spatial separation of signaling cascades is achieved by cells expressing different receptors on the surfaces, as well as inside the cell. Transmembrane signal transduction is the dominant component of cellular signaling, as it enables a cell to convert an extracellular signal into one or more intracellular signals and responses. There are three main classes of membrane proteins that enable transmembrane signal transduction.<sup>10</sup>

ICRs are the first membrane class and are responsible for sensing neurotransmitter molecules or voltage gradients across the membrane. As ICRs bind to the signal molecules or sense the membrane potential, they undergo conformational changes that open or close a channel and allow specific ions to cross through the plasma membrane.<sup>10</sup>

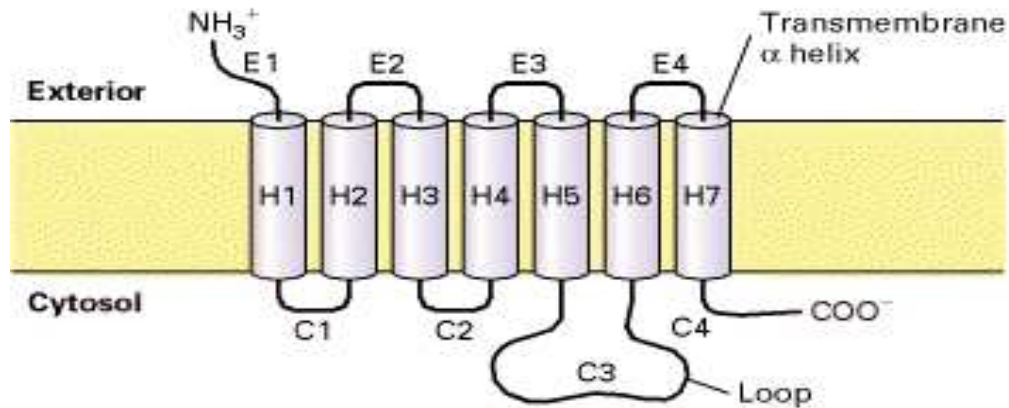
ECRs are a diverse class of transmembrane proteins that contain extracellular ligand binding sites and an intracellular catalytic/enzyme-binding site with a guanylyl cyclase, phosphatase, and serine/threonine kinase or tyrosine kinase activity.<sup>10</sup>

The last class forms the largest superfamily of membrane proteins; GPCRs. By activation from a set of extracellular signals, GPCRs undergo conformational changes, which are

transmitted to cytoplasmic G-proteins and  $\beta$ -arrestins for downstream signal transmission and potential diversification for a physiological response. GPCRs are integral membrane proteins. They have an extracellular N-terminus with seven transmembrane helices connected by loop regions and use their N-terminus, extracellular loops and extracellular facing transmembrane portions to sense their signals. They interact with the most diverse set of signals from sensory signals (vision, taste, smell, pheromones, etc) to large signal molecules, such as other proteins.<sup>10</sup>

### **G-Protein Coupled Receptors**

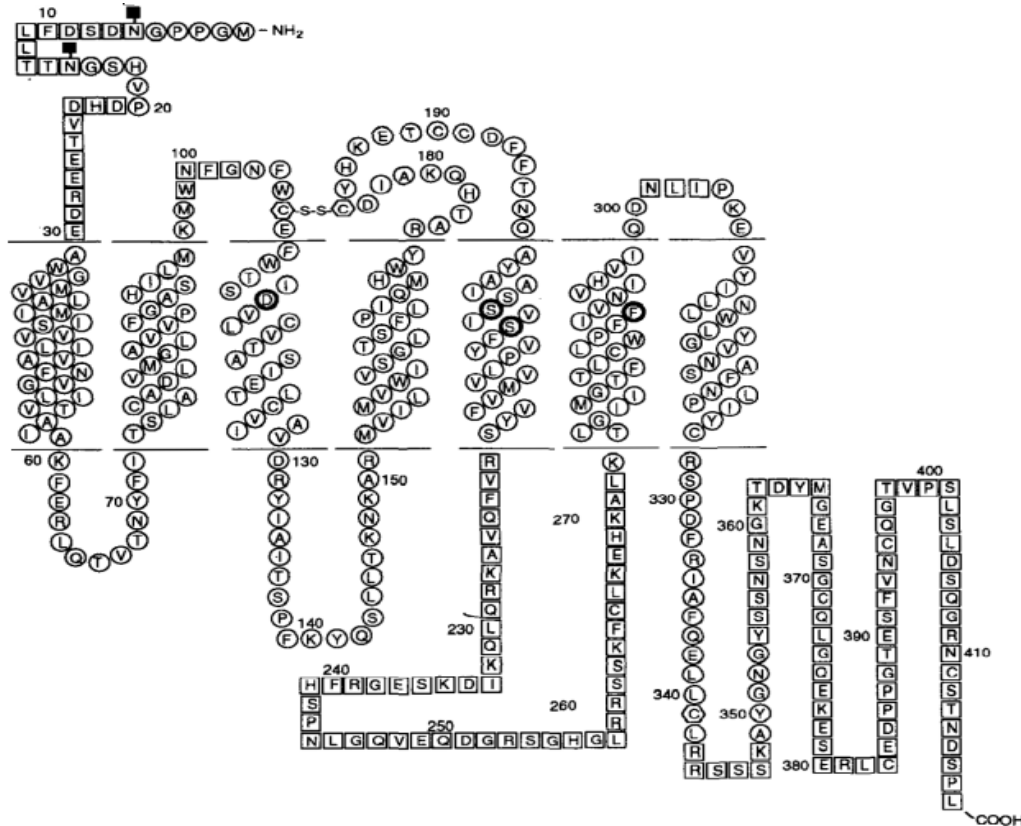
Cells have the ability to process large quantities of information transmitted by extracellular signals, such as hormones, neurotransmitters, autacoids, growth factors, odorants, and by physical signals, such as light. A variety of cell-surface receptors mediate their intracellular actions by a pathway that involves activation of one or more G-proteins.<sup>11-12</sup> GPCRs are integral membrane proteins. They have seven transmembrane helices that are connected by three extracellular loops and three intracellular loops, as illustrated below in Figure 6.<sup>13</sup> GPCRs constitute the largest family of cell surface receptors, with more than a thousand different receptors involved in signal transmission.<sup>12</sup>



**Figure 6.** Schematic diagram of the general structure of G-protein linked receptors. All receptors of this type contain seven transmembrane  $\alpha$ -helical regions. The loop between  $\alpha$  helices 5 and 6, and in some cases the loop between helices 3 and 4, which face the cytosol, are important for interactions with the coupled G-protein. E1-E4 = extracellular loops; H1-H7 = transmembrane domains; C1-C4 = cytosolic loops.<sup>13</sup>

G-protein coupled receptors have a common signaling pathway. The agonist binds to its specific receptor on the cell's surface and causes a conformational change in the receptor. This conformational change allows the interaction with the heterotrimeric G-protein in the cell membrane, forming a high-affinity agonist-receptor G-protein complex. The receptor G-protein interaction catalyzes guanine nucleotide exchange on  $\alpha$  subunit of the G-protein. This leads to the dissociation of the G-protein complex into  $\alpha$  and  $\beta\gamma$  subunits. The GTP-bound form of the G-protein is the activated form. This activated form also dissociates from the receptor and activates an effector protein that modulates levels of intracellular second messengers and transduces signal. The dissociation of the G-protein from the receptor also reduces the affinity of the receptor for the agonist. There have been at least 15 different G-proteins  $\alpha$  subunits, 5  $\beta$  subunits, and 5  $\gamma$  subunits identified to date. These proteins mediate the inhibition of a diverse collection of enzymes and ion channels, including adenylyl cyclase, guanylyl cyclase, phospholipases C and A<sub>2</sub> and Ca<sup>2+</sup> and K<sup>+</sup> channels.<sup>12</sup>

G-protein coupled receptors share a structural homology, reflecting their common mechanism of action.<sup>11</sup> Past molecular cloning efforts have revealed that this large protein receptor family is characterized by seven hydrophobic stretches of 20-25 amino acids. They are predicted to form transmembrane  $\alpha$  helices and are connected by alternating extracellular and intracellular loops; an example is shown in Figure 7.<sup>12</sup>



**Figure 7.** Primary structure of the hamster  $\beta$ AR showing the proposed topology of the seven transmembrane helices. The extracellular domain is at the top of the figure.<sup>12</sup>

G-protein coupled receptors contain a number of conserved cysteine residues. It is apparent that some of these conserved residues play a role in receptor structure. There are two highly conserved Cys residues in the second and third extracellular loops of the receptors, as shown previously in Figure 7. If a Val is substituted for either of the conserved cysteines, a destabilization of the tertiary structure of both rhodopsin and  $\beta$ AR occurs. This destabilization suggests that these two Cys residues have an important role in maintaining the active

conformation of the receptor. Analysis of rhodopsin has shown that these two residues are involved in an intramolecular disulfide bridge that links the second and third intracellular loops of the receptors. This puts constraint on the conformation of the extracellular domain of the protein. The invariance of the Cys residues suggest that the disulfide bridge is a common structural feature to the entire family of G-protein coupled receptors.<sup>12</sup>

GPCRs are integral membrane proteins that have seven transmembrane helices connected by three extracellular loops and three intracellular loops. With ~ 800 GPCRs identified, they form the largest superfamily in the human genome to date.<sup>10</sup> There are three additional groups or classes of GPCRs, class II, III, and IV receptors. These three groups do not show the conservation of amino acids characteristic of class I receptors.<sup>11</sup> GPCRs can further be organized into six families: rhodopsin, secretin, glutamate, frizzled, taste 2 and adhesion.<sup>10</sup>

The rhodopsin family is also known as Class A or Family 1. This is a diverse family dominating the human GPCRs with ~670 out of ~800 total. This family can further be categorized into four subfamilies:  $\alpha$ ,  $\beta$ ,  $\gamma$ , and  $\delta$ . The  $\alpha$  subfamily includes the light-sensing rhodopsin receptor, biogenic amine receptors, and cannabinoid and prostanoid receptors as well as others. The  $\beta$  subfamily is mainly consistent of peptide-binding proteins. The  $\gamma$  subfamily receptors bind to peptides or lipid-like molecules, such as these examples, chemokine, angiotensin, somatostatin and opioid receptors. The  $\delta$  subfamily is dominated by olfactory receptors (~388 out of ~670) and also contains purinergic and glycoprotein-binding receptors.<sup>10</sup>

Next is the secretin/adhesion family, also known as Class B or Family 2. Secretin receptors bind to peptide hormones, while the adhesion receptors bind to extracellular matrix molecules based on the knowledge of receptors thus far.<sup>10</sup>

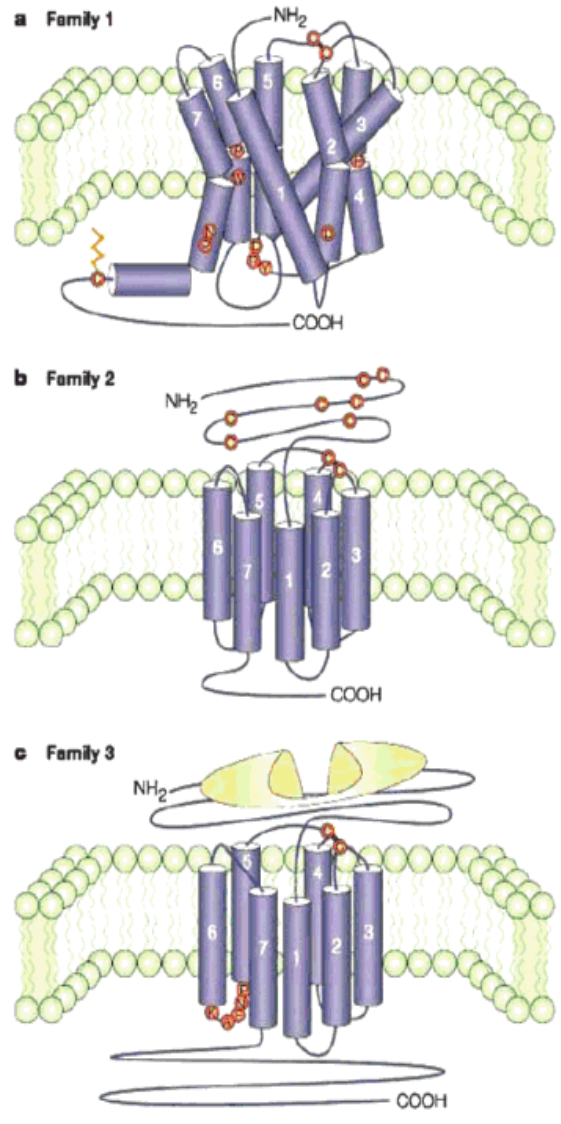


The glutamate family, or Class C or Family 3, consists of metabotropic glutamate receptors,  $\gamma$ -aminobutyric acid B receptors, sweet and umami taste receptors and calcium-sensing receptor. One of the two taste receptor monomers combines with a third monomer to form functional heterodimers for sweet taste or umami taste.<sup>10</sup>

The frizzled family consists of ~10 frizzled receptors that bind to Wnt glycoproteins and a smoothed receptor that seems to function without binding to a ligand. The Taste 2 family consists exclusively of ~25 bitter taste receptors, which share the sensing of different bitter tastants with a different subset of receptors. These taste receptors have recently also been found in the gastrointestinal tract. Their function in the gut is not known, but their activation has been shown to activate gut hormonal receptors.<sup>10</sup>

The structural topologies of these different classes appear to be similarly based on structural and sequence analysis. Figure 8 shows Families 1,2, and 3. Family 1 receptors share some common sequence motifs, as well as some conserved prolines that are usually found in the middle of many transmembranes producing kinks in their helices. Small molecule ligands typically bind in the extracellular facing half of the transmembrane regions, while peptides and proteins typically bind to the extracellular loops and the N-terminus. A highly conserved disulfide bridge is found between the cysteines in the ECL2 and top of TM3. Family 2 receptors do not share any sequence motifs with Family 1. This family of receptors has a long N-terminal ectodomain that bind to ligands and contain numerous conserved cysteines. Cysteine conservation helps the long N-terminus to form a stable tertiary structure. Family 3 receptors have a long N-terminus and long C-terminus. Most of the Family 3 receptors use their long N-terminus to bind to endogenous ligands forming a binding pocket, sometimes referred to as a venus fly trap.<sup>10</sup>

G-protein-coupled receptors (GPCRs) can be divided phylogenetically into six families (see the [GPCR database](#) online). Schematic representations of receptor monomers showing some key structural aspects of the three main families are shown. Family 1 (panel a; also referred to as family A or the rhodopsin-like family) is by far the largest subgroup and contains receptors for odorants, small molecules such as the catecholamines and amines, some peptides and glycoprotein hormones. Receptors of family 1 are characterized by several highly conserved amino acids (some of which are indicated in the diagram by red circles) and a disulfide bridge that connects the first and second extracellular loops (ECLs). Most of these receptors also have a PALMITOYLATED cysteine in the carboxy-terminal tail. The recent determination of the crystal structure of rhodopsin has indicated that the transmembrane (TM) domains of family 1 receptors are 'tilted' and 'kinked' as shown. Family 2 or family B GPCRs (panel b) are characterized by a relatively long amino terminus that contains several cysteines, which presumably form a network of disulfide bridges. Their morphology is similar to some family 1 receptors, but they do not share any sequence homology. For example, the family 2 receptors also contain a disulfide bridge that connects ECL1 and ECL2, but the palmitoylation site is missing, the conserved prolines are different from the conserved prolines in the family 1 receptors and the DRY (aspartic acid, arginine, tyrosine) motif adjacent to TM3 is absent. Little is known about the orientation of the TM domains, but — given the divergence in amino-acid sequence — it is probably quite dissimilar from that of rhodopsin. Ligands for family 2 GPCRs include hormones, such as **glucagon**, **gonadotropin-releasing hormone** and **parathyroid hormone**. Family 3 (panel c) contains the metabotropic glutamate, the  $\text{Ca}^{2+}$ -sensing and the  $\gamma$ -aminobutyric acid ( $\text{GABA}_B$ ) receptors. These receptors are characterized by a long amino terminus and carboxyl tail. The ligand-binding domain is located in the amino terminus, which is often described as being like a 'Venus fly trap'. Except for two cysteines in ECL1 and ECL2 that form a putative disulfide bridge, the family 3 receptors do not have any of the key features that characterize family 1 and 2 receptors. A unique characteristic of the family 3 receptors is that the third intracellular loop is short and highly conserved. Although the structure of the amino terminus is well characterized, similar to the family 2 receptors, little is known about the orientation of the TM domains.



**Figure 8.** Structural topology of the receptors within each family: Family 1, Family 2, and Family 3.<sup>10</sup>

G-protein coupled receptors function at the heart of inter- and intra-cellular regulation, and interact with extracellular signals, usually by binding to small signaling molecules. This ligand binding induces a change in conformation of the receptor that is transmitted to the cytoplasmic face of the protein, enabling a coupling of the cytoplasmic face with an intracellular heterotrimeric G-protein. The intracellular G-protein acts as an intracellular signal by activating

or inhibiting intracellular enzymes. This model of cellular communication is so successful that GPCRs are now used to enable the senses of taste, smell, and vision, and to control a myriad of intracellular signaling systems. Around 1,000 of these receptors are thought to be in the human genome. Diseases such as blindness, obesity, inflammation, depression, and hypertension among others, can be linked to malfunctioning GPCRs. Interestingly, around half of the drug targets in the pharmaceutical industry are also GPCRs.<sup>14</sup>

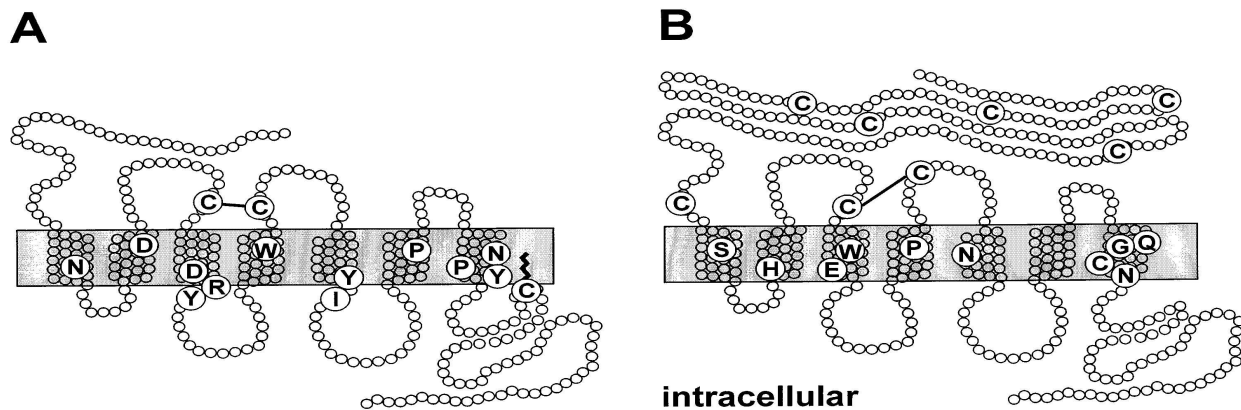
Hormones and neurotransmitters are extracellular signaling molecules that interact with a class of membranous receptors characterized by a uniform molecular architecture of seven transmembrane  $\alpha$ -helices linked by extra- and intracellular peptide loops. Binding of diverse agonists to heptahelical receptors leads to activation of a limited repertoire of heterotrimeric G-proteins, forwarding the signal to intracellular effectors such as enzymes and ion channels. Proper functioning of a G-protein coupled receptor is based on structural determinants, which are ultimately responsible for receptor folding, trafficking and transmembrane signaling.<sup>11</sup>

An essential prerequisite for the function of multicellular organisms is a fine-tuned communication between the individual cells. These cells have the ability to process an enormous amount of information related to them by hormones, neurotransmitters, autocaids, growth factors, and odorants, which are all extracellular signals, as well as physical signals such as light. Most of these signals do not enter the cell, but affect membranous receptors which depend on heterotrimeric G-proteins as transducer molecules.<sup>11</sup>

Even though GPCRs have a large diversity, essentially the global structures are quite uniform. The  $\alpha$ -helically arranged TMs are connected by three alternating extra- (e1-e3) and intracellular (i1-i3) peptide loops. Integral membrane proteins, such as GPCRs, are partially buried in the non-polar environment of the lipid bilayer. The correct integration and orientation

is guided by complex translocation machinery residing in the endoplasmic reticulum (ER). Following an initial translocation of the N-terminal receptor portion into the ER lumen, the folding procedure takes place in two stages.<sup>11</sup>

In stage I, hydrophobic  $\alpha$ -helices are established across the lipid bilayer and protein folding is predominantly driven by the hydrophobic effect. The hydrophobic amino acid residues are oriented towards the core of the TM bundle to reduce the exposure to the lipid environment. The TMs form a tightly packed, ring-like structure once they are inside the plasma membrane. Stage II, consists of a functional tertiary structure that is formed by establishing specific helix-helix interactions. The identification of such specific interhelical contact sites provides information about the relative orientation of the different helices towards each other.<sup>11</sup>



**Figure 9.** Schematic membrane topology and key structural features of class I and class II GPCRs. Seven transmembrane domains are connected by three extra- and three intracellular loops. Structural elements such as the DRY motif at the cytoplasmic end of the TM3 and NPXXY motif in TM7 are highly conserved in the rhodopsin-like GPCR family, the class I GPCRs (A). The glucagon/secretin-like family, class II GPCRs (B), is characterized by five conserved cysteines within the large extracellular domain. Further conserved amino acid residues of both receptor families (in the one-letter code) are shown in enlarged circles.<sup>11</sup>

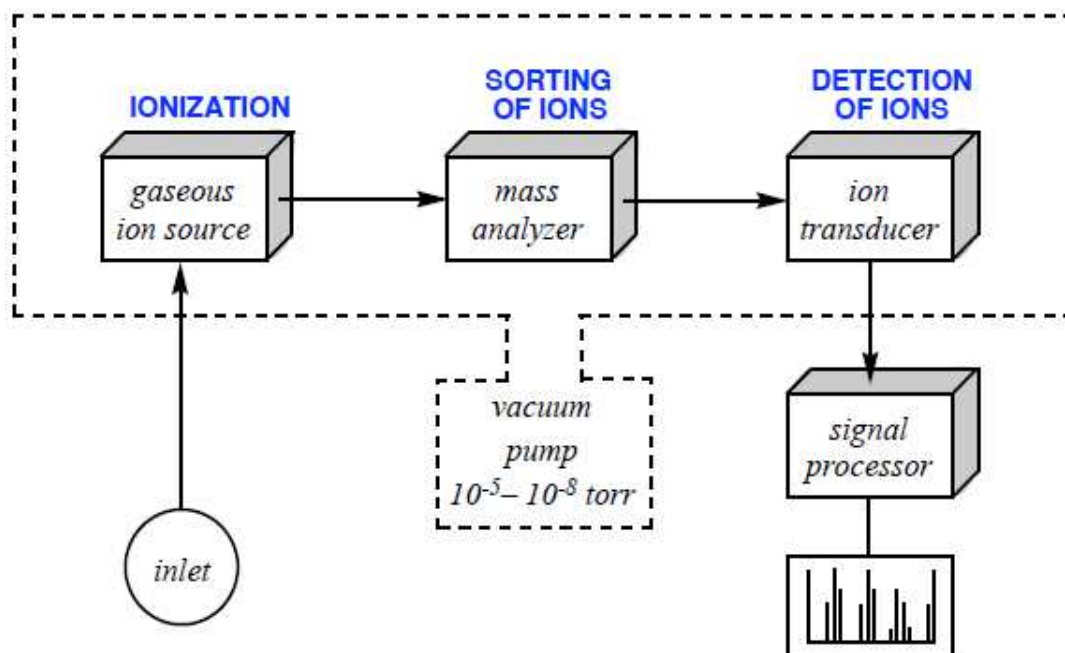
GPCR trafficking is defined as the process of ensuring proper ER insertion. Post-translational modification and transport is realized as a complex biological phenomenon. Multiple membrane targeting signals are localized within the proximal (TM1-5) and distal (TM6-

7) transmembrane portion as shown for the  $\alpha_{2A}$  adrenergic receptor in polarized renal epithelial cells. Small changes to GPCRs, far away from the C-terminus, can often lead to complete retention in the cell interior. Mutations of charged amino acid residues can cause protein retention and degradation in the ER when placed within a TM of an integral plasma membrane protein. It is assumed that the proper installment of the structure is needed to pass a type of intracellular quality control.<sup>11</sup>

## INSTRUMENTATION

### Mass Spectrometry

Mass spectrometry is an important method for characterizing proteins and a technique used frequently in proteomics. A mass spectrometer is an instrument that uses analytical properties for measuring  $m/z$  of ions.<sup>15</sup> A basic schematic of a mass spectrometer consists of four parts: a sample inlet, an ion source, a mass analyzer, and a detector (Figure 10). A sample inlet includes high-performance liquid chromatography (HPLC), gas chromatography (GC), or capillary electrophoresis. Examples of ion sources are MALD, ESI, fast atom bombardment (FAB), electron impact or direct chemical ionization. The mass analyzers include a TOF, a quadrupole, an ion trap, and a magnetic sector or ion cyclotron Fourier transform, the latter which is also a detector. Conversion dynode, electron multiplier, microchannel plate and arrays are all common detectors.



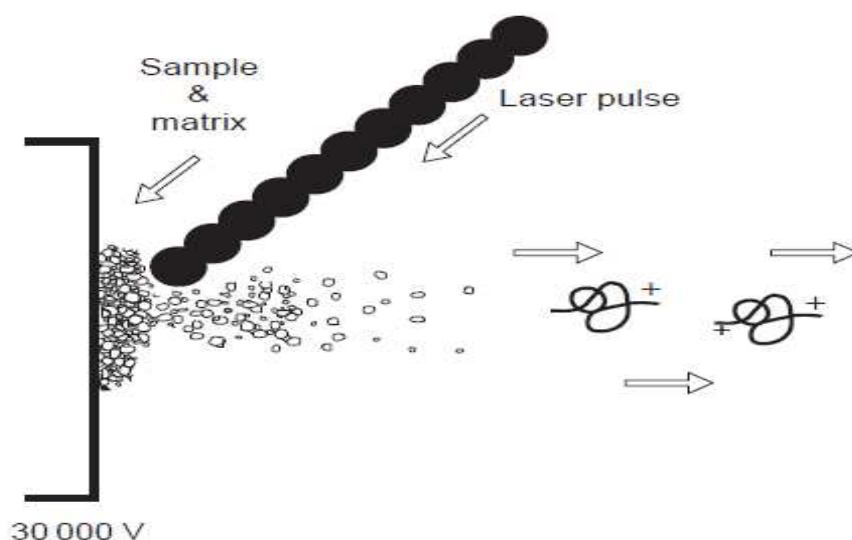
**Figure 10.** A basic schematic of a mass spectrometer.<sup>15</sup>

A sample is introduced using an inlet. Once inside the instrument, an ionization source applies an electrical charge converting the molecules into ions. Two primary methods of ionization of whole proteins are electrospray (ESI) and matrix-assisted laser desorption/ionization (MALDI). Ions are then electrostatically propelled into the mass analyzer, generally including three types: a linear time-of-flight (TOF), a TOF reflectron, and a Fourier transform mass analyzer (FTMS). The ions are separated based on their differing  $m/z$  within the analyzer. The detector converts the ion energy into electrical signals that are transmitted to a computer program producing a mass spectrum (a graph charting the intensity v.  $m/z$ ).<sup>16</sup> For the purpose of this paper, further detail will surround Matrix Assisted Laser Desorption Ionization Time-of-Flight (MALDI-TOF).

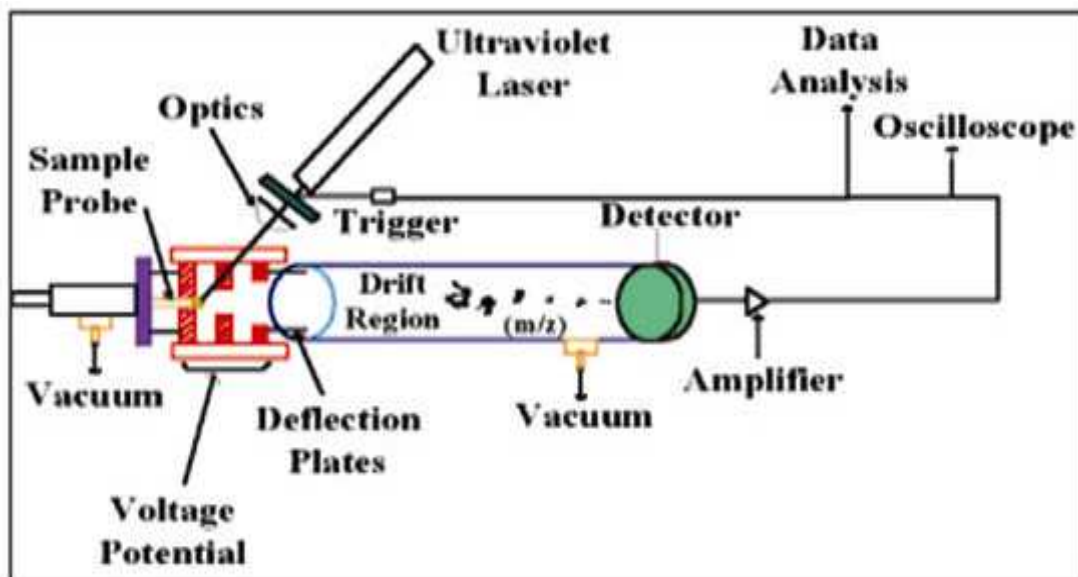
### **MALDI-TOF**

MALDI MS was first introduced by Hillenkamp and Karas in 1988 and has become a well-known analytical tool for analyzing peptides, proteins, and numerous biomolecules such as

oligonucleotides, carbohydrates, natural products, and lipids. MALDI is a soft ionization technique in which the energy from the laser is spent in volatilizing the matrix rather than degrading the polymer (Figures 11 and 12). This technique is based upon ultraviolet absorbing matrix. The matrix and the polymer are mixed at the molecular level in an appropriate solvent. The solvent helps prevent aggregation of the polymer. The sample-matrix mixture is placed on the sample probe tip, under vacuum conditions. The solvent is removed, leaving co-crystallized polymer molecules homogenously dispersed within the matrix molecules. When the pulsed laser beam is tuned to the appropriate frequency, the energy is transferred to the matrix which is partially vaporized, carrying intact polymer into the vapor phase and charging the polymer chains.<sup>17</sup>



**Figure 11.** A schematic of the principle of MALDI.<sup>17</sup>



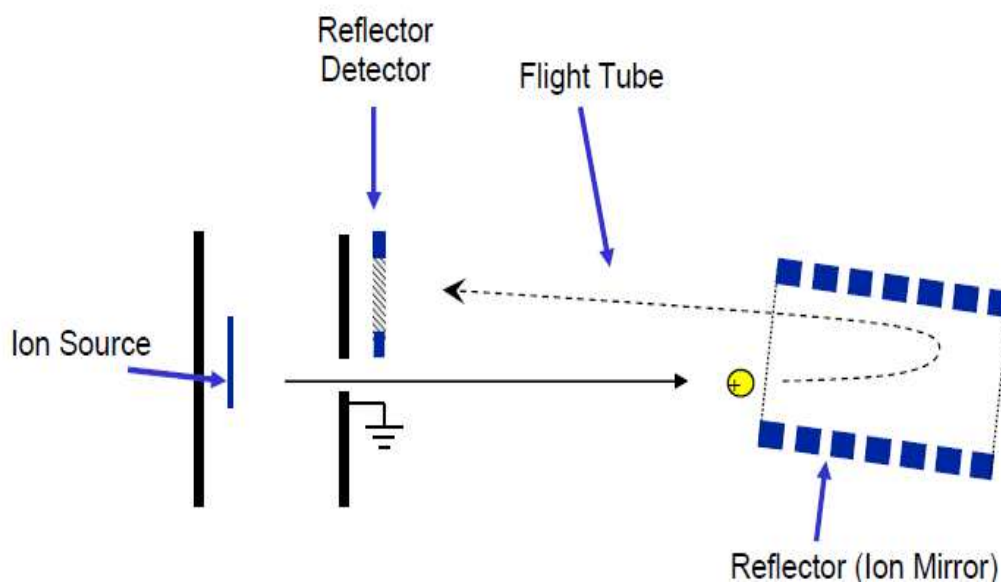
**Figure 12.** A more detailed schematic of the principles of MALDI-TOF.<sup>18</sup>

A linear time of flight (TOF), as shown previously in Figure 12, is based on the acceleration of a set of ions to a detector where all of the ions are given the same energy, but vary in mass. The ions reach the detector at different times because of differing masses. Smaller ions will have faster velocities and reach the detector first, while larger ions have slower velocities, and therefore are the last to be detected. The arrival time to the detector is dependent upon the mass, charge, and kinetic energy (KE) of the ion. KE is defined as  $\frac{1}{2}mv^2$  or velocity  $v = (2KE/m)^{1/2}$ . Ions will travel a given distance,  $d$ , within a time,  $t$ , where  $t$  depends on their mass-to-charge ratio.<sup>17</sup>

A TOF reflectron is shown in Figure 13. A reflectron is a device in the flight tube of a TOF mass spectrometer that acts as an electrostatic mirror, with a series of rings. It extends the flight length of the ions. This allows the reflector to compensate for the initial energy spreads of the ions with the same mass. Ions with a range of kinetic energies travel through the flight tube and spend different amounts of time in the reflectron. The reflectron corrects the kinetic energy



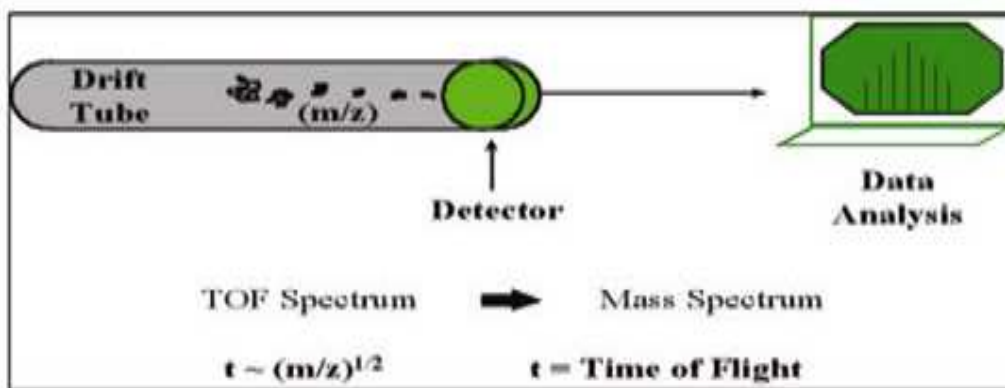
distributions by using a constant electrostatic field that reflects the ion beam(s) toward the detector. Ions with higher energies spend a longer time getting to the detector, while ions with lesser energies of the same  $m/z$  have a shorter distance to the detector. The purpose of the reflectron is for a higher overall mass resolution.<sup>19</sup>



**Figure 13.** Schematic for TOF Reflectron.<sup>19</sup>

The distribution of molecules emanating from a sample is imparted to identical translational kinetic energies after being subjected to the same electrical potential energy difference. The ions travel the same distance down an evacuated field free drift tube. The smaller ions arrive at the detector in a shorter time than the more massive ions. Separated ion fractions at the end of the drift tube are detected by an appropriate recorder that produces a signal upon impact of each ion group. The data generated from the successive laser shots add up to provide a TOF mass spectrum. The mass spectrum is a recording of the signal as a function of time. The time of flight (TOF) for a molecule of mass,  $m$ , and charge,  $z$ , to travel this distance is proportional to  $(m/z)^{1/2}$ . This relationship can be used to calculate the mass of the ions, which is

the conversion of a TOF mass spectrum to a conventional mass spectrum and can be achieved as shown in figure 14.<sup>18</sup>



**Figure 14.** Conversion of a TOF spectrum to a mass spectrum. TOF measures a particle's  $m/z$ . An ion with a known charge and unknown mass enters the mass spectrometer and is accelerated by an electrical field with a known strength. The acceleration results in any given ion having the same kinetic energy as any other ion given that they all have the same charge. The velocity of the ion depends on the  $m/z$ .<sup>17</sup>

MALDI-TOF has the ability to provide highly accurate molecular weight information on intact molecules and has been highly utilized in protein and peptide analyses. The ability to generate such accurate information can be extremely useful for protein identification and characterization.<sup>17</sup> Recently, mass spectrometry has been combined with protease digestion to enable peptide mass fingerprinting. Peptide Mass Fingerprinting combines enzymatic digestion, mass spectrometry, and computer facilitated analysis for protein identification.

### **Peptide Mass Fingerprinting and Mascot**

PMF is an analytical technique used to identify proteins by matching fragment peptide masses to theoretical masses that are found in a known database. An unknown protein is cleaved into smaller fragments by using a proteolytic enzyme, such as trypsin, to generate peptides. The fragment masses are then determined using MALDI-TOF. The premise is every unique protein has a set of unique set of peptides, therefore, unique peptide masses. These fragment masses are

compared to a database with known protein sequence masses. Successful identification is accomplished by matching the observed peptide masses to the theoretical masses derived from a sequence database. Identification relies on finding a sufficient number of peptides, 5+, from the same protein at high mass accuracy.<sup>18,20</sup>

Proteins are first isolated using two-dimensional gel electrophoresis, SDS-PAGE, or liquid chromatography. These proteins are enzymatically or chemically cleaved. Protein samples can be derived from SDS-PAGE and are then subject to some chemical modifications. Disulfide bridges in proteins are reduced and cysteine amino acids are carboxymethylated chemically or acrylamidated during the gel electrophoresis. The proteins are cut into several fragments using proteolytic enzymes, such as trypsin. Trypsin is the most commonly used endoprotease and has the capability of generating peptides with the average size of 8-10 amino acids, which are ideal for mass spectrometry. Trypsin cleaves the peptides on the C-terminal side of the relatively abundant amino acids arginine and lysine. The peptides generated using trypsin need to be purified, concentrated and desalted before analysis on the MALDI. This is achieved using reversed-phase microcolumns, such as ZipTips, which are prepacked. By using one of these microcolumns, the signal-to-noise ratio can be somewhat improved. The sample droplet is then placed on a MALDI plate, allowed to dry, and then analyzed using the instrument.<sup>18,20</sup>

The MALDI-TOF detects the sample droplet and records a mass spectrum with monoisotopic peaks. These peaks are then labeled, which is accomplished by using a mass spectra analysis program such as Mascot, MS-Fit, or Profound. All three programs should yield similar results using the same set of parameters. A successful PMF is having the identified protein found in the database.<sup>18</sup>

<b>Protease</b>	<b>Amino Acid Specificity</b>	<b>Exceptions</b>
Trypsin	X- <b>Lys</b> /-Y, X- <b>Arg</b> /-Y	Does not cleave if Y = Pro
Endoproteinase Lys-C	X- <b>Lys</b> /-Y	Does not cleave if Y = Pro
Clostripain	X- <b>Arg</b> /-Y	
Endoproteinase Asp-N	X- <b>Asp</b> /-Y, X- <b>cysteic acid bonds</b>	Does not cleave if Y = Ser
CNBr	X- <b>Met</b> /-Y	Does not cleave if Y = Ser, Thr, or Cys
Glu-C (V8 Protease (E))	X- <b>Glu</b> /-Y, X- <b>Asp</b> /-Y	Does not cleave if Y = Pro
Pepsin	X- <b>Phe</b> /-Y, X- <b>Leu</b> /-Y, X- <b>Glu</b> /-Y	Does not cleave if Y = Val, Ala, Gly
Endoproteinase Arg C	X- <b>Arg</b> /-Y	Does not cleave if Y = Pro
Thermolysin	X-/ <b>Phe</b> -Y, X-/ <b>Ile</b> -Y, X-/ <b>Leu</b> -Y, X-/ <b>Ala</b> -Y, X-/ <b>Val</b> -Y, X-/ <b>Met</b> -Y	Does not cleave if Y = Pro
Chymotrypsin	X- <b>Phe</b> /-Y, X- <b>Tyr</b> /-Y, X- <b>Trp</b> /-Y, X- <b>Leu</b> /-Y	Does not cleave if Y = Met, Ile, Ser, Thr, Val, His, Glu, Asp
Formic Acid	X-/ <b>Asp</b> -Y	

**Figure 15.** Protease Specificity. Proteolysis experiments can use any of a number of enzymes to perform digestion. The cleavage specificity of some of the different enzymes is denoted by a slash (/) before or after the amino acid responsible for specificity. Combinations of proteases can be used to reduce specificity and to mimic other proteases. For example, Lys-C and clostripain together are specific for the same sites as trypsin.<sup>16</sup>

For this study the Mascot program was utilized, which contains quite a bit of information for doing peptide mass fingerprinting. The probability mowse score is crucial. The difference between the random and significant protein scores needs to be as high as possible. A protein with a score  $>100$  is considered identified. A good rule to use is that a protein is identified with a score of  $>100$ . A score between the highest random value, usually 60-75, and 100 displays a good candidate protein. The second criterion which must be considered is the full protein summary report. If one clicks the accession number of the hit, more detailed information of the protein will be displayed. The nominal mass and pI value must match up with the experimental data obtained from gel electrophoresis. If not, protein fragments or adducts must be investigated. The sequence coverage and masses matched are also very important pieces of information. By multiplying the SC and MM, the reliability of the protein is revised and considered to be identified with a value  $> 300$ . The difference between the numbers of mass values matched should be as small as possible. If the not, the same sample may contain other protein components. After sorting the peptides by increasing mass, the matched and observed masses should be compared. The most intense peaks will be assigned to the identified protein. Any unassigned peaks could be from tryptophan oxidation, methylation of aspartic acid, and glutamic acid rich peptides.<sup>18</sup>

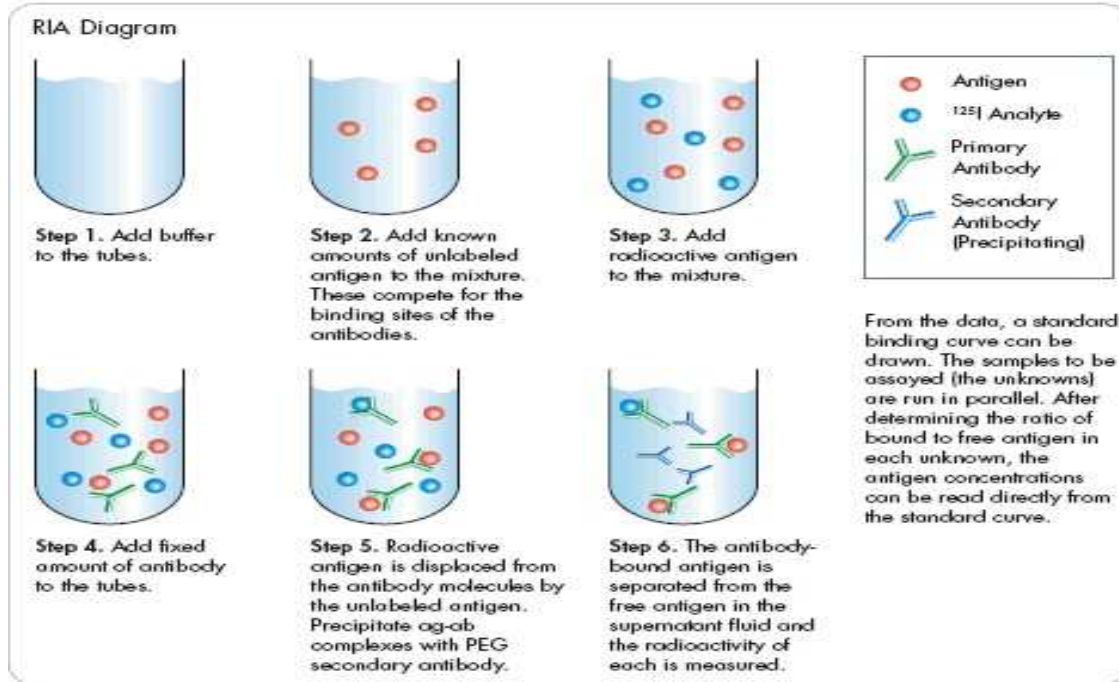
### **Radioimmunoassay**

Diabetes has always been a problem and researchers have continuously looked for new understanding of the disease. People who have diabetes have problems metabolizing sugars due the fact their bodies cannot make enough insulin. Insulin is the pancreatic hormone essential for metabolizing sugars. Diabetics take injections of insulin as treatment for supplementing the insulin their bodies do not naturally make. Two researchers, Rosalyn Yalow and Solomon

Berson, decided to find out what actually happened once insulin was injected into the body. Their theory was that diabetes might occur as a consequence of an unusually rapid degradation of insulin by the enzyme insulinase. To test their hypothesis, they used [<sup>125</sup>I]insulin as a tracer. This allowed them to detect the [<sup>125</sup>I]insulin distribution through the bodies of both diabetics and non-diabetics by the radiation emitted.<sup>21-22</sup>

Yalow and Berson observed that diabetics who had been receiving animal insulin for a longer period of time would keep the [<sup>125</sup>I]insulin in their bodies much longer than non-diabetics.<sup>21-22</sup> In those days, insulin for injection was derived from cattle. Like any other foreign object, the body flagged the cattle insulin and attacked using antibodies. This is the normal response for any normal functioning immune system. Yalow and Berson realized that the diabetics, who had taken insulin from cattle previously, developed an immune response forming antibodies. These antibodies were binding to the radio-labeled insulin, slowing down the process and making it harder for the body to use the cattle insulin. Because of this discovery, Yalow and Berson decided that human insulin was better suited for diabetic treatment, as the human body would not fight off human insulin. Bacteria were engineered to produce copious amounts of human insulin. Today nearly of all the insulin used for treatment is from genetically modified bacteria.<sup>22</sup>

After their discovery, the pair turned to kinetics of reaction of insulin.<sup>21</sup> They coined the term radioimmunoassay (RIA) for this technique.<sup>21</sup> This method is used to quantify minute amounts of substance by using radioactive-labeled material.<sup>22</sup> The basic principle uses competitive binding as found in figure 16 below.



**Figure 16.** Schematic for a RIA.<sup>22</sup>

A known quantity of the hormone of interest is tagged with a radioactive antigen, also called a tracer.<sup>22,24</sup> The tracer competes with a non-radioactive antigen for a certain number of antibody or receptor binding sites. When the amount of the unlabeled antigen is increased, the amount of tracer bound to the antibodies decreases. The target is radioactively labeled, such as using [<sup>125</sup>I]insulin, and bound to its specific antibodies. A known amount of antibody is to be added. A sample, such as blood, serum or plasma, is added to begin the competition. The competition for antibodies releases a certain amount of labeled antigens, which is proportional to the ratio of labeled and unlabeled antigens. As the concentration of unlabeled antigen is increased, the more it binds to the antibodies and displaces the labeled variant. The labeled and unlabeled antigens are then separated out and the radioactivity of the free, unbound antigens is measured.<sup>24</sup> By subtracting the amount of radioactivity in the labeled antigens from the unlabeled antigens, one is able to quantitate the amount of hormone in the blood, serum, or plasma sample.<sup>22</sup>

Experimentally, in the past, a number of protein receptors were determined by using different methodologies. The insulin receptor was found similarly in both the liver and fat cell membranes using radioreceptor assay. At the time of this report in 1971, the interaction of insulin with receptors in liver cells could only be studied by measuring the specific binding of labeled hormone to broken-cell preparations. These binding properties could not be directly compared with the properties of an intact system, since it was impossible to accurately measure the insulin binding properties from liver cells. Therefore, this report illustrates the basic properties of these binding interactions as well as the interactions observed in fat cell membranes of the liver. In the previously mentioned report, large agarose beads were covalently bound to insulin and tested on isolated adipose tissue cells, which suggested the insulin receptors would be located in the superficial regions of the cell. The study resulted in identifying the location in the plasma membrane of the fat cells.<sup>25</sup>

The methodology was very straightforward. Liver membranes were prepared in a 0.25 M sucrose solution from homogenates using differential centrifugation, a common process for fractioning cells, and resulted in the separation of major nuclear and mitochondrial elements. The protocol used for insulin preparation can be found in a report by Pedro Cuatrecasas from June 1971. Insulin was iodinated at 24 °C by adding 4.0 ug of crystalline porcine zinc-insulin, in 10  $\mu$ L of 0.1 M sodium phosphate buffer, pH 7.4, to 100 uL of 0.25 M sodium phosphate buffer, pH 7.4, containing 1.0 mCi of Na<sup>125</sup>I. 20  $\mu$ L of chloramine-T was then added and after 40 seconds, 20  $\mu$ L of sodium metabisulfite was added. 8 mL of 0.1 M sodium phosphate 0.1 % albumin buffer was immediately added. The radioactivity precipitable by trichloroacetic acid was then determined by adding one 25 mg talc tablet, dispersing it in the solution. The solution was centrifuged for 10 minutes at 2000 rpm, resulting in a talc pellet, which was re-suspended and



centrifuged 4 times with 0.1 M sodium phosphate buffer, pH, 7.4. The [ $^{125}$ I]insulin was eluted from the talc by adding 3 mL of 0.37 N HCl containing 6 % bovine albumin. The suspension was centrifuged for 30 minutes at 2500 rpm and the supernatant was then neutralized by adding 0.5 mL of 1 N NaOH.<sup>26</sup>

The standard binding assay consisted of incubating fat cells at 24 °C for 20 minutes in 0.5 mL of KRB-1 % albumin containing [ $^{125}$ I]insulin. Equilibrium was achieved within 20 minutes. To the cells was added 3 mL of ice-cold KRB-0.1 % albumin, immediately filtered, and washed the cells with another 10 mL, under reduced pressure, on cellulose acetate EAWP Millipore filters. The filters were removed, cut in halves, and placed in counting vials that contained 1 mL of 10 % sodium dodecyl sulfate. Samples were shaken for 30-40 minutes at 24 °C, and 10 mL of the counting solution was added. The choice of membrane used was critical, because insulin can adsorb strongly to most cellulose derivative.<sup>26</sup>

By subtracting the amount that was not displaced by high concentrations of native insulin from the total radioactive uptake, specific binding is determined. This is proportionally related to the concentration of iodoinsulin and independent of time or temperature of the incubation period.<sup>25</sup> Bound insulin was detected at very low concentrations in the liver cells.<sup>25</sup>

Since radioimmunoassay was first used to measure insulin in plasma membranes in 1960 by Yalow and Berson, as well as described previously, it has since been used for a number of polypeptide hormones, such as ACTH. ACTH appears to bind to specific receptors on the outer surface of cells. Adenyl cyclase, an enzyme bound to the plasma membrane, is activated by the receptor hormone complex. Adenyl cyclase rapidly converts adenosine triphosphate to cyclic adenosine monophosphate, which appeared to be the intracellular mediator for ACTH action. In 1970, a fast, sensitive assay was developed based on the  $^{125}$ I-labeled and unlabeled ACTH

competing for binding sites specific for ACTH receptors. This method allowed ACTH in plasma to be easily measured in a few hours, as well as being applicable to a variety of other hormones.<sup>27</sup>

Monoiodo [<sup>125</sup>I]ACTH was prepared by labeling ACTH with <sup>125</sup>I and separating the iodinated ACTH from the uniodinated ACTH by chromatography on carboxymethyl cellulose. Monoiodo [<sup>125</sup>I]ACTH retains at least half of the biological activity of native ACTH and adrenal extracts, containing both ACTH receptors and the ACTH-sensitive adenylyl cyclase. ACTH-sensitive adrenal tumors were transplanted into mice in order to obtain adrenal receptors. These tumors were homogenized and the fractions containing all of the ACTH-sensitive adenylyl cyclase were separated by centrifugation. The particles were then disrupted in a French pressure cell in the presence of phosphatidylethanolamine and fluoride, and then centrifuged again. After dialysis, the supernatant was collected and used as the ACTH receptor extract. Binding to the adrenal receptors was highly specific. Adrenal extracts containing ACTH-sensitive cyclase bound the hormone, while other tissue extracts were inert. While biologically active [<sup>125</sup>I]ACTH was bound, other labeled hormones were unreactive. The extent of the ACTH derivatives competitiveness for receptors appeared related to its biological activity.<sup>27</sup>

A clinical study was then done using sample tubes made of flint glass and siliconized. Sample tubes were prepared with 200 to 250 μL of adrenal extract, [<sup>125</sup>I]ACTH, and either 50 to 200 μL of the patients' plasma or 50 to 200 μL of plasma from patients who had hypopituitarism, where unlabeled ACTH was added to a grand total of 1 mL of volume. All procedures were done at 1 °C. Diluent used was 0.01 M MgCl<sub>2</sub>, 0.001 M dithiothreitol, 0.1 % human serum albumin. After an hour, ~ 10 mg of Quiso was added. The Quiso adsorbed the free [<sup>125</sup>I]ACTH but not the receptor bound. The samples were centrifuged for five minutes and the

radioactivity in the supernatant and the Quso were separately counted. In the absence of unlabeled ACTH, the [<sup>125</sup>I]ACTH was bound to the adrenal receptors.<sup>27</sup>

In the clinical study, plasma was taken from 19 different patients. Three of these patients had pituitary ACTH deficiencies; hence ACTH was not detectable. Two of the patients were on corticosteroids that caused ACTH suppression. In three of seven normal adults, ACTH was not detected at all. Four other patients had ACTH plasma levels of up to 88 pg/mL. In the patients with hyperadrenocorticism, ACTH plasma levels ranged from 200 pg/mL up to 10,000 pg/mL.<sup>27</sup>

This study proved successful in the attempt to determine ACTH receptors in human plasma. This was the first time a target tissue receptor was used for measuring a polypeptide hormone. Radioreceptor assays are advantageous based on the specificity for biologically active ACTH. Sensitivity to the receptor is an additive bonus. This study provided a direct method for studying the interaction of hormones with their target tissues.<sup>27</sup>

## **EXPERIMENTAL**

This experiment was designed to identify cellular proteins capable of binding to the C-peptide produced from the post-translational cleavage of the proinsulin peptide. A C-peptide affinity column was developed by attaching C-peptide to benzyloxybenzylbromide polystyrene derived resin beads through its N-terminus. Bovine kidney, heart, and thymus tissue samples were lysed by allowing cells to swell with hypotonic buffer, removing the cytoplasmic protein fraction, and then releasing the nuclear proteins using a high salt buffer. The protein lysates were then incubated with the C-peptide affinity beads to isolate the proteins capable of binding to the C-peptide. The isolated proteins were separated by SDS-PAGE, and the proteins of interest were excised from the gel, digested with trypsin, and identified using MALDI-TOF mass

spectrometry peptide mapping. By identifying C-peptide binding-proteins, we hope to gain a better understanding about the cellular processes involving C-peptide which could help to unlock the answers as to why the C-peptide is not degraded after its purpose of folding is complete. Upon the understanding of its purpose in the body, novel treatments could be developed to help diabetes patients especially those suffering from Type-1.

### **Preparation of Pre-clear Beads**

Benzyloxybenzylbromide polystyrene derived resin beads were swelled in dichloromethane (DCM) for one hour, then they were washed three times with dimethylformamide (DMF) and three times with lysis buffer. The pre-clear beads were stored in lysis buffer at 4 °C.

### **Preparation of C-peptide Affinity Beads**

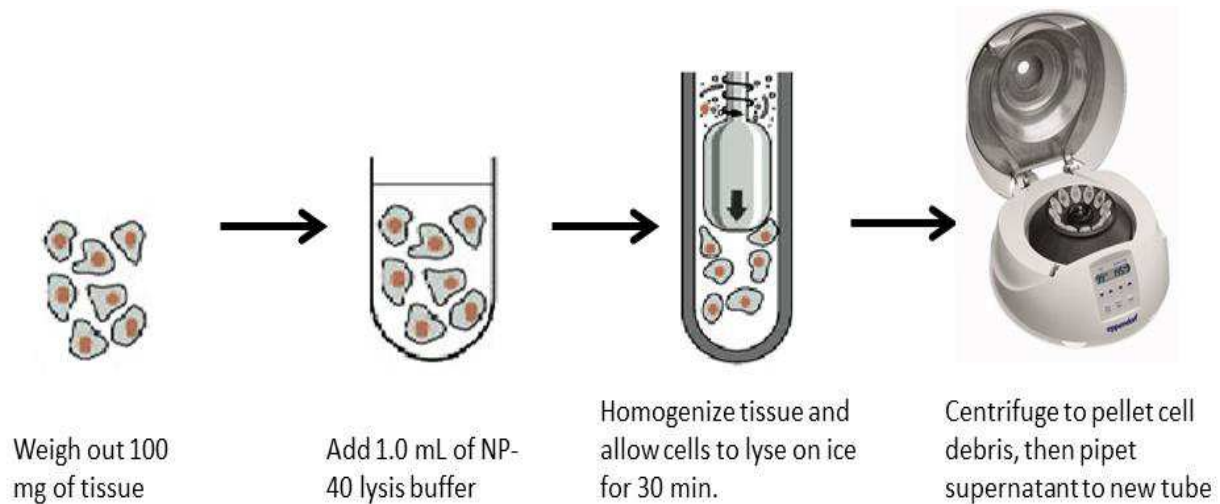
The human C-peptide residues 3-33 was purchased from Sigma. The N-terminus of the peptide was attached to a benzyloxybenzylbromide polystyrene derived resin. This process left the C-terminal residues accessible for protein binding, since researchers in other labs had determined that the last 5 amino acids at the C-terminus of the peptide were the most important for biological activity. Briefly, the resin beads were swelled in DCM for one hour, then they were washed with DMF. The beads were then incubated with the C-peptide dissolved in DMF overnight at 4 °C with stirring. The C-peptide would attach to the resin beads through an S<sub>N</sub>2 type of reaction displacing the bromine atom. The N-terminus of a peptide and the side chain of lysine residues would act as nucleophiles; however, there are no lysine residues in the amino acid sequence of C-peptide making the N-terminus the only point of attachment to the resin beads. The C-peptide beads were then washed three times with DMF, three times with DCM, and three times with lysis buffer prior to storage in lysis buffer at 4 °C.

## Tissue Lysis

Fresh bovine kidney, heart, and thymus tissue obtained from a slaughter house was used.

### A. NP-40 Lysates

Bovine tissue was lysed using a NP 40 lysis buffer (50 mM Tris-HCl, 150 mM NaCl, 1% NP 40, pH 8.0) and manual grinding with mini grinders (Bio-Rad Ready Prep Mini Grinders). Protease inhibitors were added to the lysis buffer. Briefly, 100 mg of tissue was homogenized in 1.0 mL of NP 40 lysis buffer by manual grinding for several minutes and then allowed to sit on ice for 30 min to ensure optimal lysis of the tissue. The tissue lysate was then centrifuged at 12,000 G for 30 min at 4 °C, and the supernatant lysate was transferred to a clean microcentrifuge tube. Pre-clear beads were added to the lysate to remove any proteins that had an affinity for beads themselves prior to incubation of the lysate with C-peptide beads.



**Figure 17.** Preparation of NP-40 Protein Lysates.

## **B. Cytoplasmic and Nuclear Tissue Lysates**

Cytoplasmic and nuclear tissue lysates were prepared using the CelLytic NuCLEAR extraction kit (Sigma) following the manufacturers protocol for protein extraction from tissues located online at: <http://www.sigmaaldrich.com/content/dam/sigma-aldrich/docs/Sigma/Bulletin/nxtractbul.pdf>

Briefly,

1. Weigh out about 100 mg of tissue.
2. Prepare hypotonic lysis buffer by adding 10  $\mu$ L of 0.1 M DTT and 10  $\mu$ L of protease inhibitor to 1.0 mL of isotonic lysis buffer (10 mM HEPES, 1.5 mM MgCl<sub>2</sub>, 10 mM KCl, pH 7.9)
3. Rinse tissue twice with phosphate buffered saline (PBS).
4. Resuspend tissue in 1 mL of hypotonic lysis buffer .
5. Homogenize the tissue until more than 90 % of the cells are broken.
6. Centrifuge the disrupted cells at 11,000 x g for 20 minutes.
7. Transfer the supernatant (cytoplasmic fraction) to a new tube.
8. Prepare extraction buffer by adding 1.5  $\mu$ L of 0.1 M DTT and 1.5  $\mu$ L of protease inhibitor to 147  $\mu$ L of buffer (20 mM HEPES, 1.5 mM MgCl<sub>2</sub>, 0.42 M NaCl, 0.2 mM EDTA, 25% (v/v) Glycerol, pH 7.9).
9. Resuspend nuclei pellet in 140  $\mu$ L of extraction buffer, homogenize tissue, and shake gently for 30 minutes.
10. Centrifuge for five minutes at 20,000 x g, then transfer supernatant (nuclear fraction) to a new tube.

## **Affinity Column Protocol**

Benzyloxybenzylbromide polystyrene pre-clear beads were added to the protein lysates in order to pre-clear the lysate of proteins capable of binding to the unmodified resin beads. After incubation for one hour at room temperature (with tube rotation), the solution was centrifuged at 5,000 x g for two minutes to pellet the beads, and the supernatant was transferred to a new tube. The beads were then washed once with PBST and twice with PBS. Finally, 30  $\mu$ L of SDS-PAGE loading buffer containing DTT was added to the beads.

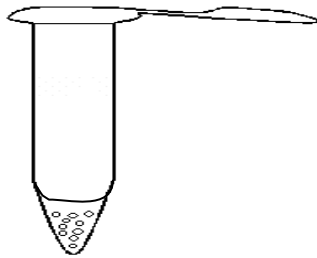
The affinity C-peptide beads were added to the supernatant from the pre-clear step described above. The solution was incubated for two hours at room temperature (with tube rotation) or overnight at 4 °C (with tube rotation). The solution was centrifuged at 5,000 x g for two minutes, and the supernatant was removed and discarded. The beads were then washed once with PBST and twice with PBS. Finally, 30  $\mu$ L of SDS-PAGE loading buffer containing DTT was added to the beads.

## **SDS-PAGE**

Proteins isolated from the protein lysates using the pre-clear and C-peptide beads were separated by SDS-PAGE prior to mass spectrometric analysis. Proteins were eluted from the beads with the addition of SDS-PAGE sample loading buffer containing DTT. The samples were heated at 100 °C for five minutes and then loaded directly onto the gels (12% polyacrylamide, BioRad). Briefly, proteins were separated for 30 min at 200 V using Tris/Glycine/SDS running buffer. Proteins were visualized with Bio-Safe Coomassie Stain (BioRad).

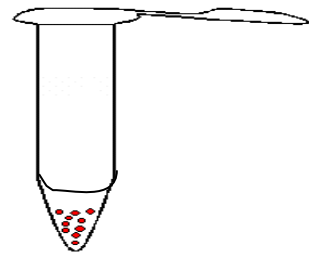
<b>Legend</b>	
○	Beads with nothing attached.
●	Beads with peptide attached.
●	Beads with peptide and any proteins attached.

1.) Cell lysate goes into an Eppendorf tube with 30  $\mu$ l of beads to pre-clear.



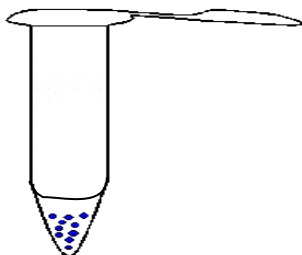
2.) Incubate for two hours.

3.) Remove supernatant and add to the next tube.



4.) This second tube now has cell lysate on the beads with peptide attached. Incubate for 2 hours.

5.) Next, remove the supernatant and discard.



6.) Wash the remaining beads twice with PBS.

7.) Add 30  $\mu$ l of SDS-PAGE loading buffer with DTT into the eppendorf tube.

8.) Heat in boiling water for 10 minutes.

9.) Now load 12  $\mu$ l of Precision Plus Protein Standards Kaleidoscope onto the first lane of the SDS-PAGE gel.

10.) From here load 27  $\mu$ l of the sample with loading buffer onto the third lane of the to the SDS-PAGE gel.

11.) Run gel to completion at 200V.

**Figure 18.** Schematic representation of Affinity Column Protocol.



## **Preparation of Tryptic Digests of Protein Samples from SDS-PAGE Bands**

Protein bands from the SDS-PAGE gels were excised, reduced with DTT, treated with iodoacetamide, and digested with trypsin overnight at room temperature following the procedure at <http://msf.ucsf.edu/protocols.html>.

Briefly, the protein bands of interest were excised from the gel and chopped into small pieces. The gel pieces were washed with 100  $\mu$ L of 25 mM  $\text{NH}_4\text{HCO}_3$ /50% ACN three times. Next, 30  $\mu$ L of 10 mM DTT in 25 mM  $\text{NH}_4\text{HCO}_3$  was added to the gel pieces and the reaction was allowed to proceed at 56° C for one hour. The solution was removed, and 30  $\mu$ L of 55 mM iodoacetamide in 25 mM  $\text{NH}_4\text{HCO}_3$  was added to the gel pieces and allowed to react for 45 minutes at room temperature in the dark. The solution was removed, and the gel pieces were washed with 100 $\mu$ L of 25 mM  $\text{NH}_4\text{HCO}_3$  in 50% ACN three times. The gel pieces were overlaid with 30  $\mu$ L of 25 mM  $\text{NH}_4\text{HCO}_3$  buffer containing between 0.1 and 0.5  $\mu$ g Promega trypsin. Protein digestion was carried out overnight at room temperature and terminated by the addition of acetic acid. The resulting solution of peptide fragments were extracted from the gel pieces, purified using a  $\text{C}_{18}$  ZipTip, and spotted onto the MALDI sample plate for mass spectral analysis.

## **Analysis of Tryptic Digests of Proteins by MALDI-TOF Mass Spectrometry**

The tryptic digests of the proteins were purified with  $\text{C}_{18}$  ZipTips (Millipore) and analyzed on a Bruker Autoflex MALDI-TOF mass spectrometer (Billerica, MA). Prior to analysis, the peptide preparations (0.6  $\mu$ L) were each mixed with a 50% aqueous acetonitrile solution (0.6  $\mu$ L) of saturated  $\alpha$ -cyano-4-hydroxycinnamic acid containing 0.05 % TFA as matrix, spotted onto a stainless steel sample plate and allowed to air dry. The mass spectra were recorded in positive ion mode with a 60 nsec delay in the  $m/z$  range from 500 to 3500. Typically, ~1000 spectra

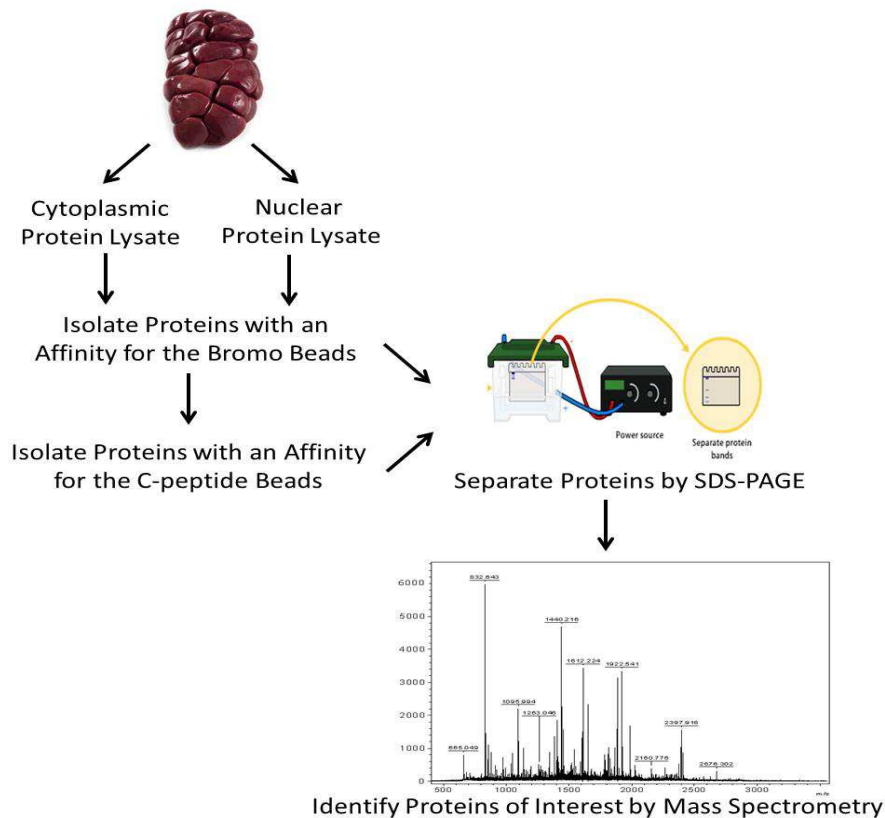
were accumulated with 50 laser shots for each sample spot analyzed. The resulting peptide mass fingerprint spectra obtained for each protein analyzed was used to identify the protein by entering the peptide molecular weights obtained into the mascot search engine (Matrix Science Ltd. (2008) MASCOT peptide mass fingerprint).

[http://www.matrixscience.com/cgi/search\\_form.pl?FORMVER=2&SEARCH=PMF](http://www.matrixscience.com/cgi/search_form.pl?FORMVER=2&SEARCH=PMF)

## **RESULTS AND DISCUSSION**

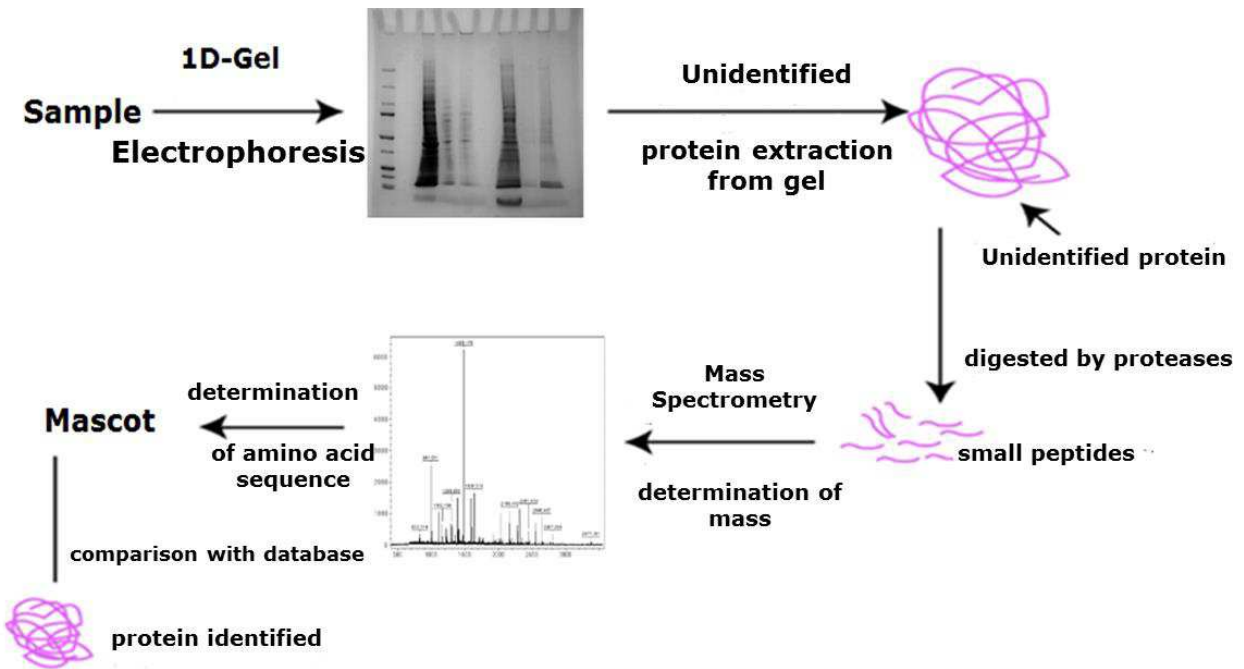
C-peptide is produced, processed, and secreted with insulin, and appears to exert a biological effect separate from that of insulin. The physiological significance of the C-peptide in normal physiology has not been fully determined. Both insulin and C-peptide secretion are lost in Type I diabetics; however, these patients only receive insulin therapy without C-peptide replacement. The loss of C-peptide does not appear to produce immediate negative effects in these patients, but they do eventually develop microvascular dysfunction, such as retinopathy, neuropathy, nephropathy, or impairments in wound healing,<sup>28-29</sup> and C-peptide replacement reverses microvascular dysfunction leading to neuropathy.<sup>30</sup> Previous studies conducted by Wahren's group have shown that C-peptide is internalized and located to the cytosol of Swiss 3T3 (fibroblast cell line) and HEK-293 (human embryonic kidney) cells, and transport of the C-peptide into the nucleus was also observed.<sup>31</sup> The results from this experiment suggest that C-peptide may function as an intracrine peptide hormone, denoting intracellular hormones giving rise to cellular effects within the cell of synthesis or a target cell.<sup>32</sup> This experiment was designed to gain a better understanding of the biological effects of the C-peptide by examining what proteins an affinity column with C-peptide attached would pull out of cell lysates. Since the C-peptide was shown to be internalized in the cytosol and nucleus of kidney cells, we started with cytoplasmic and nuclear tissue lysates obtained from bovine kidney tissue.

The basic experimental protocol for this experiment is summarized in Figure 19. C-peptide affinity beads were developed by attaching C-peptide via its N-terminus to benzyloxybenzylbromide polystyrene derived resin beads. Bovine kidney tissue was then differentially lysed cytoplasmic and nuclear protein lysates using the CellLytic NuCLEAR extraction kit (Sigma). The lysates were pre-cleared with unmodified benzyloxybenzylbromide beads (bromo beads) for one hour at room temperature. The lysate was removed and then incubated with C-peptide beads for two hours at room temperature. The bromo beads and the C-peptide beads were washed with PBS, and the bound proteins were then eluted from the beads, separated by reducing SDS-PAGE, digested with trypsin, and analyzed via MALDI-TOF mass spectrometry.



**Figure 19.** Basic experimental protocol for isolating proteins from bovine kidney tissue using C-peptide affinity beads.

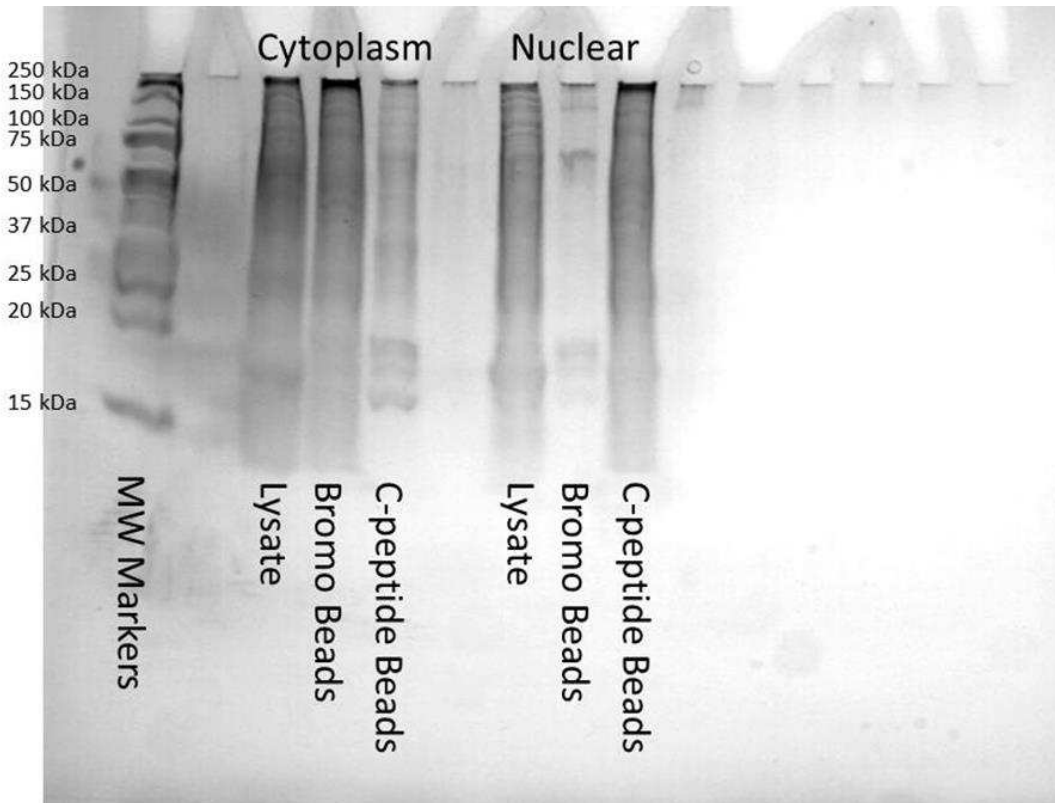
The protein bands excised from the SDS-PAGE gels were identified by mass spectrometry through the Peptide Mass Fingerprint (PMF) technique. In PMF, an unknown protein is cleaved into smaller peptide fragments using a specific protease, most commonly trypsin. The molecular weight of each fragment peptide is then accurately measured by MALDI-TOF mass spectrometry. These peptide masses are then compared to a computer database of known protein sequences in Mascot. In this experiment, the proteins isolated were identified via excising a protein band from the gel, digesting the protein with trypsin, and analyzing the resulting peptide fragments with MALDI-TOF mass spectrometry. The basic experimental protocol for identifying the proteins by the PMF technique is shown below in Figure 20.



**Figure 20.** Overview of the Peptide Mass Fingerprint technique for identifying proteins by mass spectrometry.

## **Bovine Kidney Trial 1**

The SDS-PAGE gel of the bovine kidney proteins isolated using the bromo beads and the C-peptide beads can be seen in Figure 21. There were quite a few protein bands present in the cytoplasmic and nuclear lysate lanes from both the bromo beads and the C-peptide beads. The lysates were precleared with the bromo beads prior to incubation with the C-peptide beads. Just in case the bromo beads did not isolate all of the proteins capable of binding to the beads themselves, the proteins isolated from the bromo beads were eluted to get an idea of what proteins were able to bind to these beads. In addition, a few  $\mu\text{L}$  of proteins from the lysates themselves were also loaded as a control to make sure that there were indeed proteins in the lysates. By analyzing the protein bands present in the gel, it was determined that the separation needed to run longer in order to more fully separate the proteins. In addition, many of the proteins in the lanes appear to be smeared indicating a possible overloading of protein onto the gel. It also seems as if the beads were not washed sufficiently with PBS to remove proteins bound nonspecifically to the beads prior to bound protein elution from the beads. Therefore, it was decided to repeat this experiment and add an additional washing of the beads with PBST prior to protein elution; also, to let the SDS-PAGE separation run longer in order to more fully separate the proteins from one another.

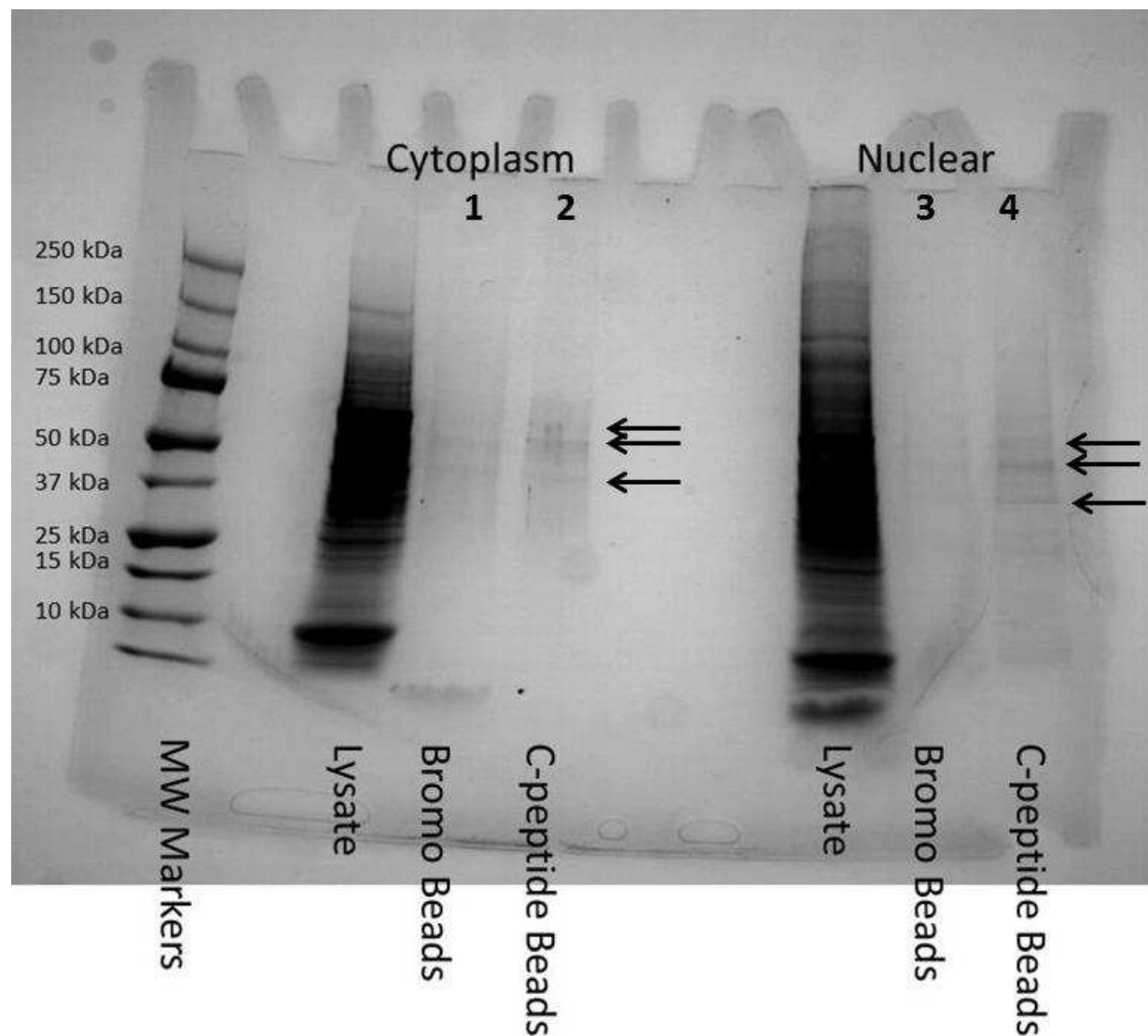


**Figure 21.** SDS-PAGE gel of proteins isolated from bovine kidney tissue in trial 1. The first lane contains molecular weight protein standards.

### **Bovine Kidney Trial 2**

The SDS-PAGE gel of the bovine kidney proteins isolated using the bromo beads and the C-peptide beads in trial 2 of the experiment can be seen in Figure 22. Again, there were quite a few protein bands present in the cytoplasmic and nuclear lysate lanes from both the bromo beads and the C-peptide beads. The additional wash step with the PBST was performed for both the bromo beads and the C-peptide beads prior to eluting the proteins from the beads with gel loading buffer. In addition, less sample was loaded into each lane of the gel. There were distinct protein bands within the 37 kDa-75 kDa molecular weight range from both the bromo beads and the C-peptide beads, and many of the protein bands appeared to be overlapping between the two bead types. Just because the protein bands appeared in the same general area of the gel for both

bead types does not mean that they were necessarily the same protein. There are many different proteins in a cell that have similar molecular weights, and they would not be separable from one another using only SDS-PAGE separation. Therefore, it was decided to excise the protein bands within this mass range from both bead types isolated from both the cytoplasmic and nuclear lysates.



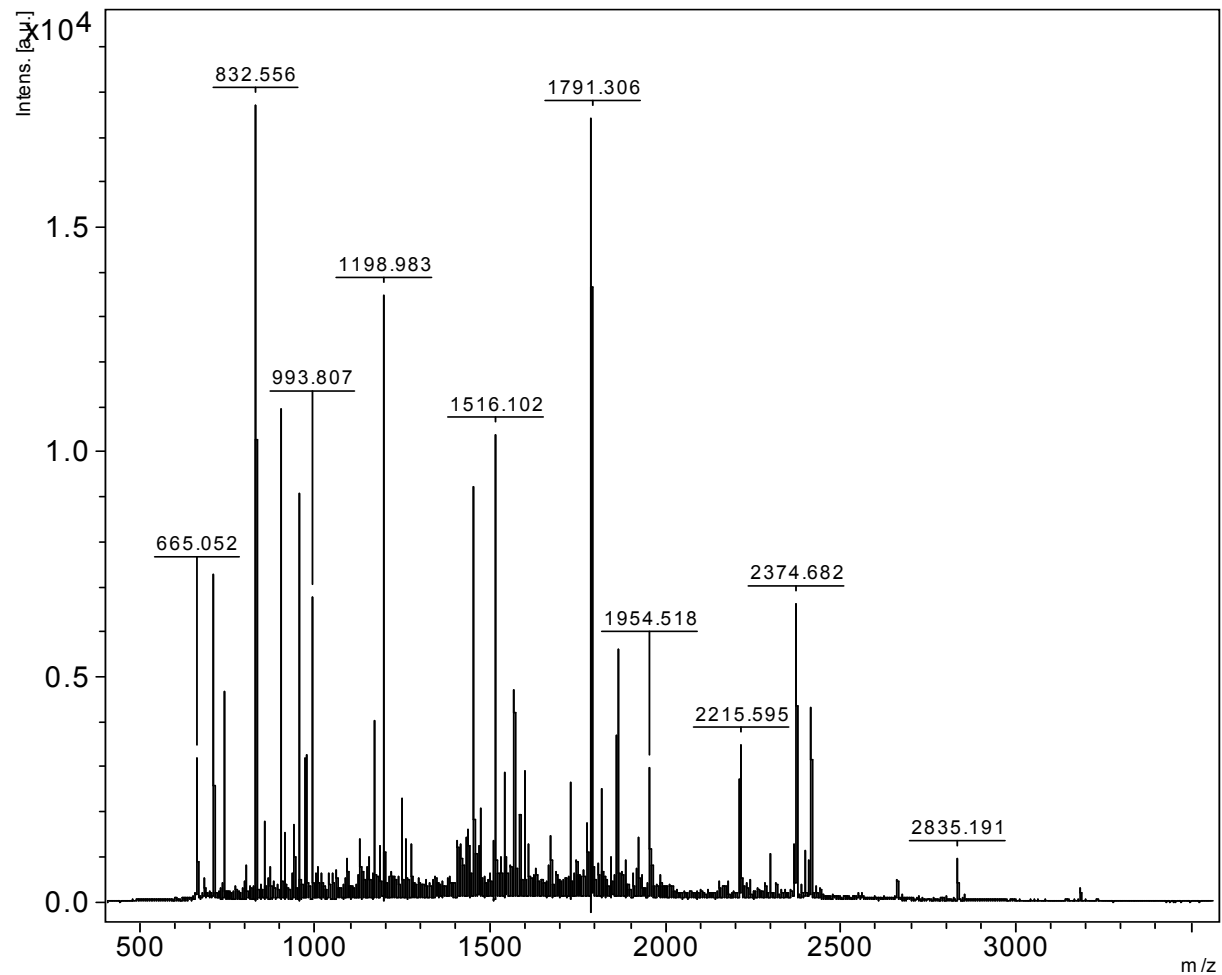
**Figure 22.** SDS-PAGE gel of proteins isolated from bovine kidney tissue in trial 2. The first lane contains molecular weight protein standards.

The MALDI-TOF mass spectrum generated from the tryptic digest of the protein at about 45 kDa in lane 4 of the SDS PAGE gel (Figure 22) is shown in Figure 23. The molecular weights of the tryptic peptides generated from this protein were entered into the Mascot search engine, and the protein was identified as actin. MALDI-TOF mass spectra were also obtained for the tryptic peptides produced from the corresponding protein bands at 45 kDa in lanes 1-3 of the gel, and the peptides generated confirmed the identity of these proteins as actin as well. Therefore, actin was isolated using both the bromo beads and the C-peptide beads, so it is not a unique C-peptide binding protein. Actin is the most abundant protein in most eukaryotic cells participating in many cellular functions ranging from cell motility and the maintenance of cell shape to the regulation of transcription.<sup>33</sup> Actin also participates in more protein-protein interactions than any other known protein, making it a “sticky” protein. Apparently that actin is nonspecifically binding to the polystyrene beads and the walls of the microcentrifuge tubes, and this interaction is then disrupted by the SDS when the tubes and beads are heated in the gel loading buffer solution.

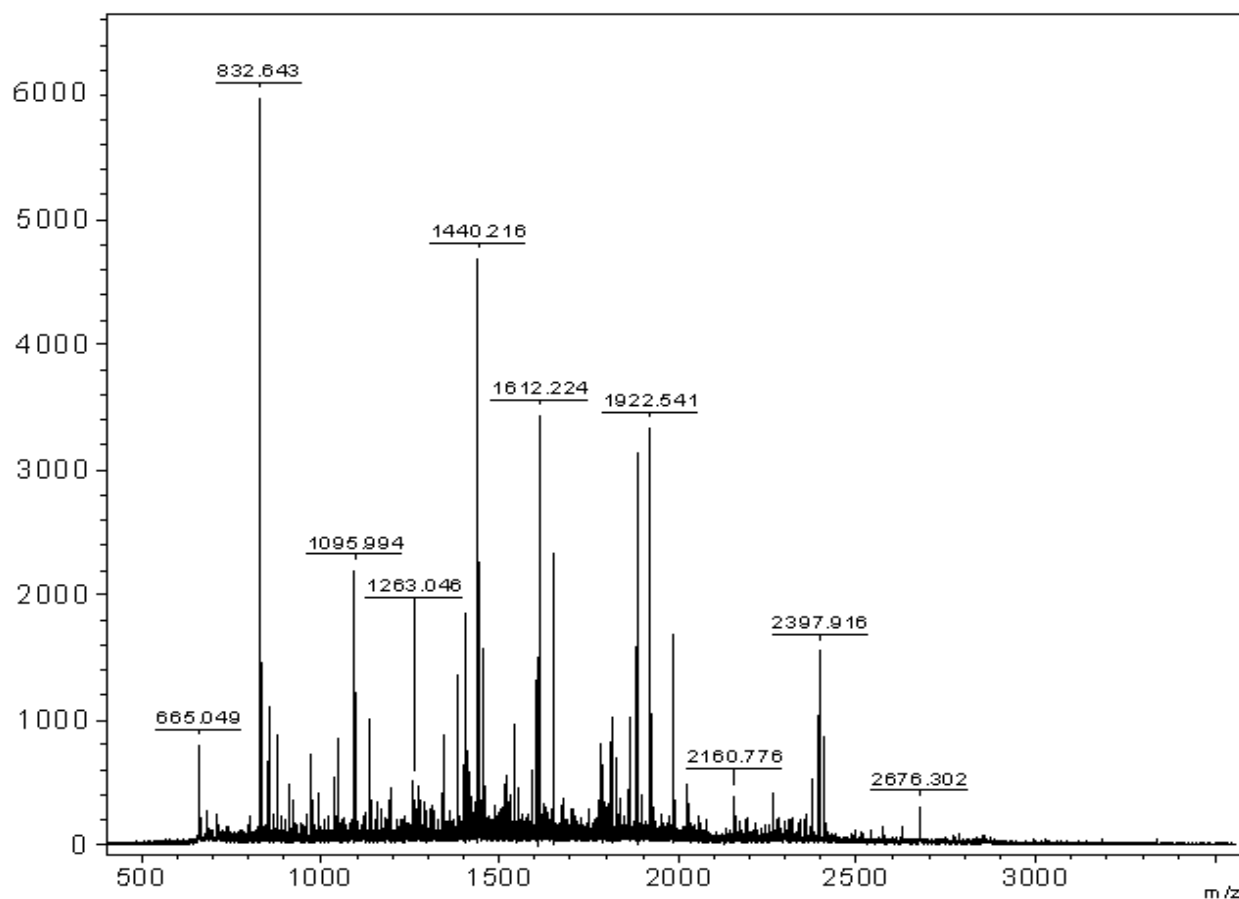
The protein band excised from the ~51 kDa position in lane 4 of the SDS-PAGE gel (Figure 22) was identified as a mixture of the  $\alpha$  and  $\beta$  subunits of F<sub>1</sub>-ATPase. The MALDI-TOF mass spectrum generated from the tryptic digest of this protein is shown in Figure 24. This protein is one of the most abundant proteins in every organism and is responsible for synthesizing adenosine triphosphate (ATP). ATP is used to power and sustain virtually all cellular processes needed to survive and reproduce. MALDI-TOF mass spectra were also obtained for the tryptic peptides produced from the corresponding protein bands at ~51 kDa in lanes 1-3 of the gel, and the peptides generated confirmed the identity of these proteins as the  $\alpha$



and  $\beta$  subunits of  $F_1$ -ATPase. Since this protein is so abundant in cells, it appears that it is just nonspecifically binding to the polystyrene beads and the wall of the tubes just like actin.



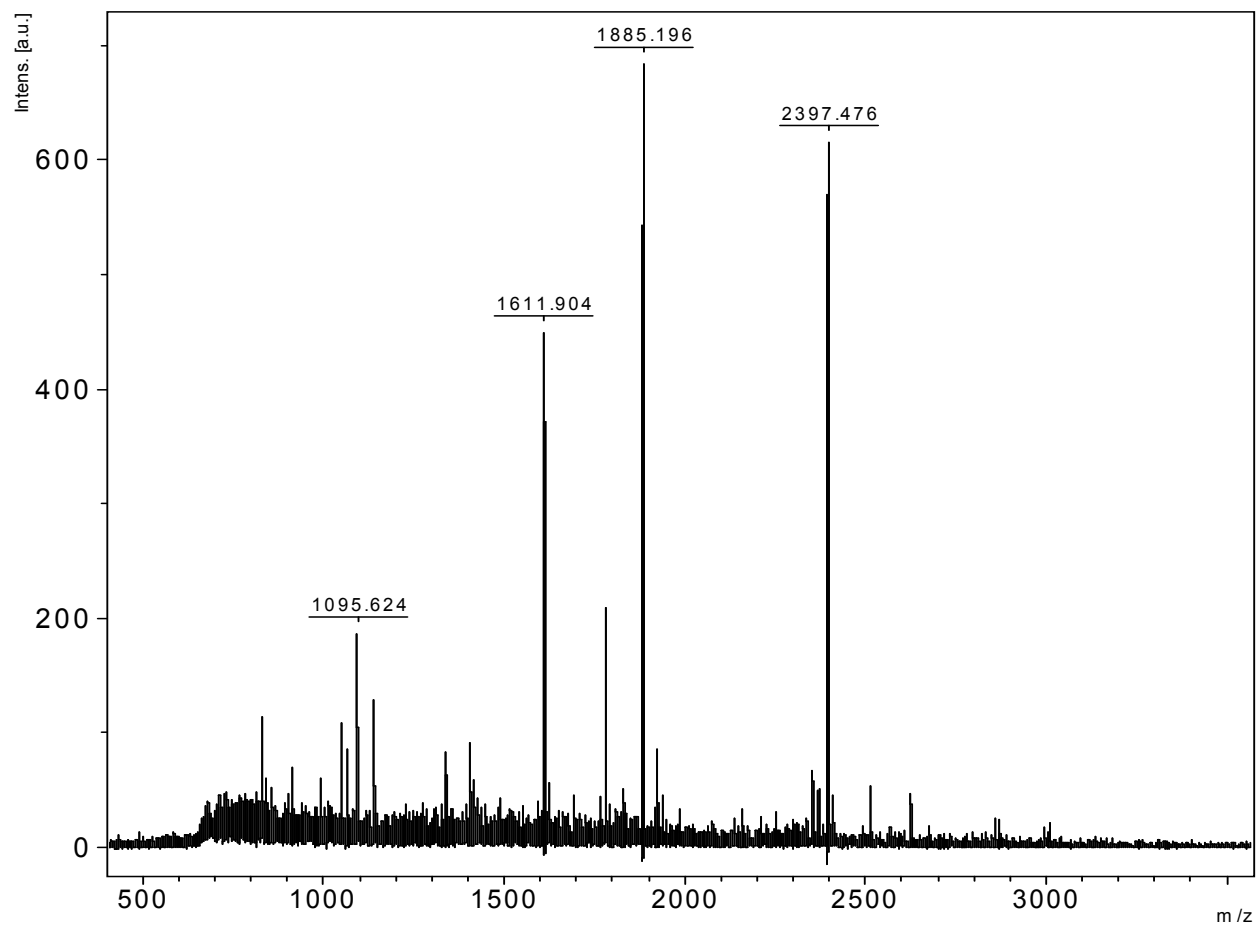
**Figure 23.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at 45 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 22. The protein was identified as actin.



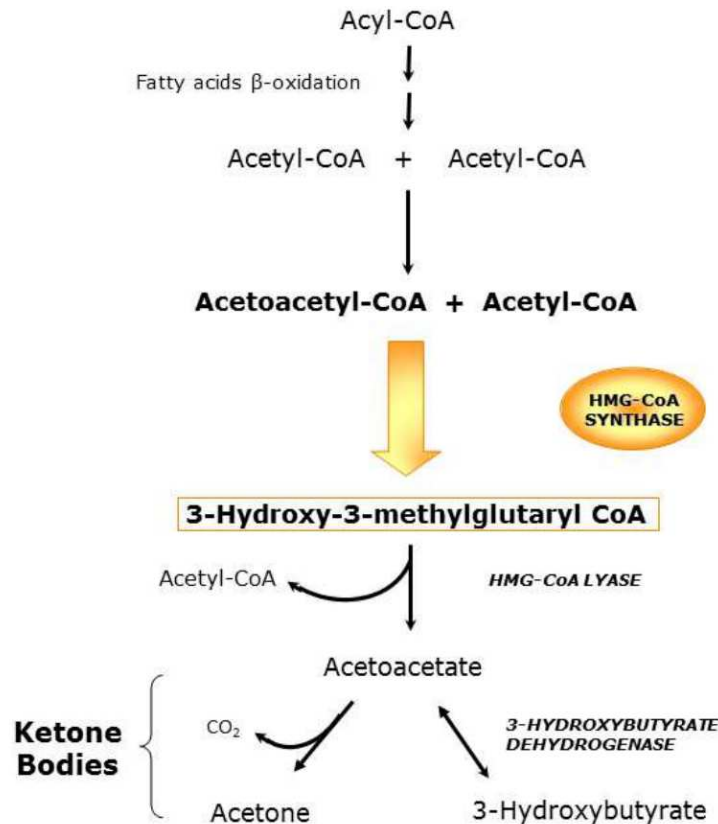
**Figure 24.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at 51 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 22. The protein was identified as a mixture of the  $\alpha$  and  $\beta$  subunits of  $F_1$ -ATPase.

The MALDI-TOF mass spectrum generated from the tryptic digest of the protein at about ~57 kDa in lane 4 of the SDS PAGE gel (Figure 22) is shown in Figure 25. Based upon the molecular weights of the tryptic peptides generated from this protein, the protein was identified as the mitochondrial isoform of HMG-CoA synthase. The mitochondrial form of this enzyme is responsible for the biosynthesis of ketone bodies (Figure 26). The main tissue where ketogenesis takes place is the liver; however, it has also been found in kidney, adipose tissue, and intestine.<sup>34</sup> Ketone bodies act as an alternative glucose fuel in a number of tissues such as heart, muscle and brain, and have a critical role during metabolic stress and starving. In patients with diabetes, this

enzyme is overactive, producing too many ketone bodies which can lead to ketoacidosis, coma, and death. Ketoacidosis is the lowering of the blood pH due to the overproduction of the acidic ketone bodies which contain carboxylic acid functional groups. A MALDI-TOF spectrum for the protein band at ~57 kDa in lane 2 of the gel was obtained which identified the protein as mitochondrial HMG-CoA synthase, but a corresponding MALDI-TOF spectrum for this protein in the bromo lanes 1 and 3 of the gel was not observed. It is difficult to determine if there is any protein band present in lanes 1 and 3 for this protein, and even if there was, the concentration of the protein was far too low to obtain a MALDI-TOF mass fingerprint spectrum. Therefore, this protein is a candidate for specifically binding to the C-peptide beads. C-peptide may play a role in inhibiting the production of excess ketone bodies when glucose levels are high in the bloodstream.



**Figure 25.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~57 kDa in lane 4 of the SDS-PAGE gel for bovine kidney shown in Figure 21. The protein was identified as the mitochondrial isoform of HMG-CoA synthase.

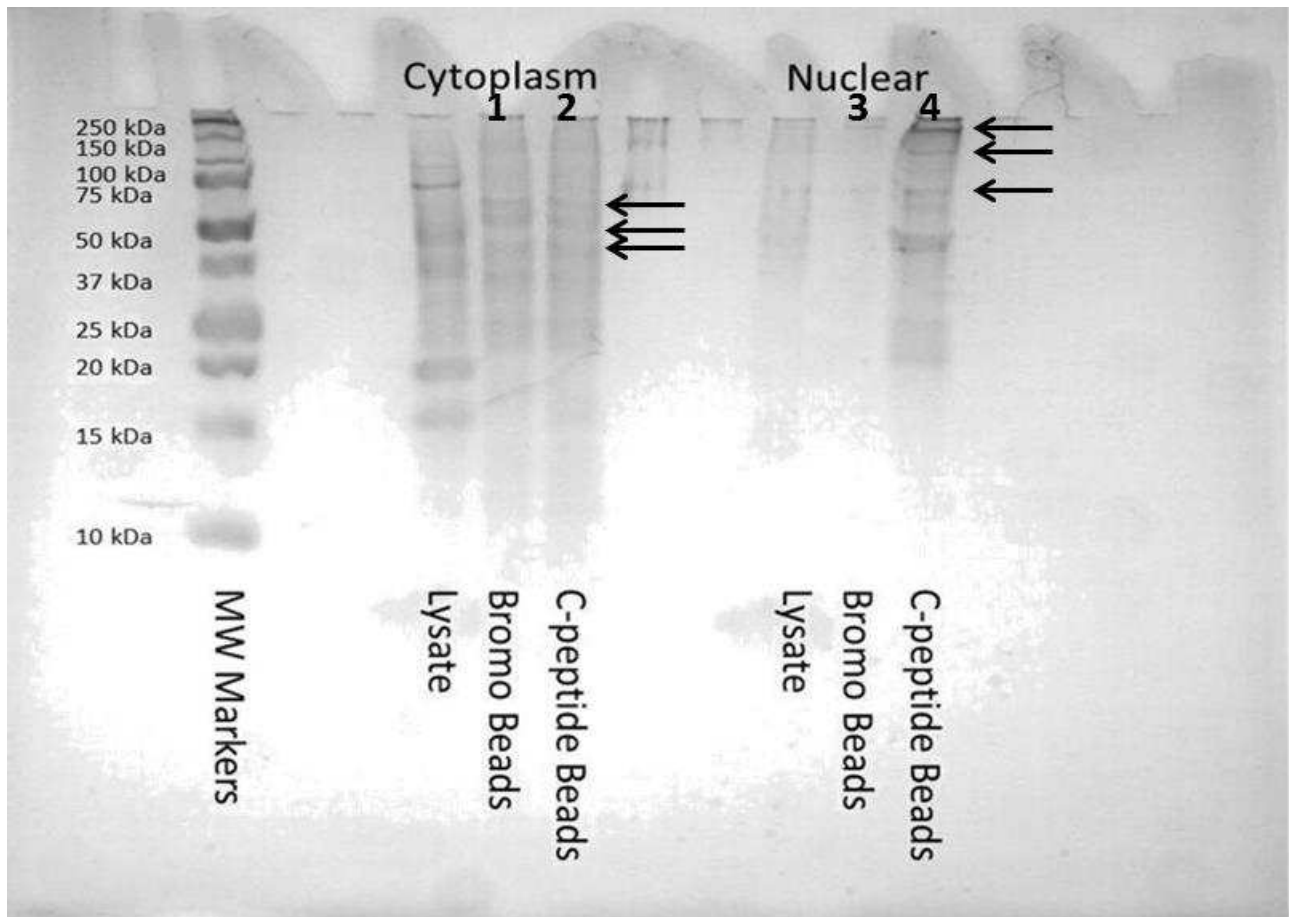


**Figure 26.** Ketone bodies synthesis pathway.<sup>35</sup>

### **Bovine Heart**

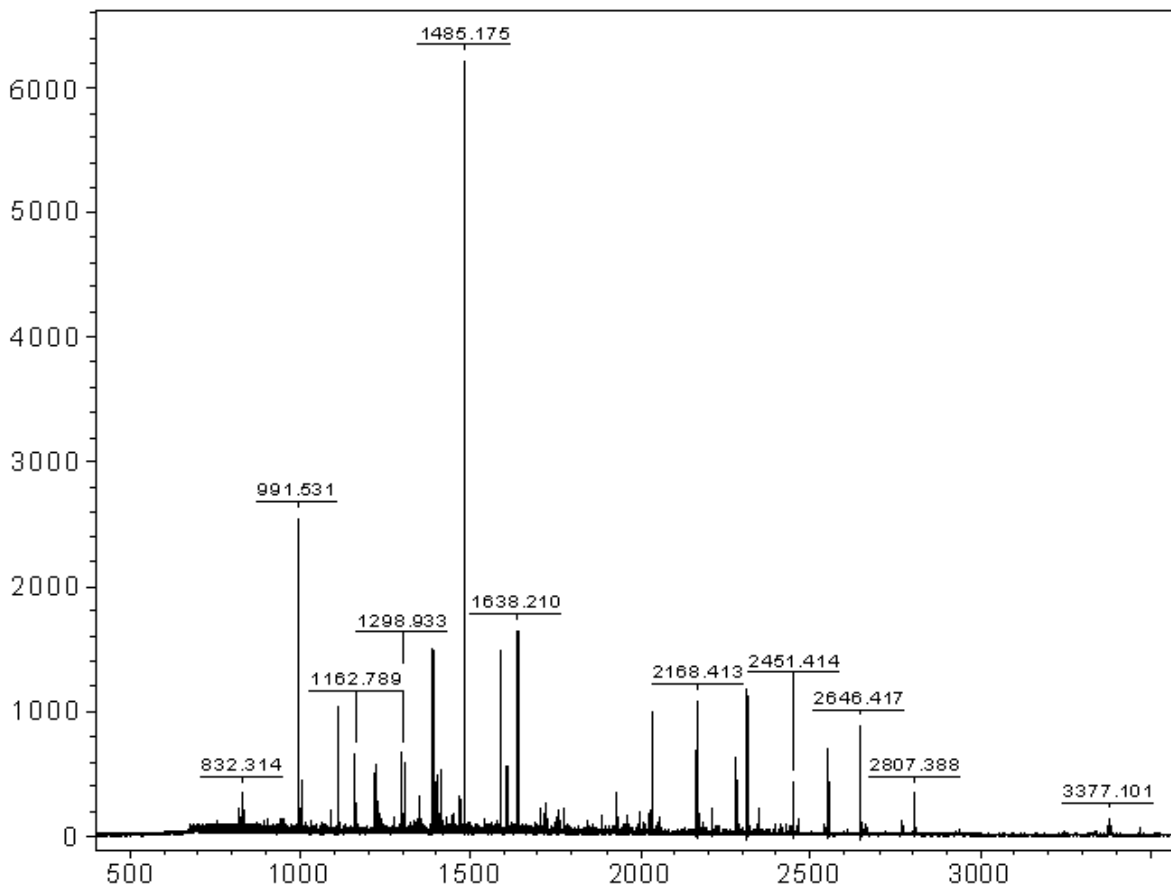
The previous experiment was repeated (Figure 19), and the only change was the use of bovine heart tissue instead of kidney tissue. Heart tissue was chosen for study because muscle tissues are one of the main target organs of the insulin peptide. C-peptide does not appear to directly alter glucose metabolism,<sup>36</sup> but it does appear that C-peptide and insulin-initiated signaling cascades interact.<sup>37</sup> Therefore, the heart may also be a target organ for the C-peptide. The SDS-PAGE gel of the bovine heart proteins isolated using the bromo beads and the C-peptide beads can be seen in Figure 27. There were distinct protein bands within the 37 kDa-75 kDa molecular weight range from both the bromo beads and the C-peptide beads for the cytoplasmic protein lysates. These protein bands were excised from the gels, digested with

trypsin, and identified by mass spectrometry using the mass fingerprint technique. The proteins were identified as actin,  $\alpha$  and  $\beta$  subunits of F1-ATPase, and HMG-CoA synthase. In this experiment, there does seem to be a protein band in the bromo bead lane for the cytoplasmic fraction that corresponds to the HMG-CoA synthase protein band in the C-peptide lane; however, the MALDI-TOF spectra for this protein did not confirm the presence of HMG-CoA synthase in the bromo lane, so this protein is still a candidate for specifically binding to the C-peptide beads. There were also distinct protein bands within the 50 kDa-250 kDa molecular weight range from the C-peptide beads for the cytoplasmic protein lysates.



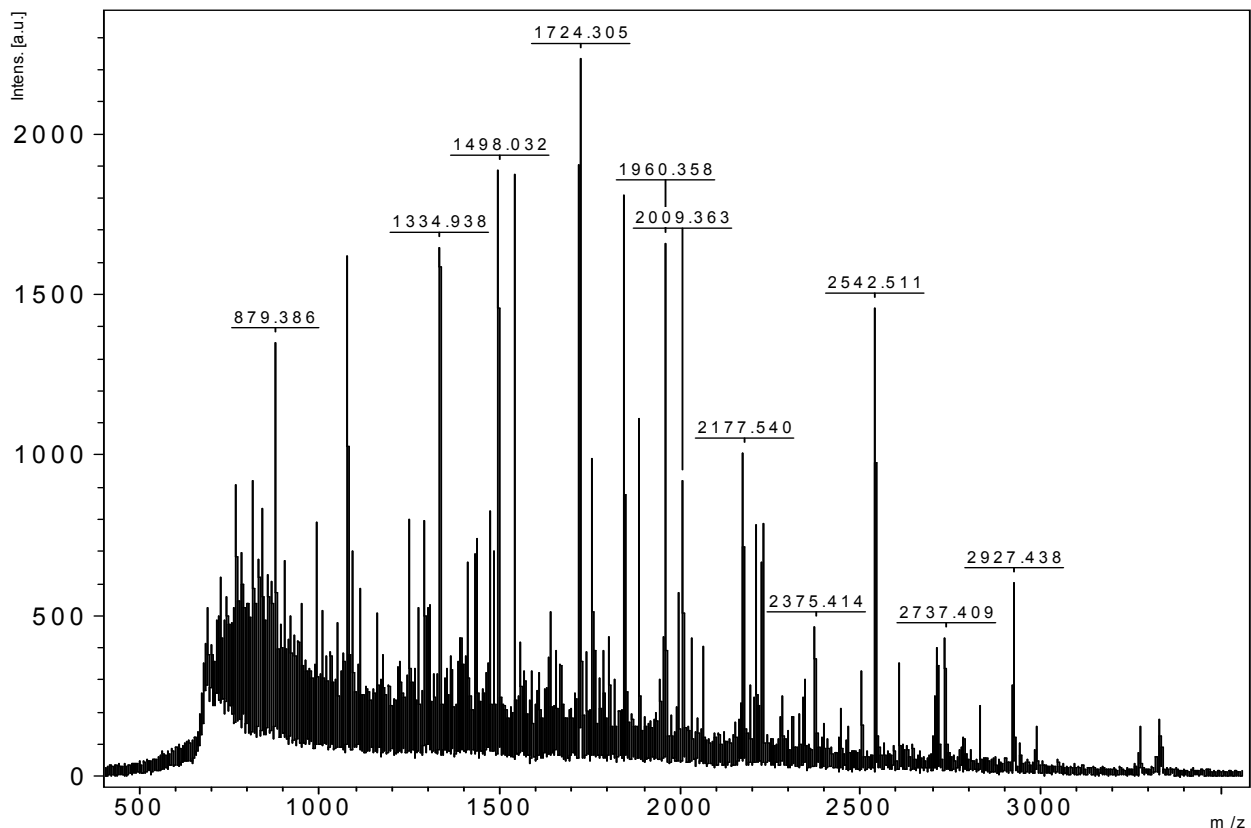
**Figure 27.** SDS-PAGE gel of proteins isolated from bovine heart tissue. The first lane contains molecular weight protein standards.

The MALDI-TOF mass spectrum generated from the tryptic digest of the protein at about ~150 kDa in lane 4 of the SDS PAGE gel (Figure 27) is shown in Figure 28. Based upon the molecular weights of the tryptic peptides generated from this protein, the protein was identified as carbamoyl phosphate synthase. This protein is an enzyme that catalyzes the ATP dependent synthesis of carbamoyl phosphate from glutamine or ammonia and bicarbonate. It represents the first committed step in pyrimidine and arginine biosynthesis. Unfortunately, this protein was also identified from the bromo beads lane in the nuclear protein lysates from the first trial using bovine kidney tissue (Figure 21). Therefore, this protein is also not a good candidate for specifically binding to the C-peptide beads.



**Figure 28.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~ 150 kDa in lane 4 of the SDS-PAGE gel for bovine heart shown in Figure 27. The protein was identified as carbamoyl phosphate synthase.

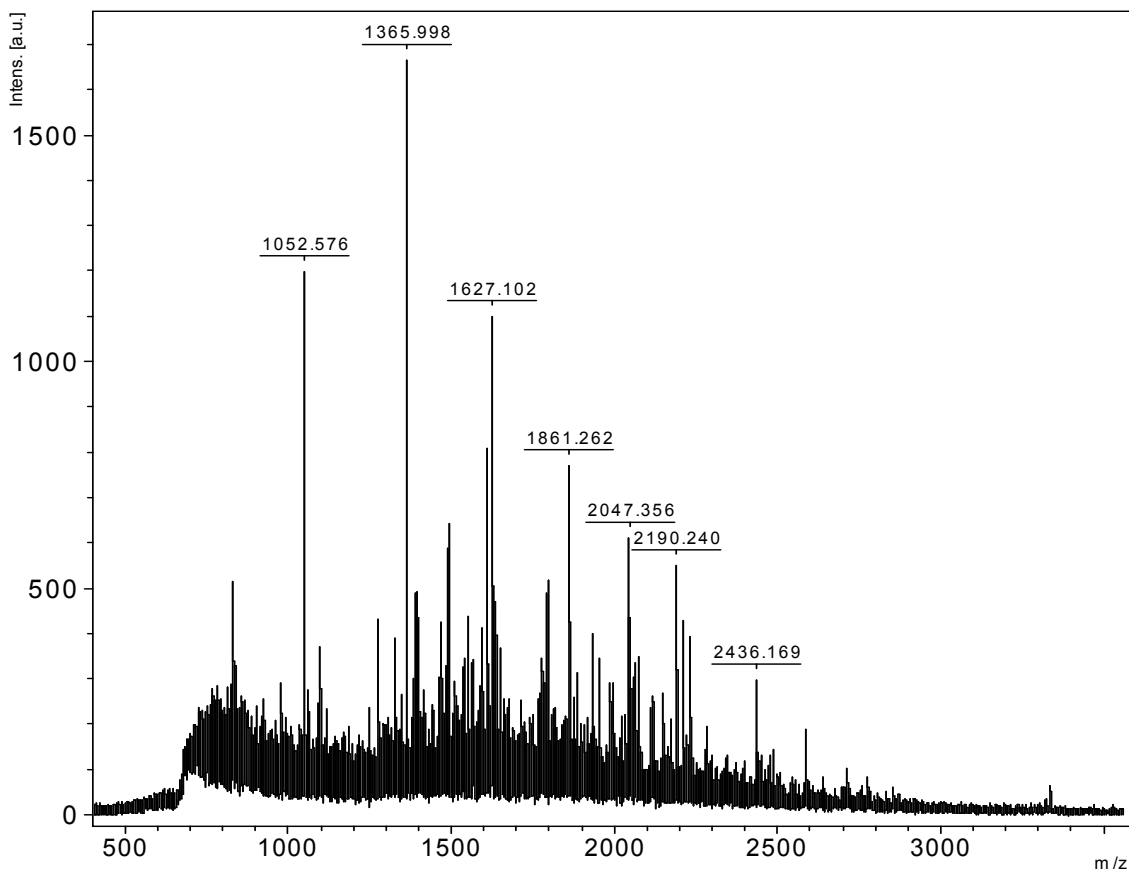
The protein band excised from the ~250 kDa position in lane 4 of the SDS-PAGE gel (Figure 27) was identified as fatty acid synthase. The MALDI-TOF mass spectrum generated from the tryptic digest of this protein is shown in Figure 29. This protein has not been previously identified from the bromo lanes of any gels, and appears to be specifically isolated from the C-peptide beads. Fatty acid synthase is a multi-enzyme protein that catalyzes the synthesis of palmitic acid from acetyl-CoA and malonyl-CoA in the presence of NADPH. The identification of this enzyme as a C-peptide binding protein is interesting, because C-peptide may be influencing the activity of this enzyme. It could activate this enzyme to convert the excess glucose from the bloodstream metabolized to acetyl-CoA into a fatty acid for storage.



**Figure 29.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~250 kDa in lane 4 of the SDS-PAGE gel for bovine heart shown in Figure 27. The protein was identified as fatty acid synthase.



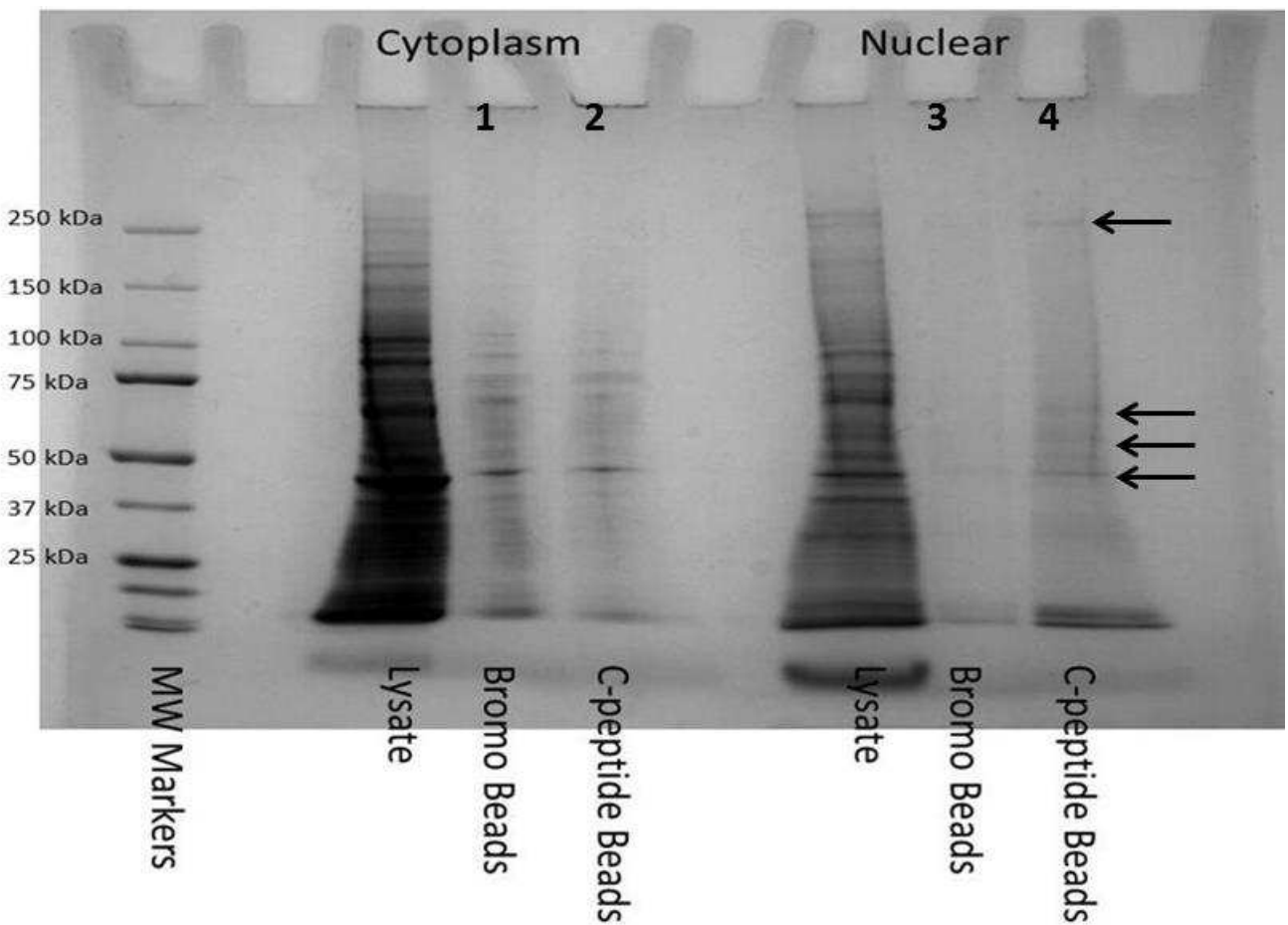
The MALDI-TOF mass spectrum generated from the tryptic digest of the protein at about ~80 kDa in lane 4 of the SDS PAGE gel (Figure 27) is shown in Figure 30. Based upon the molecular weights of the tryptic peptides generated from this protein, the protein was identified as long chain fatty acid CoA ligase. This protein has not been previously identified from the bromo lanes of any gels, and appears to be specifically isolated from the C-peptide beads. In general, fatty acyl-CoA ligases are enzymes that activate long-chain fatty acids for both synthesis of cellular lipids and degradation via beta-oxidation. There are several different isoforms of this enzyme, and this isoform may activate fatty acids for the synthesis of triacylglycerol destined for intracellular storage.



**Figure 30.** MALDI-TOF peptide fingerprint mass spectrum for the protein excised from the band at ~80 kDa in lane 4 of the SDS-PAGE gel for bovine heart shown in Figure 27. The protein was identified as long chain fatty acid CoA ligase.

## Bovine Thymus

The previous experiment was repeated (Figure 19), and the only change was the use of bovine thymus tissue instead of kidney tissue. There was no specific reason for using this tissue type other than that it was easily obtained from the slaughter house. The SDS-PAGE gel of the bovine thymus proteins isolated using the bromo beads and the C-peptide beads can be seen in Figure 31. The protein bands from the cytoplasmic lysate are not well resolved, and this could be because the lanes (especially the lysate only lane) were overloaded with protein. It was decided to concentrate on the protein bands in the nuclear lysate.



**Figure 31.** SDS-PAGE gel of proteins isolated from bovine thymus tissue. The first lane contains molecular weight protein standards.

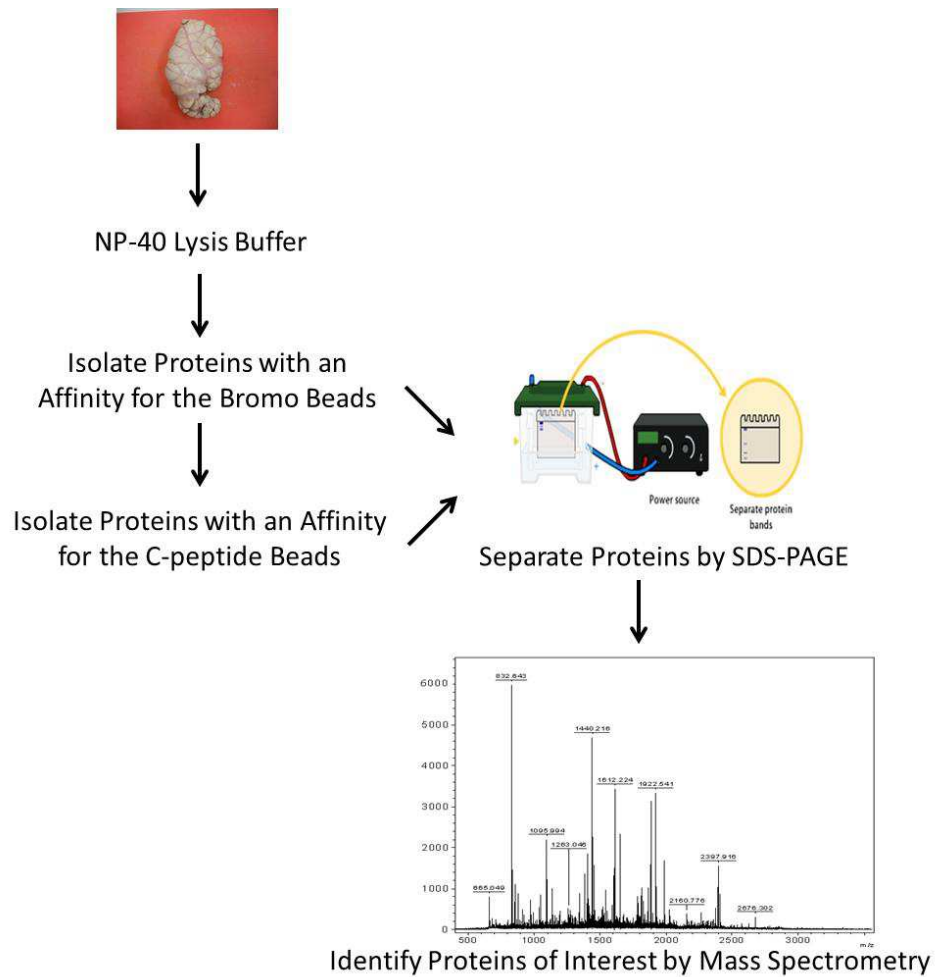
MALDI-TOF peptide mass fingerprint spectra were obtained for the protein bands in lane 4 of the gel (Figure 31) present at about 45 kDa, 52 kDa, 58 kDa, and 250 kDa. The mass spectra confirmed that the proteins were actin, F1-ATPase, HMG-CoA synthase, and fatty acid synthase just as previously identified in the heart and kidney lysates. The presence of actin in the bromo lane (lane 3) of the gel was also confirmed by mass spectrometry.

### **NP-40 Protein Lysates**

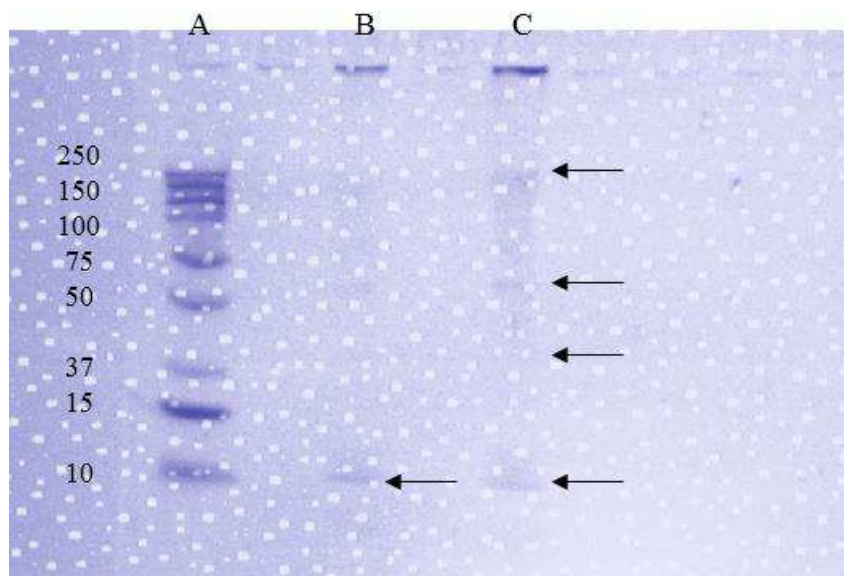
Multiple studies have shown that C-peptide binds to many different cell types, including human skin fibroblasts, kidney tubule cells, and endothelial cells.<sup>6, 38-39</sup> The C-peptide has been shown to initiate signaling cascade consistent with C-peptide interacting with a GPCR, and the actions of C-peptide in several cell systems were pertussis toxin sensitive.<sup>7,40</sup> There was concern that a cell membrane bound receptor protein in our lysates would be missed, because previous lysates focused on cytoplasmic and nuclear proteins, not cell membrane bound proteins. Therefore, it was decided to use an alternate method to lyse the tissue samples that included a detergent that would help solubilize membrane bound proteins. A mild detergent was needed that would solubilize the proteins without unfolding their tertiary structure. It was assumed that the proteins would need to be folded into their native tertiary structure in order to interact and bind to C-peptide. Therefore, the mild detergent NP-40 was chosen to lyse the tissue.

The basic experimental protocol for this experiment is summarized in Figure 32. The only difference between this experiment and the previous experiment (Figure 19) was the way in which the tissue samples were lysed. The SDS-PAGE gel of the bovine thymus proteins isolated using the C-peptide beads can be seen in Figure 33. The lysates were precleared with the bromo beads prior to incubation with the C-peptide beads. The protein bands marked with the arrows on the gel were excised, digested with trypsin, and analyzed by mass spectrometry to obtain a

peptide mass fingerprint spectrum for each protein. The molecular weights of the tryptic peptides are summarized in Table 1. When the molecular weights were entered in Mascot, it was determined that all four proteins were contaminated with keratin. Only the protein at ~40 kDa could be positively identified as actin. There was a protein band at ~250 kDa that may have been fatty acid synthase, but the keratin contamination made identification of this protein impossible.



**Figure 32.** Basic experimental protocol for isolating proteins extracted from bovine thymus using NP-40 lysis buffer with the C-peptide affinity beads.



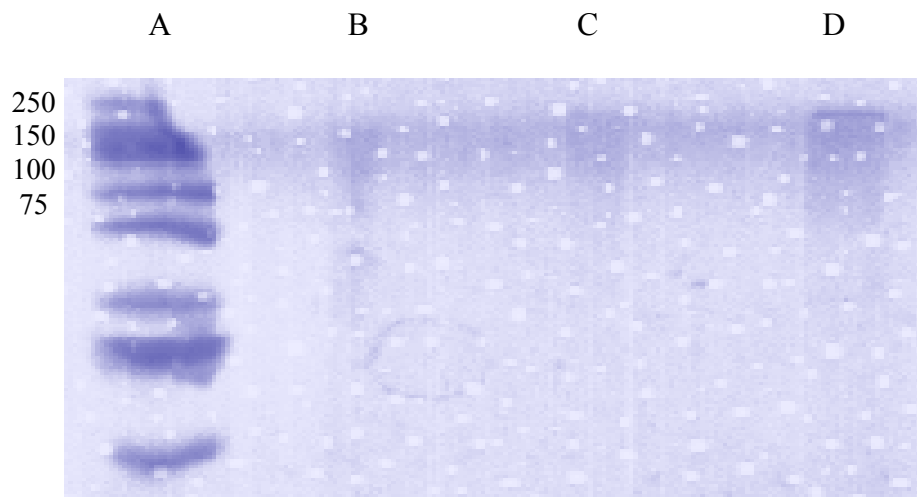
**Figure 33.** SDS-PAGE gel of proteins isolated from bovine thymus tissue using NP-40 lysis buffer. Lane A contains molecular weight protein standards, lane B is bovine thymus proteins isolated with C-peptide beads, and lane C is more concentrated bovine thymus proteins isolated with C-peptide beads.

Tryptic Peptide Mass Fingerprint for Protein at ~10 kDa	Tryptic Peptide Mass Fingerprint for Protein at ~40 kDa	Tryptic Peptide Mass Fingerprint for Protein at ~60 kDa	Tryptic Peptide Mass Fingerprint for Protein at ~150 kDa	
Measured in Da	Measured in Da	Measured in Da	Measured in Da	
989.6	1198.8	1060.8	1060.8	1791.2
1274.8	1277.8	1277.8	1179.8	1839.2
1266.0	1475.8	1476.0	1234.9	1852.2
1475.9	1516.9	1941.2	1278.0	1994.4
1635.0	1791.1	2211.5	1303.0	2150.5
1639.0	1954.2	842.6	1307.9	2211.5
1744.0	1994.2	1480.0	1476.0	2399.5
1994.2	2211.4	2283.5	1639.2	2705.8
2211.2	-	-	1708.1	2873.0
-	-	-	3053.2	-

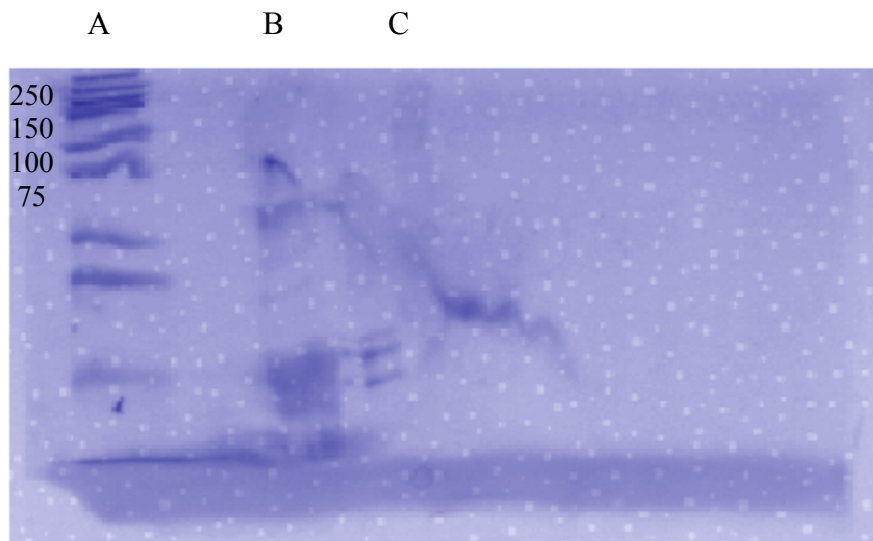
**Table 1.** Molecular weights of tryptic peptides obtained from the mass spectra of the four protein bands excised from the gel shown in Figure 32.

This experiment was repeated with bovine thymus tissue, and we also lysed rat liver

tissue and human cheek cells were also lysed. The SDS-PAGE gel of the proteins isolated using the C-peptide beads can be seen in Figures 34 and 35. Again, good enough mass fingerprint spectra for any protein bands were unattainable. These experiments with the NP-40 lysis buffer need to be repeated in future studies. Not any good protein bands on the SDS-PAGE gels were obtained using this procedure. The lysis protocol using NP-40 detergent may need to be adjusted to ensure that the proteins remain folded in their native tertiary structures.



**Figure 34.** SDS-PAGE separation of proteins eluted from the C-peptide beads using human cheek lysate, rat liver lysate and bovine thymus lysate. Lane A is the protein markers. Lane B is the cheek lysate. Lane C is the rat liver lysate. Lane D is the bovine thymus lysate.



**Figure 35.** SDS-PAGE separation of proteins eluted from the C-peptide beads using rat liver lysate and bovine thymus lysate. Lane A is the protein markers. Lane B is the rat liver lysate. Lane C is the bovine thymus lysate.

## CONCLUSIONS

In all the experiments, one common theme persisted. Most of the proteins isolated with the C-peptide beads tended to be the same molecular weights, even when different tissue types were used. A number of these reoccurring protein bands that were shared among the different gels were excised, and the proteins were identified using the peptide mass fingerprint technique. Table 2 is a list of the proteins identified indicating their location within the cell and their general cellular function.

<b>Name</b>	<b>Weight in kDa</b>	<b>Cellular Location</b>	<b>Cellular Function</b>
Actin	~45	Cytoplasmic	muscle contraction, cell motility, cell division and cytokinesis, vesicle and organelle movement, cell signaling, and establishment and maintenance of cell junctions and cell shape, etc.
CPS I	~165	Mitochondria	urea cycle-removes excess ammonia from cells
F <sub>1</sub> -ATPase	~51	Mitochondria	catalyze synthesis of ATP during oxidative phosphorylation
Long chain fatty acid CoA ligase	~80	Mitochondria	lipid biosynthesis
Fatty acid synthase	~250	Cytoplasmic	central enzyme in de novo lipogenesis. FAS is a target for SREBP and is upregulated by LXR activation; it is also one of the accepted markers for insulin resistance, SREBP and LXR activation.
HMG-CoA synthase	~57	Cytoplasmic	Ketone body synthesis. This enzyme condenses acetyl CoA with acetoacetyl CoA to form HMG CoA, which is the substrate for HMG CoA reductase.

**Table 2.** Table of proteins identified in this study indicating their cellular location and function.

\* All information obtained from [www.abcam.com](http://www.abcam.com) and [www.sbet.com](http://www.sbet.com).



Most of these proteins were isolated from the nuclear lysate fractions, however, most of the identified proteins are normally cytoplasmic proteins. The fact that they were isolated from the nuclear fractions most likely indicates that differential lysis of the cells was not entirely successful. The procedure used for lysis works best with cultured cells, but tissue was being used in these studies. Some of these proteins may also be located in the nucleus of cells under special conditions, but most likely there were cytosolic proteins mixed in with the nuclear protein lysates.

The isolation of both fatty acid synthase and long chain fatty acyl-CoA ligase indicates that C-peptide may play a role in stimulating the production of fatty acids from excess glucose and converting those fatty acids to triacylglycerides for storage inside muscle cells. C-peptide does not appear to directly alter glucose metabolism<sup>36,41</sup>, but it does interact with insulin-initiated signaling cascades which is important for the normal functioning of both peptides. Accumulating evidence suggests that lipid metabolism is as or more important to diabetes than carbohydrate metabolism. Insulin is an important regulator of fatty acid synthase. Insulin upregulates the expression of fatty acid synthase in murine cell lines and primary human adipocytes and increases fatty acid synthase enzymatic activity.<sup>42</sup> In addition, insulin resistance has been shown to impair the process of de novo lipogenesis in adipose tissue. C-peptide has a half-life in serum of about 30 minutes, which is much longer than insulin's half-life of about 3 minutes.<sup>43</sup> Perhaps C-peptide helps to fine tune the effect of insulin on fatty acid synthase and extend the effects to a longer time frame. In addition, C-peptide may be more targeted to muscle cells than liver cells. C-peptide escapes first-pass metabolism by the liver indicating that it may not enter liver cells as readily as other cell types. The results of this study indicate that C-peptide may be more involved in modulating lipid metabolism within cells. It may play a role in

determining the fate of the excess glucose that enters cells by stimulating the production of fatty acids and conversion of those fatty acids to triacylglycerides for short-term intracellular storage instead of sending the fatty acids to the adipose tissue for long-term storage.

### **FUTURE**

Even though some interesting results were obtained for this experiment, there is still a lot of work to be done on this project. The experiments using the NP-40 lysis buffer need to be repeated, and a better protocol for this lysis needs to be established to ensure the proteins remain folded in their native tertiary structure. Sequencing a few of the peptides in order to confirm protein identifications, especially for the fatty acid synthase protein and the long chain acyl-CoA ligase protein, is needed before this work can be published. In addition, it would be best to repeat the entire cytoplasmic and nuclear lysis procedure (Figure 19) using human tissue or human cell lines since human C-peptide was used to make the C-peptide affinity beads.

## LITERATURE REFERENCES

1. Nussey S, Whitehead S. Endocrinology: An Integrated Approach. Oxford: BIOS Scientific Publishers; 2001. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK22/>
2. Wahren, J.; Shafqat, J.; Johansson, J.; Chibalin, A.; Ekberg, K.; and Jörnvall, H. Molecular and Cellular Effects of C-peptide-New Perspectives on an Old Peptide. *Experimental Diab. Res.* **2004**, *5*, 15-23.
3. Steiner, D. The Proinsulin C-peptide-A Multirole Model. *Experimental Diab. Res.* **2004**, *5*, 7-14.
4. National Center for Biotechnology Information. PubChem Compound Database; CID=16132309, <https://pubchem.ncbi.nlm.nih.gov/compound/16132309> (accessed July 25, 2015).
5. Wahren, J.; Ekberg, K.; and Jörnvall, H. C-peptide is a bioactive peptide. *Diabetologia.* **2007**, *50*, 503-509.
6. Rigler, R.; Pramanik, A.; Jonasson, P.; Kratz, G.; Jansson, O.T.; Nygren, P.; Stahl, S.; Ekberg, K.; Johansson, B.L.; Uhlen, S.; Jörnvall, H.; Wahren, J. Specific Binding of Proinsulin C-peptide to Human Cell Membranes. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 13318-13323.
7. Kitamura T, Kimura K, Jung BD, et al. Proinsulin C-peptide rapidly stimulates mitogen-activated protein kinases in Swiss 3T3 fibroblasts: requirement of protein kinase C, phosphoinositide 3-kinase and pertussis toxin-sensitive G-protein. *Biochemical Journal.* **2001**; *355(Pt 1)*:123-129.
8. Johansson, J.; Ekberg, K.; Shafqat, J.; Henriksson, M.; Chibalin, A.; Wahren, J.; and Jörnvall, H. Molecular effects of proinsulin C-peptide. *Biochem. Biophys. Res. Commun.* **2002**, *295*, 1035–1040.
9. Wahren, J. C-peptide: new findings and therapeutic implications in diabetes. *Clin. Physiol. Funct. Imaging* **2004**, *24*, 180–189.
10. Abrol, Ravinder., and W. A. Goddard 3rd. "G protein-coupled receptors: Conformational "gatekeepers" of transmembrane signal transduction and diversification." *Extracellular and intracellular signaling* 2011: 188-229.
11. Schöneberg, T.; Schultz, G.; Gudermann, T. Structural basis of G protein-coupled receptor function. *Molecular and cellular endocrinology* **1999**, *151.1*, 181-193.

12. Strader, C. D.; Fong, T.M.; Tota, M.R.; Underwood, D. Structure and function of G protein-coupled receptors. *Annual review of biochemistry* **1994**, *63.1*, 101-132.
13. Lodish, H.; Berk, A.; Zipursky, S.L.; et al. G Protein –Coupled Receptors and Their Effectors. In *Molecular Cell Biology*. 4th edition. Freeman, W.H. New York, 2000. Available from: <http://www.ncbi.nlm.nih.gov/books/NBK21718/>
14. Yeagle, P. L.; Albert, A. D. G-protein coupled receptor structure. *Biochimica et Biophysica Acta (BBA)-Biomembranes*. **2004**, *1768(4)*, 808-824.
15. Lee, S. Modern Mass Spectrometry, MacMillan Group Meeting, Princeton, NJ, 2005.
16. Siuzdak, G. *The Expanding Role of Mass Spectrometry in Biotechnology*, 2<sup>nd</sup> ed.; MCC Press, 2006.
17. Lewis, J.K.; Wei, J.; and Siuzdak, G. Matrix-Assisted Laser Desorption/Ionization in Peptide and Protein Analysis. *Encyclopedia of Analytical Chemistry*. **2000**. 5880-5894.
18. Jagtap, R. N., and Ambre, A. H. Overview literature on matrix assisted laser desorption ionization mass spectroscopy (MALDI MS): basics and its applications in characterizing polymeric materials. *Bulletin of Materials Science*. **2005**. *28.6*: 515-528.
19. Matrix Assisted Laser Desorption-MALDI.  
<https://www2.unine.ch/files/content/sites/saf/files/shared/documents/MALDIversus2.pdf> (accessed November 26, 2015).
20. Thiede, B.; Höhenwarter. W.; Krah, A.; Mattow, J.; Schmid, M.; Schmidt, F.; and Jungblut, P.R. Peptide mass fingerprinting. *Methods*, January 2005, 237-247.
21. Yalow, R.S. Radioimmunoassay: Historical Aspects and General Considerations. In *Radioimmunoassay in Basic and Clinical Pharmacology*; Patrono, C.; Peskar, B.A.; Eds.; Springer-Verlag Berlin Heidelberg, 1987; Vol 82; p 1-6.
22. Chemical Heritage Foundation. Rosalyn Yalow and Solomon Berson.  
<http://www.chemheritage.org/discover/online-resources/chemistry-in-history/themes/pharmaceuticals/diagnosing-diseases/yalow-and-berson.aspx> (accessed July 16, 2015).
23. Neuroscience Core Facility. Biological Assays. <http://sites.gsu.edu/neuroscience-core/training/assays/>. (accessed Nov. 11, 2015).
24. Antibodies Online. Radioimmunoassay. <http://www.antibodies-online.com/resources/17/1215/Radioimmunoassay+RIA/>. (accessed Nov. 11, 2015).

25. Cuatrecasas, P.; Desbuquois, D.; and Krug, F. Insulin-receptor interactions in liver cell membranes. *Biochemical and biophysical research communications*. **1971**, *44.2*: 333-339.
26. Cuatrecasas, P. Insulin-receptor interactions in adipose tissue cells: direct measurement and properties. *Proceedings of the National Academy of Sciences*, June 1971, 1264-1268.
27. Lefkowitz, R. J.; Roth, J.; and Pastan, I. Radioreceptor assay of adrenocorticotrophic hormone: new approach to assay of polypeptide hormones in plasma. *Science*. **1970**, *170*, 633-635.
28. Wahren, J.; Kallas, A.; Sima, A.A. The clinical potential of C-peptide replacement in type 1 diabetes. *Diabetes*. **2012**, *61*, 761-722.
29. Yosten, G.L.; Maric-Bilkan, C.; Luppi, P.; Wahren, J. Physiological effects and therapeutic potential of proinsulin C-peptide. *Am J. Physiol. Endocrinol Metab*. **2014**, *307*, E955-E968.
30. Ekberg, K.; Brismar, T.; Johansson, B.L.; Jonsson, B.; Lindström, P.; Wahren, J. Amelioration of sensory nerve dysfunction by C-peptide in patients with type 1 diabetes. *Diabetes*. **2003**, *52*, 536–541.
31. Lindahl, E.; Nyman, U.; Melles, E.; Sigmundsson, K.; Ståhlberg, M.; Wahren, J.; Obrink, B.; Shafqat, J.; Joseph, B.; Jörnvall, H. Cellular internalization of proinsulin C-peptide. *Cell Mol. Life Sci* **2007**, *64*, 479-486.
32. Re, R. N. The cellular biology of angiotensin: paracrine, autocrine and intracrine actions in cardiovascular tissues. *J. Mol. Cell. Cardiol*. **1989**, *21*, 63-69.
33. Dominguez, R.; Holmes, K.C. Actin structure and function. *Annu Rev Biophys*. **2011**, *40*, 169-186.
34. Thumelin, S.; Forestier, M.; Girard, J.; Pegorier, J.P. Developmental changes in mitochondrial 3-hydroxy-3methylglutaryl-CoA synthase gene expression in rat liver, intestine and kidney. *The Biochemical Journal*. **1993**, *292*, 493-496.(1993).
35. Arnedo, M.; Ramos, M.; Puisac, B.; Gil-Rodríguez, M.C.; Teresa, E.; Pié, A.; Bueno, G.; Ramos, F.J.; Gómez-Puertas, P.; Pie, J. Mitochondrial HMG–CoA Synthase Deficiency. In *Advances in the Study of Genetic Disorders*; Ikehara, K.; InTech, 2011.
36. Forst, T.; Rave, K.; Pfuetzner, A.; Buchholz, R.; Pohlmann, T. Lobig, M.; Heinemann, L. Effect of C-Peptide on Glucose Metabolism in Patients With Type 1 Diabetes. *Diabetes Care*. **2002**, *25*, 1096-1097.

37. Richards, J.P.; Yosten, G.L.; Kolar, G.R.; Jones, C.W.; Stephenson, A.H.; Ellsworth, M.L.; Sprague, R. S. Low O<sub>2</sub>-induced ATP release from erythrocytes of humans with type 2 diabetes is restored by physiological ratios of C-peptide and insulin. *Am J Physiol Regul Comp Physiol* **2014**, *307*, R862-R868.
38. Henriksson, M.; Pramanik, A.; Shafqat, J.; Zhong, Z.; Tally, M.; Ekberg, K.; Wahren, J.; Rigler, R.; Johansson, J.; Jörnvall, H. Specific binding of proinsulin C-peptide to intact and to detergent-solubilized human skin fibroblasts. *Biochem Biophys Res Commun.* **2001**, *280*(2), 423-427.
39. Pramanik, A.; Ekberg, K.; Zhong, Z.; Shafqat, J.; Henriksson, M.; Jansson, O.; Tibell, A.; Tally, M.; Wahren, J.; Jörnvall, H. et al. C-peptide binding to human cell membranes: importance of Glu27. *Biochem Biophys Res Commun.* **2001**, *284*(1),:94–98.
40. Walcher, D.; Aleksic, M.; Jerg, V.; Hombach, V.; Zieske, A.; Homma, S.; Strong, J.; Marx, N. C-Peptide Induces Chemotaxis of Human CD4-Positive Cells. *Diabetes.* **2004**, *53*, 1664-1670.
41. Wahren, J.; Larsson, C. C-peptide: Newfindings and therapeutic possibilities. *Diabetes Res Clin Pract.* **2015**, *107*, 309-319.
42. Claycombe, K.J.; Jones, B.H.; Standridge, M.K.; Guo, Y.; Chun, J.T.; Taylor, J.W.; Moustaid-Moussa, N. Insulin increases fatty acid synthase gene transcription in human adipocytes. *Am J Physiol.* **1998**, *274*, R1253–1259.
43. Faber, O.K.; Kehlet, H.; Madsbad, S.; Binder, C. Kinetics of human C-peptide in man. *Diabetes.* **1978**, *27*, 207–209.

## APPENDIX A



Office of Research Integrity

December 7, 2015

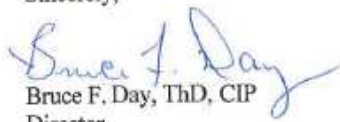
Christina Newsome  
Master's Candidate  
Chemistry Department  
Marshall University

Dear Ms. Newsome:

This letter is in response to the submitted thesis abstract entitled "*Investigation into the Biological Function of Proinsulin C-Peptide.*" After assessing the abstract it has been deemed not to be human subject research and therefore exempt from oversight of the Marshall University Institutional Review Board (IRB). The Institutional Animal Care and Use Committee (IACUC) Chair has also deemed this not to be animal research requiring their approval. The information in this study is not considered human subject or animal research as set forth in the definitions contained in the federal regulations. If there are any changes to the abstract you provided then you would need to resubmit that information to the Office of Research Integrity for review and determination.

I appreciate your willingness to submit the abstract for determination. Please feel free to contact the Office of Research Integrity if you have any questions regarding future protocols that may require IRB review.

Sincerely,

  
Bruce F. Day, ThD, CIP  
Director

**WE ARE... MARSHALL.**

One John Marshall Drive • Huntington, West Virginia 25755 • Tel 304/696-4303  
A State University of West Virginia • An Affirmative Action/Equal Opportunity Employer

## APPENDIX B EXPERIMENTAL PROTOCOLS

### A. Making Beads Pre-clearing

1. Beads swell in DCM for 1 hour.
2. Wash beads with DMF two times.
3. Let beads sit in lysis buffer until further use.

### B. Making Beads with Peptide Attached

1. Beads swell in DCM for 1 hour.
2. Wash beads with DMF two times.
3. Remove supernatant and discard.
4. Add in 150  $\mu$ L of 0.1 M NaOH to 150  $\mu$ L of H<sub>2</sub>O and then add a small amount of peptide.
5. Incubate for 1-2 hours.
6. Remove supernatant and wash beads with DCM three times, DMF three times, and lysis buffer three times.

### C. Cell Lysate Preparation

1. Add harvested cells to 1 mL of lysis buffer in a clean siliconized tube, making sure to be kept on ice.
2. Add in the Protease Cocktail Inhibitors to the tube.  
  
1  $\mu$ L of aproptin  
  
1 $\mu$ L of pepstatin
3. Vortex for about 5 minutes.
4. Sonicate for about 5 minutes
5. Put onto shaker, on ice, for about 30 minutes to an hour to let the lysate



incubate.

6. Centrifuge and remove the supernatant to keep for the following pre-clearing process part D.

**D. Pre-clearing Process**

1. Take the cell lysate and add to 30  $\mu\text{L}$  of pre-clearing beads that are in the lysis buffer.
2. Incubate for 1 hour.
3. Remove supernatant and keep for part E.

**E. Experimental Procedure**

1. Take the pre-cleared cell lysate and add to 30  $\mu\text{L}$  of the beads with peptide attached.
2. Incubate 1-2 hours.
3. Remove supernatant and keep to retrieve peptide at a later date.
4. Wash beads twice with PBS.
5. Add 30  $\mu\text{L}$  of SDS-PAGE loading buffer with DTT to the beads.
6. Heat for 10 minutes in boiling water.
7. Load 12  $\mu\text{L}$  of Precision Plus Protein Standards kaleidoscope onto the first lane of the SDS-PAGE electrophoresis gel.
8. By skipping a lane(s), load 27  $\mu\text{L}$  of each sample(s) onto the gel. Add enough Tris/Glycine/SDS Buffer to cover the gel and the surrounding area.
9. Run the gel to completion ~45 minutes to an hour at 200 V and keep for analyzing.

**F. Steps to Prepare for Analyzing**

1. Make sure to wear gloves and never touch the gel with your hands, disassemble the electrophoresis apparatus. Take the gel out and be careful not to tear it.
2. Wash the gel a few times with distilled water to wash off all of the SDS-Page Buffer Solution.
3. Add enough Coomassie Blue Stain to cover the gel.
4. Let it sit on the shaker overnight.
5. Wash the gel with distilled water to remove the stain.
6. Unless used immediately, cover with enough distilled water to cover the gel and put into freezer for further analysis.

**G. In Gel Digestion Procedures**

1. Prepare the following solutions:
  - 25 mM  $\text{NH}_4\text{HCO}_3$  (100 mg/50 mL)
  - 25 mM  $\text{NH}_4\text{HCO}_3$  in 50% ACN
  - 50% ACN/ 50% formic acid
  - 12.5 mg/ $\mu\text{L}$  trypsin in 25 mM  $\text{NH}_4\text{HCO}_3$  (freshly diluted)
2. Dice each gel slice into small pieces and place into clean dry Eppendorf tube.
3. Add ~100  $\mu\text{L}$  (or enough to cover) of  $\text{NH}_4\text{HCO}_3$  in 50% ACN and vortex for a few minutes.
4. Using gel loading pipet tip, extract and discard the supernatant.
5. Repeat steps 3 and 4 once or twice more.

-For low level proteins (<1pmol) especially those separated by 1-D SDS-PAGE, reduction and

alkylation is recommended.

The following procedures are performed after step 5.

Prepare the following fresh solutions 5a-5f.

- a. 10 mM DTT (1.5 mg/mL) in 25 mM  $\text{NH}_4\text{HCO}_3$  (1.5mg/mL)  
55 mM iodoacetamide in 25 mM  $\text{NH}_4\text{HCO}_3$  (10 mg/ $\mu\text{L}$ )
  - b. Add 25  $\mu\text{L}$  (or enough to cover) 10 mM DTT in 25 mM  $\text{NH}_4\text{HCO}_3$  to dried gels. Vortex and spin briefly, allow 1 hour at 56 °C for reaction to proceed.
  - c. Remove supernatant and discard. Add 25  $\mu\text{L}$  55 mM iodoacetamide to the gel pieces. Vortex and spin briefly. Allow reaction to proceed in the dark at room temperature for 45 minutes.
  - d. Remove supernatant and discard. Wash gels with  $\sim 100$   $\mu\text{L}$  25 mM  $\text{NH}_4\text{HCO}_3$ . Vortex for 10 minutes and spin.  
Remove supernatant and discard. Dehydrate gels with  $\sim 100$   $\mu\text{L}$  (or enough to cover) of 25 mM  $\text{NH}_4\text{HCO}_3$  in 50 % ACN. Vortex for 5 minutes and spin. Repeat once.
  - e. Proceed with trypsin digest.
6. Add trypsin solution to just barely cover the gel pieces. Estimate the gel volume and add about 3 times the volume of trypsin solution. This volume will vary from sample to sample, but on average  $\sim 5$ -25  $\mu\text{L}$  is sufficient.
  7. Rehydrate the gel pieces at 4 °C for 10 minutes. Spin. Add 25 mM  $\text{NH}_4\text{HCO}_3$  as needed to cover the gel pieces
  8. Spin briefly and incubate at 37 °C (room temperature) for overnight.

## H. Extraction of Peptides

### \*\*KEEP ALL LIQUIDS\*\*

1. Transfer the digest solution (aqueous extraction) into a clean 0.65 mL siliconized tube.
2. To the gel pieces add 30  $\mu$ L (or enough to cover) of 50 % ACN/ 5 % formic acid.
3. Vortex for 20-30 minutes, spin, and sonicate. Repeat.
4. Vortex the extracted digests, spin, and speedvac to reduce the volume to around 10  $\mu$ L.
5. Either proceed with C<sub>18</sub> ZipTip (Millipore) cleanup or analyze using LC-MS.

## I. C<sub>18</sub> ZipTip

When analyzing low levels of protein, concentrate the peptides by eluting them from ZipTips. Use 3  $\mu$ L of elution solution in a clean 0.65 mL siliconized tube. Utilize 1  $\mu$ L of the unseparated digest for analysis by MALDI.

You have tubes filled with your samples of extracted peptides. In each tube, using the same C<sub>18</sub> ZipTip do the following:

1. Pull up 10  $\mu$ L of ACN into the C<sub>18</sub> ZipTip and discard. Repeat.
2. Pull up 10  $\mu$ L of 80 % ACN into the ZipTip and then discard. To dilute to 80 % use 800  $\mu$ L of ACN and 200  $\mu$ L of TFA to get the solution.
3. Pull up 10  $\mu$ L of 0.1 % TFA into the ZipTip and discard. Repeat.
4. Pull up 10  $\mu$ L of sample and discard into its same tube 10 times.
5. Pull up 10  $\mu$ L of 0.1 % TFA into the ZipTip and discard. Repeat twice.

6. Pull up 5  $\mu$ L of 80 % ACN into the ZipTip and discard into a clean siliconized tube. Keep for analysis by MALDI-TOF.

**J. CellLytic™ Nuclear NuCLEAR™ Extraction Kit from Sigma**

**Nuclear Protein Extraction from 100 mg of tissue**

**Calculate accordingly for different tissue weight**

1. Dilute the 1 M DTT solution to 0.1 M with deionized, sterile water.
2. Prepare 1x lysis buffer, hypotonic
  - a. To 1 mL of 1x lysis buffer  
Add: 10  $\mu$ l of the prepared 0.1 M DTT solution  
10  $\mu$ l of the protease cocktail inhibitor
3. Rinse tissue twice with PBS and discard PBS.
4. Resuspend tissue in 1 mL of the 1x lysis buffer (containing DTT and protease inhibitors).
5. Homogenize the tissue until more than 90 % of the cells are broken.
6. Centrifuge the disrupted cells for 20 minutes at 10,000-11,000 x g.
7. Transfer the supernatant to a fresh tube. (This portion is called the cytoplasmic fraction.)
8. Add 1.5 mL of the prepared 0.1 M DTT solution and 1.5 mL of the protease cocktail inhibitor to 147 mL of the extraction buffer.
9. Resuspend the crude nuclei pellet in ~140 mL extraction buffer containing DTT and protease inhibitor
10. Shake gently in refrigerator for 30 minutes.
11. Centrifuge for 5 minutes at 20,000-21,000 x g.

12. Transfer the supernatant to a clean, chilled tube. (This portion is called the nuclear fraction.)

**K. Making Lysis Buffer**

**To make 40 mL of Buffer:**

1. 20 mL H<sub>2</sub>O
2. 0.08 mL EDTA
3. 0.10 mL Tris
4. 4 mL Glycerol
5. 0.04 mL NP40
6. pH needs to be around 7.5
7. QS with H<sub>2</sub>O at 40 mL