

**SYNTHESIS AND CHARACTERIZATION OF INSULIN RECEPTOR PARTIAL
AGONISTS AS A ROUTE TO IMPROVED DIABETES THERAPY**

Sara J. Brandt

Submitted to the faculty of the University Graduate School
in partial fulfillment of the requirements for the degree

Doctor of Philosophy

in the department of Molecular and Cellular Biochemistry

Indiana University

April 2015

Accepted by the Graduate Faculty, Indiana University, in partial fulfillment of the
requirements for the degree of Doctor of Philosophy.

Doctoral Committee

Richard D. DiMarchi, Ph.D.

Charles E. Dann III, Ph.D.

Martha G. Oakley, Ph.D.

Michael S. VanNieuwenhze, Ph.D.

April, 2015

Copyright © 2015

Sara J. Brandt

Acknowledgements

I would like to thank Dr. Richard DiMarchi for the opportunity and privilege of working in this laboratory. In addition, I would like to thank every scientist, postdoctoral researcher, and graduate student in the lab, all of whom have contributed to my work. A special debt of gratitude is owed to Dr. Vasily Gelfanov, Mr. Dave Smiley, Mr. Jay Levy, and Dr. Pengyun Li, all of whom have devoted considerable time to training me as a scientist. I would also like to thank Dr. Jonathan A. Karty and Mrs. Angela M. Hansen from the IU mass spectrometry facility, as well as our collaborators at the University of Cincinnati, Dr. Nikki Parker and Dr. Diego Perez-Tilve.

I would like to thank my thesis committee, past and present, which includes Dr. Erin Carlson, Dr. Charles Dann III, Dr. Martha Oakley, and Dr. Michael S. VanNieuwenhze, for their guidance both in the classroom and in committee meetings.

I would very much like to thank all of my teachers who have served as mentors and inspiration. In particular, I would like to thank Dr. Alvin Crumbliss and Dr. Ken Lyle from Duke University, and Dr. Andrew Bednarik and Mr. Gregory Horton from Fairfield High School. Thank you for a lifetime of inspiration.

Most importantly, I would like to thank Mom, Dad, and Katherine for their love and support. Thank you.

Sara J. Brandt

**SYNTHESIS AND CHARACTERIZATION OF INSULIN RECEPTOR PARTIAL
AGONISTS AS A ROUTE TO IMPROVED DIABETES THERAPY**

Insulin-dependent diabetes requires the daily administration of insulin to regulate blood glucose. Unfortunately, insulin possesses a narrow therapeutic index which represents a risk for overdose and life-threatening hypoglycemia. This research investigates the synthesis and biological characterization of insulin-based analogs that activate the insulin receptor with high potency, but varying degrees of maximal activity. These analogs are dimeric peptides that consist of native insulin and a covalently bound insulin receptor antagonist. Structure-activity analysis revealed a key amino acid within the antagonist, and mutations at this single position control the maximal activity of the heterodimer. These analogs may represent a route to a safer insulin therapy, through selection of an optimized analog that has diminished activity relative to the native hormone.

Richard D. DiMarchi, Ph.D.

Charles E. Dann III, Ph.D.

Martha G. Oakley, Ph.D.

Michael S. VanNieuwenhze, Ph.D.

Table of Contents

Table of Contents	vi
List of Figures and Tables	viii
Chapter 1: Introduction	1
Diabetes Mellitus	2
Metabolism and Insulin Signaling	4
Type 1 Diabetes	15
Type 2 Diabetes	17
Consequences of the Disease	19
Prevalence of the Disease	22
Management of Diabetes Mellitus	24
Properties of Insulin	29
Receptor/Ligand Interactions	32
References	49
Chapter 2: Synthesis, Characterization, and Optimization of	
Insulin Receptor Antagonists	56
Abstract	57
Introduction	58
Results	65
Introduction of Additional Residues for Chemical Conjugation	72
PEGylation of Peptide #4-6	81
Discussion	83
Methods	86
Historical Perspective	86
Current Research Protocols	94
References	106

Chapter 3: Biosynthetic Heterodimers	108
Abstract	109
Introduction	110
Results	117
Heterodimers, with fully <i>lys-C</i> removable domains.....	117
Heterodimers, where <i>lys-C</i> cleavage results in antagonists positioned at the N-terminus of the A chain	120
Heterodimers, where <i>lys-C</i> cleavage results in antagonists positioned at the C-terminus of the B chain.....	124
Discussion	129
Methods	134
References	137
Chapter 4: Semisynthetic Heterodimers.....	139
Abstract	140
Introduction	141
Results	147
Discussion	156
Methods	163
References	169
Chapter 5: Future Directions.....	167
References	180
Table of Abbreviations	181
Curriculum Vitae	

List of Figures and Tables

Chapter 1: Introduction

Figure 1: Primary sequence of native, human insulin.....	6
Figure 2: Primary sequence of proinsulin	6
Figure 3: Sequence homology between human proinsulin and IGF-1	7
Figure 4: Cartoon illustrating the extended conformation of the insulin receptor	10
Figure 5: The insulin receptor signaling pathway	12
Figure 6: Number (in millions) of civilian, non-institutionalized adults with diagnosed diabetes	23
Figure 7: Stylized insulin dose-response.....	30
Figure 8: Typical sigmoidal dose response for native insulin.....	31
Figure 9: An analysis of the binding of truncated human IgG protein.....	35
Figure 10: An analysis of the binding of bradykinin	36
Figure 11: Scatchard plots for radiolabeled human growth hormone and porcine insulin	37
Figure 12: Scatchard plot and radiolabeled dissociation plot of insulin	38
Figure 13: Dose-response curve analyzing the displacement of radiolabeled insulin ..	38
Figure 14: Crystal structure of the insulin receptor ectodomain	40
Figure 15: A “top down” view of the insulin receptor ectodomains.....	40
Figure 16: Schematic illustrating a stylized binding curve of radiolabeled insulin	41
Figure 17: Phosphorylation assay demonstrating the activation of insulin receptor isoform B	42
Figure 18: Schematic of insulin homodimers	43
Figure 19: Stimulation of lipogenesis by insulin and insulin dimers	44
Figure 20: Binding assay and fat cell assay of monomeric and B ²⁹ -B ²⁹ dimerized insulin	44
Table 1: Binding motifs identified by Schaffer, et al.	46

Chapter 2: Synthesis, Characterization, and Optimization of Insulin Receptor Antagonists

Figure 1: Phosphorylation assay demonstrating the lack of agonism and antagonism of peptides #1-3	65
Figure 2: Phosphorylation assay demonstrating the lack of agonism and antagonism of peptides #4 and #5	65
Figure 3: Phosphorylation assay demonstrating the agonism of peptides #7 and #10	66
Figure 4: Phosphorylation assay demonstrating the antagonism of peptides #7 and #10	66
Figure 5: Phosphorylation assay demonstrating the agonism and antagonism of peptide #9 , in its oxidized form	67
Figure 6: Phosphorylation assay demonstrating the agonism and antagonism of peptide #9 , in its reduced form	68
Figure 7: Phosphorylation assay demonstrating the agonism and antagonism of peptides #9 (oxidized) , #9 (reduced) , and #11	68
Figure 8: Phosphorylation assay demonstrating the lack of agonism and antagonism of peptides #4-6 and #6-4	70
Figure 9: Phosphorylation assay demonstrating the ability of peptides #4-6 and #6-4 to antagonize 1 nM DP8 at the IRA isoform	71
Figure 10: Phosphorylation assay demonstrating the inability of peptide #12 to agonize or antagonize native insulin at both insulin receptor isoforms.....	71
Figure 11: Phosphorylation assay demonstrating the agonism and antagonism of peptides #4 and #6 , with cysteine placement variations.....	73
Figure 12: Phosphorylation assay demonstrating the lack of agonism and the antagonism of peptide #4-6 cysteine variants	74
Figure 13: Phosphorylation assay demonstrating the lack of agonism and the antagonism of peptide #6-4 cysteine variants	75
Figure 14: Phosphorylation assay demonstrating the antagonism of peptides #4-6-Lys and #6-4-Lys	77

Chapter 2: Synthesis, Characterization, and Optimization of Insulin Receptor Antagonists (cont.)

Figure 15: Phosphorylation assay demonstrating the agonism and antagonism of N-terminal shortened antagonists	78
Figure 16: Phosphorylation assay demonstrating the antagonism of #4(AX)-6-Lys peptides, where X is the position of the alanine mutation	79
Figure 17: Phosphorylation assay demonstrating the antagonism of C-terminal shortened antagonists	80
Figure 18: Phosphorylation assay demonstrating the antagonism of peptide #4-6 and PEGylated peptide #4-6	81
Figure 19: Benzyloxycarbonyl (Cbz) reversible protecting group.....	88
Figure 20: Acid promoted deprotection of tert-butyloxycarbonyl protected amino acids	88
Figure 21: Deprotection of an Fmoc-protected amino acid	89
Figure 22: Carbodiimide activation of a deprotonated amino acid.....	89
Figure 23: A deprotonated amino acid reacts with a carbodiimide coupling reagent....	90
Figure 24: Reaction of a triazole base coupling reagent (HOBt) with a carbodiimide activated amino acid	90
Figure 25: Reaction of DEPBT with deprotonated amino acid, resulting in a reactive ester	91
Figure 26: Base promoted racemization due to intramolecular cyclization.....	92
Figure 27: Chloromethylpolystyrene, the Merrifield resin	92
Figure 28: <i>tert</i> -butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl (PAM) resin	92
Figure 29: BHA and MBHA resins	93
Figure 30: Wang resin	93
Figure 31: PAL and Rink amide resin	93
Figure 32: Analytical HPLC and MALDI of crude peptide #Cys-4.....	99
Figure 33: Purification of peptide #Cys-4.....	93

Chapter 2: Synthesis, Characterization, and Optimization of Insulin Receptor Antagonists (cont.)

Figure 34: Analytical HPLC and MALDI of purified peptide #Cys-4	100
Figure 35: LC/MS analysis of crude peptide #6	102
Figure 36: Reverse phase HPLC purification of peptide #6	102
Figure 37: LC/MS analysis of pure peptide #6	103
Table 1: Binding motifs identified by Schaffer, et al, where X represents an unspecified amino acid.....	59
Table 2: Sequence of peptides.....	63
Table 3: IC ₅₀ values of peptides #9, 11	68
Table 4: IC ₅₀ values of peptides #4-6 and #6-4.....	70
Table 5: IC ₅₀ values of the #4-6-Lys N-terminal alanine scan	79
Table 6: IC ₅₀ values of C-terminal shortened #4-6-Lys antagonists.....	80
Table 7: IC ₅₀ values of peptide #4-6 and PEGylated peptide #4-6	81
Table 8: Analytical HPLC traces demonstrating the stability of the PEGylated peptide #4-6.....	82
Table 9: Protecting groups for reactive side chains in Boc and Fmoc chemistries.....	98
Table 10: Description of modes of Fmoc synthesis	101
Table 11: Fmoc synthesis of a peptide of <i>n</i> amino acids in length using previously described modes	101

Chapter 3: Biosynthetic Heterodimers

Figure 1: Phosphorylation assay demonstrating the relative potency and activity of native human insulin, native human IGF-1, and DP8	112
Figure 2: Phosphorylation assay demonstrating the agonism of DP38	118
Figure 3: Phosphorylation assay demonstrating the agonism of DP72	118
Figure 4: Phosphorylation assay demonstrating the agonism of DP38”	119
Figure 5: Phosphorylation assay demonstrating the agonism of DP72”	119
Figure 6: Schematic representing the transition from a single chain molecule to a two chain heterodimer via treatment with <i>lys-C</i>	121

Chapter 3: Biosynthetic Heterodimers (cont.)

Figure 7: Phosphorylation assay demonstrating the agonism of the single chain DP66 and the <i>lys-C</i> treated, two chain form of DP66	121
Figure 8: Phosphorylation assay demonstrating the agonism of the single chain DP68 and the <i>lys-C</i> treated, two chain form of DP68	121
Figure 9: Phosphorylation assay demonstrating the agonism of the single chain DP37 and the <i>lys-C</i> treated, two chain form of DP37	123
Figure 10: Phosphorylation assay demonstrating the agonism of the single chain DP51 and the <i>lys-C</i> treated, two chain form of DP51 ...	123
Figure 11: Phosphorylation assay demonstrating the agonism of the single chain DP70 and the <i>lys-C</i> treated, two chain form of DP70 ...	123
Figure 12: Schematic representing the transition from a single chain molecule to a two chain heterodimer, via treatment with <i>lys-C</i>	125
Figure 13: Phosphorylation assay demonstrating the agonism of the single chain DP67 and the <i>lys-C</i> treated, two chain form of DP67 ...	125
Figure 14: Phosphorylation assay demonstrating the agonism of the single chain DP69 and the <i>lys-C</i> treated, two chain form of DP69 ...	125
Figure 15: Phosphorylation assay demonstrating the agonism of the single chain DP36 and the <i>lys-C</i> treated, two chain form of DP36 ...	127
Figure 16: Phosphorylation assay demonstrating the antagonism of the two chain form of DP36 by titrating <i>lys-C</i> treated DP36 against constant 1 nM insulin	127
Figure 17: Phosphorylation assay demonstrating the agonism of the single chain DP50 and the <i>lys-C</i> treated, two chain form of DP50 ...	127
Figure 18: Phosphorylation assay demonstrating the agonism of the <i>lys-C</i> treated, two chain forms of DP36 and DP50.....	128
Figure 19: Phosphorylation assay demonstrating the agonism of the single chain DP71 and the <i>lys-C</i> treated, two chain form of DP71 ...	128
Figure 20: Anion exchange purification of peptide DP50	136

Chapter 3: Biosynthetic Heterodimers (cont.)

Table 1: Insulin, IGF-1 and DP8 sequence comparison.....	111
Table 2: Heterodimers with fully <i>lys-C</i> removable C domains	115
Table 3: Heterodimers, where <i>lys-C</i> cleavage results in antagonists positioned at the N-terminus of the A chain	116
Table 4: Heterodimers, where <i>lys-C</i> cleavage results in antagonists positioned at the C-terminus of the B chain	116

Chapter 4: Semisynthetic Heterodimers

Figure 1: Native chemical ligation	142
Figure 2: Schematic of expressed protein ligation	143
Figure 3: Click chemistry	143
Figure 4: The Staudinger ligation.....	144
Figure 5: The traceless Staudinger ligation	144
Figure 6: Reaction scheme illustrating the chemical modification of the insulin B ²⁹ residue to generate an analog with a trityl-protected thiol group.....	145
Figure 7: Phosphorylation assay demonstrating the agonism of the B ²⁹ modified insulin	148
Figure 8: Phosphorylation assay demonstrating the activity of #Insulin-Cys-4 and #4-Cys-Insulin	148
Figure 9: Phosphorylation assay demonstrating the agonism of #Insulin-Cys-6	150
Figure 10: Phosphorylation assay demonstrating the agonism and antagonism of #6-Cys-Insulin	150
Figure 11: Phosphorylation assay demonstrating the agonism of #6-Cys-Insulin and #6(des1-5)-Cys-Insulin	150
Figure 12: Phosphorylation assay demonstrating the agonism of the #6-Cys-Insulin alanine scans	151
Figure 13: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(I2)-Cys-Insulin	152

Chapter 4: Semisynthetic Heterodimers (cont.)

Figure 14: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(dL2)-Cys-Insulin	152
Figure 15: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(V2)-Cys-Insulin	153
Figure 16: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(F2)-Cys-Insulin	154
Figure 17: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(W2)-Cys-Insulin	154
Figure 18: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(Y2)-Cys-Insulin	155
Figure 19: Phosphorylation assay demonstrating the decrease in potency but high maximal activity of #6(Q2)-Cys-Insulin	155
Figure 20: Comparative EC ₅₀ values of #6(X2)-Cys-Insulin heterodimers, where X represents the amino acid at position two	158
Figure 21: Comparative maximal activities of #6(X2)-Cys-Insulin heterodimers, where X represents the amino acid at position two	158
Figure 22: Reaction of S-trityl-β-mercaptopropionic acid with diisopropylcarbodiimide and N-hydroxysuccinimide	164
Figure 23: Reaction scheme illustrating the modification of the B ²⁹ residue of insulin	164
Figure 24: MALDI analysis demonstrating the specificity of the B ²⁹ modification	164
Figure 25: Reaction scheme demonstrating the activation of the B ²⁹ modified insulin.....	165
Figure 26: LC/MS analysis of crude peptide #6-Cys	166
Figure 27: LC/MS analysis depicting the same crude peptide #6-Cys after stirring in 2% acetic acid for 12 hours	167

Chapter 5: Future Directions

Figure 1: Blood glucose concentrations, in normal, lean mice in response to 1:1, 1:10, and 1:100 ratios of human insulin: #4-6 antagonist.....	173
Figure 2: Blood glucose concentrations, in normal, lean mice in response to 3, 10, 30, and 300 nmol/kg doses of the #4-6 antagonist	173
Figure 3: Blood glucose concentrations, in normal, lean mice, measured in response to the #4-6 and PEGylated #4-6 antagonists, over the course of seven days	175
Figure 4: Body weight in grams in normal, lean mice, measured in response to the #4-6 and PEGylated #4-6 antagonists, over the course of seven days	175
Figure 5: Blood glucose concentrations of STZ diabetic mice in response to human insulin, #6(L2)-Cys-Insulin and #6(A2)-Cys-Insulin	177
Figure 6: Penicillamine	178
Figure 7: Phosphorylation assay demonstrating the agonism and antagonism of #6(L2)-Pen-Insulin	178
Figure 8: Blood glucose concentrations of STZ diabetic mice in response to human insulin and #6(L2)-Pen-Insulin	179

Chapter 1
Introduction

Diabetes Mellitus

Diabetes mellitus was identified as early as 1550 BC, when ancient Egyptians described the symptoms and recommended a treatment of boiled wheat and lead[1]. Diabetes is also referenced in 5th century Indian medical texts, 9th century Arabic texts, and 16th century European writings, including those of the famous Swiss physician Paracelsus[1, 2].

The classic symptoms of diabetes are extreme thirst, polyuria, weight loss, vision deterioration, and, in severe forms, ketoacidosis[3], the accumulation of keto acids in the blood. Left untreated, diabetes is fatal.

The cause of diabetes was unknown until 1889, when Oskar Minkowski and Josef Von Mering removed the pancreas from a dog and discovered that the dog subsequently displayed the typical symptoms of diabetes[1]. Nearly 3000 years after the first recorded reference to diabetes, the pancreas was established as the origin of the disease. As research progressed, pancreatic extracts from healthy animals were used to lower blood glucose in diabetic animals, and it was believed that the decrease in blood glucose was due to a hypothetical hormone, named “insuline[4].”

The existence of insulin was confirmed in 1921, when insulin was discovered and purified[4]. In January of 1922, one of the first children treated with insulin was 6 year old Theodore Ryder. Until this moment in history, a diagnosis of diabetes was a death sentence. For Ryder, insulin was a lifesaving reprieve, and he lived until 1993[5].

Today, we have a much better understanding of the etiology and pathophysiology of diabetes. Diabetes is caused by an inability to properly regulate blood glucose due to a defect in insulin secretion or insulin resistance, resulting in chronic hyperglycemia. In the United States, the National Institutes of Health (NIH) has three diagnostic criteria for diabetes mellitus[6]:

1. Blood glucose concentration of 126 mg/dL after an 8 hour fast
2. Blood glucose concentration of 200 mg/dL or higher two hours after a 75 g oral glucose challenge
3. A random blood glucose concentration of 200 mg/dL or higher

All of these symptoms come from either the inability of the pancreas to produce insulin, or a systemic resistance to insulin action, usually caused by chronic hyperglycemia. In order to introduce the pathophysiology of diabetes, this chapter first describes normal metabolism and insulin signaling, followed by a description of the disease.

Metabolism and Insulin Signaling

Metabolic control of blood glucose is vital for survival. Glucose is used by skeletal muscle as an energy source, and is stored as glycogen for further use[1]. Adipose tissue stores excess glucose as fatty acids and triglycerides[1]. The human brain is especially dependent on glucose as its primary source of fuel[7, 8]. Although the brain represents only 2% of human body mass, the brain utilizes approximately 25% of all metabolic glucose[8]. In times of starvation or extremely low blood glucose the brain can derive energy from ketone bodies[8].

The pancreas regulates blood glucose mainly through the secretion of two hormones, insulin to lower blood glucose, and glucagon to raise blood glucose. These two hormones are produced in the pancreas in substructures referred to as “Islets of Langerhans.”

Within the islets, insulin is produced in the β -cells and glucagon is produced in the α -cells[9]. Insulin is released from the pancreas in response to a rise in blood glucose concentrations, such as those experienced after a meal.

Ingested food is broken down into its most basic components, including glucose. Glucose then passes through the gut and is absorbed into the blood stream, resulting in an increase in blood glucose concentration. As the blood glucose concentration rises, glucose diffuses into the pancreatic β -cells through the glucose receptor GLUT2 [10]. The increase in intracellular glucose stimulates an increase in cellular metabolism, resulting in an increase in ADP conversion to ATP[10]. The increase in ATP closes ATP-sensitive potassium channels in the cell membrane, which depolarizes the membrane[10]. In

response to this depolarization, calcium-channels open, flooding the cell with calcium, which in turn stimulates insulin release[10].

Insulin released from the pancreas travels through a hepatic duct to the liver, which means that the liver receives higher doses of insulin than any other organ. In the liver, insulin promotes glycolysis and glycogen synthesis, and inhibits glycogenolysis, gluconeogenesis, and ketogenesis[9]. In skeletal muscle, insulin signals for increased glucose uptake, increased glucose oxidation, and increased glycogen synthesis[9]. In adipose tissue, insulin signals to increase glucose uptake, decrease lipolysis and increase lipogenesis[9]. In more simplistic terms, insulin signals to store blood glucose, and not to replenish it.

Insulin is a two-chain peptide hormone. It consists of 51 amino acid residues and three disulfides (Figure 1). It is synthesized as a single chain precursor called proinsulin, which contains a “C-peptide” between the A and B chains (Figure 2). Proinsulin is metabolically inactive, although it can bind to the insulin receptor[11]. The mature hormone is created when proinsulin is cleaved post-translationally by prohormone convertase 1 (PC-1), prohormone convertase 2 (PC-2), and carboxypeptidase E (CPE)[12]. First, PC-1 cleaves the junction between the B and C chains, leaving two basic residues at the C terminus of the B chain[13]. These residues are removed by CPE[13]. Finally, PC-2 cleaves the junction between the A and C chains, yielding the mature hormone, insulin[13]. The C-peptide is released along with insulin, and is believed to have some role in glucose metabolism, although it has been shown that it does

not bind to the insulin receptor[14]. Incomplete cleavage of proinsulin can lead to heritable metabolic disorders, such as familial hyperproinsulinemia[12, 15], which can result in mild diabetes[12]. Mature insulin is stored in the pancreas as zinc-stabilized hexamers[16]. The propensity of insulin to self-aggregate has pharmacological affects, which will be discussed later.

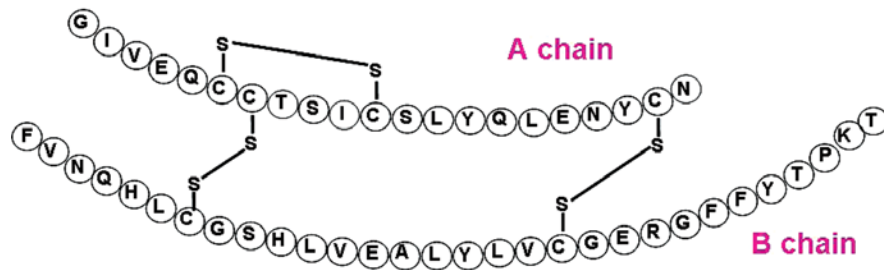


Figure 1: Sequence of native, human insulin, shown with disulfides. The C-peptide, which is cleaved post-translationally, is not shown [17].

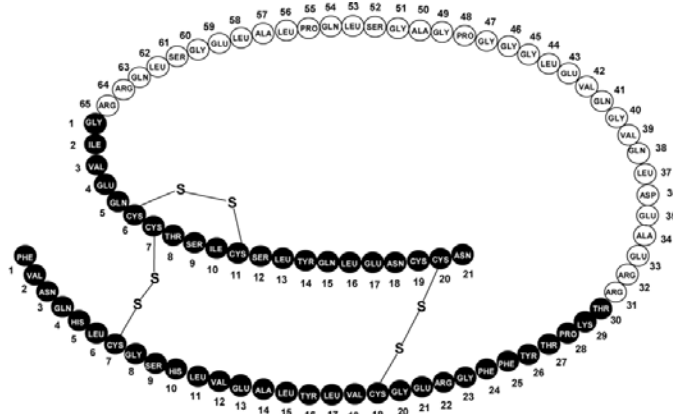


Figure 2: Sequence of proinsulin. A and B chains in black, C-peptide in white [17].

Insulin shares significant sequence homology with the hormones insulin-like growth factor 1 and 2 (IGF-1 and 2) in their A and B chains (Figure 3). However, IGF-1 has a dramatically different C-domain, which is not cleaved from the molecule. Both insulin and IGF-1 can bind to either set of receptors[18], resulting in signaling overlap. Insulin analogs with elevated binding to the IGF receptors have increased mitogenic potential in vitro[18].

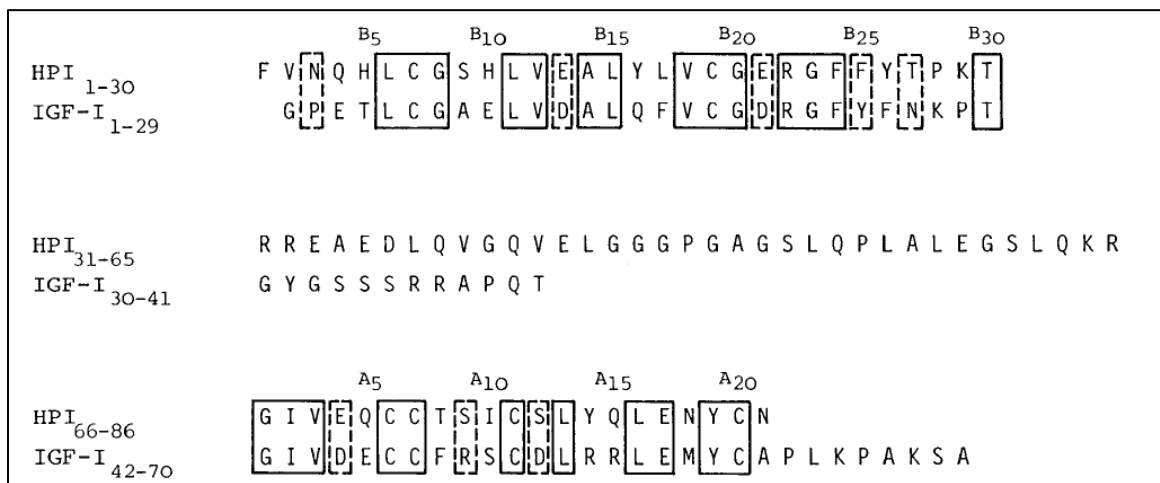


Figure 3: Sequence homology between human proinsulin (HPI) and IGF-1[19]. Mature IGF-1 retains its C-peptide (middle), insulin does not.

Native human insulin has two binding sites for the insulin receptor, referred to as site 1 and site 2. Site 1 is composed of residues GlyA¹, IleA², ValA³, GlnA⁵, ThrA⁸, TyrA¹⁹, AsnA²¹, ValB¹², TyrB¹⁶, GlyB²³, PheB²⁴, PheB²⁵, and TyrB²⁶, and represents the dimer forming face of insulin[16]. Binding site 2 is composed of residues SerA¹², LeuA¹³, GluA¹⁷, HisB¹⁰, GluB¹³, and LeuB¹⁷, and represents the hexamer forming face[16]. The secondary structure of native insulin consists of two α -helices in the A chain, which extend from residues A¹-A⁸ and A¹²-A¹⁸[20]. In the B chain, insulin has an α -helix from

residues B⁹-B¹⁹[20], as well as a β -sheet in residues B²⁴-B²⁸[21]. The B chain can exist in two physical states – tense (T), and relaxed (R). In the relaxed state the α -helix in the B chain extends to include residues B¹-B¹⁹[21]. Insulin undergoes this T-to-R transition when binding to the insulin receptor. In addition, flexibility at the C-terminus of the B chain is required, since binding to the insulin receptors causes insulin B chain residues B²⁶-B³⁰ to be displaced away from the hydrophobic core of insulin[20]. Altering the sequence of the C-terminus of the B chain has resulted in several pharmacologically relevant insulin analogs, such as “fast-acting” insulin lispro, which is the result of switching the order of the C-terminal proline and lysine residues of the B chain, disrupting the ability of the molecule to self-associate[22]. Deletion of the last five amino acids of the insulin B chain (B²⁶-B³⁰) has no effect on biological activity[23], however, deletion of the last eight amino acids of the insulin B-chain results in a peptide referred to as “desoctapeptide insulin,” which retains less than 1% of native insulin biological activity[24].

Insulin binds to its receptor, the insulin receptor (IR). The insulin receptor is part of a family of receptor tyrosine kinases (RTK), which includes the IR, the type-1 insulin-like growth factor receptor (IGF-1R) and the insulin receptor related receptor (IRR)[25]. Most RTKs function by dimerizing in the presence of their ligand, bringing the two tyrosine kinase domains in close proximity in order to participate in trans-phosphorylation, which activates the receptor and induces downstream signaling. The insulin receptor, however, is constitutively dimerized and held together by disulfide bonds in the presence or in the absence of a ligand.

The insulin receptor, from N- to C-termini, consists of several domains (Figure 4). The first is a leucine rich domain (L1), followed by a cysteine rich domain (CR), and then a second leucine rich domain (L2). The L2 domain is followed by three fibronectin-type III (Fn) domains. In the literature, the Fn domains are referred to as either FnIII-0, FnIII-1, and FnIII-2, or alternatively FnIII-1, FnIII-2, FnIII-3. The former will be used in this paper. The FnIII-1 domain contains an insert domain (ID) that is cleaved post-translationally[26], creating the alpha and beta subunits of the insulin receptor. The beta subunit is composed of the remainder of the FnIII-1 domain, followed by the FnIII-2 domain, transmembrane domain (TM), juxtamembrane domain (JM), tyrosine kinase domain (TK), and the C-terminus (CT). Alpha and beta subunits are held together by a disulfide bond between the FnIII-1 domain on the alpha subunit and the FnIII-2 domain on the beta subunit.

There are two splice variants of this receptor, which create the insulin receptor A and B isoforms. The insulin receptor A isoform lacks exon 11, a twelve residue sequence at the C-terminus of the alpha domain[26]. The physiological significance of the two different isoforms is under debate, but several key observations have hinted at specific and defined roles for each isoform. The two isoforms are not uniformly or evenly expressed in all tissues. For example, insulin receptors in liver cells consist of 75% IRB and 25% IRA receptors, whereas in adipose tissues the expression is 60% IRB and 40% IRA[27]. IRA is also implicated in mitogenic signaling and cancer[28, 29], potentially due to its affinity for non-native ligands, such as the growth factor IGF-1. Compared to IRB, IRA has a ten-

fold increase in affinity for IGF-1 and a five-fold increase in affinity for IGF-2[29, 30]. IRA can also heterodimerize with the growth factor receptor IGF1-R, and these hybrid receptors activate mitogenic pathways[31]. Therefore, any drug that aims to mimic the metabolic action of insulin should exhibit preferential binding affinity and signaling at IRB.

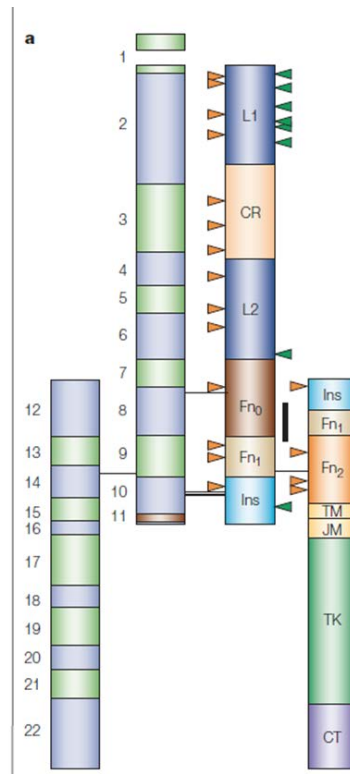


Figure 4: Cartoon illustrating the extended conformation of the insulin receptor, with the exons numbered on the left, and the domains labeled on the right[32]. Orange arrows represent N-glycosylation sites. Green arrowheads represent ligand binding sites. The black bar represents the major immunogenic region[32].

The mechanism by which insulin binding activates the receptor is unclear. Unlike most RTKs, the insulin receptor is constitutively dimerized and therefore dimerization of the receptor is not sufficient to induce signaling. In the absence of ligand, the tyrosine kinase auto-phosphorylation of the insulin receptor is likely inhibited by one or more of the following mechanisms: juxtamembrane inhibition, activation-loop auto-inhibition, and C-tail inhibition[16]. Juxtamembrane inhibition refers to the juxtamembrane domain (JM), which lies adjacent to the N-terminus of the tyrosine kinase domain. Evidence suggests this JM domain adopts a conformation that, in the absence of ligand, precludes tyrosine kinase auto-phosphorylation by sterically hindering both lobes of the kinase domain[16]. In particular, the Y984 residue in the JM domain is vital for stabilizing this conformation, and mutation of this residue leads to an increase in basal auto-phosphorylation[16]. Activation loop auto-inhibition and the C-terminus of the alpha subunit may also play a role in preventing signaling in the absence of ligand[16].

Overall, the suggested mechanism of action is that insulin binding induces enough of a conformational change to bring the two intracellular kinase domains into the correct proximity to induce trans-phosphorylation in the activation loop of the kinase domain[26]. This activation induces the activation loop to move away from the kinase catalytic site, allowing insulin receptor substrates to bind[26]. Cis-phosphorylation in the juxtamembrane domain has also been observed in response to insulin, as well as the previously described trans-phosphorylation in the tyrosine kinase domain[26]. In total, seven tyrosine residues experience phosphorylation, although not all seven tyrosine residues are required to propagate the signal. Therefore, variations in phosphorylation of

the receptor may allow for attenuation of the downstream signaling cascade. In support of this hypothesis, mutating individual tyrosines has differential effects on the metabolic and mitogenic insulin receptor signals[26]. Mutating the first tyrosine in the kinase domain leads to an increase in glycogen synthesis, but a decrease in DNA synthesis[26]. In contrast, mutating the second and third tyrosines in the kinase domain leads to a decrease in both glycogen synthesis and DNA synthesis[26].

What is clear is that the binding of the ligand induces some type of conformational change, which activates the receptor. Once phosphorylated, the insulin receptor induces a diverse signaling cascade (Figure 5).

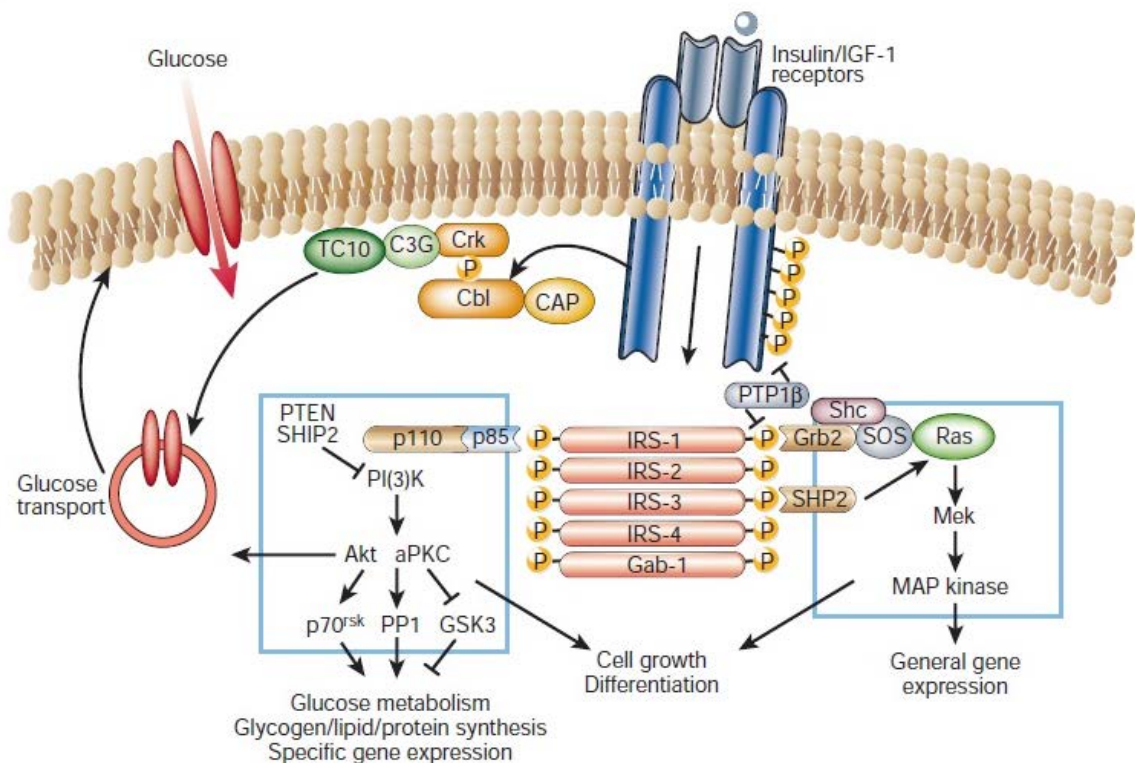


Figure 5: The insulin receptor signaling pathway[33].

Once the IR is phosphorylated, insulin receptor substrate (IRS) proteins interact with the activated receptor via their phosphotyrosine binding domains, and are phosphorylated on up to 20 tyrosine residues by the receptor[34]. The IRS proteins also contain pleckstrin homology (PH) domains that increase their affinity for the IR[34]. There are many different isoforms of the IRS proteins, referred to as IRS1-6. IRS 1 and 2 are the most abundant, but isoforms vary from tissue to tissue[34]. IRS 3 is found exclusively in fat adipocytes and neuronal tissue, and IRS 5 and 6 have very limited expression[34]. Knockout studies in mice have shown that the IRS proteins are not redundant, and that knocking out specific isoforms results in varied physiological consequences[33].

In general, the IRS proteins bind to adapter molecules that contain Src-homology 2 (SH2) domains[34], such as the adapter molecule Grb2[34]. Once phosphorylated by IRS, Grb2 complexes with Sos[34]. This complex activates Ras, which activates Raf, Mek, and the MAPK pathway[34]. The MAPK pathway ultimately leads to control of gene expression. The insulin receptor also activates the MAPK pathway by phosphorylating Shc, which activates Ras and again leads to MAPK activation. All of these signaling cascades are halted by the action of protein tyrosine phosphatase 1B (PTP1B), which dephosphorylates the insulin receptor[34].

The IRS proteins also activate phosphoinositide 3-kinase (PI3K). Activated PI3K generates PIP_3 , which in turn activates both the atypical PKCs and Akt[34]. Akt has several key roles. Akt induces the translocation of the glucose transport GLUT4 from internally sequestered vesicles to the cell surface to allow for glucose uptake[34]. Akt

also phosphorylates glycogen synthase kinase 3 (GSK-3)[34]. GSK-3, when unphosphorylated, is a repressor of glycogen synthesis[34]. When Akt phosphorylates GSK-3, it no longer suppresses glycogen synthesis and therefore glycogen synthesis proceeds[34]. Akt also inhibits FOXO1, which prevents FOXO1 from inducing gluconeogenesis[34], and activates mTOR, which leads to protein synthesis[34]. In this manner, PI3K is responsible for the majority of the metabolic signaling of the insulin receptor. This pathway can be inactivated by the proteins PTEN and SHIP2, which dephosphorylate and inactivate PIP₃[34], and by the proteins PP2A and Tribbles-3, which dephosphorylate and inactivate Akt[34].

There are also several negative feedback mechanisms that regulate insulin signaling through these pathways. The IRS proteins can be serine phosphorylated, which in turn down-regulates signaling[34]. The c-Jun-N-terminal (JNK) kinase pathway is activated by insulin as well and also phosphorylates serine residues on IRS proteins, indicating a potential negative feedback mechanism[34]. Increased serine phosphorylation of the IRS proteins is also implicated in disease states of insulin resistance[34]. SHP2 also binds to the insulin receptor[34], and activates the Ras/MAPK pathway[33], but dephosphorylates some of the residues on the IRS proteins that mediate binding with Grb2 and PI3K[34].

The insulin receptor also directly activates the membrane bound protein Cbl, which forms a complex with the adapter protein CAP[33]. CAP binds to the proline-rich regions of Cbl through its SH3 domains[33]. The Cbl:CAP complex then associates with lipid rafts in the cell membrane and activates the adapter protein CrkII[33]. CrkII activates C3G and

recruits it to the lipid rafts[33]. Once in the lipid rafts, C3G is able to activate TC10, a GTPase[33]. TC10 has been implicated in inducing the transportation of the glucose receptor GLUT4 to the outer membrane[34].

Overall, these pathways represent the main metabolic and mitogenic signaling pathways of the insulin receptor. These signaling pathways are vital to human metabolism, and the absence of or resistance to insulin results in either type 1 or type 2 diabetes.

Type 1 Diabetes

Type 1 diabetes is also called insulin-dependent diabetes, since it must be treated with exogenously administered insulin. Type 1 diabetes is caused by autoimmune T-cell degradation of β -cells in the pancreas[35], resulting in the partial or complete absence of insulin secretion[36]. Type 1 diabetes usually presents itself in childhood, but the onset of the disease can occur at any age[36]. It is believed that both environmental as well as genetic factors play a role in the progression of the disease[37]. There is a substantial genetic component to type 1 diabetes, as shown in high concordance rates in families. It has been found that if one monozygotic (identical) twin develops diabetes, there is a 30-50% chance their identical twin will develop the disease[1], and the risk increases as they age[38].

One purported genetic cause of the destruction of β -cells is variation in the human leukocyte antigen (HLA) region of chromosome 6[1, 35]. This region is responsible for producing major histocompatibility complex (MHC) proteins, which are part of the

immune system. MHC proteins are displayed on the cell surface and are responsible for presenting foreign particles to the immune system and activating an immune response[39]. One theory is that the alleles in the HLA region associated with diabetes present self-antigens to the immune system, triggering β -cell destruction[1].

Another purported genetic cause of type 1 diabetes is a single nucleotide polymorphism that affects the lymphoid-specific phosphatase, a protein expressed in human lymphocytes that mediates T-cell antigen receptor signaling pathways[40]. In addition, single nucleotide polymorphisms in the IL15RA region of the genome, which encodes for proteins that suppress cell death, have been implicated in type 1 diabetes[41]. These potential genetic causes of diabetes are indicative of the varied and multifaceted nature of the disease.

However, it is unlikely that genetic inheritance is the sole cause of diabetes. The majority of patients who have the HLA variations associated with diabetes do not develop diabetes[37]. Also, as referenced in the twin studies, identical twins have, at maximum, a 50% chance of concordant diabetes – therefore, genetics alone is not sufficient to explain the onset of disease. One theory is that a viral infection results in viral particles similar to β -cell constituents being displayed by the MHC proteins, triggering the immune system to attack β -cells[39]. It is interesting to note that there is a 10% concordance rate for diabetes between non-identical twins, indicating that they may be exposed to the same environmental stimuli[42].

Type 2 Diabetes

Type 2 diabetes is the result of genetic predisposition, usually coupled with chronic, severe hyperglycemia. Type 2 diabetes results in an intolerance or relative deficiency of insulin[36], resulting in abnormally high blood glucose. Type 2 diabetes is the most prevalent type of diabetes, accounting for over 90% of cases globally[43]. Like type 1 diabetes, there are genetic predispositions to type 2 diabetes. In one study, monozygotic twins had a 91% concordance rate, many with abnormalities in insulin secretion[1].

Non-genetic components also heavily influence type-2 diabetes. Obesity is a major risk factor for type 2 diabetes[44], and most patients with type 2 diabetes do suffer from obesity, or an increased percentage of abdominal fat[36], with one study demonstrating that obesity preceded diabetes occurrence 77% of the time[45]. Obesity is often accompanied by increased circulating leptin and cytokine levels, which have been shown to negatively affect pancreatic β -cells[44]. Adipose (fat) tissue itself is a source of inflammatory factors and free fatty acids, which can bind to the Toll-like receptor (TLR), resulting in the activation of the inflammatory transcription factor NF- κ B[46]. NF- κ B can also be stimulated by interleukin-1 β (IL-1 β), which is secreted by pancreatic β -cells as a response to chronically elevated blood glucose[44]. Low levels of IL-1 β and NF- κ B are necessary for proper β -cell function, but overstimulation by elevated blood glucose causes IL-1 β and NF- κ B to trigger β -cell apoptosis[47]. As β -cells are destroyed, the symptoms of diabetes begin to appear.

In addition to disregulating cytokines, free fatty acids (from adipose tissue) can also result in the down-regulation of PI3K and Akt, proteins involved in the insulin signaling pathway[46]. This down regulation results in desensitization to normal insulin signaling[46].

The connection between obesity and diabetes is especially troubling, since there has been a 74% increase in obese adults in the United States from 1991 to 2001[48]. Other lifestyle factors such as smoking, alcohol use, lack of physical exercise, and lack of vitamin D have all been implicated as risk factors for diabetes[45]. In isolation, smoking has been independently correlated with diabetes, however, these risk factors are more dangerous when in combination, and the co-occurrence of all of the risk factors greatly elevates the risk for diabetes[45].

In addition, ethnic backgrounds and cultural differences have been implicated as diabetes risk factors. In the United States, compared to people of Caucasian descent, people of African descent are twice as likely to develop type 2 diabetes, people of Hispanic descent are 2.5 times as likely, and people of Native American descent are five times more likely[49]. This disparity remains even when adjusting for body-mass index, physical activity, and diet[49]. Concordant with ethnic and cultural differences, low socioeconomic status has also been shown to be a risk factor for diabetes[49].

Diabetes is often described as a part of the metabolic syndrome. The term “metabolic syndrome” describes a number of frequently concordant metabolic complications, such as

obesity, insulin resistance, dyslipidaemia, and hypertension[50]. The combined effect of these disorders puts patients at increased risk for cardiovascular disease and type 2 diabetes[50].

Type 2 diabetes was previously referred to as “adult onset diabetes,” indicative of the fact that this was a progressively acquired disease that manifested later in life. However, due to an increasingly obese society, type 2 diabetes is increasingly found in young children. This is especially alarming since many type 2 diabetes therapies have not been tested or approved for children[43].

Consequences of the Disease

Diabetes leads to many short- and long-term health complications, caused by both hypoglycemia and hyperglycemia. Patients with diabetes often spend significant amounts of time at both hyper- and hypoglycemic levels of blood glucose. Insulin is administered to decrease blood glucose and reach euglycemia. However, insulin therapy comes with the constant threat of insulin overdose and subsequent hypoglycemia. Hypoglycemic complications can be generally grouped into neurogenic symptoms and neuroglycopenic symptoms[51].

Neurogenic symptoms are the psychological results of hypoglycemia, and can include anxiety, palpitations, tremors, sweating, and paresthesia, a “pins and needles” sensation in the skin[51].

Neuroglycopenic symptoms are the result of low blood glucose in the brain[51]. Examples of neuroglycopenic symptoms are cognitive impairment, behavioral changes, confusion, and, during severely low blood glucose, seizure and death[51]. A patient experiencing a severe hypoglycemic episode may need assistance from another person to return to euglycemic levels. Of patients that manage their diabetes with insulin, over 30% will experience a hypoglycemia-induced coma at least once, and 10% will suffer a diabetic coma in any given year[1]. Insulin-induced hypoglycemia results in the death of approximately 3-4% of diabetic patients[1]. Due to the immediate and deadly danger of insulin overdose, many patients with diabetes err on the side of caution when administering insulin, and as a consequence suffer from chronic hyperglycemia.

However, chronic hyperglycemia has its own long term complications. Chronic hyperglycemia over-stimulates the PKC metabolic pathway, which ultimately results in the production of reactive oxygen species (ROS)[46]. An overabundance of ROS results in an imbalance in nitric oxide bioavailability by converting NO into the highly reactive peroxynitrite ion, which is a powerful oxidant[46]. This anion has the ability to cross phospholipid membranes and results in nonspecific protein nitrosylation[46]. This can block the ability of cellular enzymes, such as NO synthase[46]. In addition, ROS can trigger vascular inflammation and the up-regulation of inflammatory factors[46]. Chronic hyperglycemia also leads to unwarranted glycosylation of enzymes and proteins, which can result in metabolic dysfunction[46]. The totality of these metabolic imbalances results in long term complications that can be broadly described as either macrovascular or microvascular. Macrovascular complications include cardiovascular complications

such as angina pectoris (chest pains), myocardial infarction (heart attack), strokes[52], and renal failure as a result of decreased angiogenesis[53]. In total, the risk of cardiovascular disease in patients with type 2 diabetes is two to five times as great as their non-diabetic peers[54]. Microvascular complications include neuropathy, poor circulation in the extremities, and retinal disease due to pathological proliferation of retinal vessels[53]. Diabetes alone is the leading cause of kidney failure, lower limb amputations, and blindness in adults in the United States[55]. It is estimated that by the year 2030, 7% of deaths in the United States will be the result of diabetes[56].

In addition to the physiological complications of the disease, there are several psychological consequences. In one study of both male and female type 1 and type 2 diabetes patients, 35% self-reported anxiety and depression[52]. It has been found that depression is roughly twice as common in patients with diabetes than in the general public[57], which can dramatically alter a patient's ability to maintain good glycemic control.

Along with the physiological consequences of the disease, there are social and economic consequences as well. In some cases, patients with diabetes are explicitly forbidden from certain careers, such as flying a commercial aircraft in the United States[58]. In other cases, the discrimination is more subtle. In a study based on young adults ages 15-34 in Sweden, it has been found that patients with type 1 diabetes will experience a significant earnings reduction[59]. Despite no difference in earnings prior to the onset of diabetes, it was found that ten years after the onset of type 1 diabetes, young adults earned 4.0% less

than their non-diabetic peers, and twenty years after the onset of diabetes, 10% less[60]. There is considerable debate as to the cause of this earnings reduction, as it may result from workplace discrimination or as the result of increased absenteeism due to decreased overall health[59].

Prevalence of the Disease

In the United States, there are 25.8 million people in a diabetic or pre-diabetic state, which accounts for 8.3% of the population[55]. In addition, diabetes is becoming more and more prevalent. The Centers for Disease Control and Prevention found that from between 1980 to 2011, the number of adults in the United States with diabetes has more than tripled (Figure 6)[61]. Diabetes is also a serious global health problem, affecting 6.4% of adults, which represents 285 million people. Over the next two decades, these numbers are projected to increase by 69% in developing countries and 20% in developed countries[62]. This epidemic means that 5.2% of deaths, globally, can be attributed to diabetes[63]. By 2030, it is projected that the total number of patients with diabetes will rise, with some estimates at 370 million[52] and others at 439 million[62]. These statistics highlight the growing epidemic of diabetes, and emphasize the need for better management of the disease.

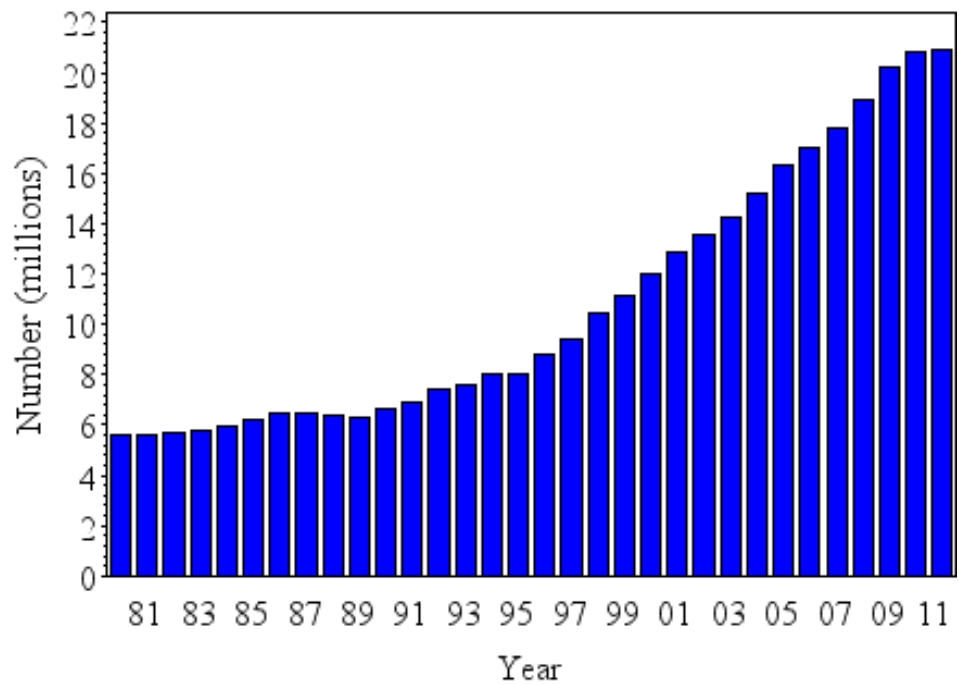


Figure 6: "Number (in Millions) of Civilian, Noninstitutionalized Adults with Diagnosed Diabetes, United States, 1980-2011." Provided by the Centers for Disease Control and Prevention [61].

Management of Diabetes Mellitus

Diabetes therapies aim to improve insulin sensitivity, or in the absence of insulin, provide exogenous insulin to mimic native metabolism.

For type 2 diabetes, the first goal may be to improve insulin sensitivity. Early intervention in diabetes mellitus is crucial. Patients with type 2 diabetes are often encouraged to change their diet and exercise habits in order to restore insulin sensitivity[64]. It has been shown that diet and exercise can have a beneficial effect on the amount of glycosylated hemoglobin (A₁C) in patients with diabetes[64]. For type 2 diabetes, pharmacological interventions are also available, such as metformin, thiazolidinediones, sulfonylurea derivatives and glucagon-like peptide 1 (GLP-1) analogs[3]. Metformin is one of the earliest prescribed diabetes medications, first appearing in Europe over 40 years ago[65]. Metformin suppresses endogenous glucose production and decreases fasting blood glucose[65].

Thiazolidinediones are a newer class of anti-diabetic therapy which improve insulin sensitivity, lower fasting and postprandial glucose concentrations, and decrease the concentration of free fatty acids[66]. The mechanism of thiazolidinediones is believed to be a “fat-stealing” mechanism, where the thiazolidinediones increase fatty acid uptake, relieving non-adipose tissues from fatty acid stress[66]. Unfortunately, this causes weight gain[66]. Another side effects of thiazolidinediones is osteopenia, a reduction in bone mineral density, and specific thiazolidinediones have been linked to liver failure, bladder cancer, and cardiovascular morbidity[67].

Another target for anti-diabetic drugs is the sulfonylurea receptor isoform SUR1, which is found in pancreatic β -cells and serves as a regulatory subunit for ATP-sensitive potassium ion channels (K_{ATP})[68]. Mutations in SUR1 are associated with chronic hyperinsulinism and neonatal diabetes[69]. SUR1 is a member of the ATP-binding cassette family and is sensitive to blood glucose levels[69]. As blood glucose increases, glucose metabolism results in increased concentration of ATP[69]. ATP binds to SUR1, and closes the K_{ATP} membrane channels, resulting in membrane depolarization, and a subsequent influx of calcium ions[69]. The influx of calcium ions stimulates insulin release[69]. Sulfonylurea-based drugs bind to SUR1 and stimulate insulin secretion through the same mechanism of membrane depolarization, influx of calcium ions, and insulin release[69, 70]. However, it has been shown *in vitro* in rodent and human β -cells that prolonged treatment with sulfonylurea derivatives can actually reduce the number of functional K_{ATP} channels[68], and induce β -cell apoptosis[68, 71, 72]. This may be the cause of clinically observed phenomenon called “secondary failure,” where sulfonylurea drugs lose their effectiveness after prolonged exposure[68].

Another important class of pharmacological diabetes therapy is GLP-1 and GLP-1 derivatives. GLP-1 is part of the incretin family of hormones, which includes GLP-1, GLP-2, glucagon, and glicentin[73]. These peptides are produced by L-cells, found mainly in the distal gut, and are released in response to the ingestion of nutrients such as glucose, triacylglycerol, and fructose[73]. GLP-1 lowers blood glucose by stimulating insulin secretion, suppressing gluconeogenesis, delaying gastric emptying and acting as

an appetite suppressor[67, 74]. Exenatide is a GLP-1 analog, and it has been shown to improve glycemic control and reduce body weight[54]. In addition, there is evidence that patients managing their diabetes with exenatide, in conjunction with other anti-diabetic therapies, have a lower risk of cardiovascular disease[54]. However, GLP-1 based therapies have their own set of side-effects and potential consequences. GLP-1 is a growth factor, and in addition to the pancreas, GLP-1 receptors are found in the thyroid, kidneys, bone, and other tissues[67]. A major concern is that the GLP-1 therapies promote pancreatitis, which may in turn lead to pancreatic cancer[67]. However, the evidence for the oncogenic effects of GLP-1 therapy is debatable, especially because it has been shown that pancreatitis is more prevalent in patients with type 2 diabetes, regardless of the type of pharmacological intervention[67].

Despite potential side effects and consequences, all of these therapies are beneficial for patients with type 2 diabetes, since insulin is still being produced by the pancreas. However, these therapies are ineffective for patients with type 1 diabetes, for whom the beta cells of the pancreas may be partially or completely destroyed. Therefore, patients with type 1 diabetes do not have the luxury of trying to restore insulin sensitivity and must rely on exogenously administered insulin. Unfortunately, exogenously administered insulin will never be a perfect mimic of native insulin. One difference between native and exogenous insulin is that exogenously administered insulin is injected subcutaneously and therefore targets all the peripheral tissues equally, as opposed to native insulin secretion through the hepatic portal, which initially targets the liver[1]. Secondly, injected

insulin is subject to a significant lag phase, during which the insulin dissociates from biologically inactive hexamers and dimers into biologically active monomers[75].

However, by using a combination of basal (long acting) and post-prandial (bolus) insulin, patients can achieve near-normal blood glucose[75]. A continuous, low level of insulin secretion is necessary to counteract endogenous gluconeogenesis in the liver[76]. In addition to basal insulin, bolus insulin is used to manage spikes in blood glucose after a meal (post-prandial)[76]. Different synthetic insulins and insulin analogs have been created to address these needs. Long-acting insulins are used to mimic basal insulin secretion. These insulins include neutral protamine Hagedorn (NPH) insulin, insulin glargine, and insulin detemir. NPH insulin has a duration of action of 8-12 hours[77]. Insulin detemir has a duration of action of 12-16 hours, and insulin glargine approaches a 24 hour duration of action[77]. As with every insulin-based therapy, there is the possibility of insulin-induced hypoglycemic events[76], and there is some evidence that insulin glargine is safer, in terms of hypoglycemic events, than insulin NPH[77].

Fast acting insulins, used to manage post-prandial blood glucose spikes, include insulin lispro and insulin aspart. These insulins are “fast-acting” because they dissociate more rapidly from a hexamer to monomers. Fast-acting insulins have been shown to reduce the risk of hypoglycaemia compared to exogenous but otherwise native human insulin[75].

It is only through a combination of slow- and fast-acting insulin that insulin replacement therapy can begin to mimic natural insulin secretion. It has been found that this basal-

bolus therapy, also referred to as “intensive” insulin therapy, is beneficial for patients with type 1 diabetes, and helps prevent retinopathy, neuropathy, and nephropathy[78]. However, this therapy is challenging due to significant variation between patients and even within patients. In one study, it was found that there was a 30% variation in interpatient insulin absorption of fast-acting insulins and 50% interpatient absorption of intermediate insulin (insulin NPH)[75]. In the same study, it was found that inpatient variability was 10-13% for fast acting, and 25% for intermediate acting[75]. This variability makes using insulin to achieve euglycemia a very daunting task.

While achieving euglycemia using exogenous insulin is difficult due to the logistics of basal and bolus injections, insulin itself is very dangerous and mismanaged insulin can be fatal, due to its very narrow therapeutic index. The therapeutic index is defined as the effective dose (ED) divided by the lethal dose (LD)[79]. The narrow therapeutic index of insulin means that patients face the possibility of unintended overdose. In fact, in 2001, the Joint Commission on the Accreditation of Healthcare Organizations, an organization within the United States, identified insulin as one of the top five most dangerous medications, on par with opiates and narcotics[80].

To summarize all of these consequences, insufficient amounts of insulin result in long term complications, while insulin overdose can be fatal immediately. Fear of hypoglycemia represents a significant factor in dose determination for patients that manage their diabetes with insulin[81, 82]. Therefore, the goal of this research is to create a “safer” insulin analog that has a reduced risk of hypoglycemia. An insulin analog

that lacks or reduces the threat of an insulin overdose would simultaneously eliminate the risk of hypoglycemia and also result in less time spent at hyperglycemic levels through increased and more rigorous glycemic control.

Properties of Insulin

This research aims to create an insulin analogue with an improved therapeutic index. Currently, exogenously administered insulin carries the danger of overdose. Normal metabolism is an interplay between insulin, which lowers blood glucose, and glucagon and other counter-regulatory hormones that raise blood glucose. Exogenously administered insulin, however, can “overpower” the normal counter-regulatory action of glucagon and other hormones and cause blood glucose to plummet to hypoglycemic or even fatally low levels (Figure 7, curve 1).

Ideally, this work will address this danger by creating an insulin analog with a shallower response gradient, resulting in a larger range of concentrations that provide glucose reduction without inducing hypoglycemia (Figure 7, curve 2).

The most desirable outcome would be the creation of an insulin analog that lowers blood glucose to normal ranges, but is subsequently opposed by the counter-regulatory hormones (Figure 7, curve 3). This would create an insulin analog where overdose is “impossible.” Not only would this eliminate the dangers of overdose, it would result in better glycemic control and subsequently eliminate the dangers of chronic hyperglycemia.

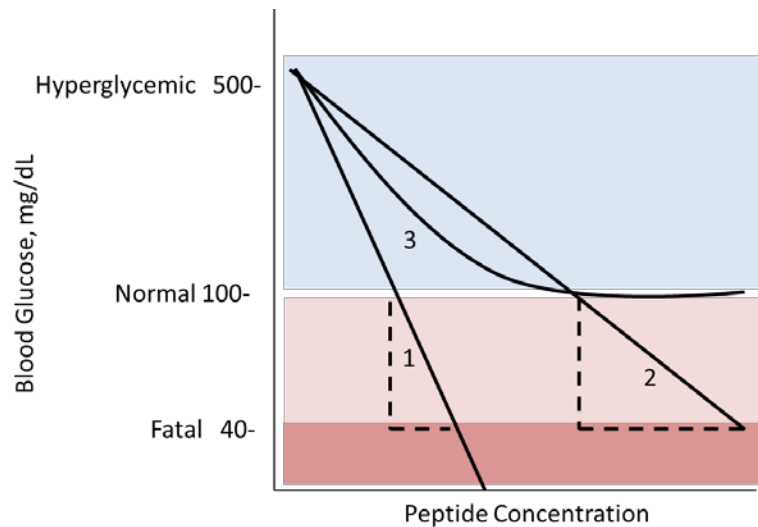


Figure 7: Stylized insulin dose-response as a function of blood glucose versus peptide concentration (S. Brandt). Curve 1 represents native insulin action. Curve 2 represents a shallower dose-response. Curve 3 represents the most ideal insulin therapy, where blood glucose is lowered but opposed by metabolic counter-regulation.

When describing the characteristics of insulin and insulin analogs, there are two important *in vitro* characteristics, potency and maximal activity, which are quantified in this research by receptor phosphorylation. Potency refers to the concentration of peptide required for a half-maximal response, also referred to as the EC_{50} value. Maximal activity represents the upper plateau in a sigmoidal shaped dose-response curve, after which additional peptide does not elicit any increase in phosphorylation. Native insulin has a very high maximal activity, although the physiological significance of this is not known, since *in vivo* insulin levels approaching the maximal activity are fatal. It is possible that an insulin analog with reduced maximal activity would result in a shallower or opposable dose-response curve. This work aims to create an analog with the same potency as native insulin, but with reduced maximal activity. A matched set of analogs, with multiple and varied maximal activities (Figure 8), would allow investigations into the effect of

maximal activity on *in vivo* activity. Maintaining the potency of these analogs at the same level as insulin will allow us to isolate the effect of maximal activity alone, and potentially allow for the rational design of insulin analogs with a specific or desired maximal activity.

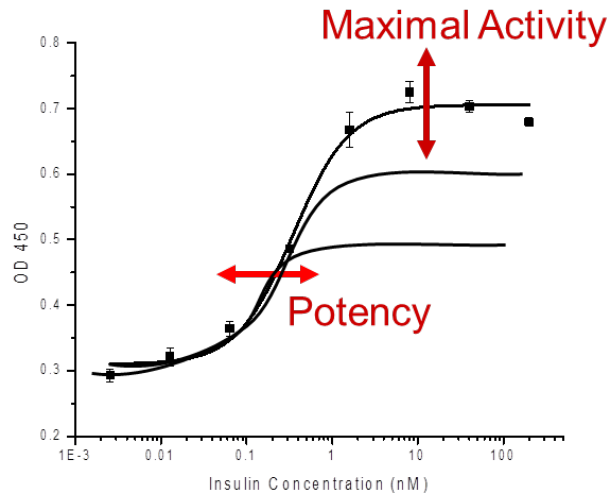


Figure 8: Typical sigmoidal dose response for native insulin (black squares) with stylized theoretical dose response curves for insulin analogs with similar potency but decreased maximal activity, relative to native insulin.

In order to create peptides with decreased maximal activity, this work will exploit a unique feature of the insulin / insulin receptor interaction: supersaturating the insulin receptor with insulin results in a decrease in receptor activation. First, this work will describe normal receptor/ligand interactions, and then progress to the unique case of insulin and the insulin receptor.

Receptor/Ligand Interactions[39]

A simple model of receptor/ligand interactions is to examine the association of free ligand (L) with that of the free receptor (R) to that of the bound complex (R*L):



The association constant is therefore

$$K_a = \frac{[R * L]}{[R][L]} \quad \text{Equation 2}$$

In biochemistry, it is often more useful to discuss the dissociation constant, which is simply the inverse of Equation 2:

$$K_d = \frac{[R][L]}{[R * L]} \quad \text{Equation 3}$$

We further define the free receptor R as a function of the total receptor concentration, R_T , minus the concentration of bound receptor, R*L:

$$R = [R_T] - [R * L] \quad \text{Equation 4}$$

Equation 4 into the dissociation constant expression, we obtain

$$K_d = \frac{([R_T] - [R^*L])[L]}{[R^*L]} \quad \text{Equation 5}$$

Equation 5 can be rearranged to yield

$$R_T = \frac{K_d[R^*L]}{L} + [R^*L] \quad \text{Equation 6}$$

Or

$$R_T = [R^*L] \left(\frac{K_d}{[L]} + 1 \right) \quad \text{Equation 7}$$

The rationale behind this rearrangement stems from a common way of analyzing receptor/ligand interactions as a fraction (Y) of bound receptor, [R*L], to total receptor, [R_T].

$$Y = \frac{[R^*L]}{[R_T]} \quad \text{Equation 8}$$

By substituting the expression for R_T obtained in, Equation 7, we can solve for the fraction of bound receptor as a function of ligand (L) and dissociation constant (K_d).

$$Y = \frac{[R^*L]}{[R^*L] \left(\frac{K_d}{L} + 1 \right)}$$

Equation 9

Which can be further simplified to

$$Y = \frac{1}{\left(\frac{K_d}{L} + 1 \right)}$$

Equation 10

Or

$$Y = \frac{L}{K_d + L}$$

Equation 11

This relationship between the fraction of receptor bound with ligand and total receptor (Y) can be analyzed by plotting Y versus the concentration of free ligand (L), which in a simple protein/ligand interaction yields a hyperbolic plot, where at $Y=0.5$, $L=K_d$ (Figure 9).

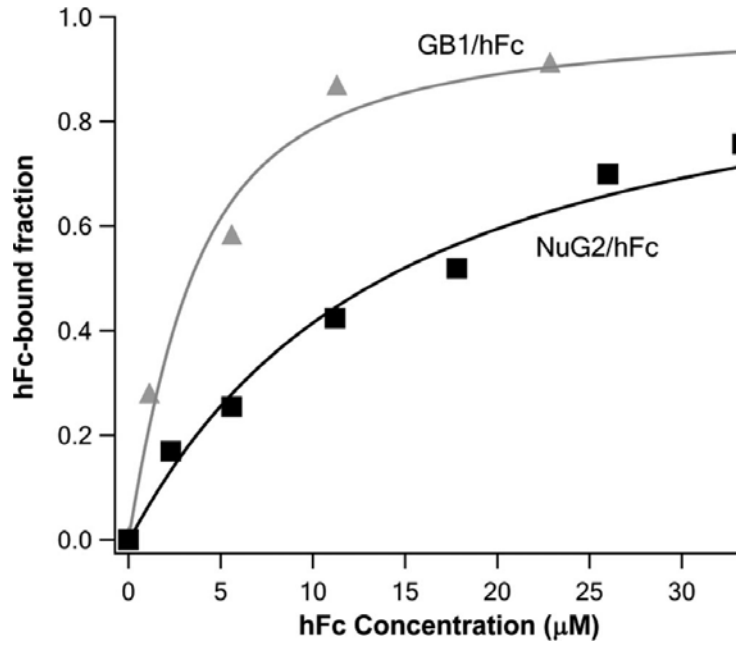


Figure 9: An analysis of the binding of a truncated human IgG protein (hFc) to its binding domain in Protein G (GB1) as well as a mutant binding domain (NuG2). The authors of this study used a single-molecule binding model to determine the K_d for hFc to GB1 ($2.2 \pm 0.4 \mu\text{M}$) and for hFc to NuG2 ($12.6 \pm 0.9 \mu\text{M}$)[83].

Another method for analyzing protein/ligand interactions is a Scatchard plot, which represents the linear form of binding analysis. First, a rearrangement of Equation 5 yields

$$\frac{[R^*L]}{[L]} = \frac{([R_T] - [R^*L])}{K_d} \quad \text{Equation 12}$$

For a Scatchard plot, the bound receptor $[R^*L]$ is customarily defined as B (bound), the free ligand L is defined as F (free), and the total receptor concentration is defined as B_{\max} , the maximal amount of bound receptor. Therefore Equation 12 is rewritten as

$$\frac{B}{F} = \frac{(B_{\max} - B)}{K_d} \quad \text{Equation 13}$$

Or in its linear form:

$$\frac{B}{F} = -\frac{1}{K_d} \cdot B + \frac{B_{\max}}{K_d}$$

Equation 14

Therefore for simple receptor/ligand interactions, a plot of the fraction of bound to free ligand versus the concentration of bound ligand will yield a straight line with a slope equal to $-1/K_d$, an x intercept equal to B_{\max} and a y intercept equal to B_{\max}/K_d (Figure 10, inset).

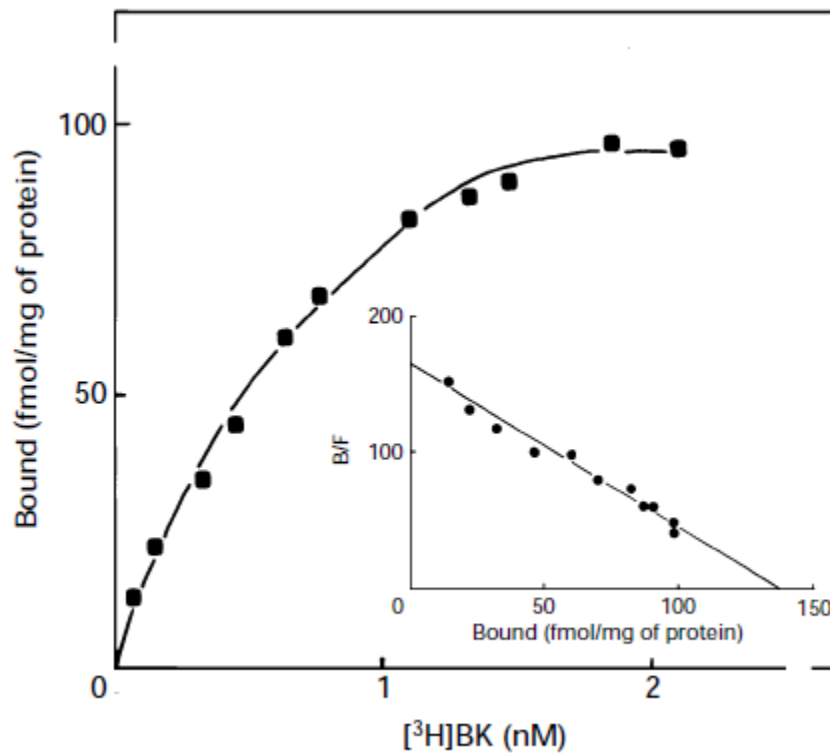


Figure 10: An analysis of the binding of bradykinin (BK) to its receptors on cell surface membranes[22]. A Scatchard plot is seen in the inset, showing the linear relationship between the B/F ratio and bound bradykinin.

When a Scatchard plot is not linear it is indicative of the presence of high and low affinity binding sites on the receptor[32]. A Scatchard plot of insulin binding to its receptor is not linear (Figure 11). This observation was initially dismissed as the result of insulin degradation products obtained in the process of lysing cells. However, after chromatographically separating insulin degradation products as well as using an inhibitor of insulin degradation (bacitracin), the Scatchard plot of insulin binding remained curvilinear[84]. In addition, the dissociation of radiolabeled insulin from its receptor is accelerated by the addition of unlabeled insulin, whereas it is not accelerated by dilution alone (Figure 12). These two observations suggest that insulin displays negative cooperativity at the insulin receptor, where the binding of the ligand decreases the affinity for additional ligands.

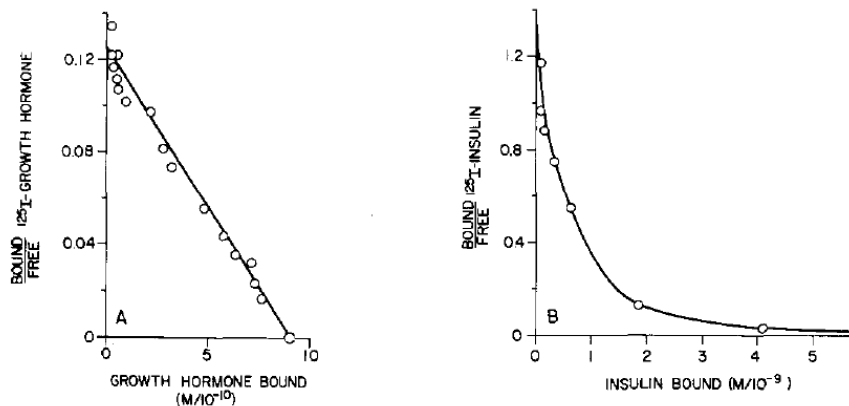


Figure 11: Scatchard plots for radiolabeled human growth hormone bound to its receptor (left) and radiolabeled porcine insulin (right) to its receptor[23]. Human growth hormone shows a linear relationship between B/F and bound ligand, whereas insulin does not. This is indicative of negative cooperativity at the insulin receptor.

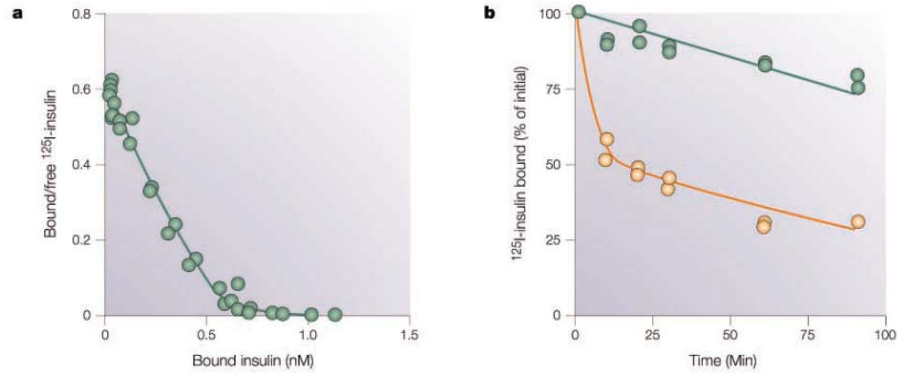


Figure 12: a: Scatchard plot demonstrating the curvilinear relationship between B/F and bound insulin[32]. b: Accelerated dissociation of radiolabeled insulin in the presence of additional unlabeled insulin (orange) compared to dissociation due to dilution alone (green)[32]. This acceleration is indicative of negative cooperativity.

A third and critical argument for negative cooperativity is the bell-shaped curve of insulin binding studies (Figure 13). It was observed that increasing concentrations of cold insulin induce dissociation of bound radiolabeled insulin (Figure 13). However, at higher concentrations the disappearance of radiolabeled insulin is diminished (Figure 13).

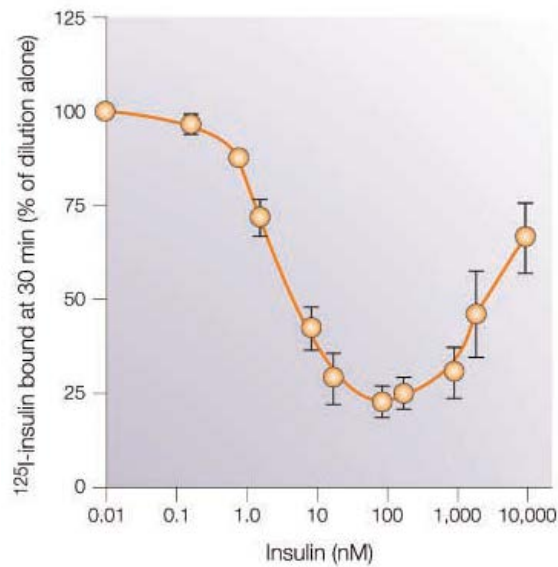


Figure 13: Dose-response curve analyzing the percent of radiolabeled insulin bound as a function of increasing amounts of cold (unlabeled) insulin[17].

Initially, this observation was thought to be the result of insulin dimerizing at high concentrations, where the dimerization somehow precluded binding to the insulin receptor[85]. However, modern crystal structures and a better understanding of the insulin receptor have led to a more appropriate interpretation.

The ectodomain of the insulin receptor exists in a folded-over conformation[86], like an inverted “V” (Figure 14). Electron micrographs of the insulin receptor, in conjunction with monoclonal antibodies, have demonstrated that the two halves of the insulin receptor are also arranged in an anti-parallel fashion[32]. This means that the L1 domain of one half of the receptor is near the fibronectin type III domains of the other half of the receptor, and vice-versa (Figure 14).

The insulin receptor has two binding sites for insulin, which have been identified as the L1 and C-terminus of the alpha domain (site 1), and the junction between the first and second fibronectin-type III domains (site 2). Because of the anti-parallel arrangement of the ectodomains, these binding sites are also arranged in an anti-parallel fashion (Figure 15). Insulin will bind to site 1 of one half of the insulin receptor, and site 2 of the other half of the receptor, cross-linking the two ectodomains.

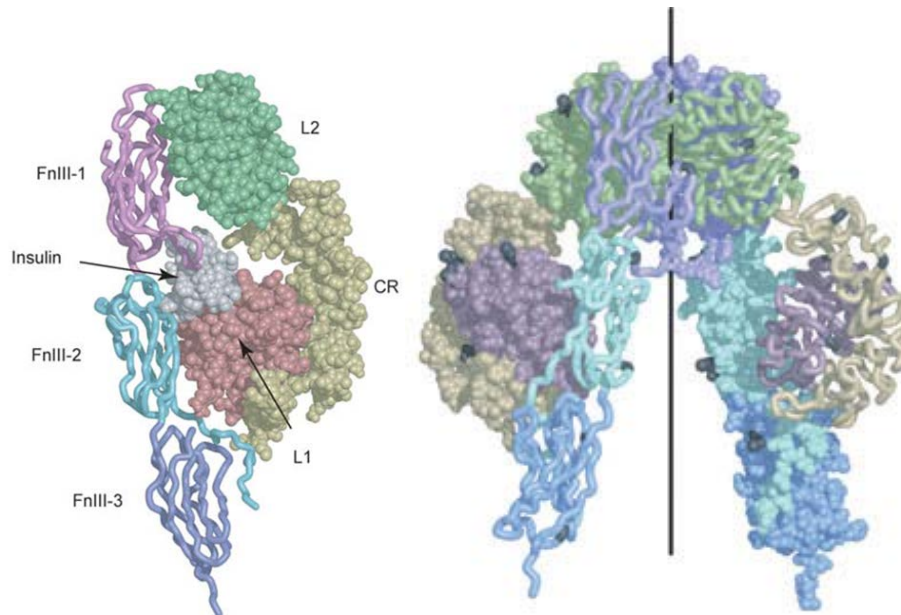


Figure 14: Crystal structure of the insulin receptor ectodomain, shown as a monomer, with insulin bound (left)[24], and as an anti-parallel dimer (right)[16].

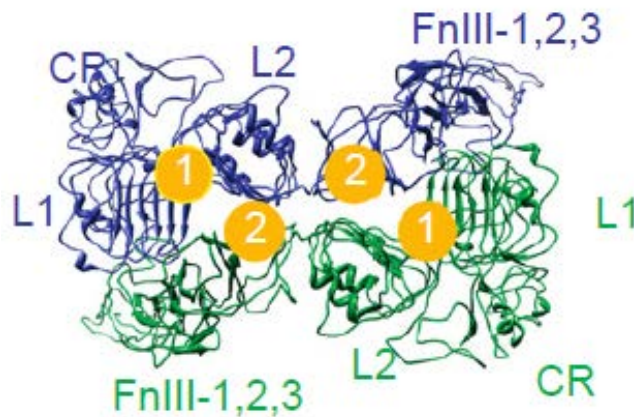


Figure 15: A “top down” view of the insulin receptor ectodomains, with one half of the receptor in green and the other half in blue[15]. Binding sites 1 and 2 of insulin are shown in yellow. The anti-parallel arrangement of the ectodomains results in an anti-parallel arrangement of binding sites.

Therefore, when a single insulin monomer binds to the insulin receptor, it occupies binding site 1 on one half of the receptor, and binding site 2 on the other half of the receptor, leaving two unoccupied binding sites available. As the concentration of insulin monomers increases, additional insulins can bind to the unoccupied sites, and the resulting signal decreases. This is proposed as the basis of the negative cooperativity observed at the insulin receptor. In regards to the binding assay, bound insulin is initially displaced by additional insulin binding to unoccupied binding sites, however, as insulin concentration increases the receptor becomes saturated with 3 insulins, one in the “original” crosslink, and two on each of the unoccupied binding sites, resulting in the maintenance of bound radiolabeled insulin at high concentrations of cold insulin (Figure 16)[25].

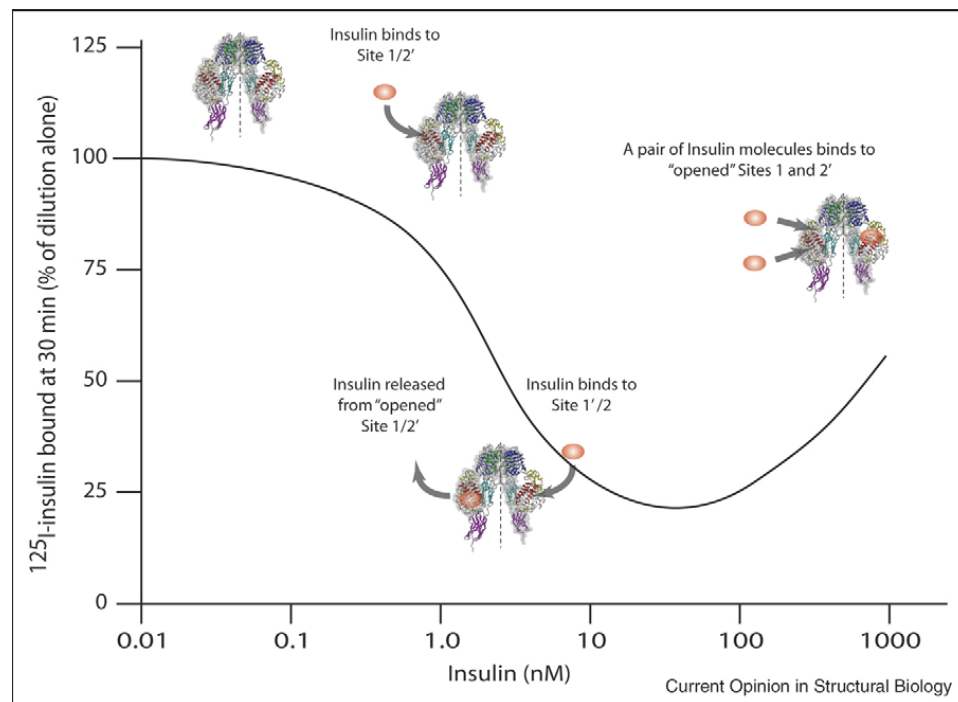


Figure 16: Schematic illustrating a stylized binding curve of radiolabeled insulin versus increasing concentrations of cold insulin[25].

The effect of “supersaturating” the insulin receptor can be seen in in vitro studies of insulin. At concentrations above 100 nM, the activation of the receptor, as measured by phosphorylation of the tyrosine kinase domain, begins to diminish (Figure 17). Therefore, if supersaturating the receptor with insulin monomers leads to diminished signal, it was hypothesized that covalently dimerized insulins might have unique and interesting activity at the insulin receptors.

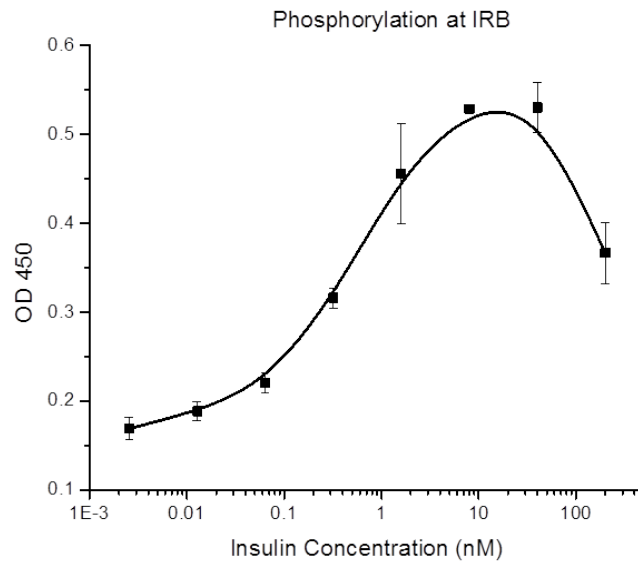


Figure 17: Phosphorylation assay demonstrating the activation of insulin receptor isoform B in response to native human insulin, where OD 450 is a measure of receptor phosphorylation. There is a decrease in signal at insulin concentrations greater than 100 nM (S. Brandt, unpublished data).

The research group of Dietrich Brandenburg explored this possibility by creating insulin homodimers, linking two identical insulins with a bis(p-nitrophenyl) ester[87]. The insulin dimers were linked at either the A¹, B¹ or B²⁹ positions (Figure 18). Therefore, the Brandenburg group created an A¹-A^{1'}, B¹-B^{1'}, B²⁹-B^{29'}, A¹-B^{1'}, A¹-B^{29'}, and B¹-B^{29'} dimers (Figure 18). All of these dimers exhibited the same maximal activity as native, monomeric insulin, as measured by an in vitro fat cell assay (Figure 19)[87]. However, the potency of these analogs, relative to insulin, varied between 1-60% (Figure 19)[87].

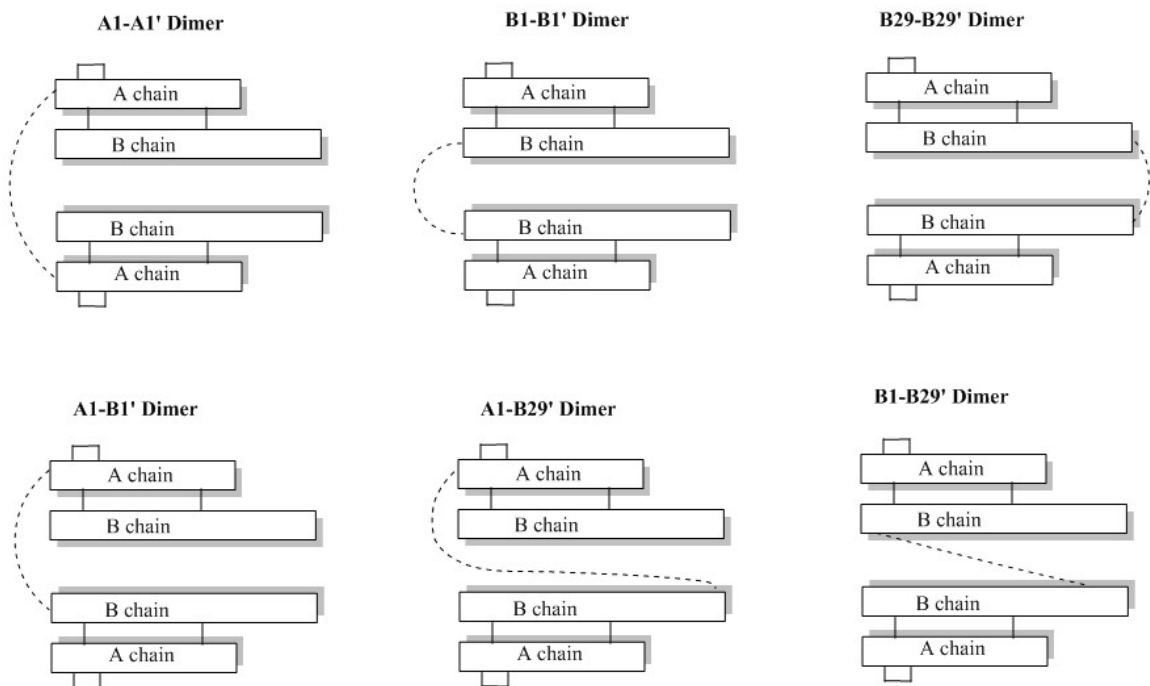


Figure 18: Schematic of insulin homodimers, dashed line representing suberoyl ester linker.

The Brandenburg group successfully demonstrated that this is not simply the result of differences in binding affinity. The B²⁹-B^{29'} linked dimer, while only mildly stimulating glucose uptake, bound to the insulin receptors with similar binding affinity as native insulin[88](Figure 20).

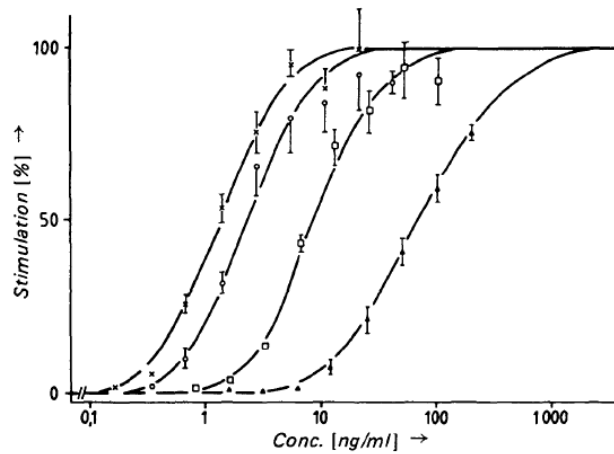


Figure 19: Stimulation of lipogenesis by insulin and insulin dimers. Insulin is represented by X's (far left), B¹-B^{29'} dimer (open circles), B¹-B^{1'} dimer (open squares) and B²⁹-B^{29'} (closed triangles).

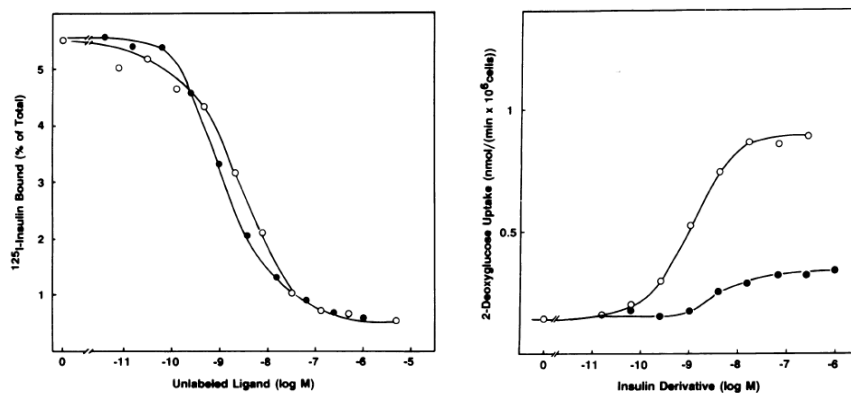


Figure 20: (Right) Binding assay demonstrating that monomeric insulin (white circles) and B²⁹-B^{29'} dimerized insulin (black circles) have similar affinities for the insulin receptor [27]. (Left) Fat cell assay analyzing glucose uptake by monomeric insulin (white circles) and B²⁹-B^{29'} dimerized insulin (black circles)[88].

This was an exciting observation because it implied that the diminished activity at the insulin receptor was not the result of a lack of binding by the dimer, but rather an inherent effect of the dimer. Therefore the Brandenburg research group demonstrated that it is possible to alter the maximal activity and potency of an insulin-based analog not by altering the binding affinity but by altering how the ligand interacts with the receptor. It is possible that the insulin dimer is a mimic of the receptor supersaturation seen at high concentrations of insulin monomers.

More recently, Dr. Yan Zhao in the DiMarchi research group investigated insulin dimers, observing the effect of linker position, linker type, and linker length. In insulin dimers linked at the B¹ position, the dimer was a full agonist of the insulin receptors, regardless of linker length[3]. In B²⁹ linked insulin dimers, Dr. Zhao observed a partial agonist with low potency[3]. In B²⁹ linked IGF-based dimers, potency and activity at the insulin receptor was restored[3]. Linking two insulin dimers together through a polyethylene glycol (PEG) C-domain resulted in a partial agonist with a bell-shaped curve, and introducing a larger PEG linker between the dimers restored activity[3].

All of these results suggest that we can alter the potency and maximal activity of an insulin dimer. Our goal, as previously stated, is to create a series of dimerized peptides that have a range of maximal activities, but maintain potency, as compared to native insulin.

Prior to this research, all insulin and insulin-mimetic dimers have been homodimers, with the two halves of the dimer being chemically identical. However, an insulin heterodimer may be equally useful at altering activity and potency at the insulin receptor. The heterodimers presented in this research will consist of insulin or a fully potent insulin analog, and an insulin receptor antagonist. The inclusion of an antagonist in a heterodimer will provide additional binding sites, but the antagonist will not possess any inherent ability to activate the receptor itself, and therefore we anticipate that these heterodimers will have diminished activity at the insulin receptor. Several potentially useful antagonists were identified in the work of Schaffer, et al. The Schaffer research group used a peptide phage library to screen for random peptides with affinity at the insulin receptor ectodomain[89, 90]. By optimizing these peptides the Schaffer group was able to identify several binding motifs (Table 1)[91]. These peptides were categorized as binding to the insulin receptor binding sites 1 and 2 (the classical binding sites of native insulin), as well as binding site 3, a region of the insulin receptor ectodomain not associated with native insulin binding[91].

Table 1: Binding motifs identified by Schaffer, et al., where X represents an unspecified amino acid.

IR Binding Site	Motif
1	FYXWF FYWW(L/D)XXL LXXLXXYF
2	Cysteine loops
3	ACVWPTYWNCG

In summary, there have been three key observations that lay the foundation of this work. First, that supersaturating the insulin receptor with insulin results in a decrease in receptor activation. Second, that insulin-insulin homodimers result in variable maximal activity, and third, that there are synthetically derived peptides capable of binding to the insulin receptor. Based on these observations, we hypothesize that it is possible to create a heterodimer capable of varied maximal activity at the insulin receptor, by dimerizing an insulin receptor agonist and antagonist.

The creation of these heterodimers requires several unique chemistries and skills. First, an appropriate peptide must be identified, characterized, and optimized for inclusion in a heterodimer. Chapter 2 will discuss the creation and characterization of several previously identified insulin receptor antagonists using synthetic chemistry. It will discuss solid phase peptide chemistry as a synthetic tool, and show through *in vitro* phosphorylation assays that we have created peptides capable of antagonizing native insulin at both the insulin receptor A and B isoforms. In addition, structure-activity relationship studies allowed for the optimization of these peptides, where structural complexity was minimized without loss of function.

Building from our optimized antagonists, Chapter 3 will discuss the creation of insulin receptor agonist/antagonist heterodimers through the use of a biosynthetic protocol, where heterodimers were expressed in *E. coli*, purified, and then tested for agonism and antagonism at both insulin receptor isoforms. A heterodimer with the desired decrease in maximal activity was identified, and it represents the first instance of a heterodimer with

altered maximal activity but unaltered potency. This breakthrough confirmed our hypothesis that heterodimers are a viable way to alter the maximal activity of insulin-based molecules. In addition, the successful creation of this heterodimer allowed for further structure-activity relationship studies into the role and effect of antagonists within a heterodimer.

Chapter 4 discusses the use of a semisynthetic approach to create and optimize a series of heterodimers, with varied maximal activities. The semisynthetic approach allowed for the creation of a set of heterodimers with tunable maximal activity. Most importantly, these heterodimers vary in a single amino acid, and therefore represent a tunable and rationally designed set of molecules, and are the first instance of heterodimers with directed maximal activity at the insulin receptors.

This research concludes with a discussion on the relevance and potential therapeutic benefits of these unique heterodimers, and emphasizes the need for continuing research into a promising field of peptide therapy.

REFERENCES

1. Pickup, J.C.W., G., ed. *Textbook of Diabetes*. Vol. 1. 1991, Blackwell Scientific Publications: Oxford.
2. Fisher, C., *All the Time in the World*, in *Warehouse 13* 2013. p. 60 min.
3. Zhao, Y., *Structural optimization of insulin and glucagon as a route to improved treatment of diabetes*. Indiana University, 2012.
4. Rosenfeld, L., *Insulin: discovery and controversy*. Clin Chem, 2002. **48**(12): p. 2270-88.
5. Bliss, M., *Theodore Ryder: The last living link to the discovery of insulin*. Practical Diabetes, 1995. **12**(4): p. 187-188.
6. *Diabetes Overview*. [cited 2013 July 15]; Available from: <http://diabetes.niddk.nih.gov/dm/pubs/overview/index.aspx>.
7. Shestov, A.A., et al., *Simultaneous measurement of glucose transport and utilization in the human brain*. Am J Physiol Endocrinol Metab, 2011. **301**(5): p. E1040-9.
8. Magistretti, P. and I. Allaman, *Brain Energy Metabolism*, in *Neuroscience in the 21st Century*, D. Pfaff, Editor. 2013, Springer New York. p. 1591-1620.
9. *Peptides in Energy Balance and Obesity*. 2009, Wallingford UK: CABI.
10. Layden, B.T., V. Durai, and W.L. Lowe Jr., *G-Protein-Coupled Receptors, Pancreatic Islets, and Diabetes*. Nature Education, 2010. **3**(9).
11. Malaguarnera, R., et al., *Proinsulin binds with high affinity the insulin receptor isoform A and predominantly activates the mitogenic pathway*. Endocrinology, 2012. **153**(5): p. 2152-63.
12. Shibasaki, Y., et al., *Posttranslational cleavage of proinsulin is blocked by a point mutation in familial hyperproinsulinemia*. J Clin Invest, 1985. **76**(1): p. 378-80.
13. Davidson, H.W., *(Pro)Insulin processing: a historical perspective*. Cell Biochem Biophys, 2004. **40**(3 Suppl): p. 143-58.
14. Zierath, J.R., et al., *C-peptide stimulates glucose transport in isolated human skeletal muscle independent of insulin receptor and tyrosine kinase activation*. Diabetologia, 1996. **39**(3): p. 306-13.

15. Gabbay, K.H., et al., *Familial hyperproinsulinemia: partial characterization of circulating proinsulin-like material*. Proc Natl Acad Sci U S A, 1979. **76**(6): p. 2881-5.
16. Ward, C. and M. Larence, *Ligand-induced activation of the insulin receptor: a multi-step process involving structural changes in both the ligand and the receptor*. Bioessays, 2009. **31**(4): p. 422-434.
17. Mayer, J.P., F. Zhang, and R.D. DiMarchi, *Insulin structure and function*. Biopolymers, 2007. **88**(5): p. 687-713.
18. Kurtzhals, P., et al., *Correlations of receptor binding and metabolic and mitogenic potencies of insulin analogs designed for clinical use*. Diabetes, 2000. **49**(6): p. 999-1005.
19. Rinderknecht, E. and R.E. Humbel, *The amino acid sequence of human insulin-like growth factor I and its structural homology with proinsulin*. J Biol Chem, 1978. **253**(8): p. 2769-76.
20. Menting, J.G., et al., *How insulin engages its primary binding site on the insulin receptor*. Nature, 2013. **493**(7431): p. 241-U276.
21. Weiss, M.A., *The structure and function of insulin: decoding the TR transition*. Vitam Horm, 2009. **80**: p. 33-49.
22. Noble, S.L., E. Johnston, and B. Walton, *Insulin lispro: a fast-acting insulin analog*. American Family Physician, 1998. **57**(2): p. 279-96, 289-92.
23. Zakova, L., et al., *Shortened insulin analogues: marked changes in biological activity resulting from replacement of TyrB26 and N-methylation of peptide bonds in the C-terminus of the B-chain*. Biochemistry, 2004. **43**(8): p. 2323-31.
24. Bromer, W.W. and R.E. Chance, *Preparation and characterization of desoctapeptide-insulin*. Biochim Biophys Acta, 1967. **133**(2): p. 219-23.
25. Lawrence, M.C., N.M. McKern, and C.W. Ward, *Insulin receptor structure and its implications for the IGF-1 receptor*. Curr Opin Struct Biol, 2007. **17**(6): p. 699-705.
26. Jensen, M. and P. De Meyts, *Chapter 3 Molecular Mechanisms of Differential Intracellular Signaling From the Insulin Receptor*, in *Vitamins & Hormones*, L. Gerald, Editor. 2009, Academic Press. p. 51-75.
27. Benecke, H., J.S. Flier, and D.E. Moller, *Alternatively spliced variants of the insulin receptor protein. Expression in normal and diabetic human tissues*. J Clin Invest, 1992. **89**(6): p. 2066-70.

28. De Meyts, P., *The insulin receptor isoform A: a mitogenic proinsulin receptor?* Endocrinology, 2012. **153**(5): p. 2054-2056.
29. Benyoucef, S., et al., *Characterization of insulin/IGF hybrid receptors: contributions of the insulin receptor L2 and Fn1 domains and the alternatively spliced exon 11 sequence to ligand binding and receptor activation.* Biochem J, 2007. **403**(3): p. 603-13.
30. Buck, E., et al., *Compensatory insulin receptor (IR) activation on inhibition of insulin-like growth factor-1 receptor (IGF-1R): rationale for cotargeting IGF-1R and IR in cancer.* Mol Cancer Ther, 2010. **9**(10): p. 2652-64.
31. Pandini, G., et al., *Insulin/insulin-like growth factor I hybrid receptors have different biological characteristics depending on the insulin receptor isoform involved.* Journal of Biological Chemistry, 2002. **277**(42): p. 39684-39695.
32. De Meyts, P., et al., *Structural Biology of Insulin and IGF-1 Receptors*, in *Biology of IGF-1: Its Interaction with Insulin in Health and Malignant States*. 2008, John Wiley & Sons, Ltd. p. 160-176.
33. Saltiel, A.R. and C.R. Kahn, *Insulin signalling and the regulation of glucose and lipid metabolism.* Nature, 2001. **414**: p. 799-806.
34. Taniguchi, C.M., B. Emanuelli, and C.R. Kahn, *Critical nodes in signalling pathways: insights into insulin action.* Nat Rev Mol Cell Biol, 2006. **7**(2): p. 85-96.
35. Noble, J.A., et al., *HLA class I and genetic susceptibility to type 1 diabetes: results from the Type 1 Diabetes Genetics Consortium.* Diabetes, 2010. **59**(11): p. 2972-9.
36. Alberti, K.G. and P.Z. Zimmet, *Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation.* Diabet Med, 1998. **15**(7): p. 539-53.
37. Leslie, R.D.G. and R.B. Elliott, *Early Environmental Events as a Cause of Iddm - Evidence and Implications.* Diabetes, 1994. **43**(7): p. 843-850.
38. Redondo, M.J., et al., *Concordance for Islet Autoimmunity among Monozygotic Twins.* New England Journal of Medicine, 2008. **359**(26): p. 2849-2850.
39. Voet, D. and J.G. Voet, *Biochemistry*. 3 ed. 2004: John Wiley & Sons.
40. Bottini, N., et al., *A functional variant of lymphoid tyrosine phosphatase is associated with type 1 diabetes.* Nat Genet, 2004. **36**(4): p. 337-8.

41. Lowe, C.E., et al., *Large-scale genetic fine mapping and genotype-phenotype associations implicate polymorphism in the IL2RA region in type 1 diabetes*. Nat Genet, 2007. **39**(9): p. 1074-82.
42. Barnett, A.H., et al., *Diabetes in Identical-Twins - a Study of 200 Pairs*. Diabetologia, 1981. **20**(2): p. 87-93.
43. Zimmet, P., K.G. Alberti, and J. Shaw, *Global and societal implications of the diabetes epidemic*. Nature, 2001. **414**(6865): p. 782-7.
44. Donath, M.Y., et al., *Islet inflammation in type 2 diabetes: from metabolic stress to therapy*. Diabetes Care, 2008. **31 Suppl 2**: p. S161-4.
45. Laaksonen, M.A., et al., *The relative importance of modifiable potential risk factors of type 2 diabetes: a meta-analysis of two cohorts*. Eur J Epidemiol, 2010. **25**(2): p. 115-24.
46. Paneni, F., et al., *Diabetes and vascular disease: pathophysiology, clinical consequences, and medical therapy: part I*. Eur Heart J, 2013.
47. Papaccio, G., et al., *A biphasic role of nuclear transcription factor (NF)-kappaB in the islet beta-cell apoptosis induced by interleukin (IL)-1beta*. J Cell Physiol, 2005. **204**(1): p. 124-30.
48. Mokdad, A.H., et al., *Prevalence of obesity, diabetes, and obesity-related health risk factors, 2001*. JAMA, 2003. **289**(1): p. 76-9.
49. Haffner, S.M., *Epidemiology of type 2 diabetes: risk factors*. Diabetes Care, 1998. **21 Suppl 3**: p. C3-6.
50. Andreassi, M.G., *Metabolic syndrome, diabetes and atherosclerosis: influence of gene-environment interaction*. Mutat Res, 2009. **667**(1-2): p. 35-43.
51. Awoniyi, O., R. Rehman, and S. Dagogo-Jack, *Hypoglycemia in Patients with Type 1 Diabetes: Epidemiology, Pathogenesis, and Prevention*. Curr Diab Rep, 2013.
52. Solli, O., K. Stavem, and I.S. Kristiansen, *Health-related quality of life in diabetes: The associations of complications with EQ-5D scores*. Health Qual Life Outcomes, 2010. **8**: p. 18.
53. Hammes, H.P., et al., *Diabetic retinopathy: targeting vasoregression*. Diabetes, 2011. **60**(1): p. 9-16.

54. Best, J.H., et al., *Risk of cardiovascular disease events in patients with type 2 diabetes prescribed the glucagon-like peptide 1 (GLP-1) receptor agonist exenatide twice daily or other glucose-lowering therapies: a retrospective analysis of the LifeLink database*. *Diabetes Care*, 2011. **34**(1): p. 90-5.
55. *National Diabetes Fact Sheet, 2011*. Available from: http://www.cdc.gov/diabetes/pubs/pdf/ndfs_2011.pdf.
56. *Diabetes Fact Sheet*. [cited 2013 July 11]; Available from: <http://who.int/mediacentre/factsheets/fs312/en/index.html>.
57. Anderson, R.J., et al., *The prevalence of comorbid depression in adults with diabetes - A meta-analysis*. *Diabetes Care*, 2001. **24**(6): p. 1069-1078.
58. Administration, F.A. *Insulin-Treated Diabetes*. August 9, 2013]; Available from: http://www.faa.gov/licenses_certificates/medical_certification/specialissuance/diabetes/.
59. Songer, T., *The earnings of people with diabetes: a signal of poor health or a sign of better things to come?* *Diabetologia*, 2010. **53**(6): p. 1017-8.
60. Steen Carlsson, K., et al., *Long-term detrimental consequences of the onset of type 1 diabetes on annual earnings--evidence from annual registry data in 1990-2005*. *Diabetologia*, 2010. **53**(6): p. 1084-92.
61. *Number (in Millions) of Civilian, Noninstitutionalized Adults with Diagnosed Diabetes, United States, 1980-2011*. March 2, 2013 [cited 2013 July 22]; Available from: <http://www.cdc.gov/diabetes/statistics/prev/national/figadults.htm>.
62. Shaw, J.E., R.A. Sicree, and P.Z. Zimmet, *Global estimates of the prevalence of diabetes for 2010 and 2030*. *Diabetes Res Clin Pract*, 2010. **87**(1): p. 4-14.
63. Roglic, G., et al., *The burden of mortality attributable to diabetes: realistic estimates for the year 2000*. *Diabetes Care*, 2005. **28**(9): p. 2130-5.
64. Boule, N.G., et al., *Effects of exercise on glycemic control and body mass in type 2 diabetes mellitus - A meta-analysis of controlled clinical trials*. *Jama-Journal of the American Medical Association*, 2001. **286**(10): p. 1218-1227.
65. Natali, A. and E. Ferrannini, *Effects of metformin and thiazolidinediones on suppression of hepatic glucose production and stimulation of glucose uptake in type 2 diabetes: a systematic review*. *Diabetologia*, 2006. **49**(3): p. 434-41.
66. Yki-Jarvinen, H., *Thiazolidinediones*. *N Engl J Med*, 2004. **351**(11): p. 1106-18.

67. Butler, P.C., et al., *A Critical Analysis of the Clinical Use of Incretin-Based Therapies*. Diabetes Care, 2013. **36**(7): p. 2118-2125.
68. Takahashi, A., et al., *Sulfonylurea and glinide reduce insulin content, functional expression of K(ATP) channels, and accelerate apoptotic beta-cell death in the chronic phase*. Diabetes Res Clin Pract, 2007. **77**(3): p. 343-50.
69. Aittoniemi, J., et al., *Review. SUR1: a unique ATP-binding cassette protein that functions as an ion channel regulator*. Philos Trans R Soc Lond B Biol Sci, 2009. **364**(1514): p. 257-67.
70. Groop, L.C., *Sulfonylureas in NIDDM*. Diabetes Care, 1992. **15**(6): p. 737-54.
71. Kim, J.Y., et al., *Exendin-4 protects against sulfonylurea-induced beta-cell apoptosis*. J Pharmacol Sci, 2012. **118**(1): p. 65-74.
72. Maedler, K., et al., *Sulfonylurea induced beta-cell apoptosis in cultured human islets*. J Clin Endocrinol Metab, 2005. **90**(1): p. 501-6.
73. Nauck, M.A., et al., *Secretion of glucagon-like peptide-1 (GLP-1) in type 2 diabetes: what is up, what is down?* Diabetologia, 2011. **54**(1): p. 10-8.
74. Toft-Nielsen, M.B., S. Madsbad, and J.J. Holst, *Determinants of the effectiveness of glucagon-like peptide-1 in type 2 diabetes*. J Clin Endocrinol Metab, 2001. **86**(8): p. 3853-60.
75. Chen, J.W., J.S. Christiansen, and T. Lauritzen, *Limitations to subcutaneous insulin administration in type 1 diabetes*. Diabetes Obes Metab, 2003. **5**(4): p. 223-33.
76. Scheiner, G. and B.A. Boyer, *Characteristics of basal insulin requirements by age and gender in Type-1 diabetes patients using insulin pump therapy*. Diabetes Res Clin Pract, 2005. **69**(1): p. 14-21.
77. Laubner, K., et al., *Daily insulin doses and injection frequencies of Neutral Protamine Hagedorn (NPH) insulin, insulin detemir and glargine in type 1 and type 2 diabetes: a multicenter analysis of 51 964 patients from the German/Austrian DPV-Wiss database*. Diabetes Metab Res Rev, 2013.
78. *The effect of intensive treatment of diabetes on the development and progression of long-term complications in insulin-dependent diabetes mellitus. The Diabetes Control and Complications Trial Research Group*. N Engl J Med, 1993. **329**(14): p. 977-86.

79. Hacker, M., W. Messer, and K. Bachmann, eds. *Pharmacology: Principles and Practice*. 2009, Academic Press: Boston.
80. 'High-alert' medications and patient safety. *Int J Qual Health Care*, 2001. **13**(4): p. 339-40.
81. Peyrot, M., et al., *Correlates of insulin injection omission*. *Diabetes Care*, 2010. **33**(2): p. 240-5.
82. Schober, E., et al., *Prevalence of intentional under- and overdosing of insulin in children and adolescents with type 1 diabetes*. *Pediatr Diabetes*, 2011. **12**(7): p. 627-31.
83. Cao, Y., et al., *A functional single-molecule binding assay via force spectroscopy*. *Proc Natl Acad Sci U S A*, 2007. **104**(40): p. 15677-81.
84. Davidson, M.B. and N. Venkatesan, *Persistence of a curvilinear Scatchard plot for insulin binding despite correcting for degradation*. *Metabolism*, 1982. **31**(12): p. 1206-9.
85. De Meyts, P., *The structural basis of insulin and insulin-like growth factor-I receptor binding and negative co-operativity, and its relevance to mitogenic versus metabolic signalling*. *Diabetologia*, 1994. **37 Suppl 2**: p. S135-48.
86. McKern, N.M., et al., *Structure of the insulin receptor ectodomain reveals a folded-over conformation*. *Nature*, 2006. **443**(7108): p. 218-21.
87. Schuttler, A. and D. Brandenburg, *Preparation and properties of covalently linked insulin dimers*. *Hoppe Seylers Z Physiol Chem*, 1982. **363**(3): p. 317-30.
88. Weiland, M., et al., *Antagonistic effects of a covalently dimerized insulin derivative on insulin receptors in 3T3-L1 adipocytes*. *Proc Natl Acad Sci U S A*, 1990. **87**(3): p. 1154-8.
89. Schaffer, L., et al., *A novel high-affinity peptide antagonist to the insulin receptor*. *Biochem Biophys Res Commun*, 2008. **376**(2): p. 380-3.
90. Schaffer, L., et al., *Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks*. *Proc Natl Acad Sci U S A*, 2003. **100**(8): p. 4435-9.
91. Pillutla, R.C., et al., *Peptides identify the critical hotspots involved in the biological activation of the insulin receptor*. *J Biol Chem*, 2002. **277**(25): p. 22590-4.

Chapter 2
**Synthesis, Characterization, and Optimization of Insulin
Receptor Antagonists**

ABSTRACT

This chapter seeks to identify, characterize, and optimize peptides to be used as antagonists in our putative agonist/antagonist heterodimers. Previously reported insulin receptor agonists and antagonists were created using synthetic Boc or Fmoc solid phase peptide chemistry. These peptides were tested in an *in vitro* phosphorylation assay for agonism at both insulin receptor isoforms, and were tested for their ability to antagonize native insulin at these receptors. The peptides consist of either a single or a pair of binding motifs known to bind to the insulin receptor. None of the single binding motif peptides were able to agonize the insulin receptors or antagonize native insulin. However, two previously reported antagonists, S597 and S661, each containing two binding motifs, were shown to be effective antagonists of the insulin receptor isoforms. Both of these peptides contained a disulfide bond, but structure/activity analyses demonstrated that the disulfide is not necessary for *in vitro* antagonism. Based on this observation, antagonists #4-6 and #6-4 were created, each consisting of two binding motifs. It was found that the relative order of the binding motifs within these peptides does not affect their *in vitro* antagonism. It was also found that within the #4-6 peptide, the N-terminal residues were necessary for antagonism, however, C-terminal truncations can maintain their antagonism. In addition, it was found that ligating a 20 kDa polyethylene glycol molecule to the N-terminus of peptide #4-6 resulted in a decrease in potency but did not affect its ability to antagonize native insulin.

INTRODUCTION

As stated in the introductory chapter, this research aims to create heterodimers, consisting of an insulin receptor agonist and antagonist. We hypothesize that the inclusion of an antagonist in a heterodimer will result in variable maximal activity. Therefore, the first step of this research is to identify and optimize an insulin receptor antagonist. Beginning with peptides originally identified by the Schaffer research group[1], structure/activity analyses were used to further optimize insulin receptor antagonists.

These antagonists are synthetic peptides, since there are no known naturally occurring insulin receptor antagonists[2]. Using synthetic strategies to create novel agonists for a given receptor has been successful in other ligand/receptor systems. For example, the research groups of Wrighton and Livnah, et al., created a synthetic agonist of the erythropoietin receptor that was capable of fully activating the receptor, without any sequence similarity to the native ligand[1]. However, the insulin receptor has historically resisted attempts at synthetic antagonists. No known native antagonists of the insulin receptor exist, and, in contrast to other receptor/ligand systems, truncated versions of the native ligand do not act as antagonists[3].

Despite these limitations, several potentially useful peptides were identified in the work of Schaffer, et al. The Schaffer research group used a peptide phage library to screen for random peptides with affinity for the insulin receptor ectodomain[3, 4]. As stated in the introductory chapter, the Schaffer group identified several motifs that bind to the insulin receptor (Table 1)[1]. The Schaffer research group posits that the site 1 motif yields

peptides capable of partially or fully activating the insulin receptor, as shown by uptake of radiolabeled glucose into mouse epididymal adipocytes[1], whereas the site 2 and 3 motifs act as antagonists to the insulin receptors[1]. A limitation of these binding motifs is that peptides containing a single binding motif had very low affinity for the insulin receptors, with dissociation constants on the order of 1 μ M[1]. However, by creating heterodimers of binding motifs, the Schaffer group created peptides with increased affinity and activity at the insulin receptors[4].

Table 1: Binding motifs identified by Schaffer, et al., where X represents an unspecified amino acid.

IR Binding Site	Motif
1	FYXWF FYWW(L/I)XXL LXXLXXYF
2	Cysteine loops
3	ACVWPTYWNCG

We recreated several of these potentially antagonistic peptides by synthesizing them using solid phase peptide chemistry. After synthesis and purification, the peptides were analyzed for agonism and antagonism in a receptor tyrosine phosphorylation assay in human embryonic kidney (HEK) cells. These cells overexpress either the insulin receptor A (IRA) or insulin receptor B (IRB) isoform, and phosphorylation of the receptors is measured using phosphotyrosine antibodies. These antibodies are conjugated to horse radish peroxidase, an enzyme that induces a color change in the presence of peroxide and a chromophore. Therefore, the optical density at 450 nm (OD 450) is a measure of receptor phosphorylation. Agonism of the peptide was measured as a function

of receptor phosphorylation in response to a range of peptide concentrations. Antagonism was measured as a function of receptor phosphorylation in response to a range of peptide concentrations, in the presence of constant 1 nM native insulin.

Single site peptides **#1**, **#2**, and **#3** (Table 2) were identified as good candidates for inclusion in a heterodimer. Peptides **#1-3** were reported to be agonists at the insulin receptor[5]. These peptides are relatively short, which simplifies their synthesis. In addition, these peptides contain a lysine residue at their C-terminus. Since the lysine side chain has unique chemical reactivity, the inclusion of this residue allows for subsequent chemical modification, a feature that is useful in the creation of heterodimers. Also chosen for investigation were the single site putative agonists **#4**, a short peptide that contains a site 1 binding motif, and peptide **#5**, a short peptide that contains a site 2 binding motif (Table 2).

In addition, we also investigated longer peptides that contain two binding site motifs, peptides **#7-9**. Peptide **#7**, known as S597 in the literature[6, 7], contains a site 2 motif and a site 1 motif (Table 2), and was shown in the literature to bind to the insulin receptor isoforms[6]. Peptide **#8**, referred to in the literature as S961[3], is composed of a binding site 1 motif and a binding site 2 motif connected by the flexible linker GGSGGS (Table 2). We investigated a variant of **#8**, peptide **#9** (Table 2). Peptide **#9**, also known as S661[3], is identical in sequence to **#8** but contains a C-terminal amide.

To investigate the significance of the disulfide bond within the site 2 motifs of these peptides, oxidized and reduced forms of **#9** were tested. In addition, mutant versions of

#7 and #9 were synthesized, where the two cysteine residues were mutated to serines. These mutants are #10 and #11 (Table 2).

To control for any steric effects or non-specific binding to the insulin receptor, we also investigated the first 31 residues of the insulin C-peptide, #12, a 33 residue peptide that does not bind to the insulin receptor[8].

Once we had characterized all of these peptides, we explored whether additional amino acids such as cysteine and lysine could be incorporated into these peptides without altering their antagonistic properties. As stated previously, the goal of this work is to create heterodimers of insulin receptor agonists with insulin receptor antagonists.

Therefore, we anticipated the need to create a covalent bond between our synthesized antagonists and subsequent agonists. One highly efficient way to promote the formation of a covalent bond is to exploit the unique reactivities of amino acids such as cysteine and lysine. Cysteine and lysine contain nucleophilic functional groups in their side chains, which can act as chemical “handles” and facilitate covalent linkages. In addition, lysine is also a cleavage site for the enzyme lys-C, a property that will be important in our biosynthetic studies, discussed in chapter 3. Therefore we created peptides with additional cysteine and lysine residues.

This research identified the #4-6 peptide as the most likely candidate for inclusion in a heterodimer. Therefore, to further explore the relationship between structure and activity in this peptide, we conducted N-terminal alanine scan and synthesized several truncation peptides (Table 2).

Finally, we investigated a 20 kDa PEGylated version of **#4-6**. Polyethylene glycol, or PEG, is a polymer made of repeating ethylene oxide units, and can be used to increase the stability, aqueous solubility, and alter the pharmacokinetics of a peptide[9]. PEG provides stability to a peptide by protecting it from proteolytic degradation[9], and since PEG itself is very water soluble, it can increase the aqueous solubility of the peptide. In addition, PEG units can be many times larger than the peptide, which means that attaching PEG greatly increases the total size of the peptide. As peptide size increases, filtration by the kidneys decreases, and therefore adding PEG can extend the amount of time the peptide spends in circulation, dramatically altering the pharmacokinetics of the peptide[9]. One disadvantage of incorporating PEG is the possibility of destroying the activity of the peptide of interest by directly or indirectly interfering with receptor binding. In order to test if PEGylation would interfere with the antagonism of **#4-6**, a 20kDa PEG was ligated to the N-terminus of peptide **#4-6** by reacting the free peptide with an N-hydroxysuccinimide (NHS) ester of a 20 kDa PEG molecule.

In summary, the research presented in this chapter demonstrates our ability to synthesize, purify, and modify peptides that are reported to be agonists and antagonists of the insulin receptor isoforms. In addition, structure activity analyses allowed us to further optimize these peptides, introduce useful chemical reactivity with additional cysteine and lysine residues, and modify an antagonist with PEG in order to potentially alter its pharmacokinetics.

Table 2: Sequence of Peptides

Reference Number	Literature Reference Name	Sequence	Synthetic Chemistry	C-terminus
1	-	DYKDLQSWGVRIWLAGLCPKK	Boc	Amide
2	-	KVRGFQGGTVWPGYEWLRNAAKK	Boc	Amide
3	-	KSMFVAGSDRWPGYGVLDWLKK	Boc	Amide
4	-	GSLDESFYDWFERQLG	Boc	Amide
Cys-4	-	CGSLDESFYDWFERQLG	Boc	Amide
4-Cys	-	GSLDESFYDWFERQLGC	Boc	Amide
5	-	SLEEEWAQIQCEVWGRGCPSTY	Boc	Amide
6	-	SLEEEWAQIQSEVWGRGSPSY	Boc	Amide
Cys-6	-	CSLEEEWAQIQSEVWGRGSPSY	Boc	Amide
6-Cys	-	SLEEEWAQIQSEVWGRGSPSYC	Boc	Amide
7	S597	SLEEEWAQIECEVYGRGCPSESFYD WFERQL	Boc	Amide
8	S961	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQCEVWGRGCPSTY	Fmoc	Acid
9	S661	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQCEVWGRGCPSTY	Boc	Amide
10	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
11	-	SLEEEWAQIESEVYGRGSPSESFYD WFERQL	Boc	Amide
12	C Peptide	RREAEDLQVGQVELGGGPGAGSLQ PLALEGSLQ	Boc	Amide
4-6	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
Cys-4-6	-	CGSLDESFYDWFERQLGGGSGGSSLE EEEWAQIQSEVWGRGSPSY	Boc	Amide
4-6-Cys	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSYC	Boc	Amide

4-6-Lys	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSYK	Boc	Amide
4(des1-4)-6	-	ESFYDWFERQLGGGSGGSSLEEEW AQIQSEVWGRGSPSY	Boc	Amide
4(des1-8)-6	-	DWFERQLGGGSGGSSLEEEWAQIQ SEVWGRGSPSY	Boc	Amide
4(des1-12)-6	-	RQLGGGSGGSSLEEEWAQIQSEVW GRGSPSY	Boc	Amide
4(A1)-6-Lys	-	ASLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
4(A2)-6-Lys	-	GALDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
4(A3)-6-Lys	-	GSADESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
4(A4)-6-Lys	-	GALDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVWGRGSPSY	Boc	Amide
4-6(des28-43)	-	GSLDESFYDWFERQLGGGSGGSSLE EE	Boc	Amide
4-6(des31-43)	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQI	Boc	Amide
4-6(des36-43)	-	GSLDESFYDWFERQLGGGSGGSSLE EEWAQIQSEVW	Boc	Amide
6-4	-	SLEEEWAQIQSEVWGRGSPSYGGS GGSGSLDESFYDWFERQLG	Boc	Amide
Cys-6-4	-	CSLEEEWAQIQSEVWGRGSPSYGGS GGSGSLDESFYDWFERQLG	Boc	Amide
6-4-Cys	-	SLEEEWAQIQSEVWGRGSPSYGGS GGSGSLDESFYDWFERQLGC	Boc	Amide
6-4-Lys	-	SLEEEWAQIQSEVWGRGSPSYGGS GGSGSLDESFYDWFERQLGK	Boc	Amide
-	DP8	GEEEEKGP EHL CGAHLVDALYLV CGDRGFYGYGSSRRAPQTGIVDEC CHRSCDLRRLNYCN	Bio- synthesized in <i>E. coli</i>	Acid
-	Insulin	GIVEQCCTSICSLYQLENYCN FVNQHLCGSHLVEALYLVCGERGF FYTPKT	Commerical ly available from Eli Lilly & Co.	Acid

RESULTS

Peptides sequences **#1-5** were tested for agonism at the insulin receptor A and B isoforms. These peptides did not display any agonism (solid lines) or any antagonism (dashed lines) at either receptor isoform (Figure 1, Figure 2).

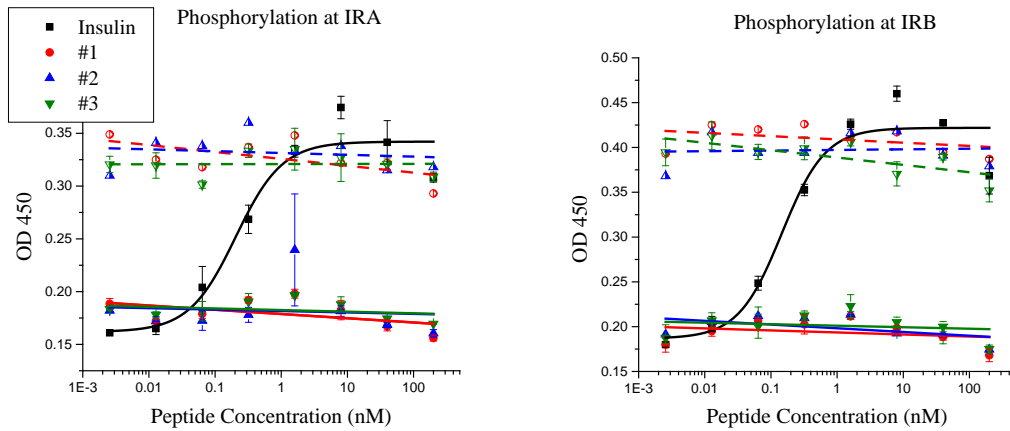


Figure 1: Phosphorylation assay demonstrating the lack of agonism (solid) and antagonism (dashed) of peptides #1-3.

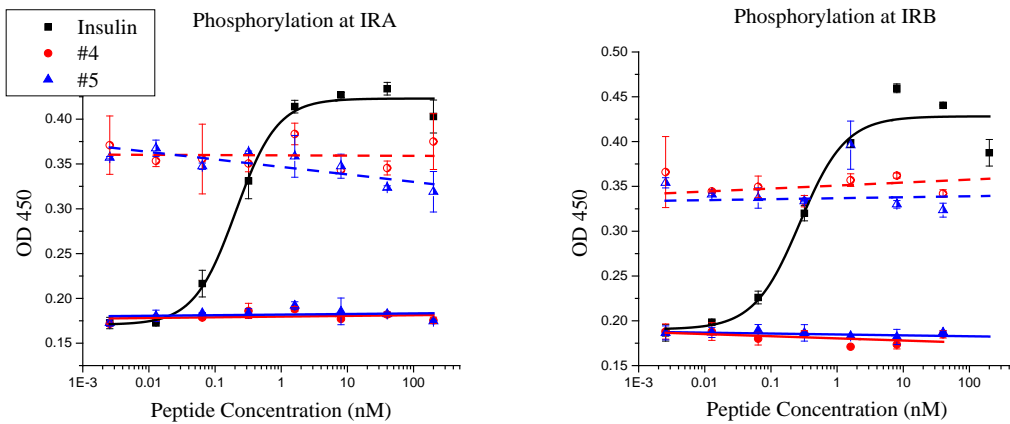


Figure 2: Phosphorylation assay demonstrating the lack of agonism (solid) and antagonism (dashed) of peptides #4 and #5.

Peptide #7 was tested for agonism at both the IRA and IRB isoforms, but did not display any significant agonism at the tested concentrations (Figure 3). In addition, the #10 mutant, in which the cysteines were replaced with serines, was also analyzed, and it was found that there was no significant change in agonism, relative to the disulfide-containing peptide (Figure 3). In light of the low activity at both receptor isoforms, peptides #7 and #10 were tested for antagonism at both receptor isoforms. Despite being reported as agonists, both #7 and #10 displayed slight antagonism at both receptor isoforms (Figure 4).

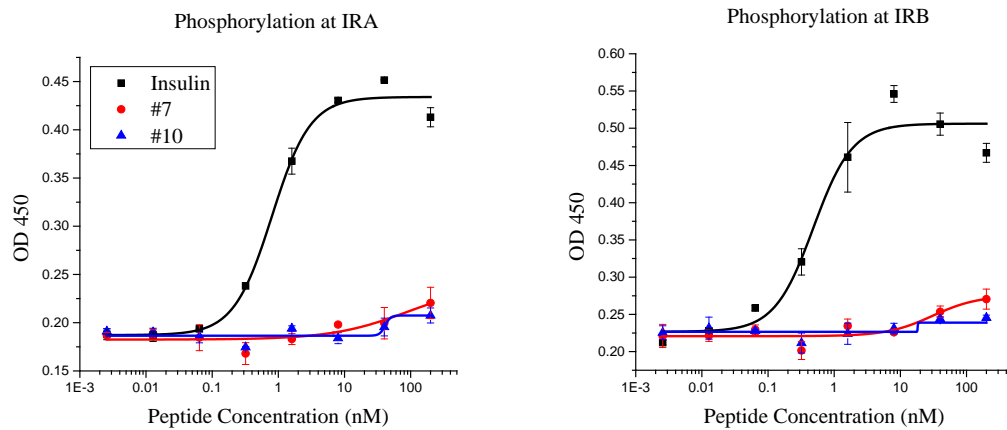


Figure 3: Phosphorylation assay demonstrating the agonism of peptides #7 and #10.

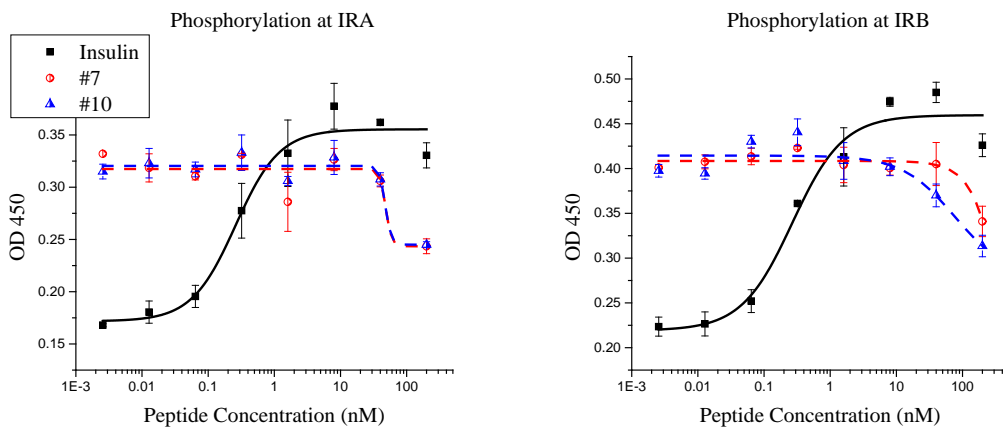


Figure 4: Phosphorylation assay demonstrating the antagonism of peptides #7 and #10.

Peptide #9, which contains a site 1 and site 2 motif, was tested for agonism and antagonism at the IRA and IRB isoforms, in both its oxidized form, where the site 2 motif contains a disulfide bond (Figure 5), and its reduced form, where the site 2 motif cysteines are present as free sulfhydryls (Figure 6). It was found that in its oxidized form, peptide #9 was able to completely antagonize 1 nM native insulin with IC₅₀ values of 7.25 nM at the IRA isoform and 3.99 at the IRB isoform (Figure 5). In its reduced form, peptide #9 was not an agonist at either receptor isoform, but was able to completely antagonize 1 nM native insulin to basal levels of phosphorylation, with IC₅₀ values of 2.75 nM at the IRA isoform and 5.08 nM at the IRB isoform (Figure 6). In addition, the mutant #11 was tested, where the site 2 motif cysteines were mutated to serines (Table 2). To illustrate the identical behavior of all three peptides, #9 (oxidized), #9 (reduced), and #11 were tested in the same phosphorylation assay. It was found that each peptide was capable of antagonizing native insulin to basal levels at both receptor isoforms (Figure 7), all with similar IC₅₀ values (Table 3).

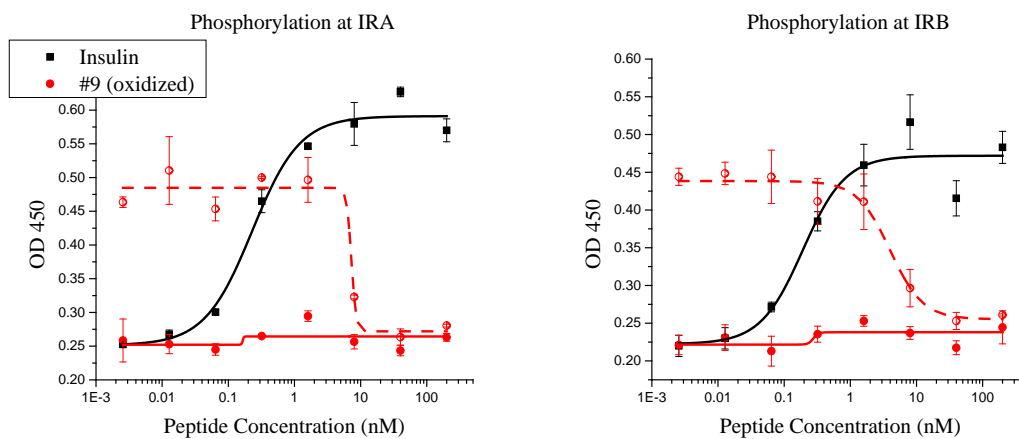


Figure 5: Phosphorylation assay demonstrating the agonism (solid) and antagonism (dashed) of peptide #9, in its oxidized form.

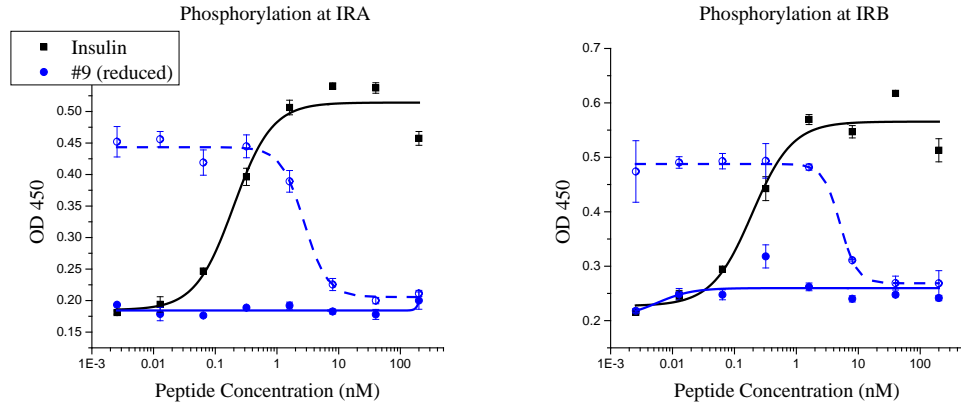


Figure 6: Phosphorylation assay demonstrating the agonism (solid) and antagonism (dashed) of peptide #9, in its reduced form.

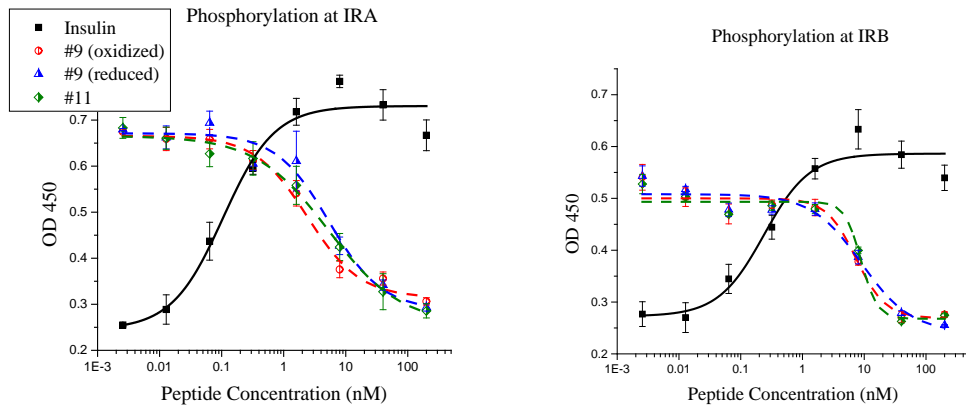


Figure 7: Phosphorylation assay demonstrating the antagonism of peptides #9 (oxidized), #9 (reduced), and #11. IC₅₀ values listed in Table 3.

Table 3: IC₅₀ values of peptides #9, 11.

Reference Number	IC ₅₀ (nM)	
	IRA	IRB
9 (oxidized)	2.41	7.35
9 (reduced)	5.38	9.04
11	5.27	8.87

As the characterization of peptides **#7-11** has shown, the disulfide bond in the site 2 motif, previously reported as necessary for antagonism, is not necessary for antagonism in our *in vitro* assay. Therefore, we created two optimized antagonists, each containing a site 1 motif, peptide **#4**, and a site 2 motif, peptide **#6**. Peptide **#6** is a variant of the **#5** motif, with the cysteine residues mutated to serines. To investigate whether the relative orientation of the binding motifs affected agonism and antagonism, we created an antagonist where the binding motifs were arranged site 1 to site 2 in an N- to C-terminal direction (**#4-6**), and an antagonist where the binding motifs were arranged site 2 to site 1 in an N- to C-terminal direction (**#6-4**) (Table 2). In both antagonists, the binding site motifs are connected by the flexible linker GGSGGS. These peptides were tested for agonism and antagonism at both receptor isoforms, and it was found that neither displayed any agonism, but both were capable of antagonizing native insulin (Figure 8), with similar IC_{50} values (Table 4). In anticipation that our biosynthetic heterodimers will include the insulin receptor agonist DP8 (Table 2), peptides **#4-6** and **#6-4** were also tested for antagonism against 1 nM DP8 at the IRA isoform and were found to have IC_{50} values of 5.33 nM and 12.33 nM, respectively (Figure 9).

To ensure that none of the antagonism observed was the result of non-specific binding or inhibition, the insulin C-peptide **#12** was also analyzed for agonism and antagonism at the insulin receptors. It was found that peptide **#12** does not display any agonism or any antagonism at either receptor isoform (Figure 10).

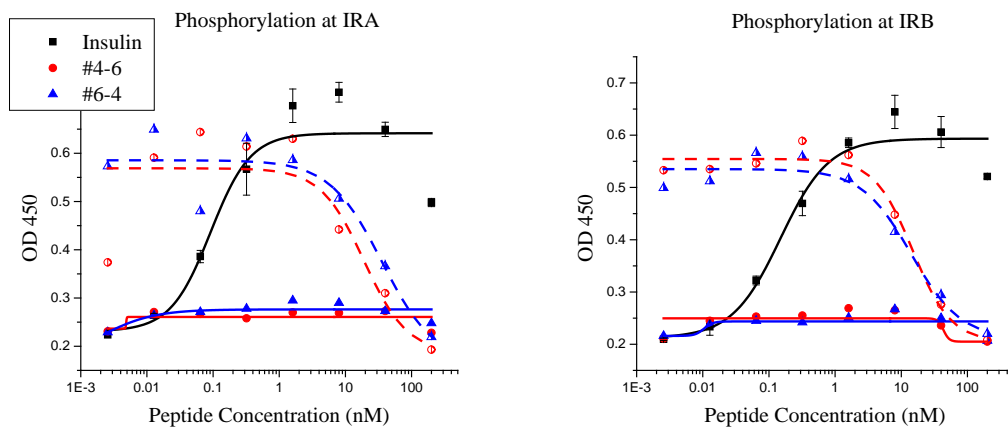


Figure 8: Phosphorylation assay demonstrating the lack of agonism (solid) and antagonism (dashed) of peptides #4-6 and #6-4. IC₅₀ values listed in Table 4.

Table 4: IC₅₀ values of peptides #4-6 and #6-4

Reference Number	IC ₅₀ (nM)	
	IRA	IRB
#4-6	19.1	14.9
#6-4	35.8	14.2

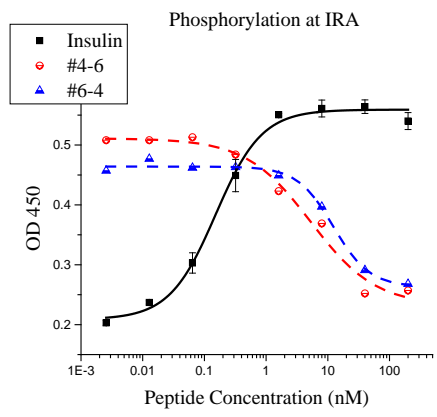


Figure 9: Phosphorylation assay demonstrating the ability of peptides #4-6 and #6-4 to antagonize 1 nM DP8 at the IRA isoform. Peptide #4-6 displays an IC₅₀ of 5.33 nM and peptide #6-4 displays an IC₅₀ of 12.34 nM.

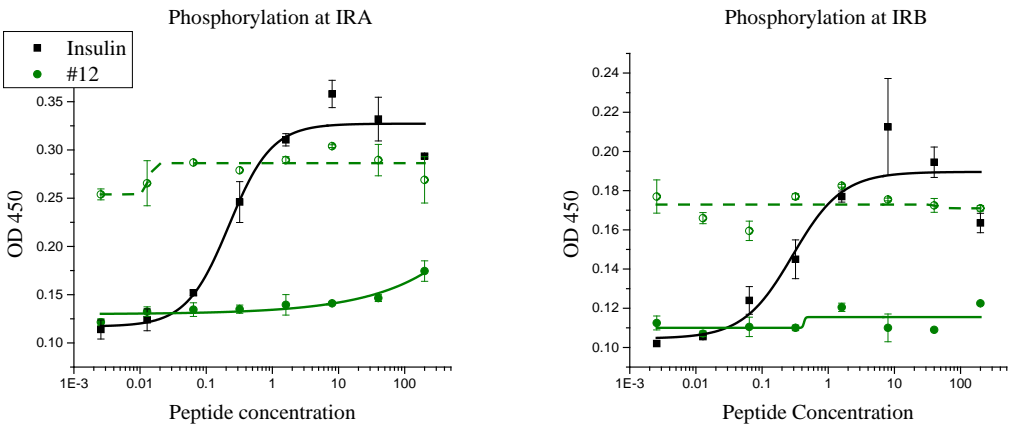


Figure 10: Phosphorylation assay demonstrating the inability of peptide #12 to agonize or antagonize native insulin at both insulin receptor isoforms.

Introduction of Additional Residues for Chemical Conjugation

The next step in this research is to incorporate unique chemical functionality into our peptides. This was accomplished through the addition of cysteine and lysine residues, since the nucleophilic side chains of these amino acids can be exploited as conjugation sites. Single site peptides **#4** and **#6** were synthesized with a cysteine at either the N- or C-terminus. These are referred to as **#Cys-4**, **#4-Cys**, **#Cys-6**, **#6-Cys** (Table 2). These peptides were found to have no agonism at either insulin receptor, and very little antagonism at either insulin receptor (Figure 11). We also investigated whether a C-terminal or N-terminal cysteine affected the antagonism of the **#4-6** and **#6-4** antagonists. It was found that the addition of cysteines at either the N- or C-terminus of these peptides did not significantly influence their antagonism (Figure 12, Figure 13).

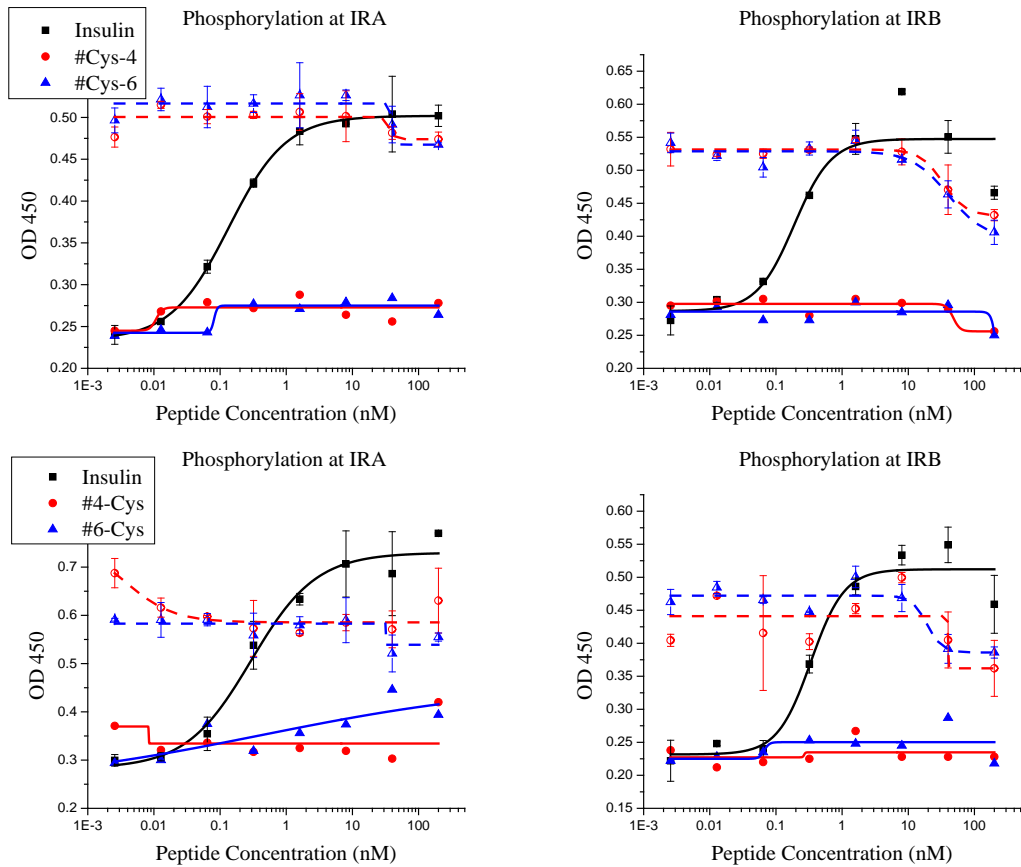


Figure 11: Phosphorylation assay demonstrating the agonism (solid) and antagonism (dashed) of peptides #4 and #6, with cysteine placement variations. None of the peptides displayed any agonism at either receptor isoform.

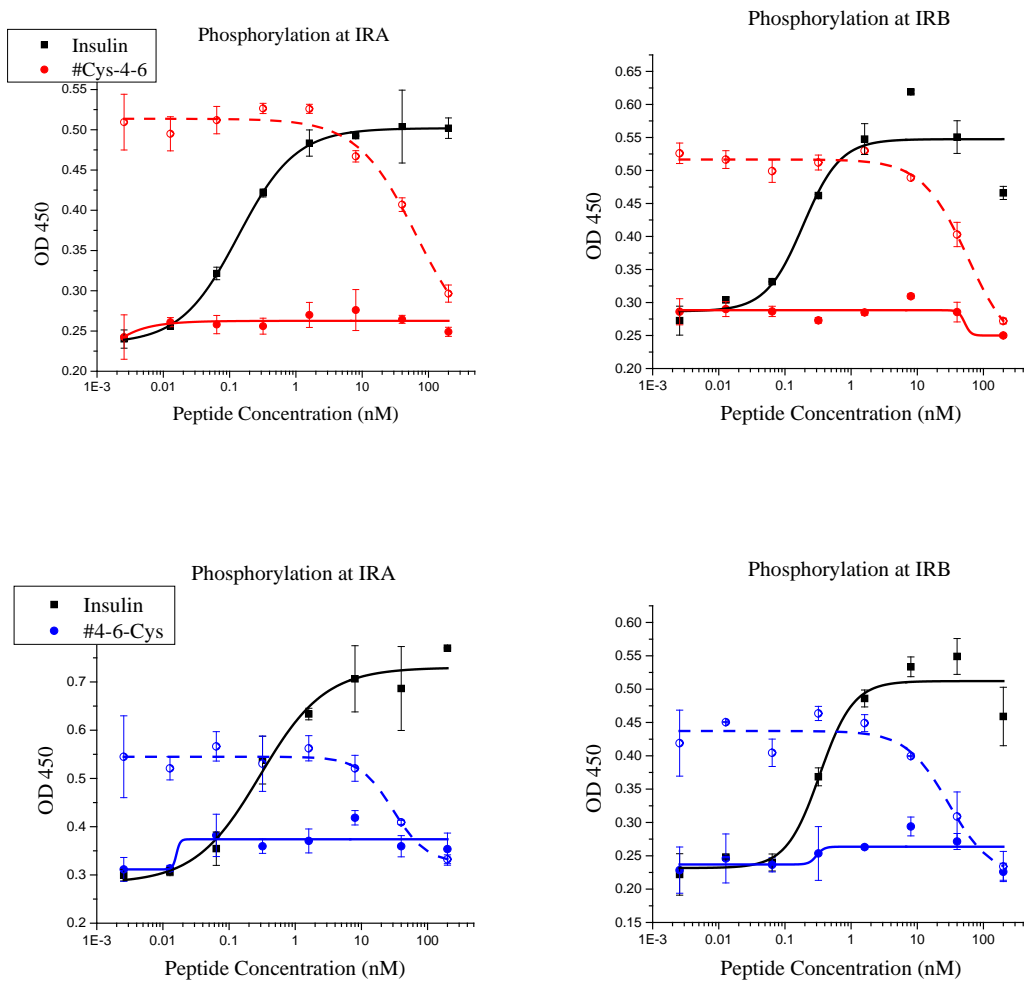


Figure 12: Phosphorylation assay demonstrating the lack of agonism (solid) and the antagonism (dashed) of peptide #4-6 cysteine variants. Peptide #Cys-4-6 displays an IC_{50} of 65.73 nM (IRA) and 57.57 (IRB). Peptide #4-6-Cys displays an IC_{50} of 30.11 nM (IRA) and 29.90 (IRB).

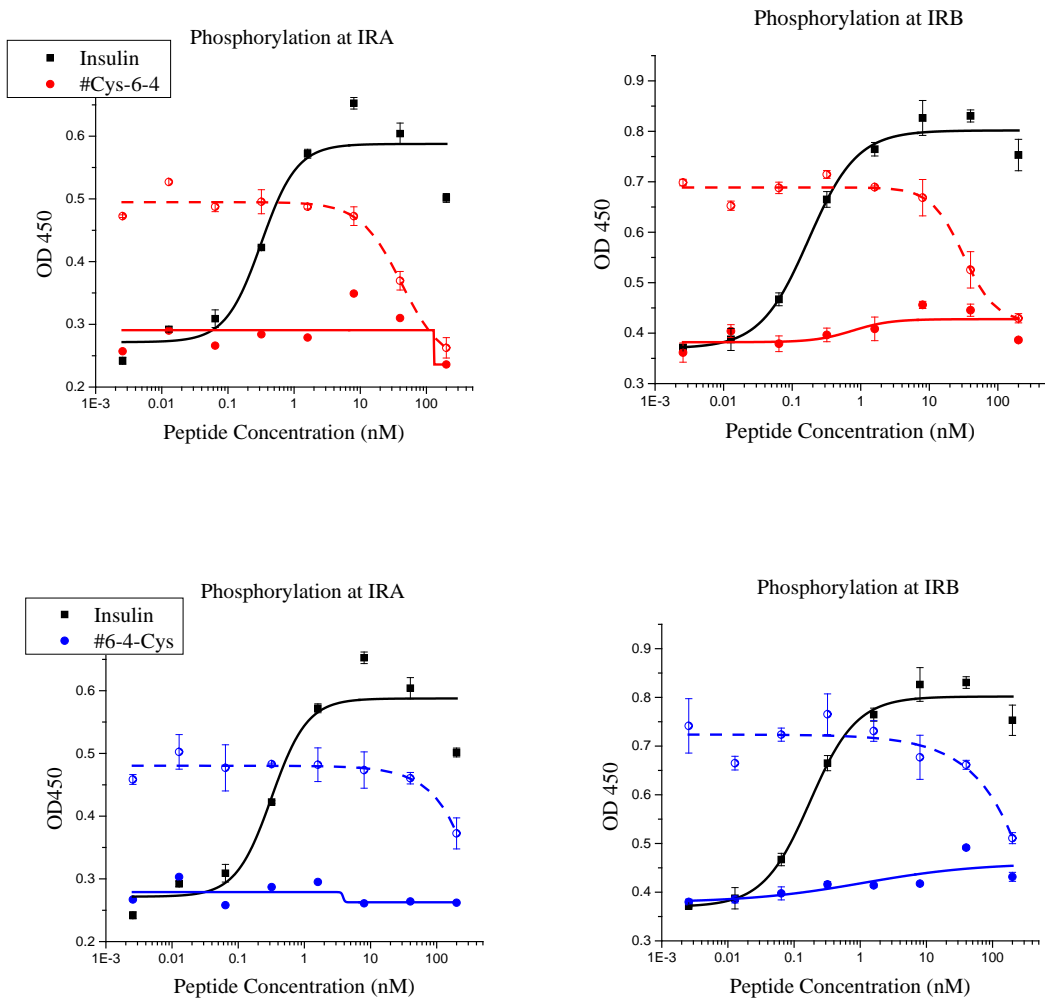


Figure 13: Phosphorylation assay demonstrating the lack of agonism (solid) and the antagonism (dashed) of peptide #6-4 cysteine variants.

The next chemical handle investigated was lysine. Anticipating that subsequent biosynthetic molecules would contain lysine at the C-terminal position of the **#4-6** and **#6-4** antagonists, we investigated whether this modification would affect the antagonism of these peptides as monomers. It was found that the C-terminal lysine had no effect on the ability of the **#4-6** and **#6-4** peptides to antagonize native insulin (Figure 14). In addition, we further investigated the **#4-6-Lys** antagonist by creating truncated versions, where the first 4, 8, and 12 N-terminal residues were deleted from the **#4** peptide (Table 2). These peptides were tested for agonism and antagonism at the insulin receptor isoforms (Figure 15). None of the peptides displayed any agonism (Figure 15). It was found that none of the truncated antagonists were able to antagonize native insulin at the concentrations tested (Figure 15). Noting that the loss of even the first four N-terminal residues of the **#4-6-Lys** peptide resulted in a loss of antagonistic activity, an alanine scan of the first four residues was conducted (Table 2). These peptides were tested for antagonism at the insulin receptors, and it was found that all of the alanine scan peptides were able to antagonize native insulin (Figure 16), with similar IC_{50} values (Table 5).

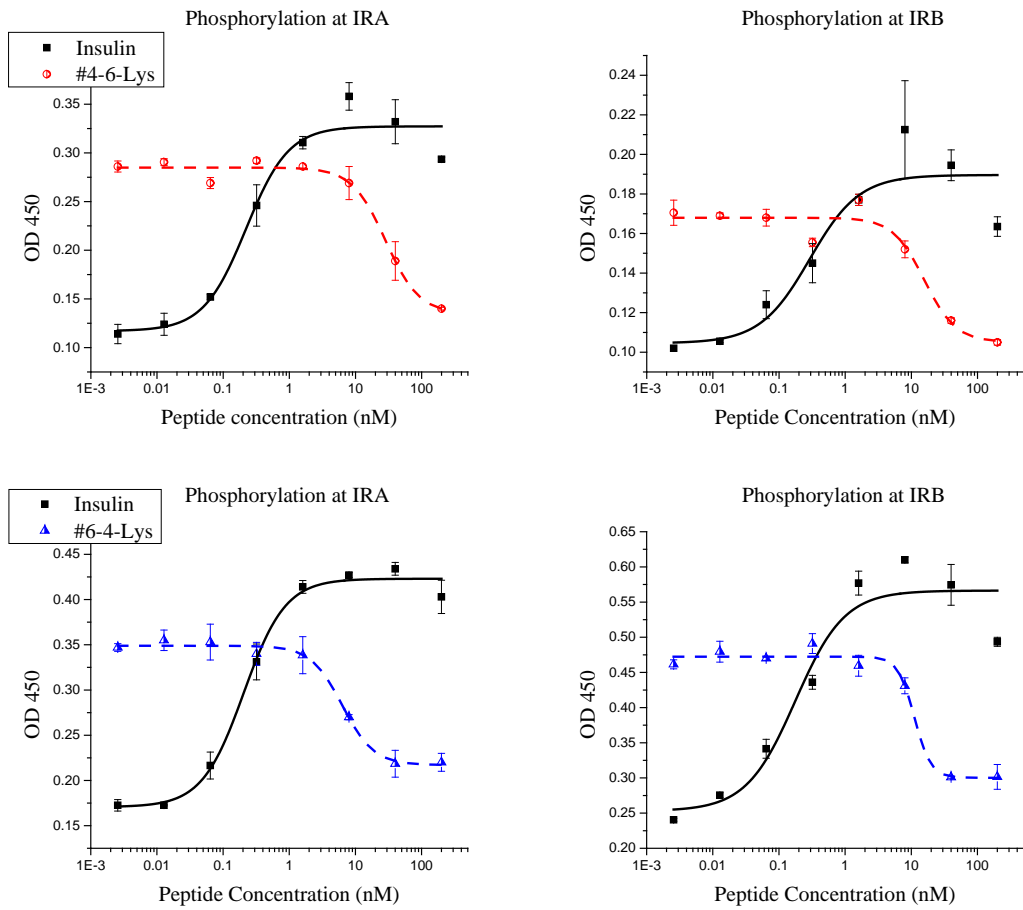


Figure 14: Phosphorylation assay demonstrating the antagonism of the peptides #4-6-Lys and #6-4-Lys.

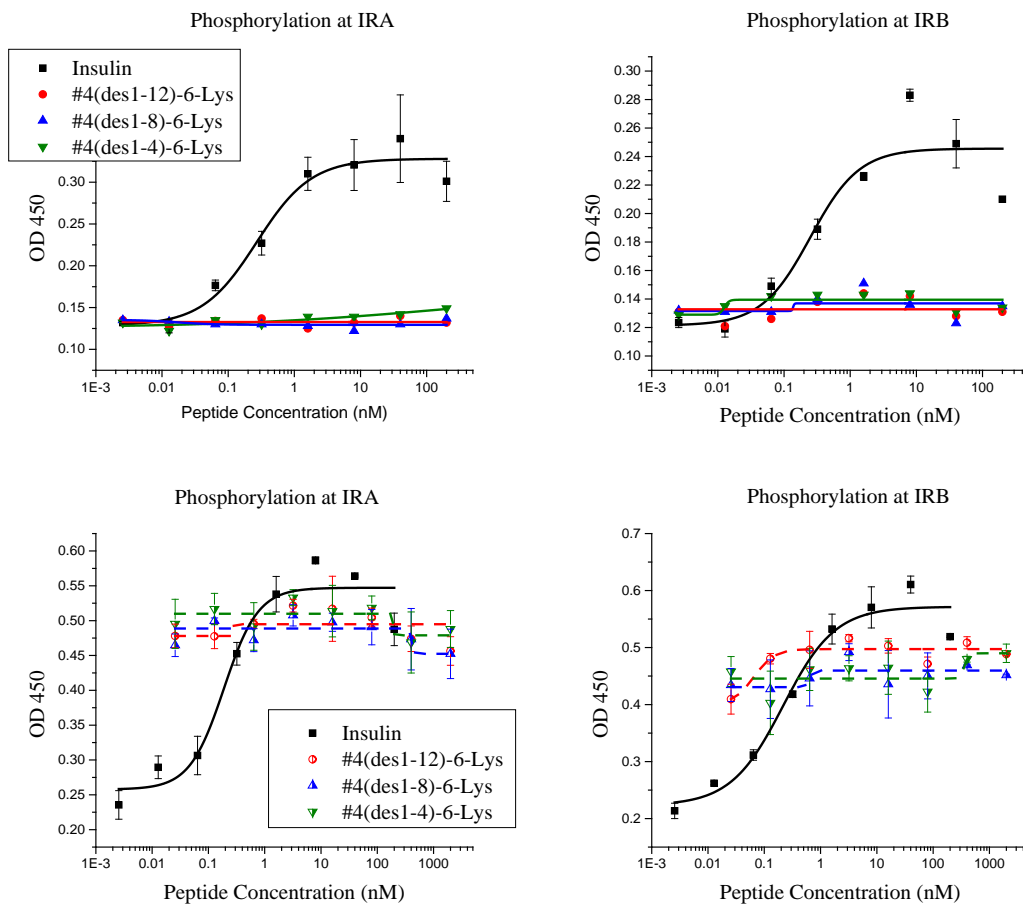


Figure 15: Phosphorylation assay demonstrating the agonism (top) and antagonism (bottom) of N-terminal shortened antagonists. None of the shortened antagonists displayed any agonism or antagonism at either receptor isoform.

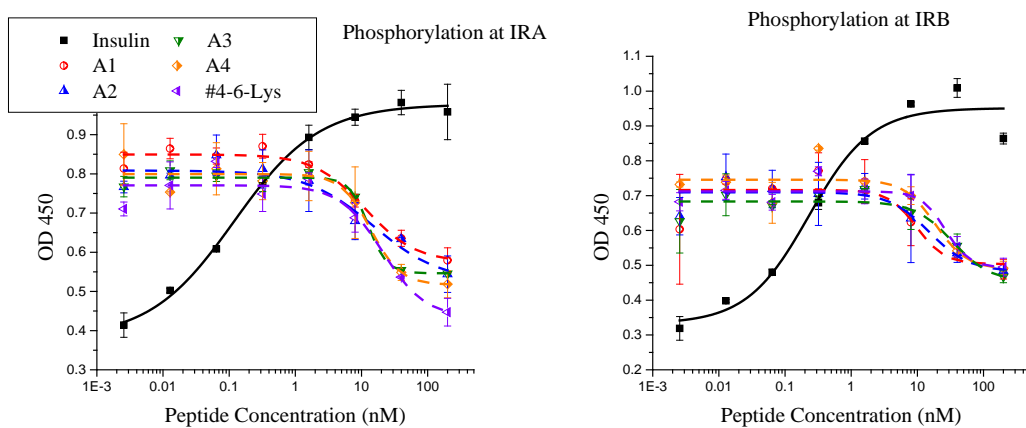


Figure 16: Phosphorylation assay demonstrating the antagonism of #4(A X)-6-Lys peptides, where X is the position of the alanine mutation.

Table 5: IC₅₀ values of the #4-6-Lys N-terminal alanine scan.

Reference Number	IC ₅₀ (nM)	
	IRA	IRB
4(A1)-6-Lys	11.3	9.74
4(A2)-6-Lys	15.2	15.1
4(A3)-6-Lys	12.7	34.1
4(A4)-6-Lys	13.8	18.8
4-6-Lys	20.7	26.4

The importance of the C-terminal residues was also explored. Peptides were synthesized composed of the first 27, 31, and 36 residues of the #4-6 antagonist (Table 2), and these peptides were tested for antagonism at the insulin receptor isoforms (Figure 17). Both the full length #4-6-Lys and the truncated #4-6(des37-43) peptides were able to antagonize native insulin at both receptor isoforms with equal potency. Peptides #4-6(des28-43) and #4-6(des32-43) appeared to have some antagonistic activity, although only at the IRB isoform and at a higher concentration than either #4-6(des37-43) or #4-6-Lys.

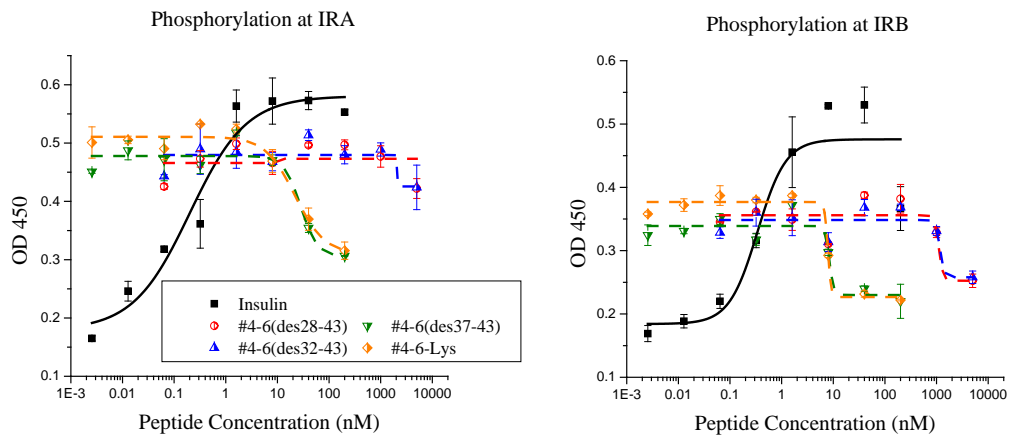


Figure 17: Phosphorylation assay demonstrating the antagonism of C-terminal shortened antagonists.

Table 6: IC₅₀ values of C-terminal shortened #4-6-Lys antagonists.

Reference Number	IC ₅₀ (nM)	
	IRA	IRB
#4-6(des28-43)	-	1110
#4-6(des32-43)	-	1160
#4-6(des37-43)	28.1	8.40
#4-6-Lys	21.8	7.86

PEGylation of Peptide #4-6

PEGylated #4-6 was created by covalently linking a 20 kDa PEG molecule to the N-terminus of the peptide. The PEGylated and unPEGylated peptides were assayed for antagonism at both insulin receptor isoforms (Figure 18). It was found that PEGylation resulted in an approximately 5 fold reduction in potency (Table 7).

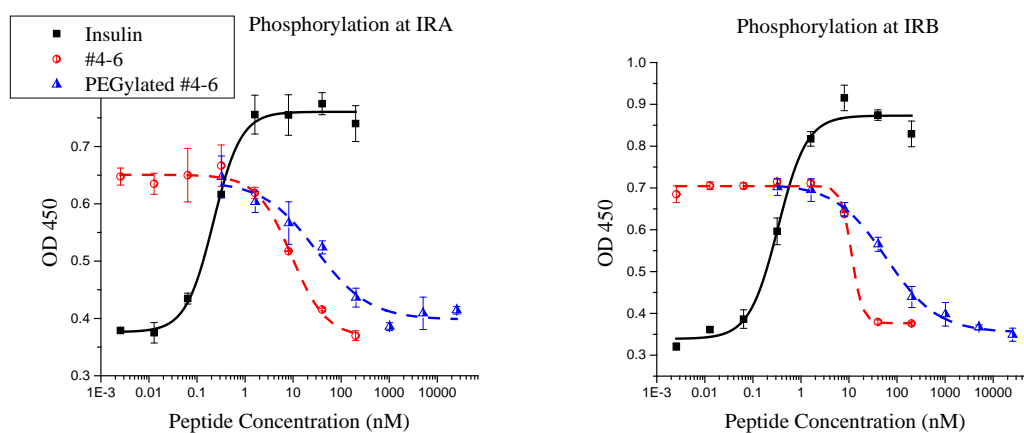


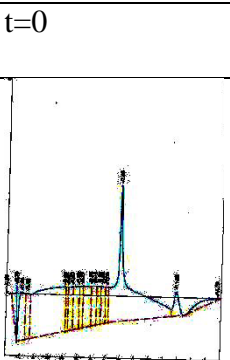
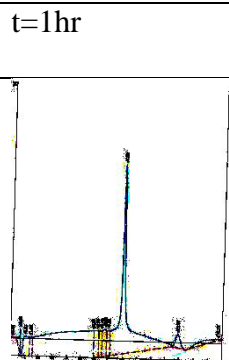

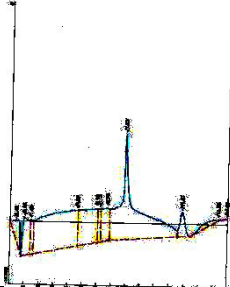
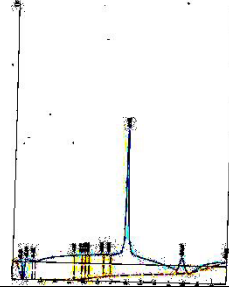
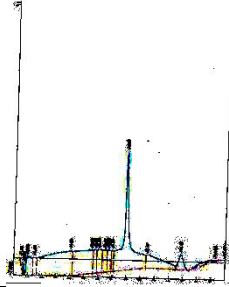
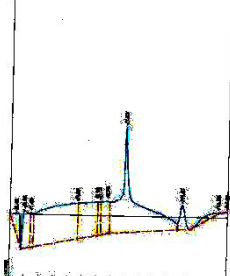
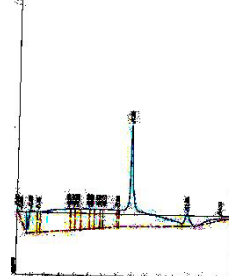
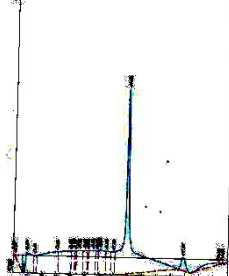
Figure 18: Phosphorylation assay demonstrating the antagonism of peptide #4-6 peptide and PEGylated peptide #4-6.

Table 7: IC₅₀ values of peptide #4-6 and PEGylated peptide #4-6.

Peptide	IC ₅₀ (nM) at IRA	IC ₅₀ (nM) at IRB
4-6	9.30	11.4
PEGylated 4-6	30.3	59.9

In order to demonstrate that the PEG modification was stable, analytical HPLC was used to monitor the peptide in ammonium bicarbonate at room temperature, phosphate buffered saline (PBS) at room temperature, and PBS at 37°C. Since PEG does not absorb light at 214 nm, but does dramatically affect the hydrophobicity of the modified peptide, the degradation of the peptide would result in two peaks, one for the PEGylated peptide and one for the unmodified peptide. Since this was not observed, we concluded that the PEGylated molecule is stable at room temperature and at 37°C (Table 8).

Table 8: Analytical HPLC traces demonstrating the stability of the PEGylated peptide #4-6.

Buffer Conditions	t=0	t=1hr	t=12 hrs
25 mM ammonium bicarbonate, pH 8, room temp.			
PBS, pH 7, room temp.			
PBS, pH 7, 37°C			

DISCUSSION

The peptides containing a single binding site motif, **#1-5**, were not capable of agonism at the insulin receptor isoforms, and were incapable of antagonizing 1 nM native insulin, and thus are unlikely candidates for inclusion in an agonist/antagonist heterodimer. The ideal antagonist would be capable of antagonizing insulin at a 1 to 1 molar ratio, however, the single binding motif peptides were incapable of antagonizing native insulin at even a 200:1 ratio. This may be the result of the low IR binding affinities of single motif containing peptides. However, peptides **#7** and **#9**, each containing a site 1 motif and a site 2 motif, were capable of antagonizing native insulin, demonstrating the importance of multivalency in ligand/receptor interactions.

Peptide **#7** was also compared to its mutant, **#10**, where the two cysteine residues were mutated to serine residues. Peptide **#10** did not display any agonism but was capable of antagonizing native insulin. These results demonstrate that the disulfide bond, previously reported as a necessary component of site 2 binding motifs, is not necessary for antagonism in our *in vitro* assay. This observation was repeated with peptide **#9**. The oxidized form of **#9**, in which a disulfide bond exists between the cysteines in the site 2 motif, did not display any agonism at the insulin receptors, but was capable of fully antagonizing native insulin. The reduced form of this peptide, where the cysteines were present as free sulfhydryls, was also able to antagonize native insulin. In addition, mutating the two cysteine in **#9** to serines, resulting in **#11**, did not affect the ability of **#11** to antagonize native insulin.

While it was possible that the antagonism we observed was the result of non-specific binding to the insulin receptor, we eliminated this concern by demonstrating that the C-peptide, #12, similar in size to the antagonists and known not to bind to the insulin receptor, was incapable of antagonizing native insulin.

The observation that the disulfide bond is not necessary for antagonism is in contrast to the conclusions drawn by Schaffer, et al[1, 3, 4]. However, the disulfide bond motif was discovered as the result of a random peptide phage display screening. It is possible that within the context of the screening process, a disulfide provided stability or affinity, but that the disulfide is not necessary for *in vitro* antagonism. Whatever advantage or selection bias the disulfide presented in the screening process is lost in our *in vitro* analysis. The most important aspect of this discovery is that it allows additional cysteine residues to be used as orthogonal chemical handles for chemical conjugation.

Having successfully shown that the disulfide bond is not necessary for antagonism, but that both a site 1 and site 2 motif are necessary, we created peptides #4-6 and #6-4, which contain a site 1 motif and the mutated site 2 motif. These peptides do not contain any cysteine residues. We have shown that #4-6 and #6-4 are both capable of antagonizing native insulin. These optimized peptides are good candidates for inclusion into agonist/antagonist heterodimers. While the most ideal antagonist would be capable of antagonizing native insulin in a 1:1 ratio, peptides #4-6 and #6-4 are effective in a 10:1 ratio. However, we anticipate that within the context of a covalently linked heterodimer, these antagonists will be more effective since they will be recruited to the insulin receptor by the binding of the insulin receptor agonist.

Having created effective, cysteine-free antagonists we then created variations of the **#4**, **#6**, **#4-6** and **#6-4** peptides, to show that the inclusion of a non-native cysteine at either terminus did not affect their agonistic or antagonistic properties. It was found that the inclusion of cysteine residues at either terminus did not affect their activity. Another potentially useful modification, lysine, was also investigated. Anticipating that subsequent biosynthetic heterodimers would include a C-terminal lysine residue, peptides **#4-6-Lys** and **#6-4-Lys** were investigated. Both peptides continued to be effective at antagonizing native insulin at both receptor isoforms, thus confirming their potential to act as antagonists in a biosynthetic heterodimer.

It was also investigated whether peptides **#4-6-Lys** and **#6-4-Lys** could be further optimized by deletion studies. It was demonstrated that within the **#4-6-Lys** peptide, the first four residues of the **#4** site 1 motif peptide are crucial for antagonism, although not in a residue specific manner. It may be that the first four amino acids help stabilize the secondary structure of the peptide. It may also be possible that the full length peptide presents the N-terminal amine in the correct position, relative to the rest of the peptide, for antagonism.

Deletion studies from the C-terminus of the **#6** site 2 motif demonstrated that only the first 36 residues are necessary for full antagonism. This may be an important for atom economy and optimization in subsequent heterodimers.

PEGylation of the #4-6 antagonist demonstrated that an N-terminally linked PEG molecule does not dramatically affect antagonism, and therefore demonstrated that PEGylated #4-6 antagonist is viable candidate for an extended-action antagonist.

The next chapter of this research will describe the creation of heterodimers through biosynthetic means. These heterodimers will consist of an insulin receptor agonist and several of the peptides characterized and optimized in this chapter.

METHODS

Historical Perspective

Whereas this research has focused on peptides as insulin receptor antagonists, peptides and proteins are ubiquitous in life, providing structural support, enzymatic activity, and hormonal signaling, and represent both therapeutic targets and therapeutic agents.

In both prokaryotic and eukaryotic systems, proteins and peptides are synthesized by ribosomes, which utilize tRNA to promote the formation of peptide bonds between the N- and C-terminus of amino acids[10]. This synthesis proceeds from the N- to C-terminus of the protein or peptide[10]. While prokaryotic systems such as *E. coli* can and have been exploited for eukaryotic peptide synthesis, this “biosynthesis” is limited to naturally occurring amino acids, and the resultant peptide must be purified from the cell lysate. Synthetic chemistries are also available for peptide synthesis. The most notable and widely used approach for synthetic peptide synthesis is referred to as solid phase peptide synthesis.

Solid phase peptide chemistry is a robust and practical methodology, but it represents the culmination of decades of research and refinement. In solid phase, the first amino acid of the peptide is anchored to an insoluble resin through its C-terminus. The next amino acid forms a peptide bond with the first, and amino acids are added sequentially until the synthesis is complete. While superficially simplistic, solid phase peptide synthesis requires four distinct and orthogonal chemistries. First, the N-terminus of each incoming amino acid requires a reversible protecting group. This protecting group ensures that incoming amino acids form a peptide bond with the nascent peptide chain and do not polymerize. Secondly, each incoming amino acid requires activation at the C-terminus, to promote the formation of the peptide bond. Third, any side chains with reactive functional groups must also be reversibly protected, to prevent interference with peptide bond formation. Finally, the resin itself requires chemical functionality such that once synthesis is complete, the peptide can be liberated from the solid support, and any remaining protecting groups removed.

All of these chemistries now exist, but their development has been a long and laborious process. Prior to their advent, the synthesis of even a dipeptide was exceedingly difficult. In 1882, Theodor Curtius published the first known peptide synthesis, which resulted in a benzoyl-glycylglycine dipeptide[11, 12]. It was not until 1901 that the first free dipeptide, glycyl-glycine, was synthesized by Emil Fischer, who went on to win the 1902 Nobel Prize in Chemistry[11]. The first major breakthrough in synthetic protein chemistry was the creation of an N-terminal reversible protecting group, the benzyloxycarbonyl group,

also called the Cbz or Z group (Figure 20)[11]. This protecting group was invented by Max Bergmann and his student Leonidas Zervas in 1922[11]. In 1957 a new reversible protecting group, *tert*-butyloxycarbonyl (Boc), was introduced by Carpino, McKay and Albertson[12]. This protecting group is acid-labile, and was used in combination with Cbz groups to synthesize peptides up to 39 residues in length[12] (Figure 21). Currently, when Boc is used as the N-terminal protecting group, the side chains are protected with benzyl-based groups (Table 9), which are removed at the end of the synthesis by treatment with anhydrous HF.

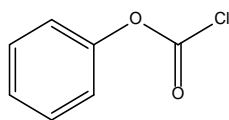


Figure 19: Benzyloxycarbonyl (Cbz) reversible protecting group.

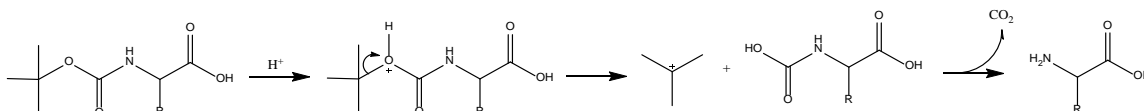


Figure 20: Acid promoted deprotection of *tert*-butyloxycarbonyl protected amino acid.

Another advance in the field of solid phase peptide chemistry was the introduction of 9-fluorenylmethyloxycarbonyl (Fmoc), a base-labile N-terminal protecting group, in 1970[13] (Figure 22). When Fmoc is used as the N-terminal protecting group, reactive side chains are protected with *t*-butyl based protecting groups (Table 9), which are acid labile and removed at the end of the synthesis by treatment with TFA.

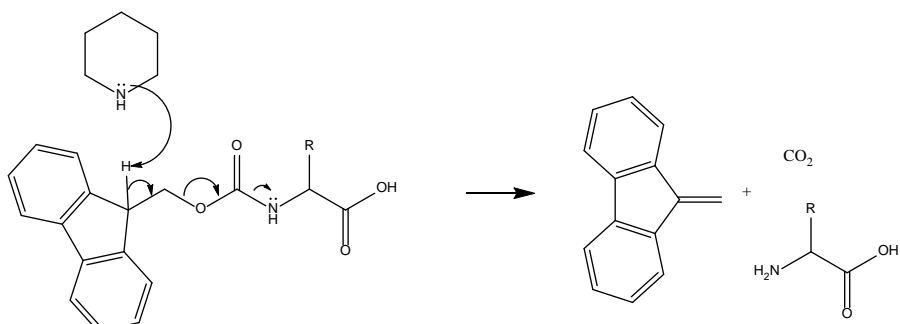


Figure 21: Deprotection of an Fmoc-protected amino acid. The Fmoc protecting group is deprotonated with a mild base, such as piperidine (shown here), which leads to an intramolecular rearrangement that generates carbon dioxide and the free amine.

In addition to N-terminal protecting groups, solid phase peptide synthesis requires coupling reagents that activate the C-terminus of the incoming amino acid. Historically, the first coupling reagents were carbodiimides. Carbodiimides react with the C-terminus of amino acids to yield a labile ester that is activated for nucleophilic attack by a free amine (Figure 23).

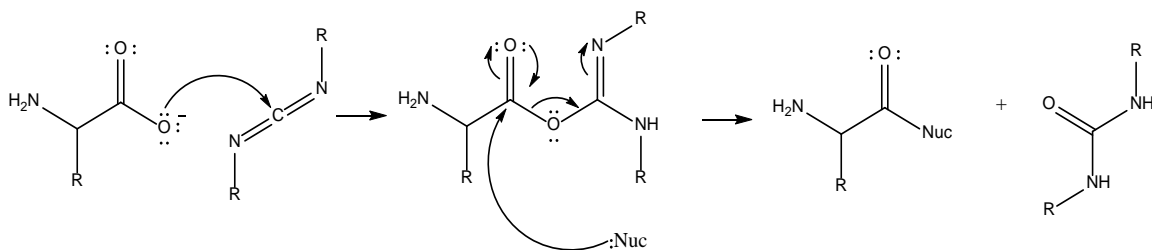


Figure 22: Carbodiimide activation of a deprotonated amino acid, followed by nucleophilic attack.

Carbodiimides have one key disadvantage: the intermediate O-acyl urea can rearrange to an inert N-acyl urea (Figure 24), which results in inefficient coupling. This disadvantage was overcome by the introduction of triazole-based coupling reagents, which react with the activated amino acid to form an ester (Figure 25)[12]. Many current solid phase peptide syntheses utilize a carbodiimide/triazole coupling scheme. In addition, there are some triazole based coupling reagents, such as 3-(Diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one, also known as DEPBT, which do not require carbodiimide activation (Figure 26).

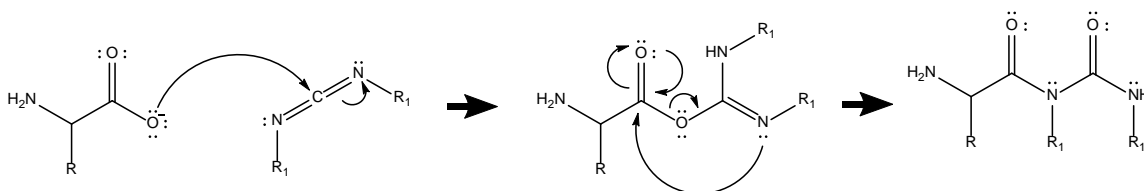


Figure 23: A deprotonated amino acid reacts with a carbodiimide coupling reagent, forming an activated O-acyl urea. This intermediate can undergo a rearrangement to produce an N-acyl urea, which is inactive.

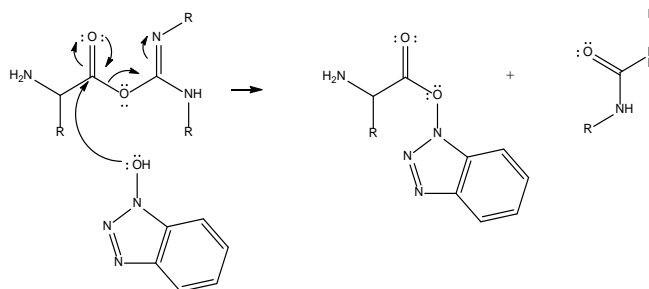


Figure 24: Reaction of a triazole based coupling reagent (HOBt) with a carbodiimide activated amino acid.

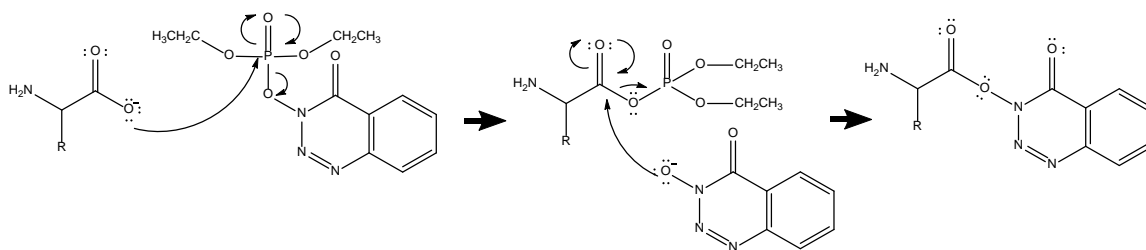


Figure 25: Reaction of DEPBT with deprotonated amino acid, resulting in a reactive ester.

Despite the introduction of a reversible protecting group and triazole-based coupling reagents, peptide synthesis remained a laborious process, due to the need to isolate products from excess reagents after every cycle, a process during which loss of product compounded. The field of peptide chemistry leapt forward in 1963, when Bruce Merrifield published his first paper introducing solid phase peptide synthesis[11, 12, 14]. Ultimately, his invention would lead to automated synthesis of large peptides with increased purity, an achievement worthy of the Nobel Prize in 1984[11].

In the Merrifield synthesis, the growing polypeptide is tethered at its C-terminus to an insoluble resin. This allows for a greater excess of reagents to drive the reaction, since excess reagents can be rinsed away while the resin is filtered and washed.

There are two possible ways to grow a peptide on a solid support – either by anchoring the peptide through the N- or through the C-terminus of the growing chain. Both methodologies were initially explored, anchoring through the N-terminus by the Merrifield research group[11], and anchoring through the C-terminus by Letsinger and Kornet[15]. Ultimately, anchoring through the C-terminus was abandoned, due to a base-promoted side reaction that results in racemized peptides (Figure 27).

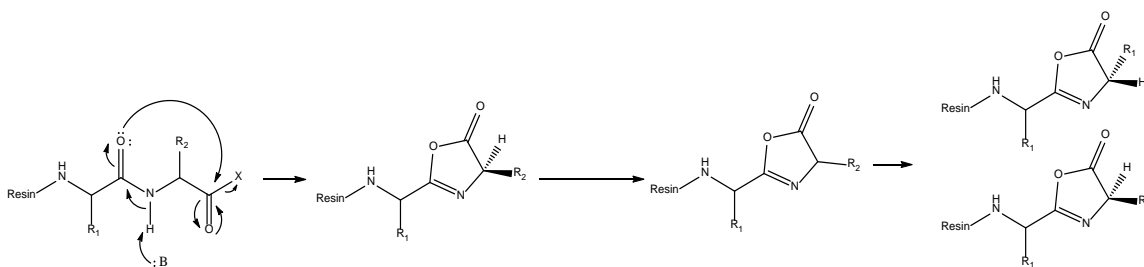


Figure 26: Base (B) promoted racemization due to intramolecular cyclization. This occurs exclusively in N- to C-terminal peptide synthesis.

Merrifield created the first insoluble resin appropriate for Boc solid phase synthesis, chloromethylpolystyrene, also referred to as the “Merrifield Resin” (Figure 28)[16]. Other Boc-compatible resins with improved stability and performance were quickly introduced. In 1974, Merrifield and Mitchell introduced the *tert*-butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl resin, the PAM resin (Figure 29), which is more stable under the acidic conditions required for Boc removal[17]. The next noteworthy resin to be developed was the benzydrylamine (BHA) resin[18], which was quickly followed by the methylbenzydrilamine (MBHA) resin[19]. These resins allow for the creation of C-terminal amide peptides, which is relevant since many biologically active peptides are present as C-terminal amides[18].

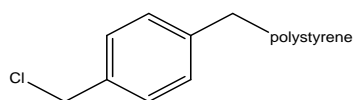


Figure 27: Chloromethylpolystyrene, the Merrifield resin.

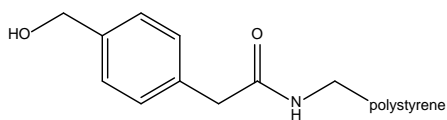


Figure 28: *tert*-butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl (PAM) resin.

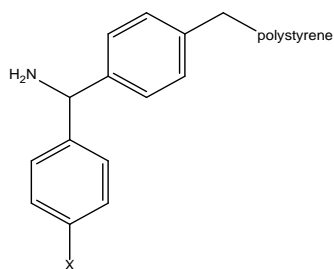


Figure 29: BHA (X=H) and MBHA (X=CH₃) resins.

Resins used in Fmoc solid phase peptide synthesis must be base-stable and acid-labile.

The first Fmoc compatible resin, the Wang resin (Figure 31)[20], results in peptides with a C-terminal acid. Other notable resins include the peptide amide linker (PAL) resin[21], and the Rink amide resin[22], both of which result in a peptide with a C-terminal amide (Figure 32).

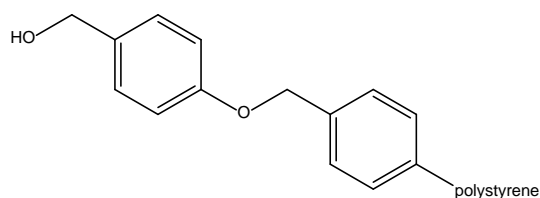


Figure 30: Wang resin.

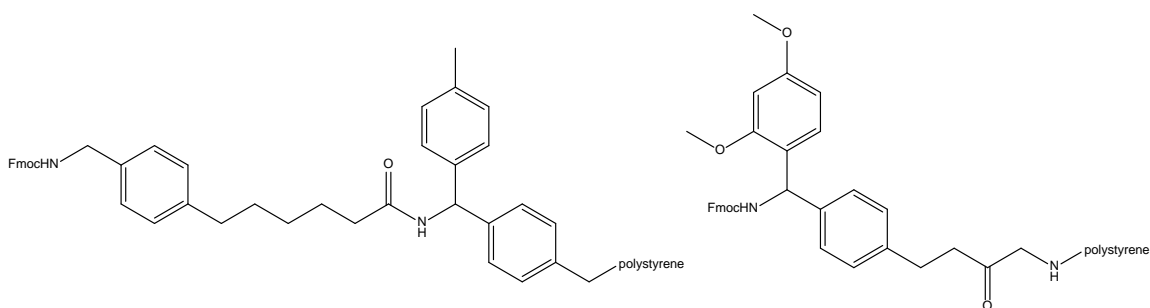


Figure 31: (Left) PAL resin. (Right) Rink amide resin.

Currently, the combination of reversible N-terminal protecting groups, effective C-terminal activation, and insoluble resins make both Boc and Fmoc effective ways to create peptides of interest.

Current Research Protocols

All peptides were synthesized using either Boc or Fmoc chemistry (Table 2). All peptides were created using automated synthesis. For Boc chemistry, synthesis was run on Applied Biosystems 430A Peptide Synthesizers. Peptides were synthesized on MBHA resin (Midwest Bio-Tech), 0.70 mmol reactive amines per gram. Protected amino acids (Midwest Biotech, Table 9) were added in 2 mmol aliquots, and 0.28 g of resin was used to achieve a 10:1 substitution ratio. All amino acids were L isomers unless otherwise specified. Standard reaction cycle consisted of 2 flow washes of dimethylformamide (DMF), 2 flow washes of 100% trifluoroacetic acid (TFA), and 1 min batch wash with TFA. Amino acid activation is performed separately, where 3.8 mL of 0.5 M DEPBT and 1 mL of diisopropylethylamine (DIEA) is added to 2 mmol of amino acid, and mixed by nitrogen bubbling for 1 minute. Activated amino acid is then transferred to the reaction vessel containing the resin and coupled for 15 minutes with vortexing. After coupling, resin is washed with 3 flow washes of DMF.

Peptides were cleaved from the resin using gaseous HF. The resin, with covalently bound peptide, was placed inside a Kel-F (polychlorotrifluoroethylene) reaction vessel, 0.5 mL of p-cresol (scavenger) was added, and the reaction vessel evacuated. Gaseous HF was

introduced to the reaction vessel and condensed by cooling the exterior of the vessel with a mixture of solid CO₂ and methanol. Once ~10 mL of HF condensed in the bottom of the reaction vessel, the reaction vessel was sealed and the resin/HF mixture was stirred for 1 hour. During this time the reaction vessel was kept at 4°C using a water/ice bath. After 1 hour, the reaction vessel was evacuated. Anhydrous diethyl ether was introduced to the vessel, resulting in the precipitation of the cleaved peptide. Ether was then filtered from the resin and free peptide precipitate. The resin and peptide mixture was then stirred in 20% acetonitrile, 1% acetic acid. Peptides dissolve in the 20% acetonitrile solution, and a second round of filtration separates the resin from the peptide in solution. The peptide solution could then be analyzed for mass and hydrophobicity, on Beckman System Gold qualitative HPLC followed by MALDI-TOF on a Bruker Autoflex III spectrometer (Figure 33), or on Agilent C8 / Quadripole LC/MS (Figure 36). The peptide solution was then lyophilized. For Boc syntheses, the percent yield of crude product varied between 10 and 30%. After lyophilization, peptides were purified using reverse phase chromatography, where the A buffer consisted of 0.1% aqueous TFA and the B buffer consisted of 90% acetonitrile, 0.1% TFA (Figure 34). Most peptides were purified on a 10-40% B gradient over 1 hour, although shallower gradients and longer run times were occasionally used to ensure pure product, as judged by LC/MS or analytical HPLC and MALDI (Figure 35). The stationary phase column used during purifications was a Phenomenex Luna silica-based C8 column. Fractions were collected in 30 second intervals and analyzed for purity using LC/MS or analytical HPLC and MALDI. Pure fractions were pooled and lyophilized. Purification typically resulted in a ~50% loss of peptide. Lyophilized peptides were dissolved in 50 mM ammonium bicarbonate buffer,

pH 8, and analyzed for absorbance at 280 nm using UV-Vis spectroscopy on a Thermo Scientific Nano-Drop 1000 spectrophotometer. Extinction coefficients for each peptide were calculated as the sum of the number of tryptophan residues and tyrosine residues multiplied by their respective molar extinction coefficients (Equation 1).

$$\text{Molar extinction coefficient} = (\text{Trp} \times 5500) + (\text{Tyr} \times 1490) \quad \text{Equation 1}$$

For Fmoc chemistry, syntheses were run on Applied Biosystems 433A Peptide synthesizers. Peptides were synthesized on H-Rink-Amide-Chemmatrix resin (PCAS BioMatrix, Inc). Resin contained 0.45 mmol reactive amines per gram, and 0.22 g was used to achieve a 10:1 substitution ratio. On the Applied Biosystems machine, the chemical reactions of Fmoc synthesis are each referenced by a mode (Table 10). Repeats of the modes (Table 11) result in peptide synthesis. For Fmoc synthesis, the percent yield varied between 15 and 25%. After synthesis, resin was cleaved from the peptide using 90% TFA, with 2.5% each of H₂O, β-mercaptoethanol, thioanisole and triisopropylsilane to function as scavengers for carbocations and free t-butyl groups, and to prevent cysteine oxidation. Cleavage reaction was mixed for 2 hours. After cleavage, the peptidyl-TFA solution was filtered from the resin. Addition of diethyl ether precipitated the peptides. Ether was removed by centrifugation and decanting. The peptide was then dissolved in 20% acetonitrile, 1 % acetic acid, analyzed by LC/MS (Figure 36) and lyophilized. After lyophilization peptide was purified with reverse phase chromatography, under the same

conditions as Boc synthesized peptides. Purification typically resulted in a 40% loss of peptide.

For peptides containing a disulfide bond, the disulfide was formed in the presence of Clear-Ox resin (Peptides International). Clear Ox resin was washed sequentially with DCM, DMF, methanol, deionized water, and 50% acetonitrile solution. Peptides with cysteine residues were then dissolved in 50% acetonitrile and added to the resin, and stirred at room temperature for 2 hours. LC/MS and Elmann's reagent were used to test for the presence of the disulfide. If successful, Clear-Ox resin is filtered from the solution and the peptide solution is lyophilized to recover the oxidized peptide.

For the the **#4-6** peptide modified by PEGylation, 2 mg of of pure peptide #4-6 was reacted with 2x molar equivalent of 20 kDa PEG-SCM (Creative PEGWorks) in 500 μ L of 50 mM sodium bicarbonate buffer, pH 9, and monitored with analytical HPLC and MALDI. Analytical HPLC demonstrated the appearance of a peak with a later retention time. Preparative HPLC was used to purify the PEGylated peptide, which was then lyophilized. After lyophilization, the peptide was dissolved in 50 mM ammonium bicarbonate buffer for phosphorylation assay analysis, as well as in PBS buffer for the stability study.

Table 9: Protecting groups for reactive side chains in Boc and Fmoc chemistries.

Amino acid / Single letter abbreviation	Side chain protecting group, Boc	Side Chain protecting group, Fmoc
Alanine / A	-	-
Cysteine / C	4-MeBzl	Trt
Aspartic acid / D	OcHx	OtBu
Glutamic acid / E	OcHx	OtBu
Phenylalanine / F	-	-
Glycine / G	-	-
Histidine / H	BOM	Trt
Isoleucine / I	-	-
Lysine / K	Cl-Z	Boc
Leucine / L	-	-
Methionine / M	-	-
Asparagine / N	Xan	Trt
Proline / P	-	-
Glutamine / Q	-	Trt
Arginine / R	Tos	Pbf
Serine / S	Bzl	tBu
Threonine / T	Bzl	tBu
Valine / V	-	-
Trptophan / W	CHO	Boc
Tyrosine / Y	Br-Z	tBu

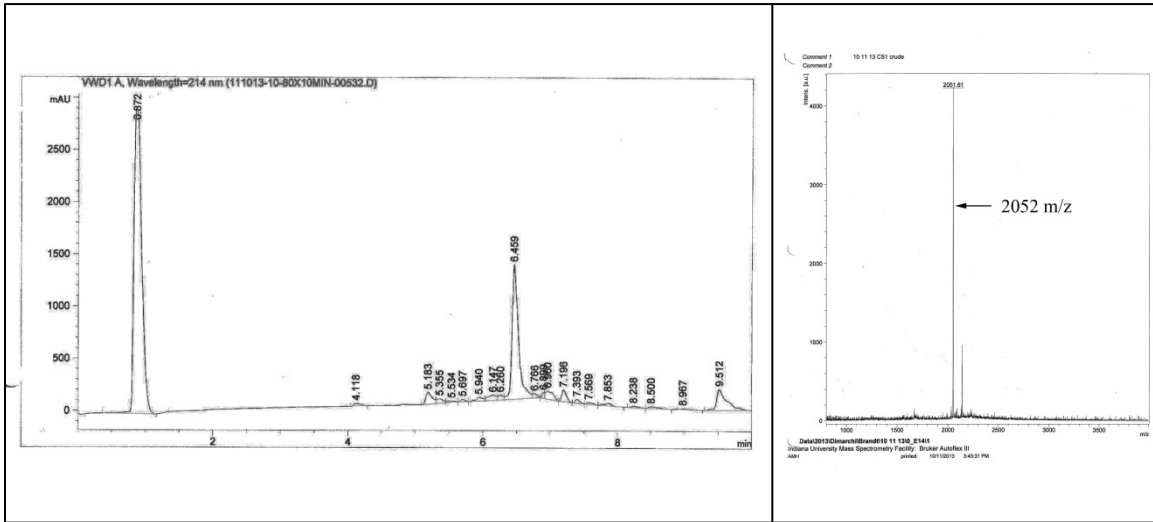


Figure 32: (Left) Analytical HPLC of crude peptide #Cys-4 synthesized via Boc chemistry. (Right) MALDI of crude peptide #Cys-4 demonstrating the correct molecular weight of 2052 has been achieved.

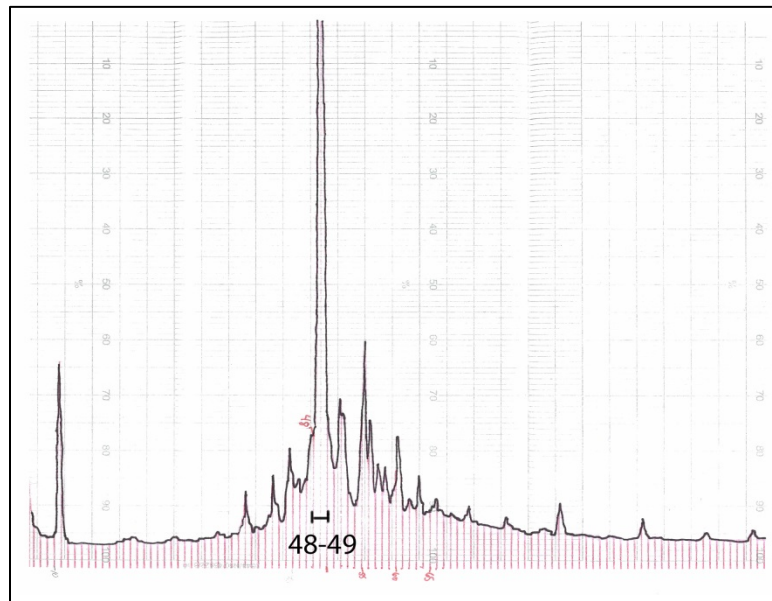


Figure 33: Purification of peptide #Cys-4 on a C8 column, 10-80%B, 75 minutes, with fractions collected every 30 seconds. Fractions 48-49 were judged to be pure by analytical HPLC and MALDI and were lyophilized. Representative of all peptides synthesized via Boc chemistry.

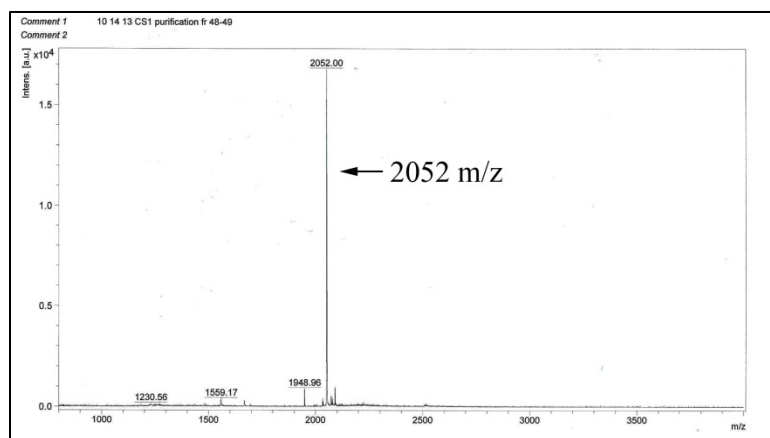
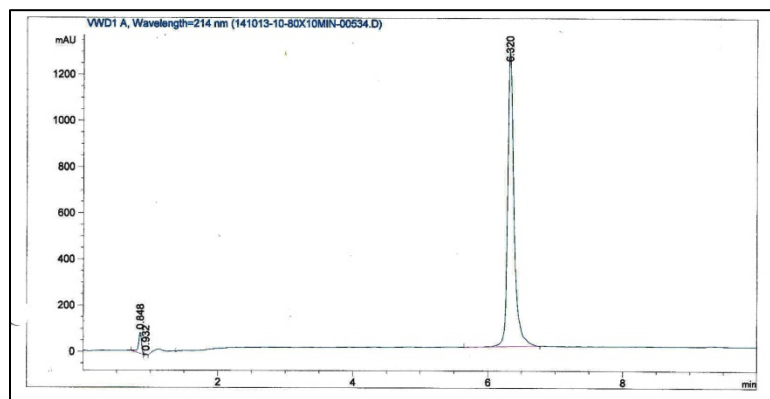


Figure 34: (Top) Analytical HPLC of purified peptide #Cys-4. (Bottom) MALDI of purified peptide #Cys-4. Representative of all peptides synthesized via Boc chemistry and purified via reverse phase HPLC.

Table 10: Description of modes of Fmoc synthesis.

Mode	Description
a	Activation: previous amino acid cartridge is ejected. 2.1 mL of NMP solvent and 1 mL of 1 M 6-Cl-HOBt is added to cartridge and mixed for 15 minutes. After mixing, 1 mL of 1 M DIC is added.
b	Piperidine deprotection: NMP wash followed by 3 minute wash of 20% piperidine / NMP, then an additional 15 min wash. Reaction vessel is then drained and washed with NMP.
c	DCM washes: 8 DCM washes
d	NMP washes: 6 NMP washes
e	Transfer and washing: transfers the activated amino acid to the reaction vessel
f	Coupling: 20 minutes of vortexing to mix resin and incoming amino acid.
g	Resin wash: reaction vessel is washed with NMP.

Table 11: Fmoc synthesis of a peptide of n amino acids in length using previously described modes.

Cycle	Repeats	Mode
1	1	d
2	1	aibde
3	$n-1$	afgbde
n	1	fffbdc

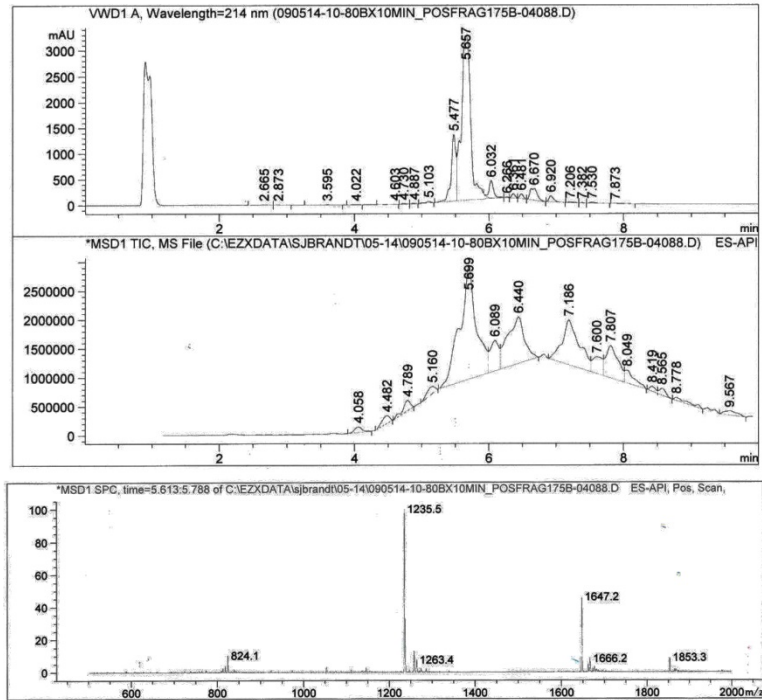


Figure 35: LC/MS analysis of crude peptide #6. Top to bottom: absorbance at 214 nm, ion counts, and observed m/z values at t=5.6 minutes. The molecular weight of peptide #6, 2470, is represented by charge states 1235 and 823. This LC/MS trace is typical of peptides synthesized using Fmoc chemistry.

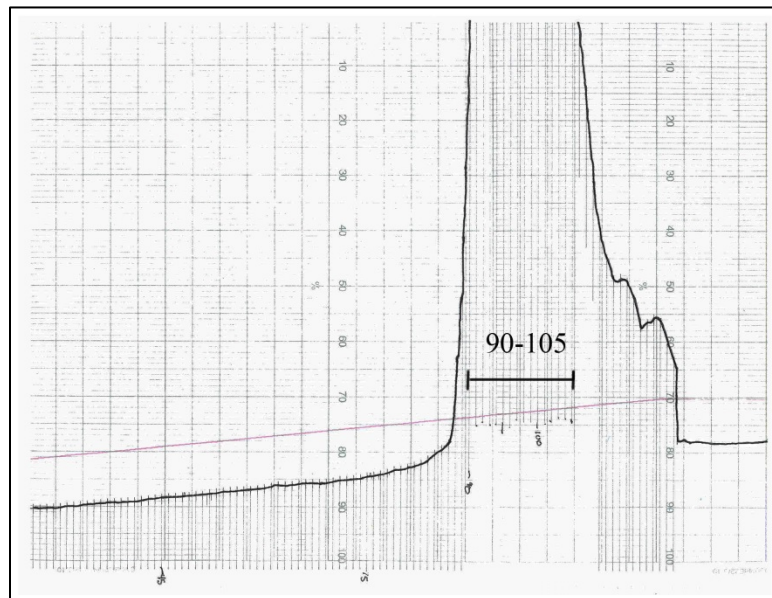


Figure 36: Reverse phase HPLC purification of peptide #6, 10-40% acetonitrile, 0.1% TFA, 1 hr, fractions collected every 30 seconds. Fractions 90-105 were judged to be pure by LC/MS. Representative of all peptides synthesized via Fmoc chemistry.

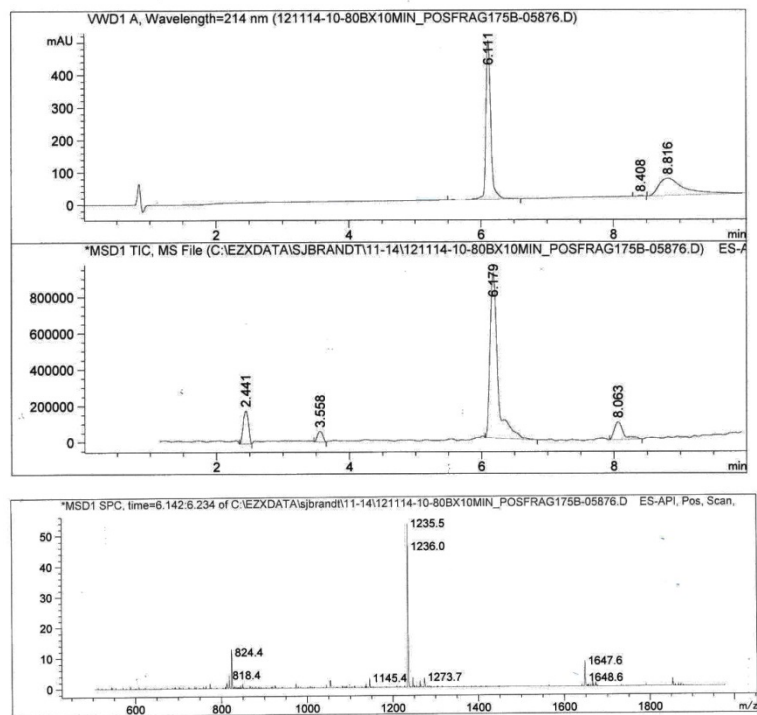


Figure 37: LC/MS analysis of pure peptide #6, as synthesized using Fmoc chemistry and purified using reverse phase HPLC. Top to bottom: absorbance at 214 nm, ion counts, and observed m/z values at 6.1 minutes. This LC/MS trace is typical of peptides synthesized by Fmoc chemistry and purified via HPLC.

In vitro analysis

Phosphorylation assays were run in human insulin receptor-transfected HEK293 cells in DMEM high glucose medium, supplemented with 10% bovine growth serum (BGS, HyClone), antibiotics (gentamicin, hygromycin B) and 10 mM HEPES. Cells were plated 4×10^4 cells/80 μ L well in DMEM in a poly-L-lysine coated 96 well tissue culture plate. Cells were checked for adherence under a microscope. The peptides to be investigated and the control peptide, native insulin, were dissolved in bicarbonate and diluted to 1 μ M concentration in DMEM 0.5% BSA. Through serial dilution, a gradient of peptides was prepared in a 96 well polypropylene plate. 20 μ L of each peptide concentration was added to the cell plate. If the assay is investigating the antagonism of the peptide of interest, 20 μ L of media is removed from the cell plate and 20 μ L 5x (5 nM) concentrated insulin is added to the cell plate, prior to the addition of peptide of interest. Next, the cells are incubated for 15 minutes at 37°C, 5% CO₂. After incubation, the cells were fixed by the addition of 10% formalin, 100 μ L/well, and incubated for 20 minutes at room temperature. The liquid contents of each well are then removed by inverting the plate. The wells are then washed 2x with 200 μ L/well PBS-0.1% Triton X-100. Next, 200 μ L blocking buffer (2% BSA in PBS-0.1% Triton X-100) is added to each well and incubated for 1 hour at room temperature. Plates are then washed twice with PBS-0.1% Triton X-100. Anti-phospho-insulin receptor 4G10-HRP (UBI) was prepared at a 1:10000 dilution in blocking buffer, and 50 μ L was added to each well and incubated 3h at room temperature or overnight at 4 °C. Plates are then washed 4x with PBS-0.1% Triton X-100. Finally, 100 μ L/well of room temperature TMB single solution is added to each plate, and incubated for 5 minutes, after which 50 μ L of stop solution (1 M HCl)

was added to each well. The optical density of each well at 450 nm was monitored using a Titertek Multiskan MCC/340 plate reader. Optical density at 450 nm was analyzed as a function of the concentration of peptide using Origin Pro 9.0 graphing software. EC₅₀ and IC₅₀ values were determined using a sigmoidal fit for all data points.

Chemical reactions and schematics displayed in this chapter were created with ChemBioDraw Ultra 13.0.

REFERENCES

1. Pillutla, R.C., et al., *Peptides identify the critical hotspots involved in the biological activation of the insulin receptor*. J Biol Chem, 2002. **277**(25): p. 22590-4.
2. Knudsen, L., et al., *Agonism and Antagonism at the Insulin Receptor*. Plos One, 2012. **7**(12).
3. Schaffer, L., et al., *A novel high-affinity peptide antagonist to the insulin receptor*. Biochem Biophys Res Commun, 2008. **376**(2): p. 380-3.
4. Schaffer, L., et al., *Assembly of high-affinity insulin receptor agonists and antagonists from peptide building blocks*. Proc Natl Acad Sci U S A, 2003. **100**(8): p. 4435-9.
5. Pillutla, R.C., et al., *Insulin and IGF-1 Receptor Agonists and Antagonists*, U.S. Patent, Editor 2007, Antyra Inc., Novo Nordisk A/S: United States. p. 232.
6. Jensen, M., et al., *Activation of the insulin receptor by insulin and a synthetic peptide leads to divergent metabolic and mitogenic signaling and responses*. J Biol Chem, 2007. **282**(48): p. 35179-86.
7. Jensen, M., et al., *Activation of the insulin receptor (IR) by insulin and a synthetic peptide has different effects on gene expression in IR-transfected L6 myoblasts*. Biochem J, 2008. **412**(3): p. 435-45.
8. Zierath, J.R., et al., *C-peptide stimulates glucose transport in isolated human skeletal muscle independent of insulin receptor and tyrosine kinase activation*. Diabetologia, 1996. **39**(3): p. 306-13.
9. Sharma, S., et al., *Protein Pegylation: Impact of PEG Quality on Thermal Stability of Protein*. Global Journal of Biotechnology and Biochemistry, 2011. **6**(2): p. 48-59.
10. Klinge, S., et al., *Atomic structures of the eukaryotic ribosome*. Trends Biochem Sci, 2012. **37**(5): p. 189-98.
11. Mitchell, A.R., *Bruce Merrifield and solid-phase peptide synthesis: a historical assessment*. Biopolymers, 2008. **90**(3): p. 175-84.
12. Kimmerlin, T. and D. Seebach, *'100 years of peptide synthesis': ligation methods for peptide and protein synthesis with applications to beta-peptide assemblies*. J Pept Res, 2005. **65**(2): p. 229-60.

13. Carpino, L.A. and G.Y. Han, *9-Fluorenylmethoxycarbonyl Function, a New Base-Sensitive Amino-Protecting Group*. Journal of the American Chemical Society, 1970. **92**(19): p. 5748-&.
14. Merrifield, R.B., *Solid Phase Peptide Synthesis. 1. The Synthesis of a Tetrapeptide*. Journal of the American Chemical Society, 1963. **85**(14): p. 2149-2154.
15. Letsinger, R.L. and M.J. Kornet, *Popcorn Polymer as a Support in Multistep Syntheses*. Journal of the American Chemical Society, 1963. **85**(19): p. 3045-&.
16. Merrifield, R.B., *Solid-Phase Peptide Synthesis. 3. An Improved Synthesis of Bradykinin*. Biochemistry, 1964. **3**: p. 1385-90.
17. Mitchell, A.R., et al., *Tert-butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl-resin, a more acid-resistant support for solid-phase peptide synthesis*. J Am Chem Soc, 1976. **98**(23): p. 7357-62.
18. Pietta, P.G. and G.R. Marshall, *Amide Protection and Amide Supports in Solid-Phase Peptide Synthesis*. Journal of the Chemical Society D-Chemical Communications, 1970(11): p. 650-&.
19. Matsueda, G.R. and J.M. Stewart, *A p-methylbenzhydrylamine resin for improved solid-phase synthesis of peptide amides*. Peptides, 1981. **2**(1): p. 45-50.
20. Wang, S.S., *p-alkoxybenzyl alcohol resin and p-alkoxybenzyloxycarbonylhydrazide resin for solid phase synthesis of protected peptide fragments*. J Am Chem Soc, 1973. **95**(4): p. 1328-33.
21. Albericio, F. and G. Barany, *Mild, orthogonal solid-phase peptide synthesis: use of N alpha-dithiasuccinoyl (Dts) amino acids and N-(iso-propyldithio)carbonylproline, together with p-alkoxybenzyl ester anchoring linkages*. Int J Pept Protein Res, 1987. **30**(2): p. 177-205.
22. Rink, H., *Solid-Phase Synthesis of Protected Peptide-Fragments Using a Trialkoxy-Diphenyl-Methylester Resin*. Tetrahedron Letters, 1987. **28**(33): p. 3787-3790.

Chapter 3
Biosynthetic Heterodimers

ABSTRACT

This chapter discusses the creation and analysis of biosynthetic heterodimers. These heterodimers are expressed in *E. coli*, and consist of an A, B, and C domain. The A and B domains are identical in sequence to DP8, an insulin receptor agonist. The C domain of these constructs contains one of the previously characterized peptides of interest. Based on the synthetic studies in chapter 2, peptides **#4**, **#6**, **#4-6**, **#6-4**, and **#12** were chosen for incorporation into heterodimers. In order to facilitate *lys-C* cleavage, these peptides contained either a lysine residue at the N-terminus, C-terminus, or both termini. Based on the position of the lysine residues, treatment with *lys-C* resulted in a two chain heterodimer, or released the C domain entirely. Heterodimers were analyzed in a phosphorylation assay for agonism and antagonism at both insulin receptor isoforms. It was found that heterodimers with fully removable C domains yielded an agonist molecule that is fully potent and active at the insulin receptor, indicating that the agonist portion of the heterodimers is formed correctly. Heterodimers with partially cleavable C domains were also investigated. Many of these peptides displayed reduced potency, or no activity. However, the **#6-4** peptide, when positioned at the C-terminus of the B chain, displayed reduced activity without a decrease in potency. This represents the first instance of a potent, reduced activity heterodimer, and confirms our hypothesis that heterodimers can be used to alter the maximal activity of insulin-based analogs.

INTRODUCTION

While solid phase peptide synthesis represents an effective way to create peptides, it is limited both in maximal size and in the formation of disulfide bonds. As stated previously, our goal is to create heterodimers consisting of an insulin receptor agonist and an insulin receptor antagonist. Solid phase synthesis is sufficient to create the antagonist portion of the heterodimer, but the agonist portion is more complex. Native insulin contains three disulfide bonds, two of which are between chains and one of which is intrachain. While it is possible to create A and B chains separately and use orthogonal protecting schemes to promote disulfide bond formation[1], this is a laborious process and requires careful separation of isoforms, since a single disulfide mis-match results in an inactive peptide.

In addition, native insulin is relatively insoluble in aqueous solutions, and therefore we chose to begin our investigations into heterodimers with a more practical insulin receptor agonist, DP8 (Table 1), which is an “insulinized” version of insulin growth factor 1 (IGF-1). Native IGF-1 displays a roughly two fold decrease in potency at the insulin receptor, relative to native insulin (Figure 1)[1]. Due to modifications such as eliminating the C domain and introducing the two key residues TyrB¹⁶ and LeuB¹⁷[1], DP8 has full activity and potency at the insulin receptor isoforms (Figure 1). In this chapter, DP8 was used as the agonist portion of our heterodimers.

Table 1: Insulin, IGF-1 and DP8 sequence comparison. DP8 represents an “insulinized” IGF-1 peptide, where the IGF-1 A and B chains have been mutated at key residues so that DP8 has full activity and potency at the insulin receptor. Native IGF-1 also contains a C chain, which is omitted in DP8.

A chain	Insulin	G	I	V	E	Q	C	C	T	S	I	C	S	L	Y	Q	L	E	N	Y	C	N									
	IGF-1	G	I	V	D	E	C	C	F	R	S	C	D	L	R	R	L	E	M	Y	C	A									
	DP8	G	I	V	D	E	C	C	H	R	S	C	D	L	R	R	L	E	N	Y	C	N									
B chain	Insulin	F	V	N	Q	H	L	C	G	S	H	L	V	E	A	L	Y	L	V	C	G	E	R	G	F	F	Y	T	P	K	T
	IGF-1		G	P	E	T	L	C	G	A	E	L	V	D	A	L	Q	F	V	C	G	D	R	G	F	Y	F	N	K	P	T
	DP8		G	P	E	H	L	C	G	A	H	L	V	D	A	L	Y	L	V	C	G	D	R	G	Y						

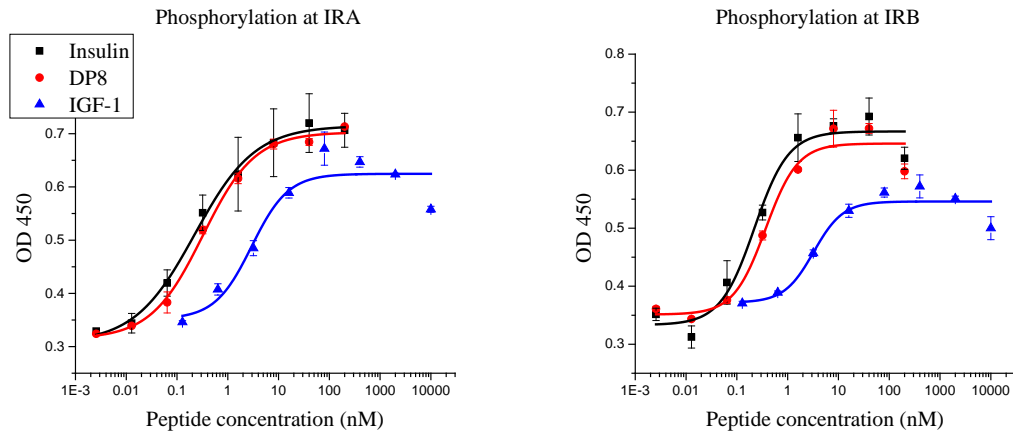


Figure 1: Phosphorylation assay demonstrating the relative potency and activity of native human insulin, native human IGF-1, and DP8.

The heterodimers presented in this chapter were produced in *E. coli*. *E. coli* has long been used for biosynthesis, due to the propensity of the bacteria to engage in lateral gene transfer. Lateral gene transfer is the transfer of genetic material between species of bacteria, as opposed to between generations of bacteria. These transfers can result in increased metabolic capabilities, pathogenicity, or antibiotic resistance. Lateral gene transfer has been recognized since as early as the 1950s and is responsible for some of the antibiotic resistance seen today[2]. The propensity of bacteria to engage in lateral gene transfer can be exploited to introduce plasmids coding for a protein of interest. In 1977, the first human protein created using *E. coli* biosynthesis was somatostatin, a 14 amino acid hormone[3]. Interestingly, somatostatin suppresses the secretion of a number of other hormones, including insulin[3]. Insulin itself was subsequently produced in *E. coli*, although initially the A and B chains were synthesized separately, in separate bacterial strains[4]. This reflects one of the original drawbacks of using *E. coli* as a vehicle for biosynthesis – the cytoplasm, where proteins are produced, is a highly reducing

environment. For proteins containing disulfide bonds, such as insulin and the insulinized IGF proteins used in this research, a reducing environment can render the proteins inactive or promote disulfide shuffling. It was shown, however, that mutations which disrupted the natural reducing pathways of thioredoxin and glutathione reductase in *E. coli* resulted in an increase in disulfide bond formation in biosynthetic peptides[5]. This research utilized a strain of *E. coli* referred to as OrigamiDE3, which contains mutations in the thioredoxin reductase and glutathione reductase pathways. Genes coding for our peptides were introduced in a plasmid, downstream of a lac promoter, which makes the cells responsive to induction with IPTG, a lactose analog[6].

A potential setback in *E. coli* biosynthesis is the propensity for recombinant proteins to aggregate, sometimes in conjunction with a lack of solubility within the cellular environment. The solution to this problem is to introduce a small ubiquitin-related modifier (SUMO) protein at the N-terminus of the recombinant proteins, which also includes a TEV protease cleavage site. This results in an increase in protein folding as well as an increase in solubility[7]. In addition, inclusion of a poly-histidine tag in the SUMO fusion protein allows for nickel-column purification of the protein of interest. After induction and sonication, the cell lysate is poured over a resin that contains nickel ions. The poly histidine tag binds reversibly to the column, while the remaining cell lysate is washed away. The protein of interest can then be eluted with a concentrated imidazole solution. After elution, the peptide is treated with TEV protease, which cleaves the SUMO/polyhisitidine tag. The peptide can then be purified and assayed *in vitro*.

In order to create heterodimers, we designed DP8 molecules that contain a peptide of our choosing as the C domain. These molecules are expressed in *E. coli* as a single chain construct. A lysine residue was included at either the N-terminus, C-terminus, or at both termini of the C domain. Since DP8 does not include any additional lysine residues, treatment with the endopeptidase *lys-C* results in two chain construct with the peptide of interest bound, via a peptide bond, to the N-terminus of the A chain, the C-terminus of the B chain, or removed from the construct entirely.

Based on our results from chapter 2, we chose the site 1 binding motif #4, the site 2 binding motif #6, the two-motif constructs #4-6 and #6-4, and the negative control #12 as C domain peptides. These peptides were placed as the C domain, with lysines at their N-, C-, or at both termini. This results in peptides with fully *lys-C* removable domains (Table 2), peptides where *lys-C* cleavage results in the antagonist positioned at the N-terminus of the A chain (Table 3), and peptides where *lys-C* cleavage results in the antagonist positioned at the C-terminus of the B chain (Table 4). These peptides were tested for agonism and antagonism as a single chain construct, as well as in their *lys-C* treated two-chain form.

We anticipate that this approach will be successful for several reasons. We will be introducing non-native C domains into DP8, but C domains occur naturally in both insulin and IGF-1, which is the molecular basis of DP8[8, 9]. Even though the C domain of insulin is removed post-translationally[10], it plays a critical role in proper disulfide

formation by maintaining unimolecularity during folding[11]. Thus, our heterodimers are a mimic of naturally occurring peptides and will likely be properly folded.

Assuming that our heterodimers fold correctly, there are still a few potential obstacles in the creation of a high potency, lowered activity dimer. Some of our dimers are analogous to split proinsulin, in which the C domain is not completely removed from the proinsulin molecule. Split proinsulin exhibits reduced potency but full activity[12]. It is possible that our heterodimers will merely exhibit split-proinsulin-like reductions in potency without reductions in maximal activity. Another potential limitation is that small C domains can diminish the ability of insulin like molecules to signal, due to a lack of flexibility in the C-terminus of the B chain[13].

In spite of these potential difficulties, we investigated the ability of *E. coli* to create biosynthetic heterodimers, and tested these dimers for activity at the insulin receptor A and B isoforms.

Table 2: Heterodimers with fully *lys-C* removable C domains

Peptide	Domains	Sequence
DP38	A: DP8 B: DP8 C: #4-6	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKGS L DESFYDWFERQLGGGSGGSLEEWAQIQSEVWGR GSPSYK GIVDECCHRSCDLRRL ENYCN
DP72	A: DP8 B: DP8 C: #12	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKRRE AEDLQVGQVELGGGPGAGSLQPLALEGSLQKGIVD ECCHRSCDLRRL ENYCN

Table 3: Heterodimers, where *lys-C* cleavage results in antagonists positioned at the N-terminus of the A chain.

Peptide	Domains	Sequence
DP66	A: DP8 B: DP8 C: #4	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKGS L DES FYD WFERQLGGIVDECCHRSCDLRRL ENYCN
DP68	A: DP8 B: DP8 C: #6	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKSLE EEWAQIQSEVWGRGSPSYGIVDECCHRSCDLRRL EN YCN
DP37	A: DP8 B: DP8 C: #4-6	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKGS L DES FYD WFERQLGGGSGGSSLEEEWAQIQSEVWGR GSPSYGIVDECCHRSCDLRRL ENYCN
DP51	A: DP8 B: DP8 C: #6-4	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKSLE EEWAQIQSEVWGRGSPSYGGSGGSGSLDES FYDWF ERQLGGIVDECCHRSCDLRRL ENYCN
DP70	A: DP8 B: DP8 C: #12	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYKRRE AEDLQVGQVELGGGPGAGSLQPLALEGSLQGIVDEC CHRSCDLRRL ENYCN

Table 4: Heterodimers, where *lys-C* cleavage results in antagonists positioned at the C-terminus of the B chain.

Peptide	Domains	Sequence
DP67	A: DP8 B: DP8 C: #4	GEEEEKGP EHL CGAHLVDALYL VCGDRGYGSLDE SFYD WFERQLGKIVDECCHRSCDLRRL ENYCN
DP69	A: DP8 B: DP8 C: #6	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYSLEEE WAQIQSEVWGRGSPSYKIVDECCHRSCDLRRL EN YCN
DP36	A: DP8 B: DP8 C: #4-6	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYGS LD ESFYD WFERQLGGGSGGSSLEEEWAQIQSEVWGRG SPSYKIVDECCHRSCDLRRL ENYCN
DP50	A: DP8 B: DP8 C: #6-4	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYSLEEE WAQIQSEVWGRGSPSYGGSGGSGSLDES FYD WFER QLGKIVDECCHRSCDLRRL ENYCN
DP71	A: DP8 B: DP8 C: #12	GEEEEKGP EHL CGAHLVDALYL VCGDRGFYRREA EDLQVGQVELGGGPGAGSLQPLALEGSLQKIVDEC CHRSCDLRRL ENYCN

RESULTS

Heterodimers, with fully *lys-C* removable C domains

DP38 and DP72 contain identical DP8 A and B chains. These peptides also include the short sequence GEEEEK, at the N-terminus of the A chain, in order to increase solubility. DP38 and DP72 each contain a peptide in the C domain, flanked by lysine residues. In DP38, the C domain is composed of the #4-6 antagonist (Table 2). In DP72, the C domain is composed of the #12 insulin C peptide (Table 2). As a single chain molecule, DP38 does not display any agonism at either receptor isoform (Figure 2). However, the single chain DP72 construct does display full activity at the insulin receptor isoforms, albeit with greatly reduced potency (Figure 3). This is reasonable, since DP72 is analogous to human proinsulin, which is known to activate the insulin receptors with reduced potency[14].

DP38 and DP72 were subsequently treated with *lys-C*. In both constructs, the peptide within the C chain is flanked by lysine residues, and therefore treatment with *lys-C* liberates the C domain. The only additional lysine in the construct is in the GEEEEK solubility sequence, and therefore treatment with *lys-C* also removes the solubility sequence. The result of the enzymatic cleavage is a two chain form of DP8, as well as either free #4-6 antagonist or free C peptide. The *lys-C* treated peptides were purified in order to isolate the two chain agonist. The agonists isolated from both these constructs (DP38” and DP72”) were fully potent and fully active at both insulin receptor isoforms (Figure 4, Figure 5).

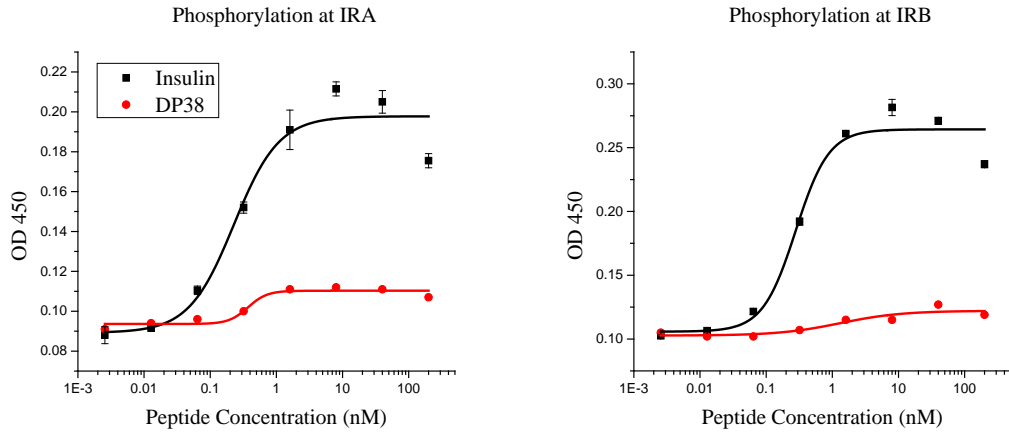


Figure 2: Phosphorylation assay demonstrating the agonism of DP38.

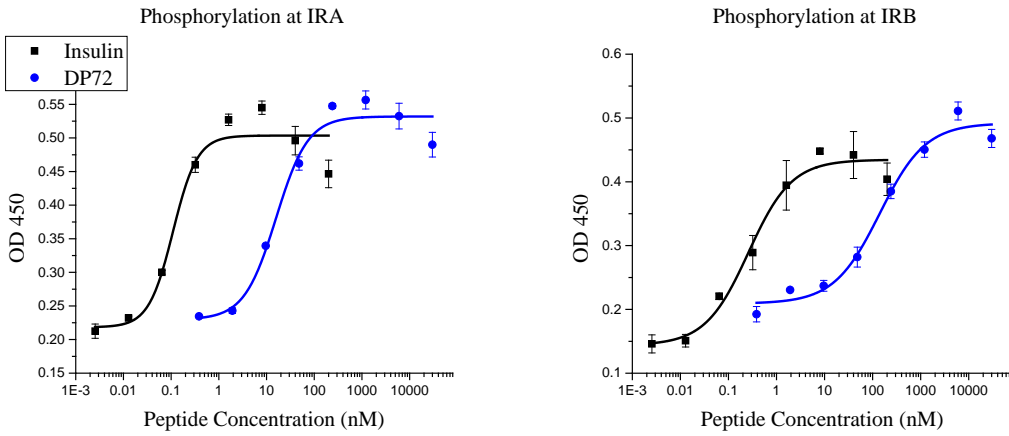


Figure 3: Phosphorylation assay demonstrating the agonism of DP72.

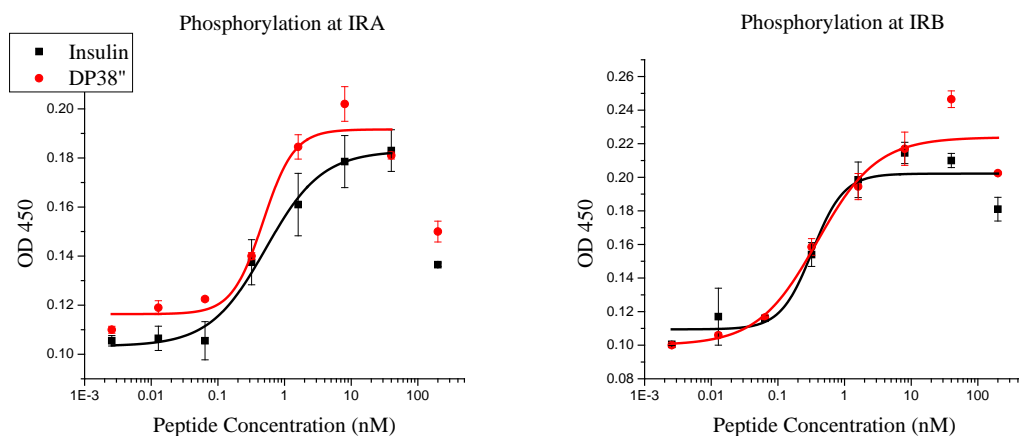


Figure 4: Phosphorylation assay demonstrating the activity of DP38''.

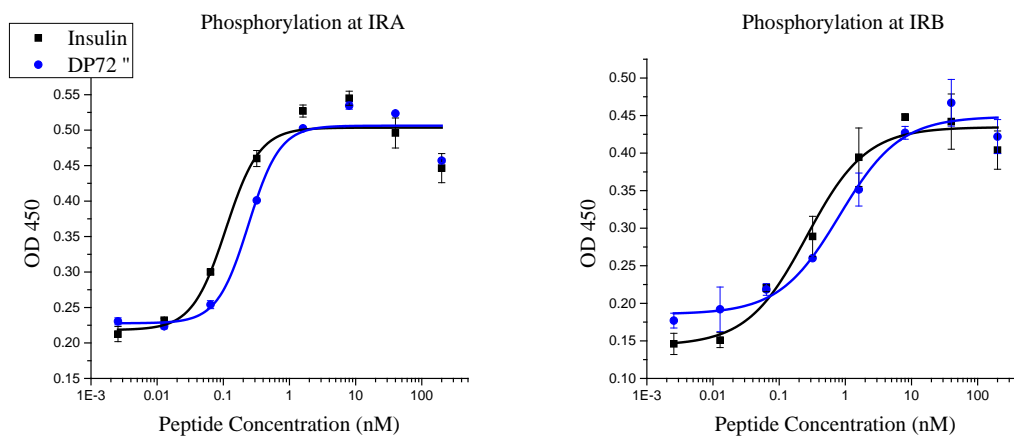


Figure 5: Phosphorylation assay demonstrating the activity of DP72''.

Heterodimers, where *lys-C* cleavage results in antagonists positioned at the N-terminus of the A chain

The next sequence of peptides was also synthesized biosynthetically as a single chain peptide, with the A and B chains consisting of DP8, and the C domain consisting of either the site 1 binding motif **#4** (DP66), the site 2 binding motif **#6** (DP68), the two site antagonist **#4-6** (DP37), the two site antagonist **#6-4** (DP51), or the **#12** insulin C-peptide (DP70) (Table 3). These peptides also contained the GEEEEEEK solubility sequence and a single additional lysine residue, such that treatment with *lys-C* results in a heterodimer where the peptide of interest is positioned at the N-terminus of the A chain (Figure 6).

DP66 contains the site 1 binding motif **#4** in the C chain, and as a single chain molecule it shows slight agonism at either insulin receptor isoform (Figure 7). Treatment with *lys-C* creates the two chain form of the molecule, with the site 1 motif **#4** bound to the N-terminus of the A chain. The two chain form is also slightly active at both insulin receptor isoforms (Figure 7).

DP68 contains the site 2 binding motif **#6** within the C domain. As a single chain molecule, DP68 displays agonism at the IRA and IRB isoforms, although with greatly reduced potency (Figure 8). The *lys-C* treated two-chain form of the molecule displays even less agonism at the receptor isoforms (Figure 8).

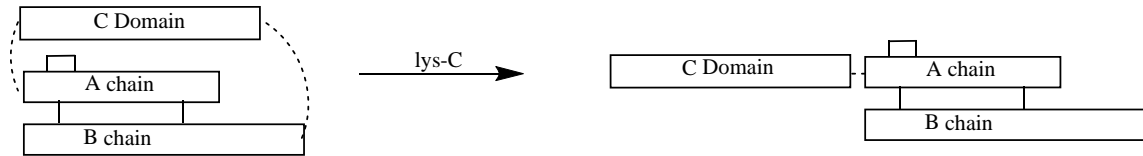


Figure 6: Schematic representing the transition from a single chain molecule to a two chain heterodimer via treatment with *lys-C*. Dotted lines represent a peptide bond, solid lines represent a disulfide bond.

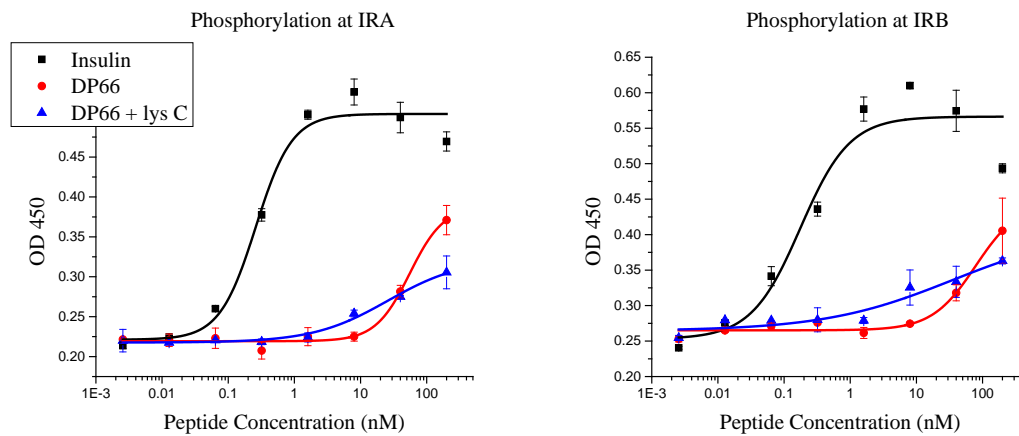


Figure 7: Phosphorylation assay demonstrating the agonism of the single chain DP66 and the *lys-C* treated, two chain form of DP66.

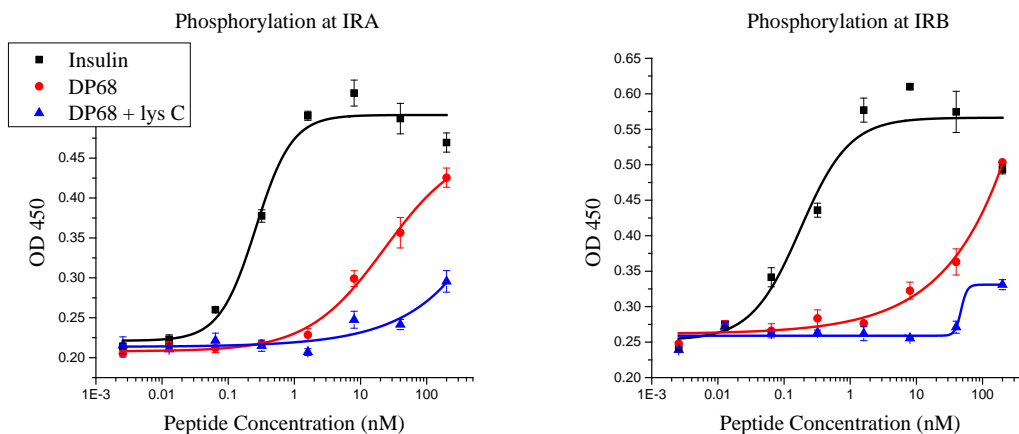


Figure 8: Phosphorylation assay demonstrating the agonism of the single chain DP68 and the *lys-C* treated, two-chain form of DP68.

The next pair of heterodimers contains the antagonists with two binding motifs as the C domain. In DP37, the C domain is antagonist #4-6 (Table 3). In DP51, the C domain is antagonist #6-4 (Table 3). Both constructs contain a single lysine residue such that treatment with *lys-C* results in a two chain molecule with the antagonist positioned at the N-terminus of the A chain.

As single chain constructs, neither DP37 nor DP51 display any agonism at either insulin receptor (Figure 9, Figure 10). Treatment with *lys-C* creates the two chain form of these molecules, but again, no agonism was detected at either insulin receptor isoform (Figure 9, Figure 10).

The final heterodimer in this series contains the #12 peptide as the C domain, with a single lysine residue such that treatment with *lys-C* results in a two chain molecule with peptide #12 positioned at the N-terminus of the A chain (Table 3). As a single chain molecule, this heterodimer displays full activity at both insulin receptor isoforms, although with reduced potency (Figure 11). Treatment with *lys-C* does not result in any increase in activity or potency (Figure 11).

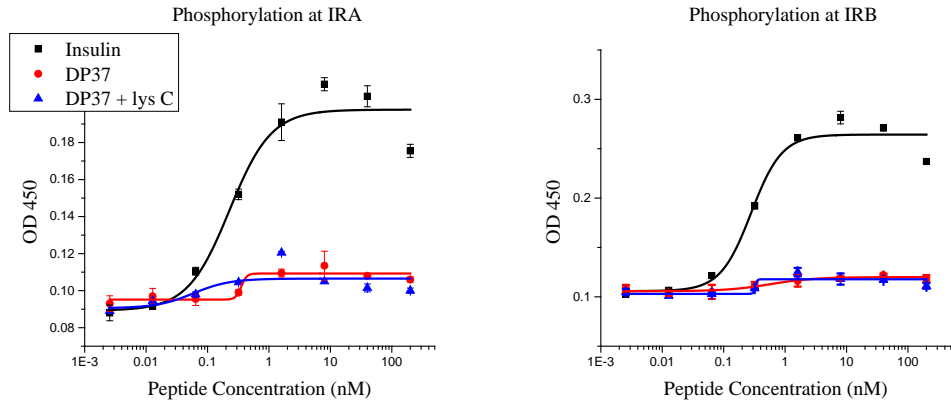


Figure 9: Phosphorylation assay demonstrating the agonism of the single chain DP37 and the *lys-C* treated, two chain form of DP37.

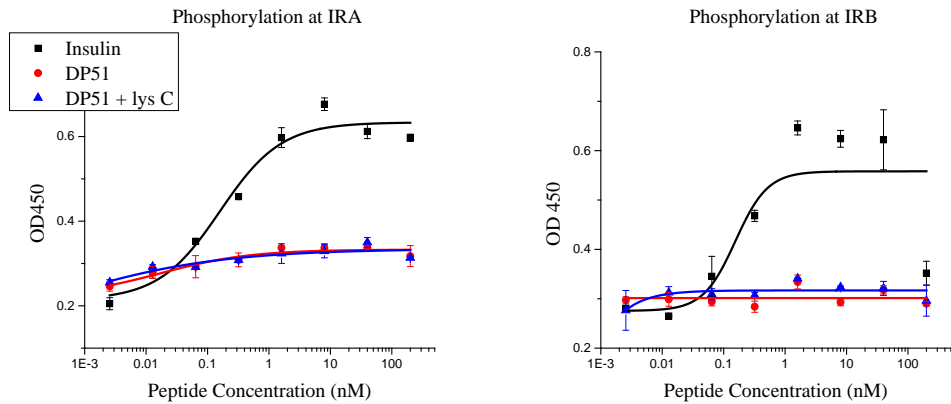


Figure 10: Phosphorylation assay demonstrating the agonism of the single chain DP51 and the *lys-C* treated, two chain form of DP51.

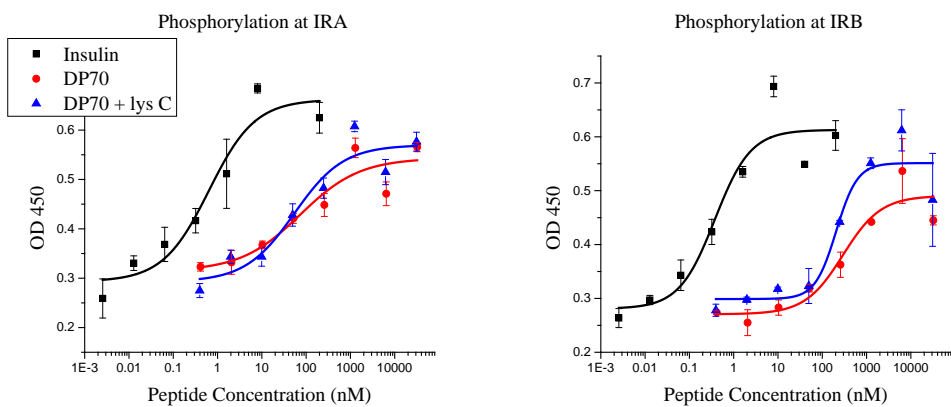


Figure 11: Phosphorylation assay demonstrating the agonism of the single chain DP70 and the *lys-C* treated, two chain form of DP70.

Heterodimers, where *lys-C* cleavage results in antagonists positioned at the C-terminus of the B chain

The next series of peptides were also synthesized as a single chain constructs, where the A and B chains consist of DP8, and the C domain consists of **#4** (DP67), **#6** (DP69), **#4-6** (DP36), **#6-4** (DP50), or **#12** (DP71) (Table 4). These peptides contained the GEEEEEEK solubility tag and a single additional lysine residue, such that treatment with *lys-C* results in a heterodimer where the peptide of interest positioned at the C-terminus of the B chain (Figure 12).

DP67 contains the **#4** site 1 binding motif in the C domain. It is identical to DP66 except for the placement of the key lysine residue. Much like DP66, as a single chain DP67 shows full activity but reduced potency at the insulin receptor A isoform, and reduced activity at the B isoform (Figure 13). However, treatment with *lys-C* creates a two chain molecule where the **#4** binding motif is positioned at the C-terminus of the B chain. As a two chain molecule, DP67 has full activity and only slightly reduced potency at both insulin receptor isoforms (Figure 13).

DP69 contains the **#6** site 2 binding motif in the C domain, and is identical to DP68 except for the placement of the single lysine residue. As a single chain peptide, DP69 shows full activity but reduced potency at both receptor isoforms (Figure 14). Treatment with *lys-C* results in a two chain molecule, with the **#6** motif positioned at the C-terminus of the B chain, but there was no change in the potency or activity, relative to the single chain molecule (Figure 14).

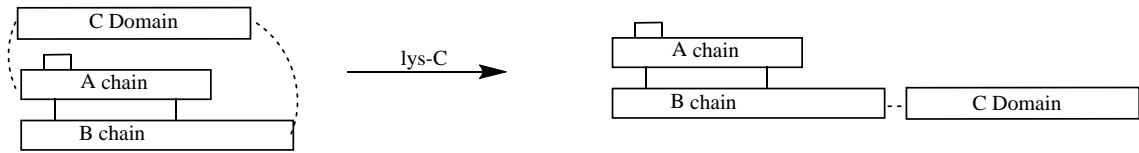


Figure 12: Schematic representing the transition from a single chain molecule to a two chain heterodimer via treatment with *lys-C*. Dotted lines represent a peptide bond, solid lines represent a disulfide bond.

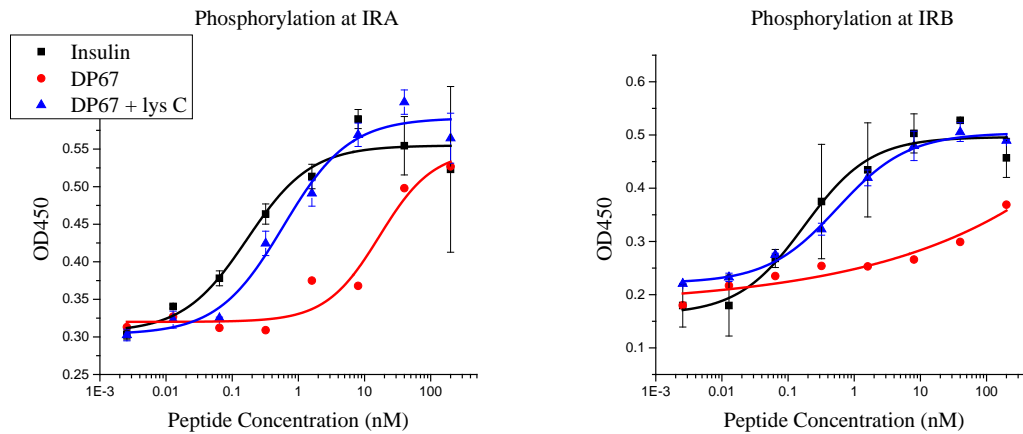


Figure 13: Phosphorylation assay demonstrating the agonism of the single chain DP67 and the *lys-C* treated, two chain form of DP67.

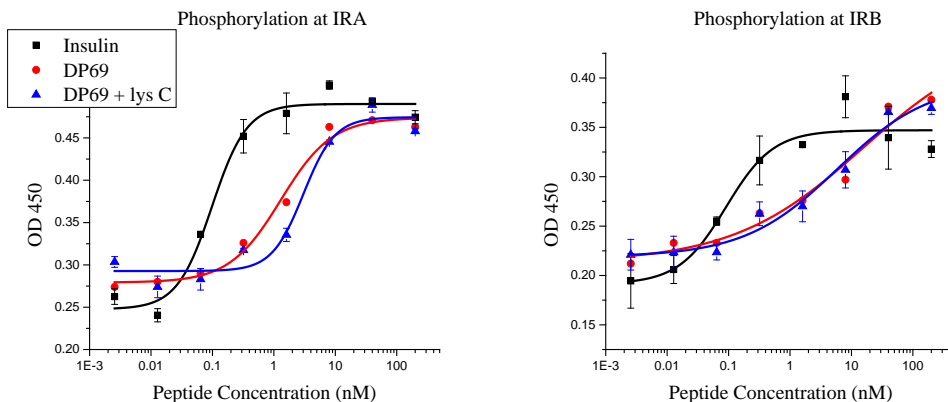


Figure 14: Phosphorylation assay demonstrating the activity of the single chain DP69 and the *lys-C* treated two chain form of DP69.

The next set of biosynthetic heterodimers contained the two binding motif antagonists in the C domain. DP36 contained the #4-6 antagonist. DP50 contained the #6-4 antagonist. DP36, as a single chain molecule, shows a small amount of agonism at both insulin receptor isoforms (Figure 15). Treatment with *lys-C* results in the two chain form of the molecule, which shows no agonism at either receptor isoform (Figure 15). It was also investigated whether the two chain form of DP36 was able to antagonize native insulin. When titrated against constant 1 nM insulin, high concentrations of *lys-C* treated DP36 show antagonism at both receptor isoforms (Figure 16).

DP50, identical to DP36 except for the relative order of the binding sites within the C domain, has radically different activity. As a single chain molecule, it is inactive at both insulin receptor isoforms (Figure 17). However, treatment with *lys-C* creates the two chain form of the molecule, and a significant increase in activity is seen at both insulin receptor isoforms (Figure 17). This observation was further validated by comparing the *lys-C* treated forms of DP36 and DP50 within the same phosphorylation assay (Figure 18). This demonstrates the significant difference in activity between these two dimers.

The final heterodimer in this series is DP71, which contains the #12 peptide in the C domain. As a single chain molecule, DP71 shows reduced potency but full activity at the insulin receptor (Figure 19). When treated with *lys-C* the two chain form of the molecule is generated, with the insulin C peptide ligated to the C-terminus of the B chain. The *lys-C* treated form of the peptide regains some, but not all of the potency of native insulin (Figure 19).

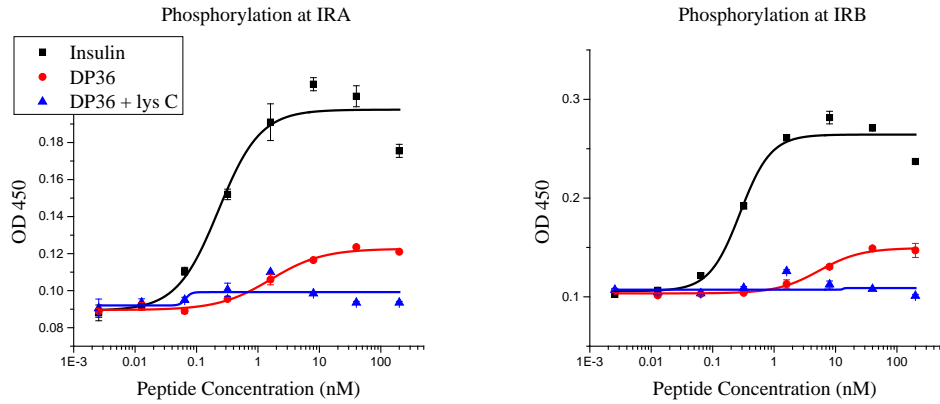


Figure 15: Phosphorylation assay demonstrating the agonism of the single chain DP36 and the *lys-C* treated, two chain form of DP36.

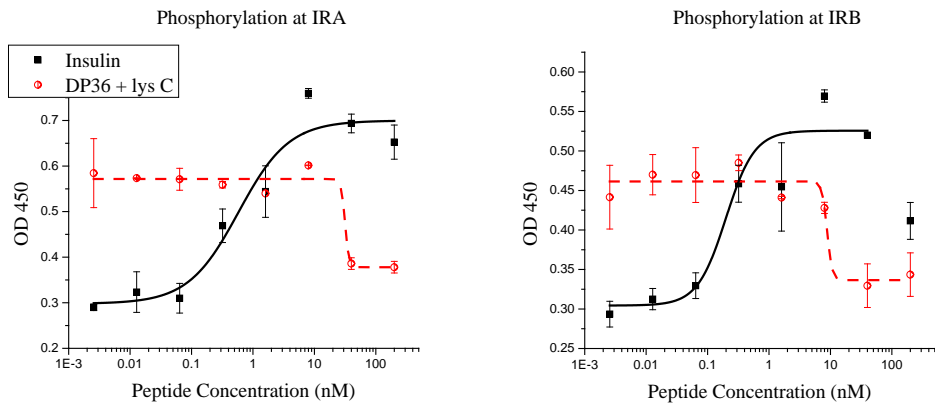


Figure 16: Phosphorylation assay demonstrating the antagonism of the two chain form of DP36 by titrating *lys-C* treated DP36 against constant 1 nM insulin.

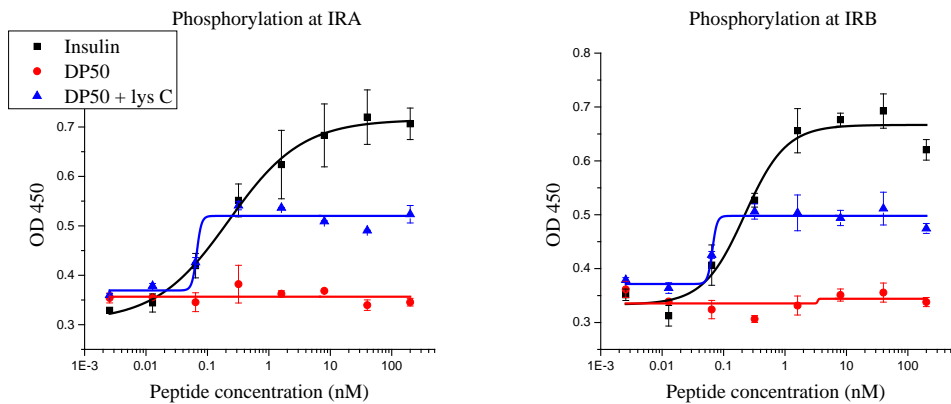


Figure 17: Phosphorylation assay demonstrating the agonism of single chain DP50 and the *lys-C* treated, two chain form of DP50.

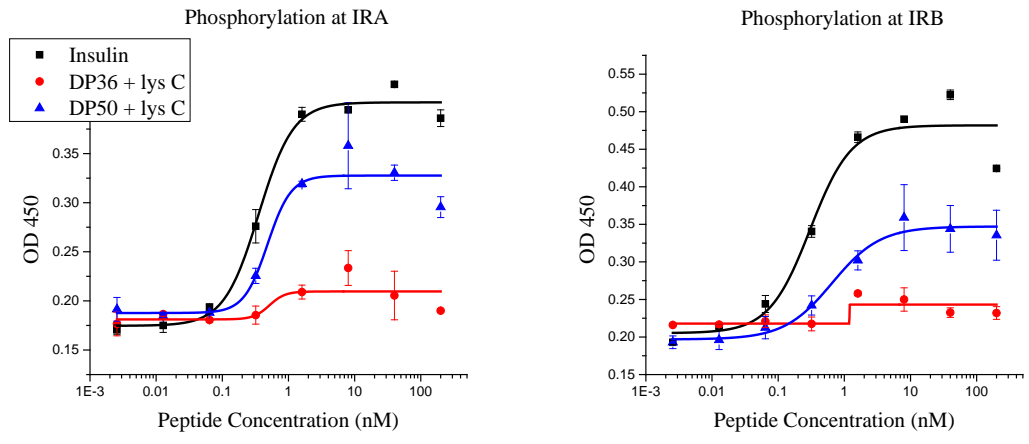


Figure 18: Phosphorylation assay demonstrating the agonism of the *lys-C* treated, two chain forms of DP36 and DP50.

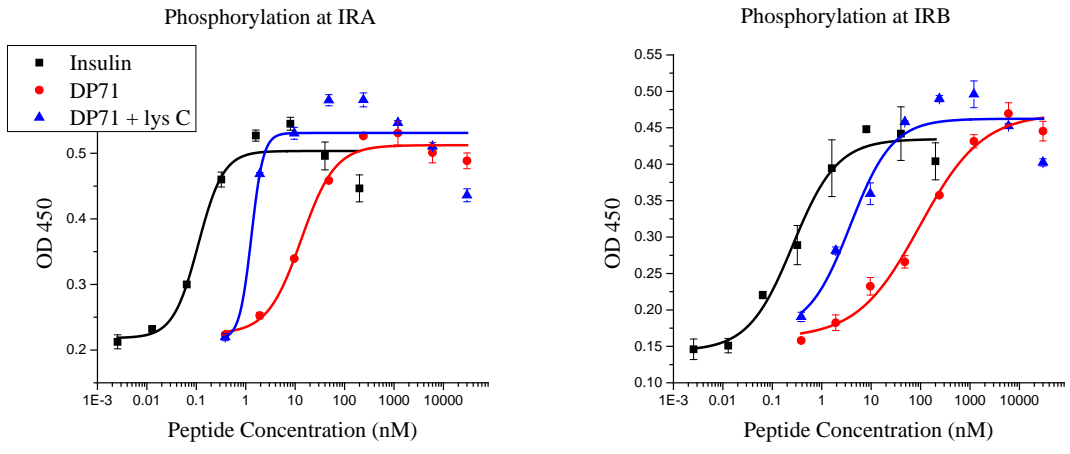


Figure 19: Phosphorylation assay demonstrating the agonism of the single chain DP71 and the *lys-C* treated, two chain form of DP71.

DISCUSSION

Heterodimers DP38 and DP72 yielded two key observations. First, and most importantly, when the C domain of these heterodimers is completely removed, the resultant agonist portion of the molecule is a fully active and potent agonist of the insulin receptor isoforms. This observation demonstrates that the agonist portion is formed correctly, and therefore biosynthesis is a viable method for creating heterodimers. Secondly, DP38 and DP72 demonstrate that activity of a heterodimer as a single chain entity can be influenced by the peptide present in the C domain. DP72 is capable of fully activating the insulin receptors, albeit with reduced potency. However, DP38, as a single chain entity, displays no activity at the insulin receptor. There are several possible explanations for this observation. First, it is possible that DP38 is also capable of activating the insulin receptors, but that a decrease in potency has put the activity outside of the range of tested concentrations. Another explanation is that the specific C domain, in this case peptide #4-6, sterically precludes binding, or does not allow sufficient flexibility in the B chain of DP8 to accommodate binding. A third possibility is that the peptide #4-6 is itself binding to the insulin receptor and precluding the agonist portion of the molecule from activating the receptor. Regardless of the mechanism of action, DP38 shows that the activity of a heterodimer can be influenced by the C domain in a sequence specific manner.

Subsequent heterodimers offer additional insight. Heterodimers DP66 and DP68 each contain a single binding motif in the C domain. As single chain molecules, they show some activity at the highest concentrations assayed. This suggests that as a single chain molecule, a single binding motif in the C domain is capable of affecting potency. The

relatively small size of the #4 and #6 motifs may also be responsible for the decrease in potency, since previous research has shown that single chain insulins with small C domains lose activity at the insulin receptor[13]. The constrained nature of the small C domain, however, is removed by treatment with *lys-C*, which results in a two chain molecule with the binding motif positioned at the N-terminus of the A chain. However, as two chain molecules, a decrease in activity is observed. It is ambiguous whether this decrease is the result of specific binding to the insulin receptor by the #4 or #6 peptide, or whether it is simply the result of hindering the binding of the agonist portion of the molecule. The N-terminal residues of the agonist A chain are critical for binding[1], at it may be that the additional flexibility granted by cleavage with *lys-C* allows the #4 and #6 peptides to disrupt binding. Regardless, these heterodimers do not display the desired characteristics of high potency but lowered maximal activity.

DP37 and DP51 contain two binding motifs within the C domain, either as #4-6 or #6-4. As single chain molecules, both constructs are inactive at the insulin receptor isoforms. Treatment with *lys-C*, which results in the antagonists positioned at the N-terminus of the A chain, does not result in any increase in agonism. As with the single-site heterodimers, these peptides are precluding insulin receptor activation, either through a sequence specific antagonism or through non-specific steric preclusion.

DP70, the heterodimer with the #12 insulin C peptide as the C domain, begins to address the question of whether the previous heterodimers were acting in a sequence specific or a non-specific manner. It has been shown that the C peptide alone does not bind to the

insulin receptor[15], and our research has shown that the **#12 C** peptide alone is neither an agonist nor an antagonist at the insulin receptor isoforms (chapter 2). Also, the C peptide is similar in length to the two binding motif antagonists used in DP37 and DP51. As a single chain construct, DP70 displays full activity at the insulin receptors, with reduced potency. This is not unexpected, because DP70 is a mimic of proinsulin, which is also known to activate the receptor with reduced potency[14]. When treated with *lys-C*, resulting in the **#12** peptide positioned at the N-terminus of the A chain, DP70 does not gain any activity or potency. *Lys-C* treated DP70 is analogous to split proinsulin, which is known to have reduced biological activity[12]. Therefore, the fact that we see reduced potency in both the single chain and two chain forms of DP70 conforms to previous observations of proinsulin and split proinsulin.

By comparing the activity of all the other heterodimers to DP70, we have evidence that lack of agonism in these heterodimers is sequence specific. It appears that single binding site motifs are capable of dramatically decreasing the potency of these molecules relative to insulin, whether the binding motif is present as the C domain or is positioned at the N-terminus of the A chain. Dimers containing the two binding site motif antagonists, **#4-6** and **#6-4**, are completely inactive at the insulin receptor, regardless of whether the antagonist is present as the C domain or is positioned at the N-terminus of the A chain, although it appears that the antagonist is slightly more effective at suppressing activity when positioned at the A chain.

Heterodimers where the peptide of interest is positioned at the C-terminus of the B chain yielded more illuminating results. DP67, which contains the #4 binding site 1 motif, was fully active but displayed decreased potency as a single chain molecule. Treatment with *lys-C*, however, restored nearly all of the potency, resulting in a two chain molecule with full potency and full activity. This demonstrates that it is the restricted nature of the single chain molecule that is responsible for the suppressive action of the binding site 1 motif #4. In contrast, DP69, which contains the site 2 motif #6, displayed full activity and reduced potency whether the #6 motif was present as the C domain, or positioned at the C-terminus of the B chain. Again, regardless of the mechanism of action, the decreased potency but full activity of these peptides makes them fall short of our aim of high potency, but lowered activity.

Heterodimers DP36 and DP50 both contain two binding site motifs in the C domain. DP36 contained the #4-6 antagonist. As a single chain molecule, it displayed no agonism. As a two chain molecule, where the #4-6 antagonist is positioned at the C-terminus of the B chain, it also displayed no agonism. The two chain form was capable of antagonizing native insulin, indicating that despite the lack of agonism, two chain DP36 is binding to the insulin receptor isoforms. This indicates that the lack of activity of DP36 is not simply due to a lack of binding.

DP50 contains the #6-4 antagonist as the C domain. As a single chain molecule, it displayed no agonism. As a two chain molecule, with the #6-4 antagonist positioned at the C-terminus of the B chain, a significant increase in maximal activity was observed. This is in contrast to DP36, and indicates that the relative orientation of the binding sites

within the antagonist portion of the heterodimer affects the maximal activity of the heterodimer. The mechanism of action of this reduced activity is unclear, and the phosphorylation assay alone does not offer mechanistic insight. However, there are several plausible explanations. First, it is possible that the antagonist portions of these heterodimers are binding to unoccupied binding sites in the insulin receptor, and that the **#4-6** orientation is more effective than the **#6-4** orientation. Secondly, it is possible that the antagonist portion of these molecules is “folding back” and interacting with the DP8 portion of the dimer. This would most likely be a non-specific hydrophobic interaction, and it may be possible that the **#4-6** antagonist binds more tightly, thus more effectively suppressing activity. Despite the ambiguity of the mechanism, the results of DP71, which contains the **#12**, non-binding peptide, demonstrate that this is a sequence-specific effect.

The sequence-specific reduced activity of DP50 is groundbreaking for several reasons. First, this represents the first instance of a heterodimer displaying full potency but reduced maximal activity at the insulin receptor isoforms. This is both a fusion of previous insulin research as well as a technical achievement. Our observations with DP50 validate our hypothesis and simultaneously open up a new avenue of research for insulin-based therapies.

Despite this achievement, there is room for advancement. The lowered activity of DP50 represents a single alternative maximal activity. Ideally, we would like to be able to direct maximal activity to a range of values. In addition, the agonist portion of our heterodimers, DP8, is not the ideal candidate for a heterodimer-based insulin therapy, since DP8 has the potential to induce mitogenic effects through the IGF-1 receptor.

Therefore, the next step in this research is to create heterodimers where native insulin functions as the agonist.

METHODS

Biosynthesis occurred in *E. coli* OrigamiDE3 cells, by transfection and induction of plasmids coding for the peptide of interest. OrigamiDE3 cells are resistant to kanamycin and tetracycline. Plasmids were created using a modified version of the Clontech “In-fusion” protocol[16], where overlapping primers and PCR were used to create the gene of interest, which was then incorporated into a plasmid. Primers were designed by Dr. Pengyun Li of the DiMarchi lab and were purchased from IDT technologies. PCR was used to amplify plasmids for transfection. DNA sequencing was performed by the IU biotechnology center to confirm that the sequence was correct. Every plasmid encoded for resistance to ampicillin. The plasmids also encoded for the protein of interest, along with an N-terminal SUMO/polyhistidine tag. The SUMO/polyhistidine/protein coding region was downstream of a *lac* promoter, in order to induce protein expression. The plasmid was then inserted into competent cells via heat shock in a 40°C water bath for 1 minute, followed by incubation on ice for 30 minutes. After transformation, these cells were introduced to 10 mL of LB media, with the antibiotics kanamycin, tetracycline and ampicillin. After 1 hour of aerobic growth at 37°C, with shaking, the entire media was added to 200 mL of LB media, with antibiotics, and shaken at 37°C overnight. Next, the 200 mL culture was added to a 1 L culture, with antibiotics, which was allowed to grow to an optical density of 0.8, after which it incubated with shaking for 1 hr at 18°C, and then induced with 400 µL of 0.2 mM IPTG (a lactose analog), to promote protein

synthesis. The induced culture was incubated with shaking overnight at 18°C, after which the culture was spun down by centrifugation and the excess media decanted and discarded. The pellet of cells was then sonicated in a 10mM imidizol buffer (50 mM sodium phosphate, 300 mM NaCl) to lyse the cells. The cell lysate was then poured over a nickel affinity column containing Ni-NTA agarose matrix (Qiagen). Due to the polyhistidine tag, the protein of interest is removed from the cell lysate by binding to the nickel column. The column was then washed with 30 mL aliquots of increasing concentrations of imidazole in 50 mM sodium phosphate, 300 mM NaCl, pH 8. Sequentially, 10, 15, and 40 mM imidazole solutions were used. The biosynthetic peptide was then eluted from the affinity column with 50 mL of 500 mM imidazole buffer. The eluent was then treated with TEV protease to remove the SUMO tag. In order to purify the peptide, the peptide/SUMO mixture was first desalted on a G-25 Sephadex matrix with 25 mM sodium bicarbonate buffer. Next, the protein mixture was purified via anion exchange with Q-Sepharose tertiary amine matrix, with an increasing gradient of 20 mM tris(hydroxymethyl)aminomethane (Tris), 300 mM NaCl, 10% glycerol (Figure 20).

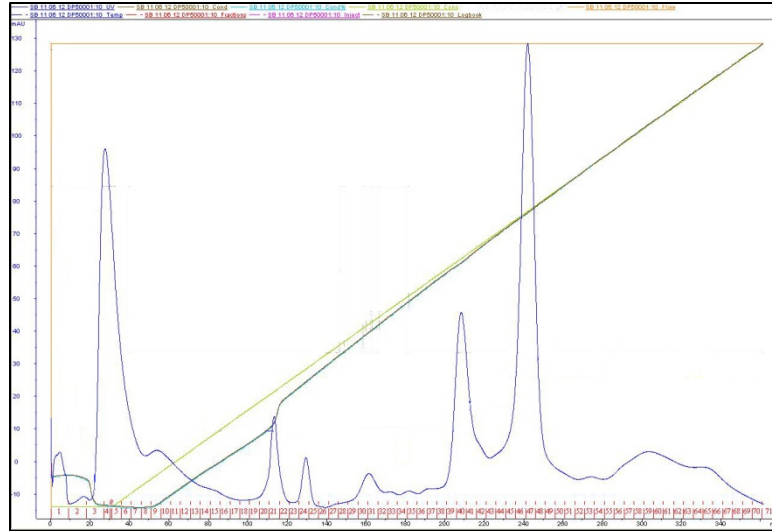


Figure 20: Anion exchange purification of peptide DP50. Fractions 46-48 were analyzed with HPLC and MALDI and found to contain peptide of the correct hydrophobicity and mass.

Peptides were analyzed for purity using analytical HPLC and MALDI. To generate the two chain form of the peptide, an aliquot of endopeptidase *lys-C* was added to concentrated (~500 μ M) peptide solutions in 50 mM ammonium bicarbonate buffer, pH 8. Cleavage of the peptide was confirmed with MALDI.

Activity of the heterodimers was investigated in vitro using a phosphorylation assay, as described in chapter 2. Optical density at 450 nm was graphed as a function of peptide concentration in OriginPro 9.0 graphing software. Schematics presented in this chapter were created using ChemBioDraw Ultra 13.0.

REFERENCES

1. Mayer, J.P., F. Zhang, and R.D. DiMarchi, *Insulin structure and function*. Biopolymers, 2007. **88**(5): p. 687-713.
2. Ochman, H., J.G. Lawrence, and E.A. Groisman, *Lateral gene transfer and the nature of bacterial innovation*. Nature, 2000. **405**(6784): p. 299-304.
3. Itakura, K., et al., *Expression in Escherichia coli of a chemically synthesized gene for the hormone somatostatin*. Science, 1977. **198**(4321): p. 1056-63.
4. Goeddel, D.V., et al., *Expression in Escherichia coli of chemically synthesized genes for human insulin*. Proc Natl Acad Sci U S A, 1979. **76**(1): p. 106-10.
5. Prinz, W.A., et al., *The role of the thioredoxin and glutaredoxin pathways in reducing protein disulfide bonds in the Escherichia coli cytoplasm*. J Biol Chem, 1997. **272**(25): p. 15661-7.
6. Novagen. *Origami(DE3) Competent Cells*. 2014 [cited 2014 June 9]; Available from: http://www.emdmillipore.com/life-science-research/origamide3-competent-cells/EMD_BIO-70627/p_FUKb.s1Oh6IAAAEjTBI9.zLX.
7. Lu, W.G., et al., *Production and Characterization of Hirudin Variant-1 by SUMO Fusion Technology in E-coli*. Molecular Biotechnology, 2013. **53**(1): p. 41-48.
8. Heath, W.F., et al., *(A-C-B) human proinsulin, a novel insulin agonist and intermediate in the synthesis of biosynthetic human insulin*. J Biol Chem, 1992. **267**(1): p. 419-25.
9. Froesch, E.R., et al., *Actions of insulin-like growth factors*. Annu Rev Physiol, 1985. **47**: p. 443-67.
10. Davidson, H.W., *(Pro)Insulin processing: a historical perspective*. Cell Biochem Biophys, 2004. **40**(3 Suppl): p. 143-58.
11. Steiner, D.F., *The proinsulin C-peptide--a multirole model*. Exp Diabetes Res, 2004. **5**(1): p. 7-14.
12. Temple, R.C., et al., *Insulin deficiency in non-insulin-dependent diabetes*. Lancet, 1989. **1**(8633): p. 293-5.
13. Rajpal, G., et al., *Single-chain insulins as receptor agonists*. Mol Endocrinol, 2009. **23**(5): p. 679-88.

14. Malaguarnera, R., et al., *Proinsulin binds with high affinity the insulin receptor isoform A and predominantly activates the mitogenic pathway*. *Endocrinology*, 2012. **153**(5): p. 2152-63.
15. Zierath, J.R., et al., *C-peptide stimulates glucose transport in isolated human skeletal muscle independent of insulin receptor and tyrosine kinase activation*. *Diabetologia*, 1996. **39**(3): p. 306-13.
16. Withers-Martinez, C., et al., *PCR-based gene synthesis as an efficient approach for expression of the A+T-rich malaria genome*. *Protein Eng*, 1999. **12**(12): p. 1113-20.

Chapter 4

Semisynthetic Heterodimers

ABSTRACT

This chapter describes semisynthetic heterodimers created by linking previously investigated insulin receptor antagonists to native insulin via a disulfide bond. Native insulin was modified at the B²⁹ residue to contain an activated thiol group, which reacts with any free sulfhydryl. Therefore, the introduction of a peptide with a cysteine residue results in disulfide bond formation and creates a heterodimeric complex with insulin. It was found that a specific insulin heterodimer, with previously characterized peptide #6 chemically conjugated through a C-terminal cysteine, displayed only 20-30% of the maximal activity of native insulin. This implies that the binding motif of peptide #6 is sufficient to suppress the ability of native insulin to activate its receptors. An alanine scan of the #6 peptide revealed that the leucine residue at position two is a key regulator of the maximal activity of the heterodimer. Mutating this leucine residue resulted in a series of peptides with similar EC₅₀ values, but varying levels of maximal activity. It was found that the maximal activity correlated with hydrophobicity of the side chain at position two. This set of heterodimers represents a tunable molecular tool for investigations of the biological significance of the *in vitro* maximal activity, as it relates to the *in vivo* pharmacology and regulation of blood glucose by insulin-based peptides.

INTRODUCTION

This chapter describes the synthetic heterodimers created by conjugating insulin receptor antagonists to native insulin. In the previous chapter, the agonist portion of the heterodimer was DP8, a biosynthetic insulin analog. DP8 was used due to its increased aqueous solubility and served as a useful proof of concept that heterodimers could display altered maximal activity. However, native insulin is the most physiologically relevant agonist of the insulin receptor, and one goal of this research is to minimize the number of chemical differences necessary to achieve altered pharmacology. Therefore, we created semisynthetic heterodimers, where the agonist portion of the molecule is native insulin. Based on the biosynthetic observations, we chose to position the antagonist at the C terminus of the B chain of insulin, as this is the orientation that resulted in variable maximal activity in the DP8-based biosynthetic heterodimers.

There are several synthetic chemistries that can be utilized to conjugate peptides to one another. The largest challenge in creation of these specific heterodimers is the identification of chemistry that is specific, irreversible, and occurs under conditions that do not disrupt the native structure of insulin, which has three disulfide bonds. The field of peptide chemistry encompasses many conjugation techniques, such as native chemical ligation, click chemistry, the Staudinger ligation, and disulfide bond formation, all of which are described below. However, due to the chemically delicate nature of native insulin, only disulfide bond formation was used to create the desired heterodimers.

The beginnings of what is now called native chemical ligation can be traced as far back as 1953 to a report by Wieland, et al[1]. They report the ligation of two peptides through an N-terminal cysteine and a C-terminal thioether. After displacement of the thiolate by the nucleophilic N-terminal amine, an intramolecular S to N acyl transfer forms a native peptide bond (Figure 1). One disadvantage of native chemical ligation is the necessity of a cysteine residue, which may not be native to a convenient ligation position. Another, more significant disadvantage is the potential for disulfide bond-scrambling under reaction conditions. Insulin, one of the targets for ligation, contains 3 disulfide bonds and the scrambling of even a single bond results in an inactive molecule.

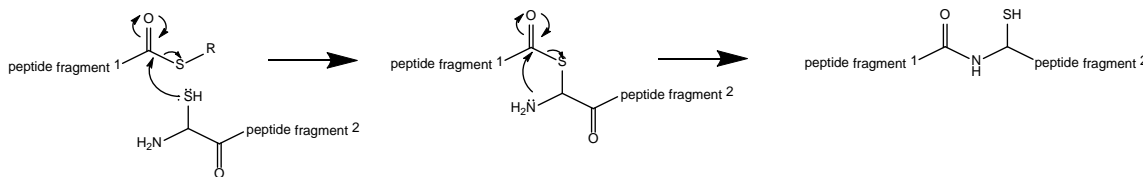


Figure 1: Native chemical ligation.

Another conjugation chemistry currently utilized is expressed protein ligation. This synthetic methodology is a mimic of a naturally occurring peptide element, an intein. An intein is a self-splicing unit of a protein, analogous to a genetic intron, in that it is not incorporated into the final product. Inteins are spliced similarly to the mechanism of native chemical ligation. Due to the constrained and twisted nature of the intein secondary structure, an intramolecular N to S acyl shift occurs, followed by a transthioesterification and S to N acyl shift, which results in the excision of the intein and the formation of an amide bond (Figure 2)[2]. This approach is often used to introduce differently labeled protein segments or unnatural amino acids. While an effective ligating

method, this chemistry is not the best for our heterodimers because it requires the synthesis of the intein, and the preponderance of thioethers could lead to disulfide bond shuffling.

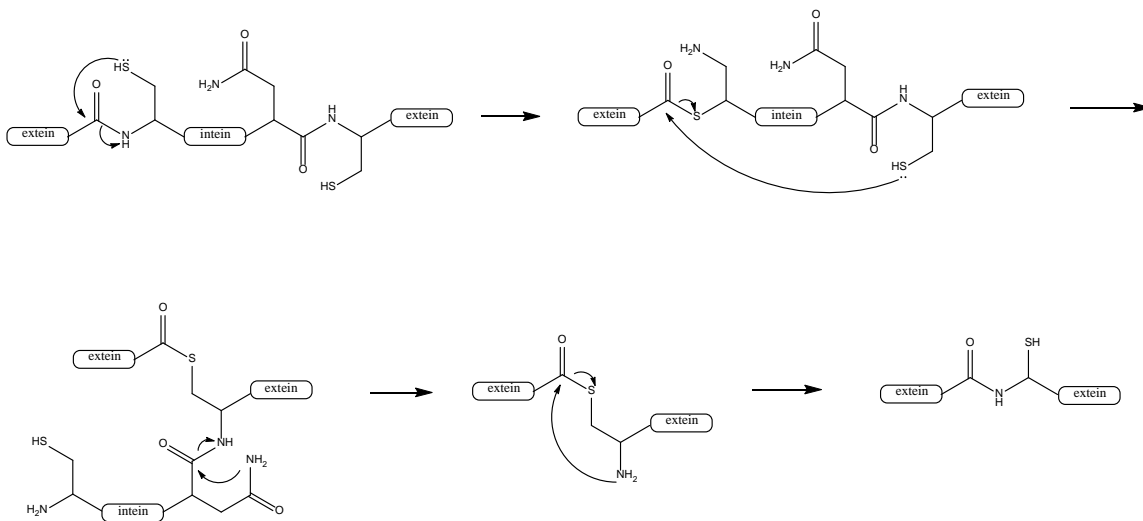


Figure 2: Schematic of expressed protein ligation.

Another relatively new but commonly used ligation method is termed “click chemistry.” It employs an alkyne and an azide to form a stable triazole ring (Figure 3). Click chemistry can be performed at room temperature and under mild conditions, although it does require a copper catalyst, which can present challenges in complete removal prior to *in vivo* study[3].

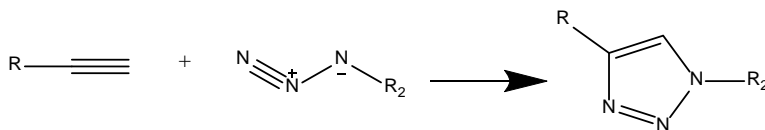


Figure 3: Click chemistry.

A fourth ligation chemistry is the Staudinger ligation, which was introduced in 2000 by Saxon and Bertozzi[4]. In this reaction, a bond is formed between an azide and a triaryl phosphine (Figure 4). The intermediate in this reaction is an aza-ylide, which is susceptible to hydrolysis. However, the inclusion of an electrophilic “trap” promotes an intramolecular cyclization that ultimately yields a stable amide bond. The triaryl phosphine can be used as a chemical handle, although this does incorporate an intervening triaryl phosphine oxide group between the two moieties[5]. A variation on this reaction can provide a “traceless” ligation, where the phosphine is hydrolyzed, leaving a simple amide bond[5]. A disadvantage of the Staudinger ligation is that it does have cross-reactivity with lysine residues[6].

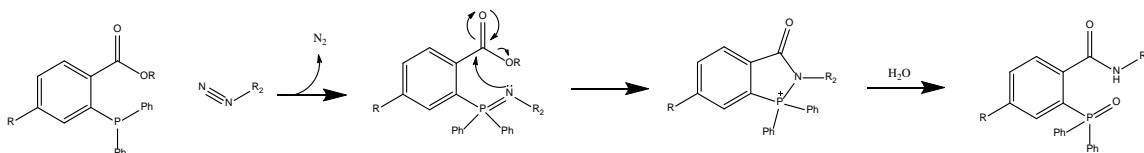


Figure 4: The Staudinger ligation.

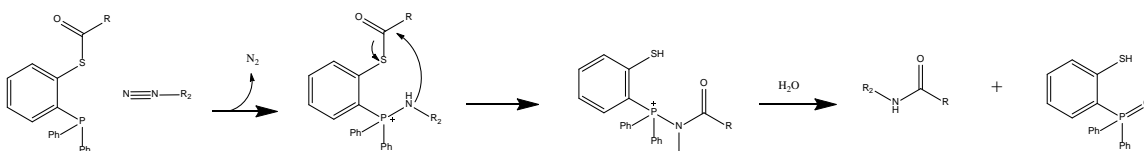


Figure 5: The traceless Staudinger ligation.

While all of these chemistries represent potential methodologies for forming heterodimers, in this research, we chose to promote the formation of a disulfide bond between native insulin and our antagonist peptides. To do so, we first chemically modified native human insulin by exploiting the unique reactivity of its only lysine residue, which is at position B²⁹. Since the B chain has only 30 residues, this lysine is

ideally situated near the C terminus of the B chain and is not involved in receptor binding. We were able to modify this residue by reacting native insulin with an N-hydroxysuccinimide (NHS) ester at an aqueous pH of 10. At this pH, only the ϵ -amine of the lysine residue reacts with the ester, sparing reactivity at the two unprotected N-terminal amines. The NHS ester contains a trityl-protected thiol at the other end of the reagent, such that native insulin can be modified to form an amide at B²⁹ that extends to include the protected thiol function group. Removal of the trityl-protecting group, followed by activation with 2,2'-dithiobis(5-nitropyridine) (DTNP) yields an insulin analog that will readily react with free thiols in aqueous solution at relatively low pH to form a disulfide bond. Therefore, an array of heterodimers could be synthesized by reacting the activated insulin with peptides that contain a single cysteine residue (Table 1). The advantage of this approach is that it minimizes the number of modifications required to covalently bond unique antagonist peptides to insulin. It is also highly specific, and compatible with peptides created by solid phase peptide synthesis. The heterodimers created using this approach were tested for the degree of biological action at the insulin receptor isoforms. Further optimization of the specific sequence of the antagonist followed by analogous conjugation to insulin led to the discovery of a library of heterodimers possessing high potency and variable maximal activity.

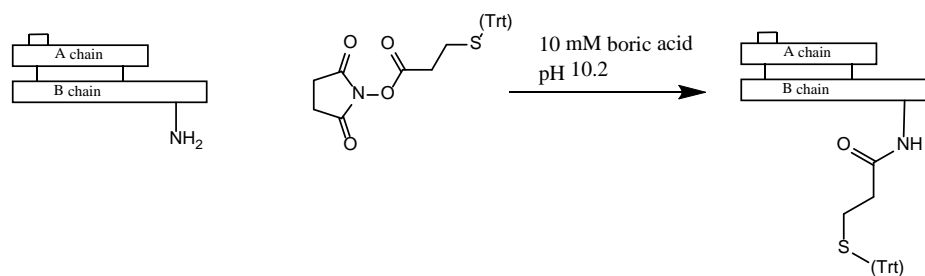


Figure 6: Reaction scheme illustrating the chemical modification of the insulin B²⁹ residue to generate an analog with a trityl-protected thiol group.

Table 1: Names and sequences of peptides containing single cysteine residues

Peptide Reference Number	Sequence
Cys4	CGSLDESFYDWFERQLG
4Cys	GSLDESFYDWFERQLGC
Cys6	CSLEEEWAQIQSEVWGRGSPSY
6Cys	SLEEEWAQIQSEVWGRGSPSYC
6(des1-5)Cys	WAQIQSEVWGRGSPSYC
6(A1)Cys	ALEEEWAQIQSEVWGRGSPSYC
6(A2)Cys	SAEEEWAQIQSEVWGRGSPSYC
6(A3)Cys	SLAEEWAQIQSEVWGRGSPSYC
6(L2)Cys*	SLEEEWAQIQSEVWGRGSPSYC
6(dL2)Cys**	SLEEEWAQIQSEVWGRGSPSYC
6(I2)Cys	SIEEEWAQIQSEVWGRGSPSYC
6(V2)Cys	SVEEEWAQIQSEVWGRGSPSYC
6(F2)Cys	SFEEEWAQIQSEVWGRGSPSYC
6(W2)Cys	SWEEEWAQIQSEVWGRGSPSYC
6(Y2)Cys	SYEEEWAQIQSEVWGRGSPSYC
6(Q2)Cys	SQEEEWAQIQSEVWGRGSPSYC

**Identical to #6Cys, renamed to emphasize the identity of the residue at position two.*

***dL refers to the d-stereochemistry of the leucine residue at position two. All other amino acids are of l-stereochemistry.*

RESULTS

Native insulin was modified through the use of a specific NHS ester to create an insulin molecule with one additional thiol functional group at the B²⁹ lysine (Figure 6). In order to ensure that this modification did not interfere with binding to the insulin receptor, the modified insulin was tested for agonism at both insulin receptors (Figure 7). It was found that the modified insulin was still capable of fully activating the insulin receptor isoforms, although with slightly reduced potency.

The modified insulin was chemically activated for thiol-conjugation by removing the trityl protecting group with TFA and reacting the resulting free thiol with DTNP. The activated insulin was purified and subsequently conjugated to peptides containing a free thiol to yield a specific disulfide bond.

The first set of heterodimers consisted of the modified insulin and the site 1 binding motif possessed by peptide #4 (Table 1). Insulin conjugated to #Cys4 (**#Insulin-Cys-4**) displays only a small change in maximal activity and a small decrease in potency when compared to native insulin (Figure 8). However in contrast the insulin conjugated to #4Cys (**#4-Cys-Insulin**) does display some degree of antagonism at both receptor isoforms at higher peptide concentrations (Figure 8).

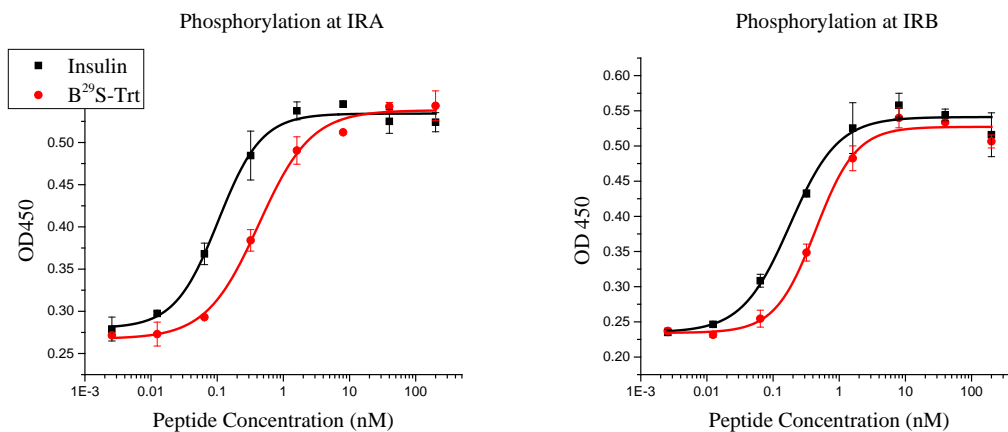


Figure 7: Phosphorylation assay demonstrating the agonism of the B²⁹ modified insulin.

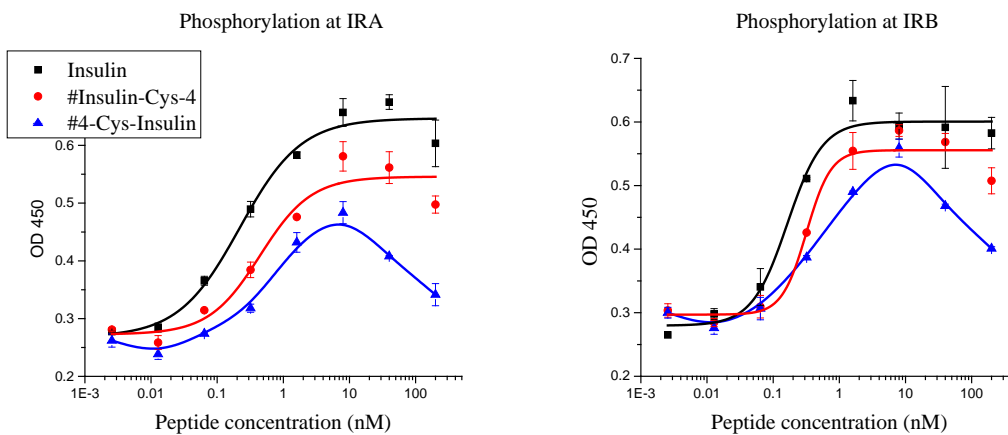


Figure 8: Phosphorylation assay demonstrating the agonism of #Insulin-Cys-4 and #4-Cys-Insulin.

The next set of heterodimers prepared and biologically characterized contains the site 2 binding motif possessed by peptide **#6** (Table 1). When the heterodimer is conjugated to **#Cys-6** (**#Insulin-Cys-6**), there appears a slight decrease in both potency and maximal activity (Figure 9). However, when insulin is conjugated to **#6-Cys**, the heterodimer is completely inactive at both insulin receptor isoforms (Figure 10). This is not merely the result of a lack of receptor binding, since this peptide, **#6-Cys-Insulin**, is capable of fully antagonizing exogenously added native insulin (Figure 10). This was a surprising result, since our previous work has shown that single binding site motifs, including **#6Cys**, are incapable of antagonizing native insulin when applied as a non-covalent addition (*in trans*). Therefore, the inclusion of this peptide into a heterodimer results in a molecule that has emergent antagonism, relative to what is observed for the individual constituent peptides.

We further investigated the activity of the **#6-Cys-Insulin** heterodimer by creating a truncated version of the **#6** motif, **#6(des1-5)-Cys**, which is devoid of the first five N-terminal amino acids (Table 1). The **#6(des1-5)-Cys-Insulin** heterodimer was tested for activity at the insulin receptor isoforms, and it was found that removal of the N-terminal pentapeptide of the **#6** binding motif restored the heterodimer to full activity, and full potency (Figure 11). This astonishing result demonstrates that the antagonistic activity of this heterodimer can be completely controlled by the first five residues of the **#6** motif.

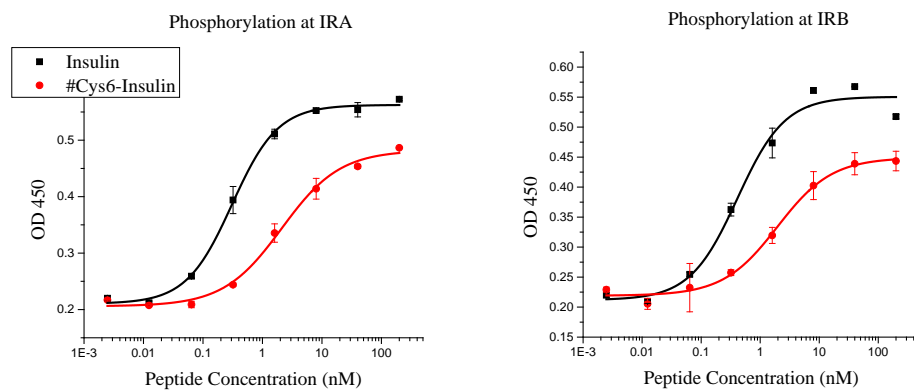


Figure 9: Phosphorylation assay demonstrating the agonism of #Insulin-Cys-6.

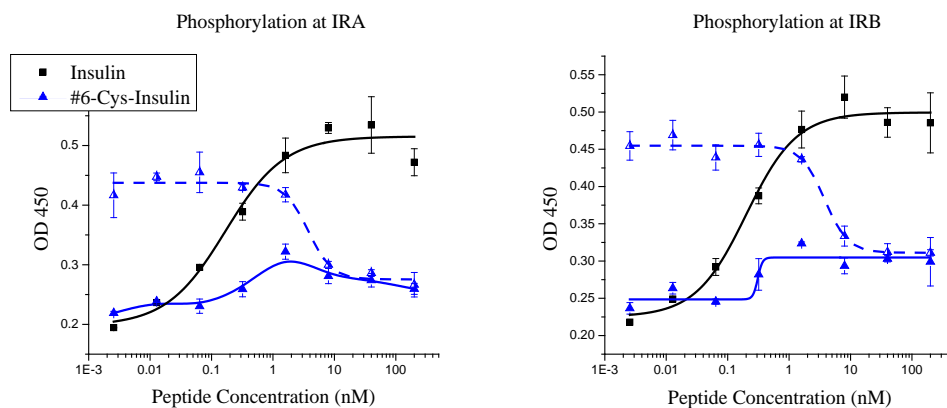


Figure 10: Phosphorylation assay demonstrating the agonism (solid) and antagonism (dashed) of #6-Cys-Insulin.

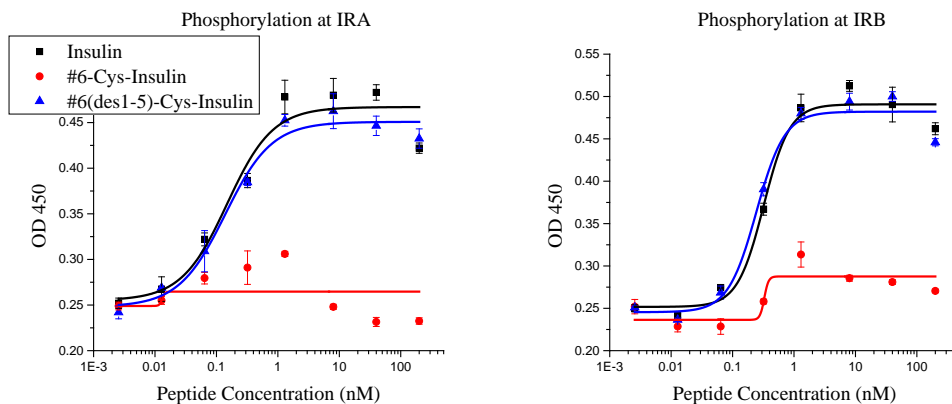


Figure 11: Phosphorylation assay demonstrating the agonism of the #6-Cys-Insulin heterodimer and the truncated version, #6(des1-5)-Cys-Insulin.

To interrogate this structure-activity relationship further, an alanine scan of the N-terminal residues of the #6 motif was employed to create heterodimers with site-specific mutation of the conjugated #6-Cys. Relative to the N-terminus of the #6 sequence, an alanine in the first position had almost no effect upon bioactivity, relative to the unaltered heterodimer (Figure 12). An alanine in the second position, however, restored almost all of the potency and activity to the heterodimer (Figure 12). An alanine in the third position seemed to shift the potency of the heterodimer, but had little effect on maximal activity (Figure 12). Based on these results, we shifted the attention to the second amino acid in the site 2 binding motif, a leucine. A series of mutations to the second position of the #6-Cys antagonist were made (Table 1) and these peptides were similarly conjugated via a disulfide to B²⁹ modified insulin. Most of these mutations affected the maximal activity of the heterodimer without a large shift in potency. It was found that leucine and isoleucine had very similar, low activities (Figure 13). However, substituting d-leucine at position two resulted in a low activity heterodimer with decreased potency (Figure 14). This suggests that there is an appropriate size, hydrophobicity and chirality that results in low activity heterodimers.

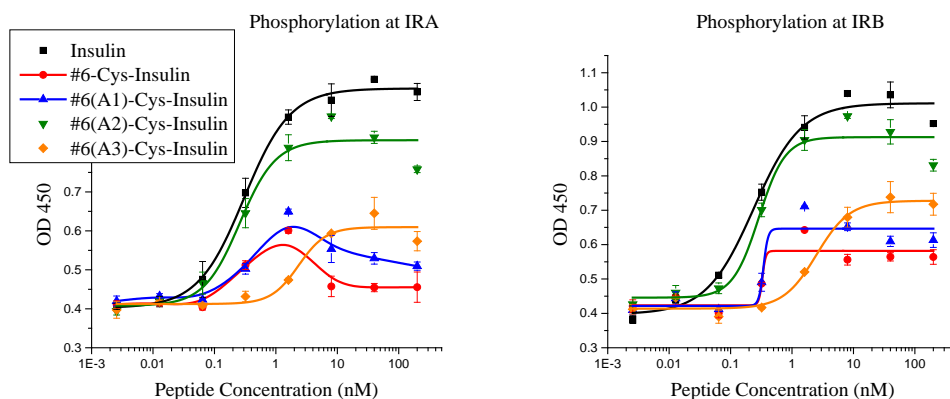


Figure 12: Phosphorylation assay demonstrating the agonism of the #6-Cys-Insulin alanine scans.

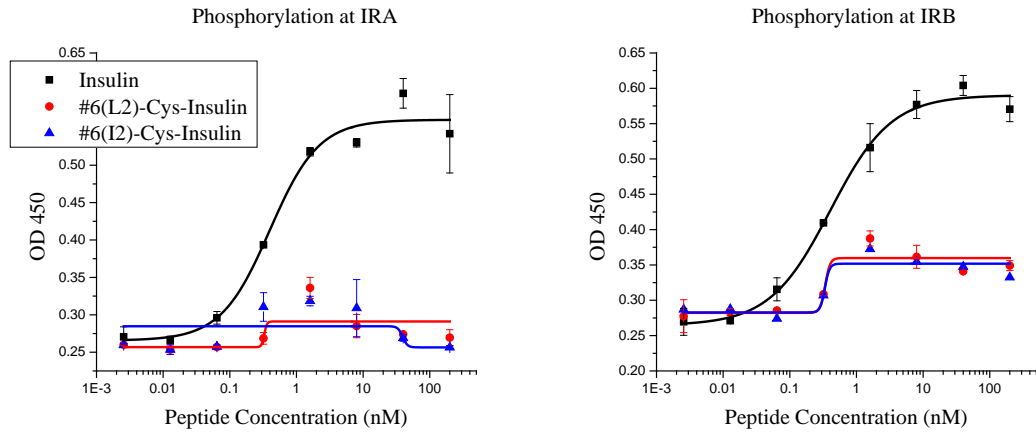


Figure 13: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(I2)-Cys-Insulin.

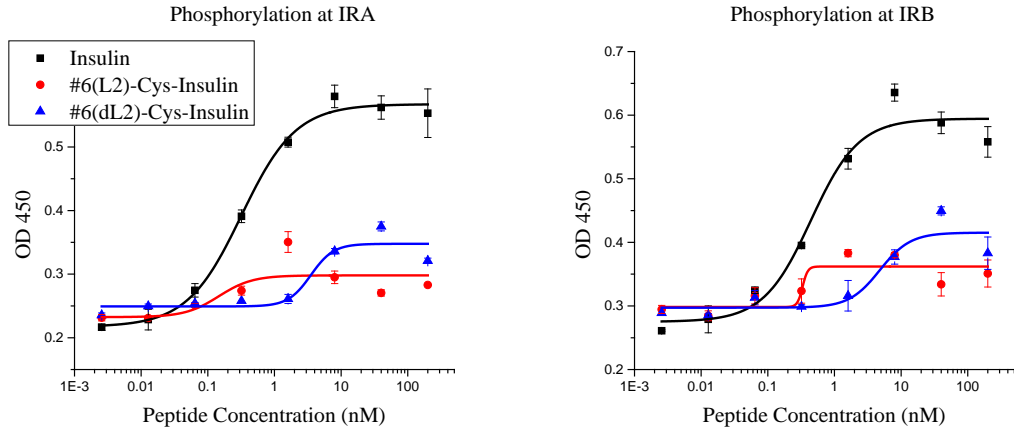


Figure 14: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(dL2)-Cys-Insulin.

Based upon the observations with leucine, isoleucine, and d-leucine, we decided to further explore hydrophobic residues at position two. Substituting to valine at position two results in a slight increase in activity (Figure 15). Substituting to phenylalanine at position two results in a further increase in activity, and begins to approach 50% activity at both receptor isoforms (Figure 16). Substituting to tryptophan at position two results in a heterodimer with approximately 60% the activity of native insulin (Figure 17), and the final substitution to tyrosine, results in a heterodimer with approximately 70% the activity of native insulin (Figure 18). The introduction of a more hydrophilic residue, glutamine, was also investigated at position two (Figure 19). This mutation resulted in a shift in potency, without a significant decrease in maximal activity. The most satisfying aspect of these modifications is that all of the heterodimers with hydrophobic residues at position two maintain a high inherent potency. This represents the discovery of a library of peptides with high inherent potency, but variable maximal activity, as was the primary goal of this research.

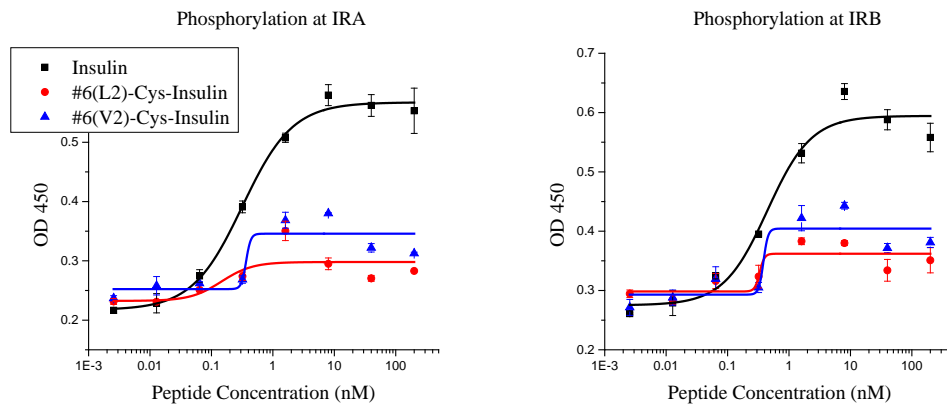


Figure 15: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(V2)-Cys-Insulin.

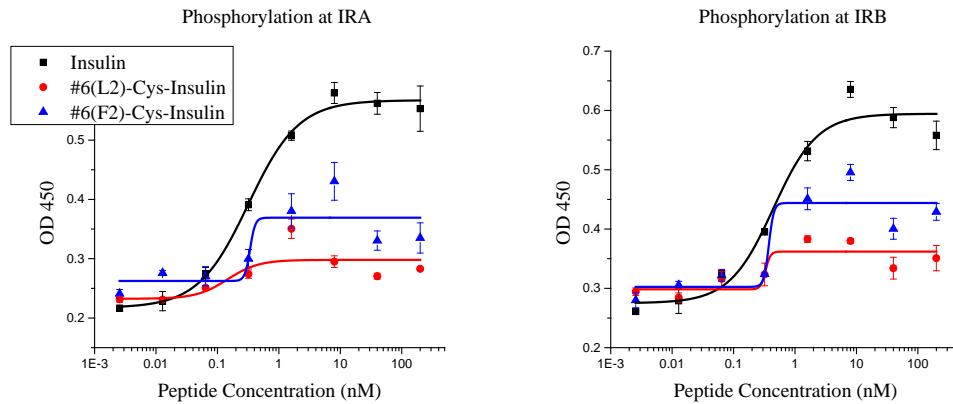


Figure 16: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(F2)-Cys-Insulin.

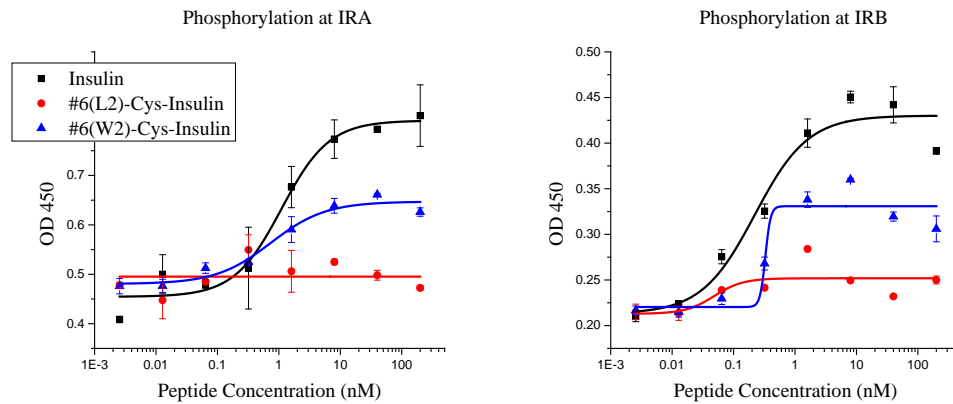


Figure 17: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(W2)-Cys-Insulin.

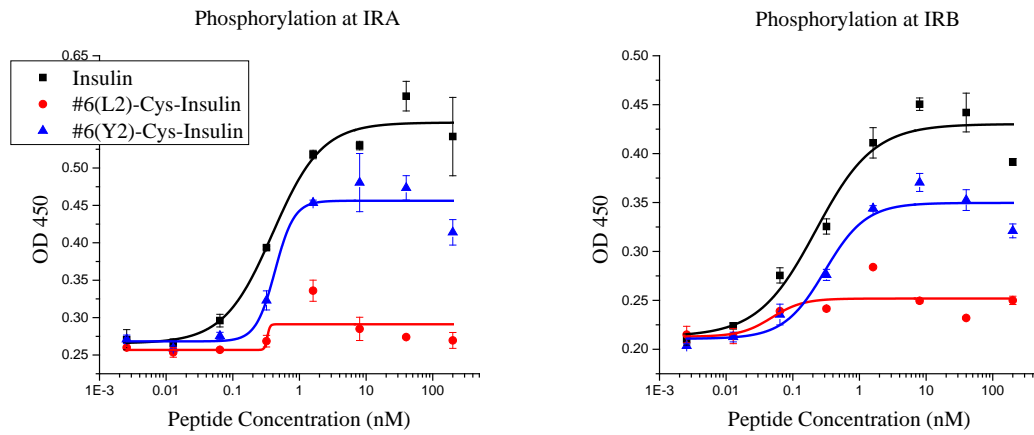


Figure 18: Phosphorylation assay demonstrating the agonism of #6(L2)-Cys-Insulin and #6(Y2)-Cys-Insulin.

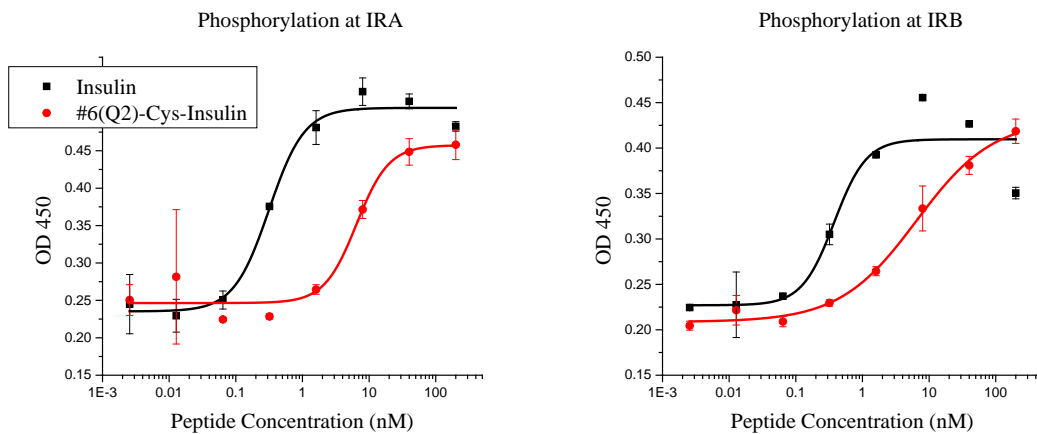


Figure 19: Phosphorylation assay demonstrating the decrease in potency but high of maximal activity of #6(Q2)-Cys-Insulin.

DISCUSSION

The first and most important observation in this chapter is that, in contrast to previous studies with monomeric peptides (chapter 2) and biosynthetic heterodimers (chapter 3), it was found that a single binding site motif, **#6**, is sufficient to suppress the activity of native insulin, when positioned at the C terminus of the insulin B chain. This was unexpected, since in chapter 2, it was demonstrated that the **#6** peptide was incapable of antagonizing native insulin on a 1:1 ratio. In chapter 3 it was demonstrated that the **#6** peptide, positioned at the C-terminus of the B chain, decreased the potency but did not affect the maximal activity of the heterodimer. Operating under the assumption that two binding motifs, such as **#4-6**, were required at the C terminus of the B chain, the heterodimer **#6-Cys-Insulin** was originally envisioned as a negative control. However, after observing the low activity of **#6-Cys-Insulin**, further investigation into the **#6** motif was conducted. It was found that removing the first five N-terminal residues resulted in a fully active molecule, indicating that the first five residues were necessary for suppressing activity. An alanine scan revealed that the most important residue, in terms of maximal activity, is the leucine at position two.

The introduction of alternate hydrophobic amino acids at position two in the **#6** motif created a set of peptides that do not vary greatly in their potency, compared to native insulin. Every peptide containing a hydrophobic residue at position two had a potency, as measured by EC_{50} values, near to the potency of native insulin (Figure 20). However, comparing the maximal activity, as a function of percent maximal activity of native insulin, we see that we have created a range of peptides from low to high maximal

activity (Figure 21). The peptides with mutations most similar to the original leucine residue (isoleucine and valine) show no significant increase in maximal activity. However, a valine residue begins to increase in maximal activity, while phenylalanine, tryptophan, and tyrosine all increase in maximal activity, with tyrosine achieving nearly 70% the activity of native insulin. An alanine mutation results in 80% of the maximal activity of insulin. This suggests that there is an appropriate size and hydrophobicity required for antagonism at the insulin receptors. Deviating too far in size, such as the much larger tyrosine and tryptophan, or the much smaller alanine, results in an increase in activity at the insulin receptors.

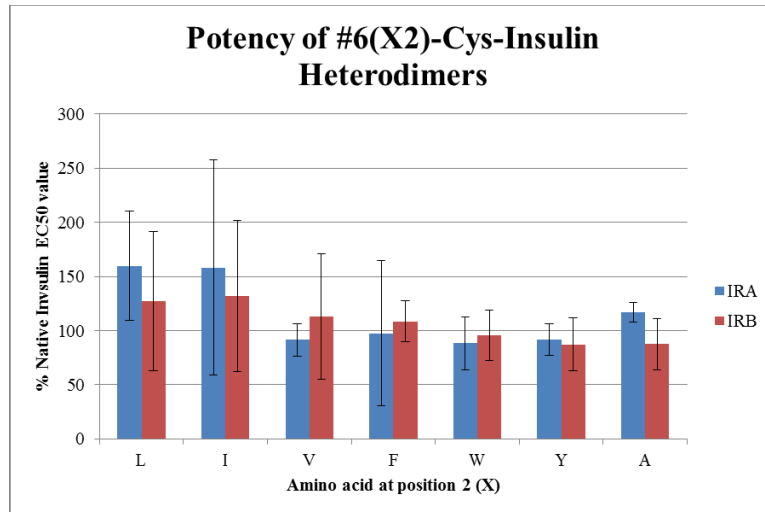


Figure 20: Comparative EC₅₀ values, as a percentage of native insulin, of #6(X2)-Cys-Insulin heterodimers, where X represents the amino acid at position two. Data shown for both receptor isoforms.

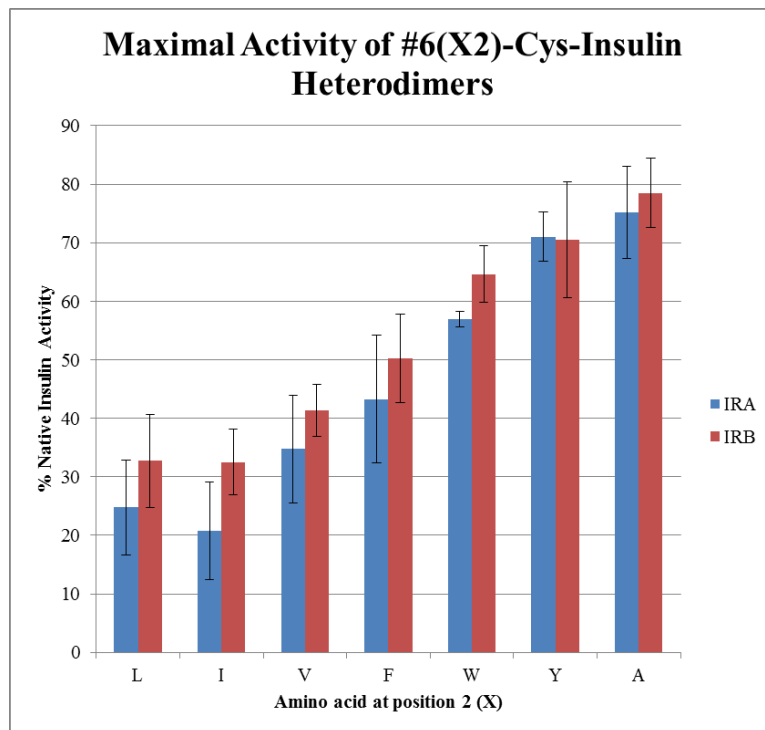


Figure 21: Comparative maximal activity, as a function of percentage of native insulin, of #6(X2)-Cys-Insulin heterodimers, where X represents the amino acid at position two. Data shown for both receptor isoforms.

The dependence on side chain hydrophobicity suggests that some form of non-specific hydrophobic binding is occurring to alter the maximal activity of these molecules at the insulin receptor isoforms. Under this hypothesis, the trend observed in the maximal activity of our mutants is easily explained. The isoleucine, leucine, and d-leucine mutants all contain the same number of methylene groups and comparable hydrophobic radii. They therefore appear to hydrophobically bind to either the insulin receptor or the conjugated insulin partner to comparably disrupt signaling. Valine, which differs by one less methylene group, is almost as effective at suppressing activity. More significant structural deviations from leucine, such as phenylalanine and tryptophan, result in higher maximal activity. Phenylalanine and tryptophan are both relatively hydrophobic residues, but they are larger and contain conjugated aromatic rings. The *in vitro* results show that these larger hydrophobic side chains are less effective at suppressing maximal activity, presumably due to a less effective hydrophobic association to whichever site is mediating this disruption in activity. When comparing phenylalanine and tyrosine, it can be seen that the inclusion of a hydroxyl group greatly impairs the suppression of activity. Even more dramatically, a glutamine side chain does not suppress activity but does decrease inherent potency. Taken together, these results show that suppression of activity is modulated by position two in the site 2 binding motif, and that hydrophobic residues at this position are the most effective suppressive agents.

These results, while powerful in illustrating the ability to create heterodimers with tunable maximal insulin receptor activity, do not address the underlying mechanism of

this phenomena. There are several reasonable possibilities, and the actual mechanism may be a combination of several of the following.

First, it is possible that the agonist and the antagonist are themselves interacting through a non-specific hydrophobic interaction. Support for this hypothesis comes from the biosynthetic and synthetic heterodimers. In our biosynthetic heterodimers, it was found that the orientations of the binding sites influenced maximal activity. Later, in our synthetic heterodimers, it was found that the leucine in position two of the #6 site 2 binding motif was critical to the degree of maximal activity. Applying the lessons learned from the synthetic heterodimers, we can examine the relative position of the leucine residue in the biosynthetic heterodimers. In DP37, where the antagonist was arranged in a #4-6 orientation, the key leucine was distal to the native insulin agonist portion of the molecule. This heterodimer exhibited very low maximal activity. In contrast, DP50 contained #6-4, where the key leucine residue is more proximal to the insulin agonist, and this heterodimer was a partial agonist. If the key interaction is a hydrophobic binding between the #6 motif and the insulin agonist, DP37 may contain sufficient flexibility to allow for the leucine residue to “fold back” upon itself. DP50, however, may not contain sufficient flexibility to completely fold back to present this hydrophobic interaction.

The same analysis could explain the difference between the synthetic heterodimers where the #6 motif was ligated through either a C- or N-terminal cysteine. In #Insulin-Cys-6, the important leucine residue is held closely to the B²⁹ residue of insulin, whereas in the

#6-Cys-Insulin, the leucine residue has more rotational freedom, relative to the insulin agonist.

This interpretation would also account for the fact that leucine, isoleucine, and d-leucine are all sufficient to suppress the activity of the heterodimer, by participating in a non-specific hydrophobic interaction. It is possible that in folding upon itself, the heterodimer is altered in its ability to bind to and signal through the insulin receptor.

However, there is also evidence to suggest that the antagonist portion of the heterodimer may be interacting with the insulin receptor itself. We created a set of heterodimers that do not vary in potency, but vary in maximal activity. The lack of activity is not simply due to a lack of binding, since these peptides have EC_{50} values similar to insulin. Since the heterodimers are binding but not signaling, it is possible that the antagonist portion of the heterodimer may be interacting with the insulin receptor and mimicking negative cooperativity. The heterodimer, as a whole, may be inefficiently crosslinking the two halves of the insulin receptor, or it may be occupying more than the usual number of receptor binding sites.

The actual mechanism may be a combination of these two hypotheses, therefore the reduction in maximal activity may be the result of interactions between the agonist and the antagonist, the receptor and the agonist, and the receptor and the antagonist.

The most important aspect of this research is the creation of a set of peptides with similar potencies but varying maximal activities at the insulin receptors. In addition, these maximal activities are tunable by a single point mutation within the antagonist portion of the heterodimer. The side chain at position two in the antagonist determines the maximal activity in a way that is both predictable and in keeping with current understanding of hydrophobicity and binding. These analogs will serve as useful tools in determining the physiological significance of maximal activity within insulin receptor analogues, and may eventually serve as the basis for tunable medicines aimed at treating diabetes mellitus.

The overarching hypothesis of this research is that heterodimers with altered maximal activity may have beneficial pharmacological properties. The next chapter will discuss the future directions of this work, specifically with the potential of these analogs to offer altered pharmacology *in vivo*.

METHODS

The heterodimers presented in this chapter were created by ligating antagonists to a modified insulin through a disulfide bond. Native insulin contains three primary amines, one at the N-terminus of the A chain, one at the N-terminus of the B chain, and a third present in the side chain of the only native lysine residue, the B²⁹ position. The difference in reactivity between the N-terminal amines and the lysine side chain amine allows for the specific functionalization of the lysine residue. Native insulin (Eli Lilly and Co), was reacted with a specific NHS ester. The ester itself was created by reaction of S-trityl- β -mercaptopropionic acid (National Biochemicals Corporation) with N-hydroxy succinimide (Chem Impex International) and diisopropylcarbodiimide (Aldrich) in a 1:1:0.9 ratio in anhydrous DMF (Aldrich) (Figure 22). After centrifuging to precipitate unwanted side products, the NHS ester was reacted in a 1:1 molar ratio with native insulin in 25 mM boric acid buffer, 50% acetonitrile, pH 10.2 (Figure 23). The modification was confirmed by LC/MS, which shows a single modification by mass. This peptide was purified using reverse phase chromatography with a silica based C8 column. The trityl protecting group was then removed with anhydrous TFA. The specificity of the B²⁹ modification was determined by subjecting the modified peptide to a trypsin digest, which removes the last eight residues of the B chain, including the B²⁹ lysine residue. The corresponding decrease in mass demonstrated that the N-terminal amines of the A and B chain remained unmodified (Figure 24).

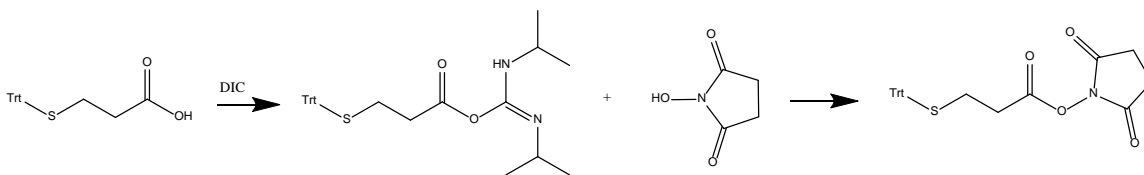


Figure 22: Reaction of S-trityl- β -mercapto propionic acid with diisopropylcarbodiimide (DIC) and N-hydroxysuccinimide.

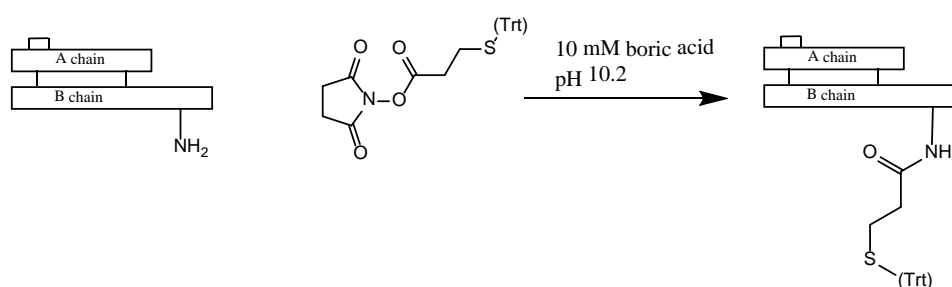


Figure 23: Reaction scheme illustrating the modification of the B²⁹ residue of insulin, resulting in a trityl-protected sulfide.

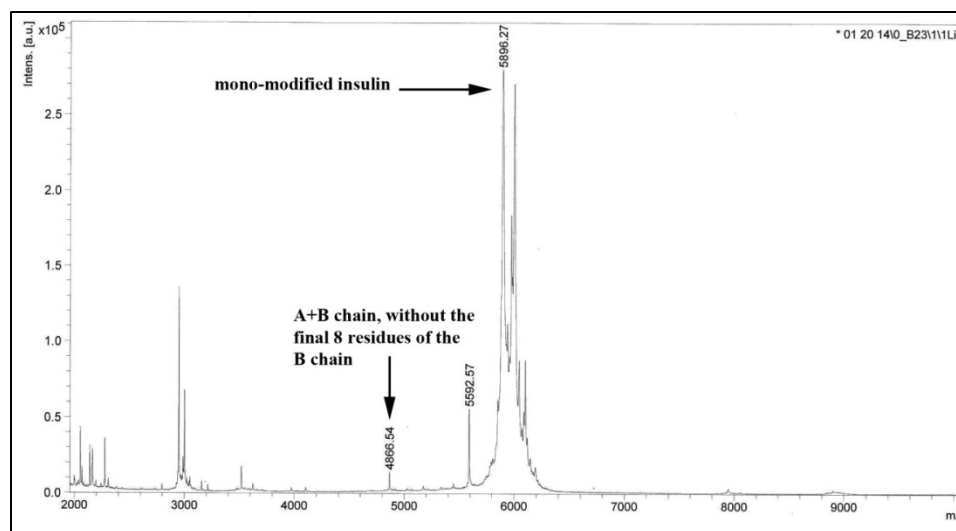


Figure 24: MALDI analysis demonstrating the specificity of the B²⁹ residue modification. The peak at 4867 represents an unmodified A chain and an unmodified B chain without the final eight C-terminal residues of the B chain. The peak at 5897 represents the mass of a mono-modified insulin. These two peaks demonstrate that the B²⁹ modification is specific to the B²⁹ lysine residue.

The trityl-protecting group of the modified insulin can subsequently be removed with anhydrous TFA and activated with 20 molar equivalents of 2,2'-dithiobis(5-nitropyridine) (DTNP), to yield an activated disulfide (Figure 25). 5% triisopropylsilane was included in this reaction to quench the trityl cations. The peptide was then purified using reverse phase chromatography. This peptide contains an activated disulfide that reacts with any free sulfhydryl group to create a disulfide bond. This enables the formation of a disulfide bond between the activated insulin and any peptide that contains a single free sulfhydryl.

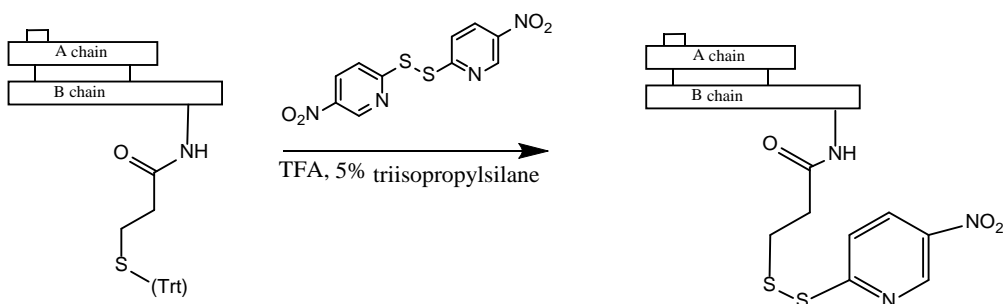


Figure 25: Reaction scheme demonstrating the activation of the B²⁹ modified insulin.

All other peptides presented in this chapter were synthesized using Fmoc chemistry on a Chemmatrix rink amide resin. Peptides were cleaved from the resin in a cleavage cocktail of 2.5% triisopropylsilane, β -mercaptoethanol, thioanisole, and H₂O in TFA. Crude peptides often contained multiple peaks as shown by LC/MS (Figure 26). This was thought to be the result of incompletely deprotected side chains, most likely due to acetyl groups on tryptophan residues. Stirring the crude peptide in a dilute acid solution, such as 2% acetic acid or 1% acetic acid, 20% acetonitrile overnight resolved the crude peptide into a single peak with the correct charge to mass ratio, as shown by LC/MS (Figure 27).

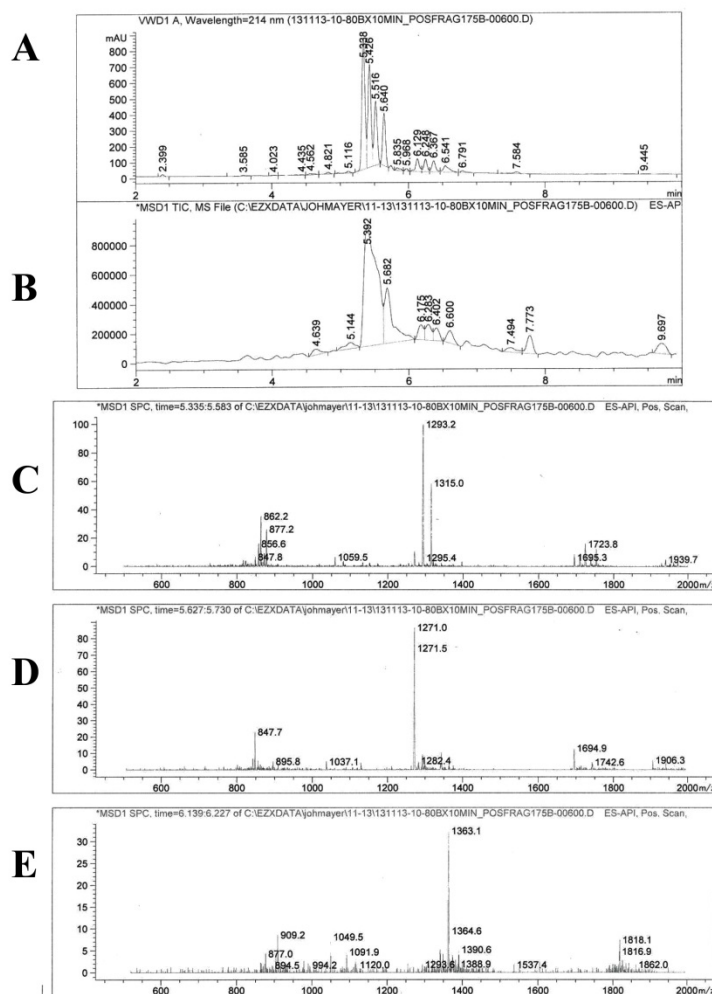


Figure 26: LC/MS analysis of crude peptide #6-Cys Fmoc synthesis. Absorbance at 214 nm (A) shows multiple peaks. Ion counts shown in (B), mass/charge ratios shown in (C-E). Peptide #6-Cys has molecular weight of 2540, therefore the charge states should appear as 1270 and 847. The presence of multiple peaks suggested incomplete side chain deprotection. The presence of the 1293 charge state suggests an acetyl group.

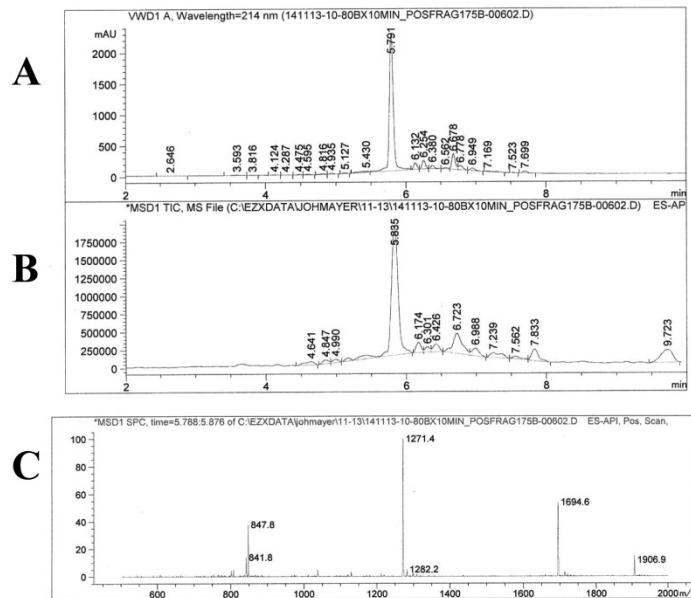


Figure 27: LC/MS depicting the same crude peptide #6-Cys as in Figure 26, after stirring in 2% acetic acid for 12 hours. Absorbance at 214 nm is depicted in (A), ion count in (B) and mass/charge ration in (C). Peptide resolved into a single peak at $t=5.7$ minutes with the correct mass to charge ratio. Peptide was then purified and used in ligation reactions.

After Fmoc synthesis and cleavage, peptides were precipitated with ether, dissolved in 20% acetonitrile, 1% acetic acid solution and stirred overnight to remove any remaining protecting groups. Peptides were then lyophilized and subsequently purified using reverse phase chromatography. LC/MS was used to confirm the purity and accuracy of the synthesis.

Ligation reactions occurred in 50 mM sodium phosphate buffer, 20% acetonitrile, pH 6.0 - 7.0. A peptide containing a single free thiol group and activated insulin were mixed in a 6:1 molar ratio and dissolved in 50 mM sodium phosphate / 20% acetonitrile buffer to a concentration of 20 mg/mL. Reaction was monitored by LC/MS. The ligated peptide was subsequently purified using reverse phase FPLC with an Amberchrome XT-20

divinylbenzyl polystyrene column. LC/MS was used to confirm the mass and analyze the purity of the ligated peptides. The ligated peptides were then dissolved in 25 mM ammonium bicarbonate buffer, pH 8, and analyzed for concentration using UV-Vis nanodrop spectroscopy. These peptides in solution were subsequently analyzed in a phosphorylation assay for activity at the insulin receptor A and B isoforms, as described in chapter 2. Optical density at 450 nm was graphed as a function of peptide concentration using Origin Pro 9.0 graphing software. Chemical reactions and schematics in this chapter were created with ChemBioDraw Ultra 13.0.

REFERENCES

1. Nilsson, B.L., M.B. Soellner, and R.T. Raines, *Chemical synthesis of proteins*. *Annu Rev Biophys Biomol Struct*, 2005. **34**: p. 91-118.
2. David, R., M.P. Richter, and A.G. Beck-Sickinger, *Expressed protein ligation. Method and applications*. *Eur J Biochem*, 2004. **271**(4): p. 663-77.
3. Moses, J.E. and A.D. Moorhouse, *The growing applications of click chemistry*. *Chem Soc Rev*, 2007. **36**(8): p. 1249-62.
4. Saxon, E. and C.R. Bertozzi, *Cell surface engineering by a modified Staudinger reaction*. *Science*, 2000. **287**(5460): p. 2007-10.
5. Saxon, E., J.I. Armstrong, and C.R. Bertozzi, *A "traceless" Staudinger ligation for the chemoselective synthesis of amide bonds*. *Org Lett*, 2000. **2**(14): p. 2141-3.
6. Kohn, M. and R. Breinbauer, *The Staudinger ligation-a gift to chemical biology*. *Angew Chem Int Ed Engl*, 2004. **43**(24): p. 3106-16.

Chapter 5
Future Directions

FUTURE DIRECTIONS

The primary objective of this thesis research was directed at the synthesis, *in vitro* biochemical characterization, and further chemical optimization of insulin heterodimers of full inherent potency but of partial maximal activity. The next step in this research is to begin to investigate the relative *in vivo* activity of these peptides. While our research lab at Indiana University is not equipped to conduct *in vivo* analyses, we benefit from a collaboration with Dr. Diego Perez-Tilve at the University of Cincinnati. The following results represent initial *in vivo* research in mouse model systems, conducted at the University of Cincinnati, to which I have provided collaborative assistance.

Mice are often used to test metabolic drugs, due to the highly conserved nature of insulin and the insulin receptor. There are numerous mouse models available, either as mimics of healthy metabolism (lean, normal mice), or as a model of type 1 or type 2 diabetes. A model of type 1 diabetes can be chemically induced in mice, most often by treatment with streptozotocin (STZ) or alloxan[1, 2]. These toxins target pancreatic beta cells in the animals to induce a diabetic state. There are also mouse model systems that exhibit spontaneous autoimmune destruction of pancreatic beta cells[1, 3], as well as virus-inducible mouse models of type 1 diabetes[1]. Type 2 diabetes is often induced in mouse model systems by offering a high-fat diet[4], although there are also less-obese model systems available[1].

Initial *in vivo* investigations were conducted in lean, normal mice. Chapter 2 of this research described the creation of an optimized antagonist, #4-6. In order to investigate

whether this antagonist is capable of antagonizing native insulin *in vivo*, pre-formulated mixtures of native human insulin and the #4-6 antagonist were created in 1:1, 1:10, and 1:100 ratios. Mice were fasted for three hours, and then given a single subcutaneous injection containing both peptides. Insulin was administered in 3 nmol/kg doses, and the #4-6 antagonist was therefore simultaneously administered in 3, 30, and 300 nmol/kg doses. Blood glucose was measured with a commercially available blood glucose meter at 0.5, 3, and 6 hours after injection. It was found that the antagonist was only capable of opposing insulin (and thereby raising blood glucose), in the highest ratio of 100:1 (Figure 1). This effect was short lived, as it was only observable at 0.5 hours after injection. At the next data point, 3 hrs after injection, there was no remaining antagonistic effect.

In order to better characterize the short-lived antagonist, the #4-6 antagonist was investigated at 3, 10, 30, and 300 nmol/kg doses, in the absence of competing insulin (Figure 2). This experiment was also conducted with a shorter 3 hour duration of investigation. In this more sensitive assay, it was determined that the antagonist was capable of raising blood glucose levels at 30 nmol/kg and above, although the antagonism was again short lived.

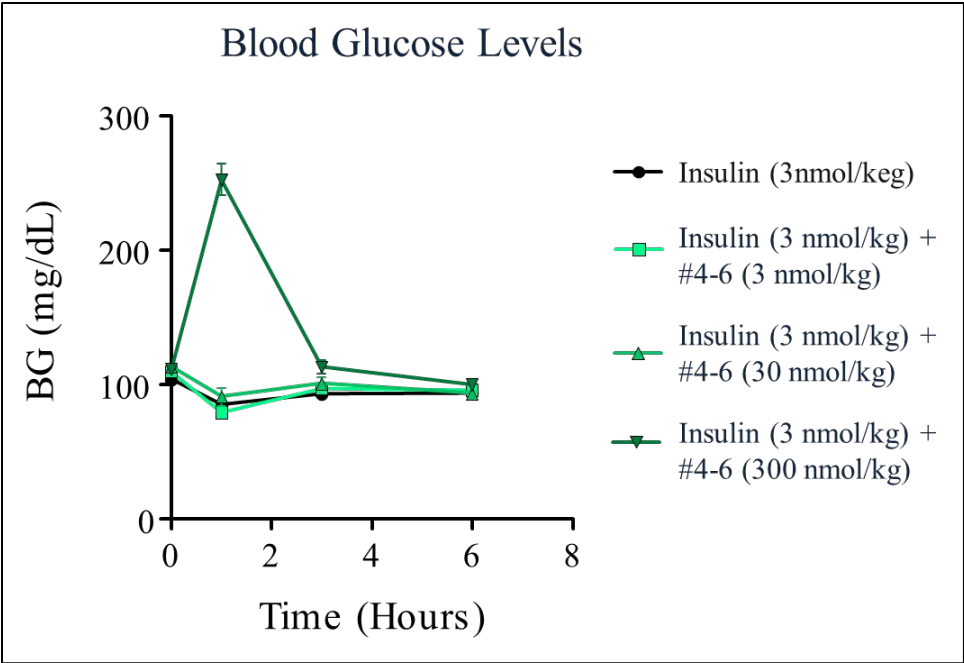


Figure 1: Blood glucose (BG) concentrations, in normal, lean mice in response to 1:1, 1:10, and 1:100 ratios of human insulin:#4-6 antagonist.

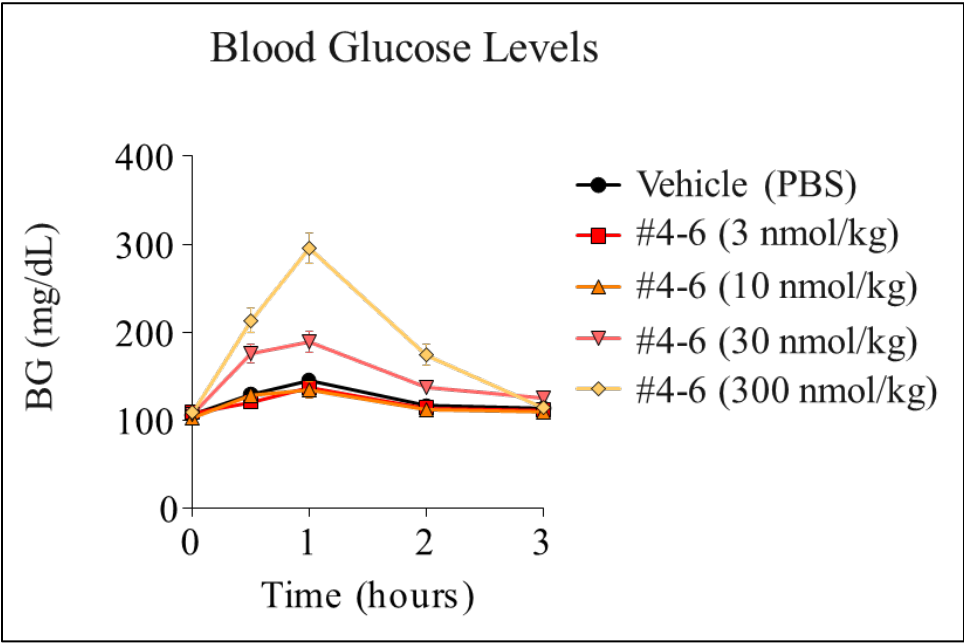


Figure 2: Blood glucose (BG) concentrations, in normal, lean mice in response to 3, 10, 30, and 300 nmol/kg doses of the #4-6 antagonist.

The *in vivo* effects of the PEGylated #4-6 antagonist, described in chapter 2, were also investigated. In our *in vitro* assay, we found that the PEGylated version was still capable of antagonizing native insulin, although with decreased potency. It was hypothesized that the increased molecular size of this peptide, due to the PEG modification, would result in increased duration of action *in vivo*. Therefore both the PEGylated and unmodified #4-6 antagonist were tested in lean, normal mice. The antagonists were delivered either as a daily subcutaneous injection, or as a continuous infusion from a surgically implanted pump. Regardless of the mechanism of delivery, each mouse was dosed with a total of 20 nmol/mouse per week, for one week.

Based on the hypothesis of PEG-dependent extended action, it was anticipated that the PEGylated peptide would antagonize native insulin, thereby creating “diabetic-like” symptoms, including an increase in blood glucose and decrease in body weight. However, it was found that the PEGylated and unmodified antagonists had no significant effect on either blood glucose (Figure 3) or body weight (Figure 4), relative to the vehicle (PBS).

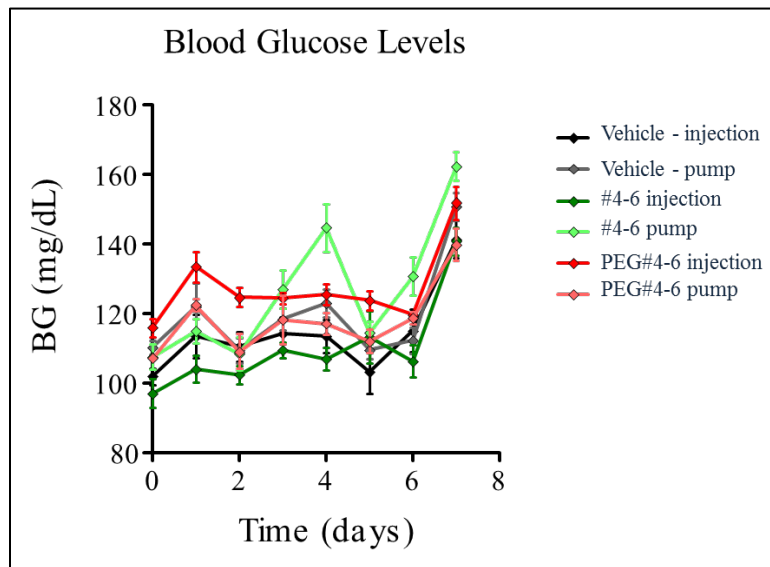


Figure 3: Blood glucose (BG) concentrations in normal, lean mice in response to the #4-6 and PEGylated #4-6 antagonists, measured over the course of seven days. All peptides were administered at 20nmol/mouse/week doses, either as a daily subcutaneous injection or through a surgically implanted pump. The spike in blood glucose at day 7 was reported as the direct result of a change in housing and diet.

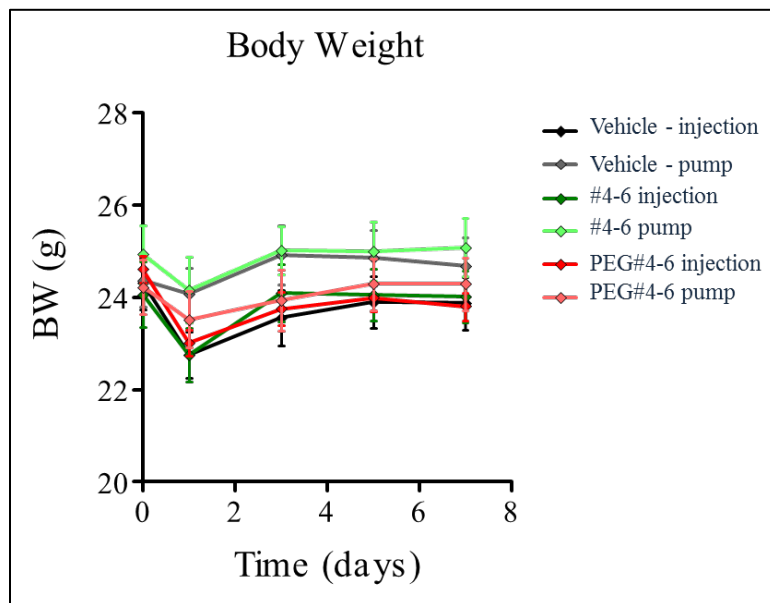


Figure 4: Body weight (BW) in grams, in lean, normal mice, measured in response to the #4-6 antagonist and PEGylated #4-6 antagonist over the course of seven days. All peptides were administered at 20 nmol/mouse/week doses, either as a daily subcutaneous injection or through a surgically implanted pump.

These results show that the antagonist, both in its PEGylated and unmodified form, is incapable of significantly affecting blood glucose and weight in lean, normal mice. The significance of these results is uncertain. We have shown that the **#4-6** antagonist, while capable of affecting blood glucose at very high doses, does not have measureable effects at the low dose of 20 nmol/week, either in PEGylated or in its unmodified form. However, this may be a consequence of the choice of model system. All of the mice tested were lean, normal mice. In mice where insulin tolerance is impaired, it may be possible to observe antagonism. The conclusions we can draw, at this time, from the *in vivo* analyses of the **#4-6** antagonist is that it is capable of effecting physiological change in the short term, but is an unlikely candidate for long term antagonism in normal mice, due to the high doses required.

The initial *in vivo* results also include two of the heterodimers synthesized and characterized in chapter 4, **#6(L2)-Cys-Insulin** and **#6(A2)-Cys-Insulin**. The peptides were tested in STZ-mice, due to their high fasting blood glucose levels. The heterodimers **#6(L2)-Cys-Insulin** and **#6(A2)-Cys-Insulin** were chosen for study, since these were the lowest and highest activity heterodimers in our library, and therefore represented the greatest possible dynamic range. Mice were fasted for 2 hours prior to the injection of native human insulin, **#6(L2)-Cys-Insulin**, or **#6(A2)-Cys-Insulin**. Insulin and **#6(A2)-Cys-Insulin** were administered at 10 nmol/kg doses. In addition, **#6(L2)-Cys-Insulin** was administered at 10, 30, and 100 nmol/kg doses. Each data point represents the average of 8 mice. It was found that the high-activity **#6(A2)-Cys-Insulin** heterodimer behaved identically to an equivalent dose of native insulin (Figure 5).

However, the low activity heterodimer, **#6(L2)-Cys-Insulin**, also behaved identically to an equivalent dose of native insulin (Figure 5). Further increases in dosing led to delayed decreases in blood glucose, to the extent that some of the animals exhibited signs of hypoglycemia (seizures, hypothermia), and had to be rescued with exogenous glucose.

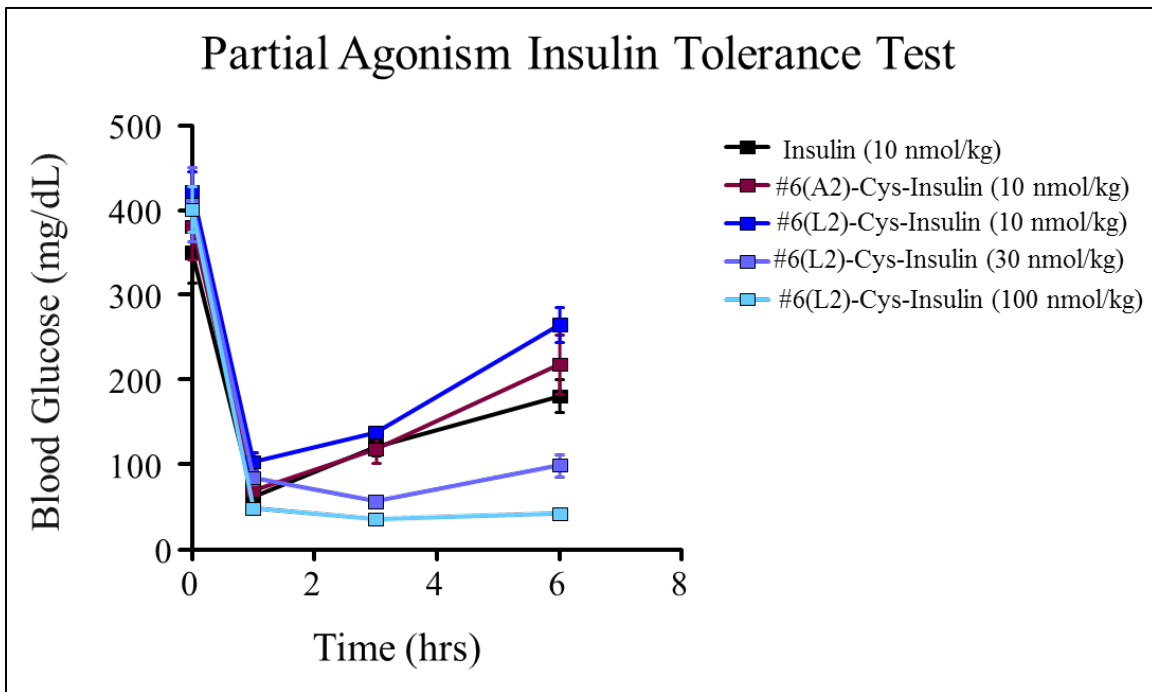


Figure 5: Blood glucose concentrations of STZ diabetic mice in response to human insulin, #6(L2)-Cys-Insulin and #6(A2)-Cys-Insulin.

These results suggest that diminished activity at the insulin receptor *in vitro* does not translate to diminished activity *in vivo*. However, this may be due to peptide instability in this *in vivo* assay. The *in vitro* phosphorylation assay does not fully mimic a plasma exposure, which can contain reduction cofactors, and enzymes that can participate in the reduction of disulfide bonds. If the insulin heterodimers are not stable *in vivo*, then the

breakage of the disulfide bond would result in a fully potent agonist, and the **#6-Cys** antagonist, which is incapable of antagonism *in vitro*, as shown previously in chapter 2.

To address the possibility of disulfide instability, a slightly modified version of the **#6(L2)-Cys-Insulin** molecule was created, where the cysteine residue has been replaced by penicillamine, which is simply a gem-dimethyl form of cysteine (Figure 6). This new heterodimer, **#6(L2)-Pen-Insulin**, behaves identically to its cysteine counterpart in the *in vitro* assays (Figure 7). The di-methylation should yield the disulfide bond additional *in vivo* stability, since it is less susceptible to enzymatic degradation, and disulfides containing a penicillamine residue are approximately three times less susceptible to reduction by glutathione, a cellular reducing agent[5].

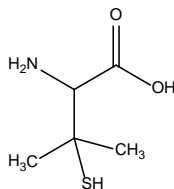


Figure 6: Penicillamine.

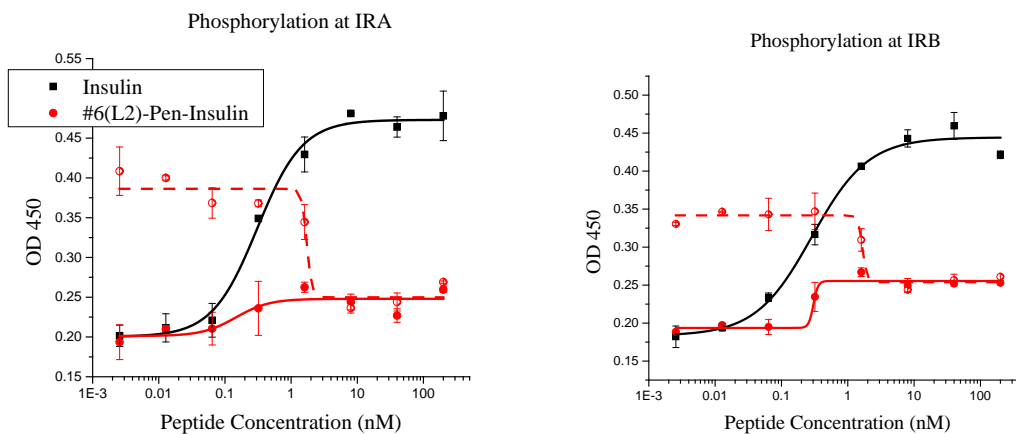


Figure 7: Phosphorylation assay demonstrating the agonism (solid) and antagonism (dashed) of #6(L2)-Pen-Insulin.

Despite this modification, the **#6(L2)-Pen-Insulin** heterodimer also lowers blood glucose in STZ mice (Figure 8). While this may indicate that diminished *in vitro* activity does not result in diminished *in vivo* activity, there are a few interesting characteristics of this study. First, despite the dramatic drops in blood glucose, none of the animals in this study displayed the typical symptoms of hypoglycemia, seizures and hypothermia. Therefore, none of the animals needed to be rescued with glucose injections. In addition, while the 25 nmol/kg dose induced low blood glucose in these animals, no additional drops in blood glucose were observed when the dose was increased from 25 to 50 nmol/kg. It can also be seen that the **#6(L2)-Pen-Insulin** heterodimer has a significantly extended duration of action, relative to native insulin. These observations may indicate that despite lowering blood glucose, the **#6(L2)-Pen-Insulin** heterodimer appears to represent a “safer” insulin therapy.

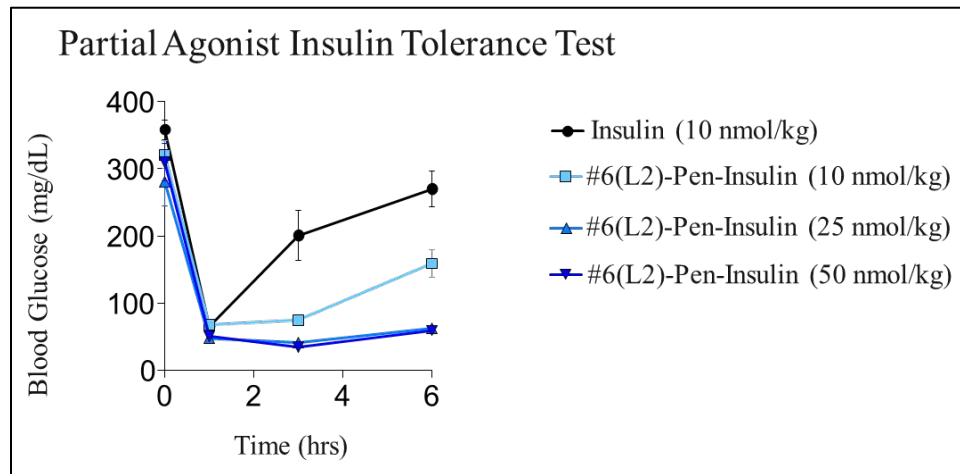


Figure 8: Blood glucose concentrations of STZ diabetic mice in response to human insulin, and #6(L2)-Pen-Insulin.

Therefore, the future direction of this work will focus on further characterization of the **#6(L2)-Pen-Insulin** heterodimer. Currently, the *in vivo* results are all the result of a single subcutaneous dose, bolus injection of peptide. This results in a sudden spike in the plasma concentration of peptide, which if consistent with native insulin time action, would peak between 30 and 60 minutes and clear from circulation within 2-4 hours. Peptides can also be delivered as an infusion, which minimizes the pharmacokinetic effects upon the pharmacodynamics effects of the administered drug. It is possible that delivering our heterodimers as an infusion might reveal a more subtle difference in activity, relative to native insulin.

In addition, a glucose clamp technique can be used to further characterize our heterodimers. A glucose clamp experiment infuses insulin and simultaneously provides exogenous glucose in order to maintain a predetermined blood glucose concentration. Therefore, the amount of exogenous glucose required is a measure of insulin sensitivity. It is possible that a glucose clamp experiment may reveal differences in our heterodimers, relative to insulin, that explain the lack of hypoglycemic symptoms (seizures, hypothermia) in mice treated with **#6(L2)-Pen-Insulin**.

Despite the indications that **#6(L2)-Pen-Insulin** may be a safer therapy, we have shown that the decrease in receptor activation we observed *in vitro* does not translate to reduced activity *in vivo*. There are two possible explanations for this phenomenon. First, it is possible that full metabolic signaling can be induced by a molecule that is only 20-30% active at insulin receptor isoforms. In living systems, the insulin receptor is expressed in

excess of the concentration required for full signaling. This is referred to as receptor reserve. Therefore, 20-30% activity at the insulin receptors may be sufficient to fully activate the insulin signaling pathway.

Another explanation for the discrepancy between our *in vivo* and *in vitro* results is that the decrease in activity observed in our *in vitro* assays may be the result of selective phosphorylation of the insulin receptor itself. The insulin receptor has multiple tyrosines that are phosphorylated in response to insulin, and it has been shown that differential mutation of these tyrosines leads to different metabolic and mitogenic effects[6]. For example, a tyrosine to phenylalanine mutation at the insulin receptor Tyr-1146 (IRA isoform numbering) does not inhibit *in vitro* metabolic responses to insulin, such as glycogen synthase stimulation[7]. However, this mutation does suppress the insulin-stimulated incorporation of radiolabeled thymidine into DNA, an *in vitro* measure of mitogenic potential[7]. It is possible that our heterodimers might also exhibit position-specific effects on receptor tyrosine phosphorylation. The phosphorylation assay utilized in this research is incapable of discriminating between the multiple tyrosines which are present on the insulin receptor. However, there are commercially available antibodies that do differentiate between insulin receptor tyrosine residues. Therefore, investigating our heterodimers with position-specific phosphotyrosine antibodies might yield more insight into the physiological significance of reduced maximal activity. If the low-activity heterodimers preferentially phosphorylate the metabolic tyrosines, it may explain the discrepancy between the *in vitro* and *in vivo* effects. In addition, it is possible that one of

our higher activity analogs may represent the “safest” maximal activity due to its tyrosine phosphorylation profile.

In addition to these specific avenues of inquiry, the heterodimers that we have created within this research open a new avenue of inquiry entirely. We have shown that peptides previously dismissed as potential antagonists, such as the #6 peptide, can function within a heterodimer to effect meaningful change in receptor activation. Therefore, other peptides that have been dismissed from this and other research may actually represent viable therapeutic tools. The most important conclusion of this research is that we have validated the hypothesis that diminished maximal activity can be achieved through the creation of agonist/antagonist heterodimers. Most satisfyingly, we have identified an antagonist that is tunable, through a single site mutation, such that we can direct an array of maximal activities.

The antagonist that is responsible for this modulated activity contains a single binding site motif. This is in contrast to a number of published works that suggest only a two-site peptide is capable of antagonism. This illustrates the dramatic effect of multivalency within a heterodimer. The observed suppression of agonistic activity is more than the sum of agonist and antagonistic properties, but instead represents the emergent activity of a molecule that modulates its own activity. These observations have implications in all multivalent systems, and demonstrate the scientific basis for the truism that we are more than the sum of our parts.

REFERENCES

1. King, A.J.F., *The use of animal models in diabetes research*. British Journal of Pharmacology, 2012. **166**(3): p. 877-894.
2. Graham, M.L., et al., *The streptozotocin-induced diabetic nude mouse model: differences between animals from different sources*. Comp Med, 2011. **61**(4): p. 356-60.
3. Van Belle, T.L., P. Taylor, and M.G. von Herrath, *Mouse models for Type 1 Diabetes*. Drug Discovery Today: Disease Models, 2009. **6**(2): p. 41-45.
4. Winzell, M.S. and B. Ahren, *The high-fat diet-fed mouse: a model for studying mechanisms and treatment of impaired glucose tolerance and type 2 diabetes*. Diabetes, 2004. **53 Suppl 3**: p. S215-9.
5. Gilbert, H.F., *Thiol/Disulfide Exchange Equilibria and Disulfide Bond Stability*. Biothiols, Pt A, 1995. **251**: p. 8-28.
6. Jensen, M. and P. De Meyts, *Chapter 3 Molecular Mechanisms of Differential Intracellular Signaling From the Insulin Receptor*, in *Vitamins & Hormones*, L. Gerald, Editor. 2009, Academic Press. p. 51-75.
7. Wilden, P.A., et al., *The Insulin-Receptor with Phenylalanine Replacing Tyrosine-1146 Provides Evidence for Separate Signals Regulating Cellular-Metabolism and Growth*. Proceedings of the National Academy of Sciences of the United States of America, 1990. **87**(9): p. 3358-3362.

TABLE OF ABBREVIATIONS

Abbreviations	Full name, description
μM	Micromolar
4-MeBzl	4-methyl-benzyl, acid labile protecting group, used to protect cysteine residues during Boc SPPS
A ₁ C	Glycosylated hemoglobin
ADP	Adenosine diphosphate
Ala, A	Alanine
Arg, R	Arginine
Asn, N	Asparagine
Asp, D	Aspartic acid
ATP	Adenosine triphosphate
B	Concentration of bound ligand
BG	Blood glucose
BGS	Bovine growth serum
BHA resin	Benzydrylamine resin
B _{max}	Total maximal amount of bound receptor
Boc	Tert-butyloxycarbonyl protecting group, acid labile, used to protect N terminal amines during Boc SPPS and lysine and tryptophan residues during Fmoc SPPS
BOM	Benzyloxymethoxy, acid labile protecting group used to protect histidine residues during Boc SPPS
Br-Z	2-Bromobenzyloxycarbonyl, acid labile protecting group, used to protect tyrosine residues during Boc SPPS
BSA	Bovine serum albumin, fraction V
Bzl	Benzyl acid labile protecting group, used to protect serine and threonine residues during Boc SPPS

C8	Hydrocarbon chain containing 8 carbon atoms, used in silica based columns for HPLC purification
Cbz, Z	Benzyloxycarbonyl acid labile protecting group
CHO	Formyl, acid labile protecting group, used to protect tryptophan residues during Boc SPPS
CR	Insulin receptor cysteine rich domain
CT	Insulin receptor C-terminus
Cys, C	Cysteine
DCM	Dichloromethane
DEPBT	3-(Diethoxyphosphoryloxy)-1,2,3-benzotriazin-4(3H)-one, coupling reagent
des	Without, removed, usually in reference to a number of residues within a peptide
DIC	Diisopropylcarbodiimide
DIEA	Diisopropylethylamine
dL	Deciliter
DMEM	Dulbecco's modified Eagle's medium
DMF	Dimethylformamide
DNA	Deoxyribonucleic acid
DPX	DiMarchi Peptide, where X represents an internal numbering system in the DiMarchi laboratory
DTNP	2,2'-dithiobis(5-nitropyridine)
EC ₅₀	The concentration at which a drug elicits half the maximal response
ED	Effective dose
ELISA	Enzyme-linked immunosorbent assay
F	Concentration of free ligand

Fmoc	9-fluorenylmethyloxycarbonyl, base labile protecting group, used to protect N terminal amines during Fmoc SPPS
Fn	Fibronectin type III domain
FPLC	Fast protein liquid chromatography
FPLC	Fast protein liquid chromatography
Gln, Q	Glutamine
GLP-1	Glucagon-like peptide
Glu, E	Glutamic acid
Gly, G	Glycine
GSK-3	Glycogen synthase kinase 3
HCl	Hydrochloric acid
HEK cells	Human embryonic kidney cells
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HF	Hydrofluoric acid
His, H	Histidine
HLA	Human leukocyte antigen
HOBt	Hydroxybenzotriazole, coupling reagent
HOBt	Hydroxybenzotriazole
HPLC	High pressure liquid chromatography
HRP	Horse radish peroxidase
IC ₅₀	The concentration at which a drug is inhibited to half the maximal response
ID	Insulin receptor insert domain
IGF-1	Insulin-like growth factor 1
IGF-2	Insulin-like growth factor 2

IL-1 β	Interleukin-1
Ile, I	Isoleucine
IR	Insulin receptor
IRA	Insulin receptor A isoform, lacks exon 11
IRB	Insulin receptor B isoform, retains exon 11
IRS	Insulin receptor substrate proteins
IU	Indiana University
JM	Juxtamembrane domain
K _a	Equilibrium association constant
K _{ATP}	ATP-sensitive potassium ion channels
K _d	Equilibrium dissociation constant
kDa	kilodalton
kg	kilogram
L	Ligand
L1	Insulin receptor leucine rich domain 1
L2	Insulin receptor leucine rich domain 2
LB	Lysogeny broth, nutrient rich media for bacterial growth
LC/MS	Liquid chromatography / mass spectrometry
LC/MS	Liquid chromatography / mass spectrometry
LD	Lethal dose
Leu, L	Leucine
Lys, K	Lysine
Lys-C	Endopeptidase Lys-C, hydrolyzes peptide bonds at the C terminal of lysine residues
M	Molar

MALDI-TOF	Matrix assisted laser desorption /ionization – time of flight
MAPK	Mitogen-activated protein kinase
MBHA resin	Methylbenzylamine resin
Met, M	Methionine
mg	Milligrams
MHC	Major histocompatibility complex proteins
mmol	Millimole
NF- κ B	Nuclear factor kappa-light-chain-enhancer of activated B cells
NHS	N-hydroxysuccinimide
NIH	National Institutes of Health
nM	Nanomolar
nm	Nanometers
nmol	Nanomole
NMP	N-methyl-2-pyrrolidone
NPH	Neutral protamine hagedorn insulin
OcHx	Cyclohexyl ester, acid labile protecting group used to protect aspartic acid and glutamic acid residues during Boc SPPS
OD 450	Optical density, 450 nm
OtBu	t-Butyl ester, acid labile protecting group used to protect aspartic acid and glutamic acid residues during Fmoc SPPS
PAL resin	Peptide amide linker resin
PAM resin	<i>tert</i> -butoxycarbonylaminoacyl-4-(oxymethyl)-phenylacetamidomethyl resin
PBS	Phosphate buffered saline
PCK	Protein kinase C
PCR	Polymerase chain reaction

PEG	Polyethylene glycol
pH	$-\log[H^+]$, indicative of the acidity or basicity of a solution
PH	Pleckstrin homology domains
Phe, F	Phenylalanine
PI3K	Phosphoinositide 3-kinase
PIP ₃	Phosphatidylinositol (3,4,5)-triphosphate
Pro, P	Proline
R	Receptor
RNA	Ribonucleic acid
ROS	Reactive oxygen species
R _T	Total receptor Concentration
RTK	Receptor tyrosine kinase
SCM	Succinimidyl carboxy methyl
Ser, S	Serine
SH2	Src-homology 2 domains
SPPS	Solid phase peptide synthesis
STZ	Streptozotocin
SUMO	Small ubiquitin-related modifier
SUR1	Sulfonylurea receptor isoform 1
tBu	t-Butyl acid labile protecting group, used to protect serine, threonine and tyrosine residues during Fmoc SPPS
TFA	Trifluoroacetic acid
Thr, T	Threonine
TK	Tyrosine kinase comain
TLR	Toll-like receptor

TMB	3,3',5,5'-tetramethylbenzidine
Tos	Tosyl acid labile protecting group, used to protect arginine residues during Boc SPPS
Trp, W	Trptophan
Trt	Trityl, acid labile protecting group used to protect cysteine residues during Boc SPPS and histidine, asparagine, and glutamine residues during Fmoc SPPS
Tyr, Y	Tyrosine
UV-Vis	Ultraviolet – visible
UV-Vis	Ultra violet – visible
Val, V	Valine
Xan	Xantyl acid labile protecting group, used to protect asparagine residues during Boc SPPS
Y	Fraction of bound receptor concentration to total receptor concentration

Sara J. Brandt

sjbrandt@indiana.edu

Education

Indiana University, Biochemistry PhD Program, 2015
Minors in Cellular and Medicinal Biochemistry

Duke University, Bachelor of Science in Chemistry, 2009
Minors in Biology, Biological Anthropology and Anatomy

Publications

Claire J. Parker Siburt, Emily M. Lin, Sara J. Brandt, Arthur D. Tinoco, Ann M. Valentine, Alvin L. Crumbliss, Redox potentials of Ti(IV) and Fe(III) complexes provide insights into titanium biodistribution mechanisms. Journal of Inorganic Biochemistry. (2010) 104 (9), 1006-1009.

Research Experience

Richard DiMarchi Lab, Indiana University, 2010-2015
Used biosynthetic and synthetic chemistries to create insulin receptor partial agonists as a basis for drug design. Research experience includes Boc and Fmoc solid phase peptide synthesis, reverse phase and size exclusion chromatography, and in-vitro ELISA-based phosphorylation assays. Experience with MALDI and ESI mass spectrometry.

Alvin L. Crumbliss Lab, Duke University, 2007-2009
Used cyclic voltammetry to analyze multiple complexes containing titanium.

Teaching Experience

Indiana University, Chemistry Department, 2010-2015
Associate Instructor
C117 Laboratory, introductory chemistry course
C483 Biological chemistry, fourth year biochemistry course
Undergraduate Tutor, 2010-2015
Private tutor for all introductory level chemistry courses

Duke University, Chemistry Department, 2010-2015
Teaching Assistant
C22 Laboratory, introductory chemistry course

Awards

Velocity Conference, 2013
One of five students selected from the life sciences to represent IU at a biotechnology conference, in conjunction with the IU Kelley School of Business.

IU Biochemistry Student Ambassador, 2012-2015
Selected to be profiled on the IU biochemistry website. Responsibilities include answering questions from prospective students and representing IU at recruiting events.

ACC Academic Honor Roll, 2005-2009

Work Experience

Educational Innovations, Inc., 2005-2007

Customer Service

Chemical Demonstrator

ChemEd Conference, University of North Texas, 2007

Biennial Conference on Chemical Education, Purdue University, 2006

Volunteer Experience

Volunteer, Bloomington Wonderlab Museum, 2010-2015

Museum host, participates in science demonstrations and live animal shows

Volunteer math and science tutor, Monroe County Public Library, 2012-2015

Duke University Chemistry Outreach Programs, 2005-2009

Co-founder, Radical Chemistry Productions

Designed and filmed pedagogy-based laboratory experiments

Volunteer tutor, EK Powe Elementary School, Durham, NC, 2006-2009