12-2012

The Path to an Orally Administered Protein Therapeutic for the Treatment of Diabetes Mellitus

Susan Clardy James
Syracuse University

Follow this and additional works at: https://surface.syr.edu/che_etd

Part of the Chemistry Commons

Recommended Citation
James, Susan Clardy, "The Path to an Orally Administered Protein Therapeutic for the Treatment of Diabetes Mellitus" (2012). Chemistry - Dissertations. 196.
https://surface.syr.edu/che_etd/196

This Dissertation is brought to you for free and open access by the College of Arts and Sciences at SURFACE. It has been accepted for inclusion in Chemistry - Dissertations by an authorized administrator of SURFACE. For more information, please contact surface@syr.edu.
Abstract

Protein therapeutics like insulin and glucagon-like peptide-1 analogues are currently used as injectable medications for the treatment of diabetes mellitus. An orally administered protein therapeutic is predicted to increase patient adherence to medication and bring a patient closer to metabolic norms through direct effects on hepatic glucose production. The major problem facing oral delivery of protein therapeutics is gastrointestinal tract hydrolysis/proteolysis and the inability to passage the enterocyte. Herein we report the potential use of vitamin B₁₂ for the oral delivery of protein therapeutics.

We first investigated the ability of insulin to accommodate the attachment of B₁₂ at the B₁ vs. B₂₉ amino acid position. The insulinotropic profile of both conjugates was evaluated in streptozotocin induced diabetic rats. Oral administration of the conjugates produced significant drops in blood glucose levels, compared to an orally administered insulin control, but no significant difference was observed between conjugates. We also report, for the first time, a dose dependent response of a B₁₂-insulin conjugate. We then explored the implications of B₁₂ conjugation on the biological activity of the potent peptide glucagon-like peptide-1 (7-36) amide, with a K₃₄R amino acid substitution. Various in vitro bioassays utilizing human embryonic kidney cells and human pancreatic islets were conducted and indicated B₁₂ has a minimally negative effect on GLP-1 biological activity. Finally we modified the structure of B₁₂ for future conjugation work. The modification of the 5’hydroxyl group of the ribose unit of B₁₂ to a carboxylic acid is predicted to benefit the field of B₁₂ conjugation significantly with the ability to produce higher yielding and more stable B₁₂ conjugates.
The Path to an Orally Administered Protein Therapeutic for the
Treatment of Diabetes Mellitus

By

Susan Clardy-James

B.S. Chemistry, Millikin University 2008

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy in Chemistry

in the Graduate School of Syracuse University

December 2012
Acknowledgements

It would not have been possible to write this thesis without the guidance and support of many kind people around me, to only some of whom it is possible to give particular mention here.

I am truly indebted and thankful to my advisor, Professor Robert P. Doyle, whose encouragement, wisdom and commitment to the highest standards inspired and motivated me to become the scientist I am today. I will forever be thankful for his patience, guidance and support throughout my time at Syracuse University. I could not have imagined having a better advisor and mentor for my Ph. D studies and am incredibly grateful to have had the opportunity to work with him.

I would like to thank all of the individuals in the Doyle lab. Each one brought a different perspective that helped me grow and learn in unexpected ways. I especially want to thank Professor Nadia Marino for the wonderful conversation about research, life and the future. I also want to thank Professor G. G. Holz, Dr. Colin Leech and Dr. Oleg Chepurny for allowing me the opportunity learn a new skill, in a new lab.

I would also like to thank my friends and family for their love and support. My dearest friend, Jessica Seely, for being there any time even with thousands of miles between us. My parents, Theresa and Richard Clardy, for always loving me and encouraging me to never stop learning. My brother, Richard Clardy Jr., for putting things into perspective and reminding that there is a world outside the lab. Last but far from least, my husband, Stephen James, for his tireless devotion and complete understanding through the good and bad times. I am without a doubt blessed to have such an amazing support system and am truly thankful to all of them.
# Table of Contents

**Chapter 1** General Introduction  
1.1 Diabetes Mellitus  1  
1.2 Injectable Peptides and Proteins for the Treatment of Diabetes Mellitus  2  
  1.2.1 Insulin  3  
    1.2.1.1 History and Discovery  3  
    1.2.1.2 Biosynthesis and Secretion of Insulin  4  
    1.2.1.3 Structure and Sequence of Bovine Insulin  8  
    1.2.1.4 The Insulin Receptor  10  
  1.2.2 Glucagon-Like Peptide-1  12  
    1.2.2.1 Discovery of the Incretin Effect  12  
    1.2.2.2 Biosynthesis and Secretion  13  
    1.2.2.3 Structure and Sequence of Human GLP-1  15  
    1.2.2.4 The Glucagon-Like Peptide Receptor  16  
  1.3 A Pursuit of a Noninvasive Peptide/Protein Treatment for Diabetes Mellitus  17  
  1.3.1 Strategies for Oral Delivery of Peptides and Proteins  19  
  1.4 Vitamin B_{12}  20  
    1.4.1 History  20  
    1.4.2 Structure and Chemistry of Vitamin B_{12}  21  
    1.4.3 Physiological Roles of B_{12}  23  
    1.4.4 The Dietary Uptake Pathway Overview  25  
    1.4.5 Haptocorrin  27
3.2 Results and Discussion

3.2.1 Synthesis and Purification of B_{12}-K34R-GLP-1 (7-36) amide (3) 80

3.2.2 MALDI-ToF Mass Spectrometry 82

3.2.3 In vitro Experimentation 83

3.3 Conclusion 90

3.4 References 90

Chapter 4 Synthesis of a B_{12} Derivative Capable of Amide Bond Formation 93

4.1 Introduction 93

4.2 Results and Discussion 95

4.2.1 Synthesis of B_{12} 5’-carboxylic acid (4) 95

4.2.2 1D and 2D NMR analysis of 4 100

4.2.3 Synthesis of B_{12} 5’-carboxylic acid-benzylamine (5) 104

4.3 Conclusion 110

4.4 References 110

Chapter 5 Experimental Procedure 113

5.1 Materials and Methods 113

5.2 Experimental Procedure 114

5.2.1 Synthesis of B_{12}-B^{1}Insulin (1) 114

5.2.2 Synthesis of B_{12}-B^{29}Insulin (2) 115

5.2.3 Synthesis of B_{12}-K34R-GLP-1(7-36) amide (3) 116

5.2.4 Synthesis of B_{12}-5’-carboxylic acid (4) 116

5.2.5 Synthesis of B_{12} 5’-carboxylic acid-benzylamine (5) 117

5.3 In vitro Experimentation of 3 118
5.3.1 Cell Cultures 118
5.3.2 Luciferase Assay 118
5.3.3 Fura-2 Assay 118
5.3.4 Insulin Secretion in Human Pancreatic Islets 119

5.4 *In vivo* Experimentation of 1 and 2 120

5.5 References 120

**Chapter 6** Future Work 122

6.1 Vitamin B₁₂ 122

6.2 Glucagon-Like Peptide-1 127

6.3 C-peptide 128

6.4 References 129

Appendix A Publications arising from this work in chronological order
List of Figures

Figure 1. The primary treatment goal for DM is to maintain glycemic control (indicated as normal glucose level). Glycemic control in a patient with diabetes can be challenging since the proper amount of insulin must be administered after meals. Not injecting enough or injecting too much insulin will result in hyperglycemia or hypoglycemia, respectively. While persistent hyperglycemia causes microvascular complications, hypoglycemia can cause diabetic coma or death.

Figure 2. The subcellular organization of the insulin secretory pathway.

Figure 3. The post-translational modifications of preproinsulin, which results in active monomeric insulin and C-peptide.

Figure 4. Glucose induced insulin secretion (GSIS) is regulated by three pathways: triggering pathway, metabolic amplify pathway and neurohormonal amplify pathway. Gs: Gs α subunit, AC: adeny cyclase.

Figure 5. A representation of the primary structure of bovine insulin including the disulfide bonds formed between the A- and B-chains and the A-chain intrachain disulfide bond.
Figure 6. A carton representation of the domain organization within the homodimeric insulin receptor. L1, L2: first and second leucine-rich-repeat domains, CR: cysteine-rich region, Fn0-1-2: fibronectin domains, Ins: Insulin insert domain, TM/JM: trans- and juxta-membrane regions, TK: tyrosine kinase domain and CT: C-terminal domain.

Figure 7. The primary structure of human GLP-1 (7-36) amide. The C-terminal amidation has been shown to increase plasma stability.

Figure 8. The secondary structure of GLP-1 (7-36) amide determined by solution NMR. Lys 26 (indicated as ball and stick) has been identified as the best position for conjugation. Protein data bank ID: 1D0R. Modified using RCSB PDB protein workshop software 3.9.

Figure 9. An illustration of the physical barriers and available pathways for oral drug delivery in the intestines.

Figure 10. Structure of B12 with a variable R group shown to the left. Three of the more common forms of B12 (cyanocobalamin, methylcobalamin and adenosylcobalamin) are shown to the right.

Figure 11. Outline of methionine synthase (Enz) catalytic cycle, which involves two methyl transfer half-reactions: homocysteine to methionine and N5-methyltetrahydrofolate to tetrahydrofolate.
Figure 12. An overview of the $\text{B}_12$ uptake pathway as outlined by a recent 2012 review. The transfer of $\text{B}_12$ from IF to TCII is currently being debated as discussed in section 1.4.7 and shown in Figure 15. Image is reproduced from reference 120 with permission of Nature Publishing Group. MRP1: multidrug resistance protein 1.

Figure 13. Ribbon diagram of IF-$\text{B}_12$ complex with the $\text{B}_12$ shown in ball and stick. Used with permission of National Academy of Science from reference 129.

Figure 14. The structure of CUB 5-8 (left, blue) interacting with IF bound to $\text{B}_12$ (right, green and pink). The interaction is calcium (red sphere) dependent. The image is reproduced from reference 125 with permission from Nature Publishing Group.

Figure 15. There are currently two theories involving the export of $\text{B}_12$. a) Seetharam et al. present $\text{B}_12$ bound to TCII in the cell while more recently b) Nexo et al. have established an alternative export system involving MRP1. Images used from references 146 (a) and 120 (b) with permission from Cambridge University Press and Nature Publishing Group, respectively.

Figure 16. A representation of the TCII-$\text{B}_12$ complex. The TCII $\alpha$-domain is shown on the bottom (red) with the $\beta$-domain on top (blue). The $\text{B}_12$ molecule is shown in ball and stick and found sandwiched between the two domains (orange). Reprinted from reference 132 with permission of National Academy of Science.
Figure 17. Conjugation can occur at any of the functional groups (indicated in shaded shapes) without loss of binding by the three uptake proteins.

Figure 18. A basic representation of the B₁₂ dietary uptake pathway.

Figure 19. The three amine groups available for conjugation on the bovine insulin structure are indicated. Conjugation at the GlyA1 has previously been shown to eliminate biological activity, therefore only the PheB1 and LysB29 are available for conjugation.

Figure 20. The activation of B₁₂ occurs at the 5’-hydroxyl group of the ribose with the aid of CDT. B₁₂ Act = CDT activated B₁₂.

Figure 21. The synthetic route of 1 and 2 are outlined.

Figure 22. The RP-HPLC spectra of a BOC protected insulin plus B₁₂ reaction monitored at 254 nm.

Figure 23. The RP-HPLC spectra of an FMOC protected insulin plus B₁₂ reaction monitored at 360 nm (Dash, indicating B₁₂ is present) and 254 nm (Solid, indicating insulin is present).

Figure 24. A representative MALDI-Tof MS spectrum of DTT reduced 1 or 2 shows mass at 4754 m/z representing [B₁₂-CN+H]^+ attached to the B chain of insulin. Absent is B₁₂ attached to the A strand, 3694 m/z.
Figure 25. (a) MALDI-Tof MS spectrum of reduced and tryptically digested 1 produces a 2286 m/z fragment consistent with $[\text{B}_{12}-\text{CN}+\text{H}]^+$ attached to the LysB29 position. (b) Reduced and tryptically digested 2 produces a 3844 m/z fragment consistent with $[\text{B}_{12}-\text{CN}+\text{H}]^+$ at the N-terminus of the B strand, as expected with PheB1 conjugation.

Figure 26. The “typical” absorption spectra of CNCbl. The predominate bands highlighted (α, β, and γ) are due to spin allowed intraligand charge transfers of the corrin ring.

Figure 27. The electronic absorption spectra of B$_{12}$-insulin conjugate contain the characteristic B$_{12}$ peaks and insulin peak.

Figure 28. Relative (% change from pre-gavage) blood glucose concentration following administration of 1 at concentrations of $10^{-9}$ M, $10^{-7}$ M and $10^{-5}$ M.

Figure 29. Relative (% pre-gavage) blood glucose concentration following administration of free insulin (n=4), 1 (n=6) and 2 (n=6).

Figure 30. MD simulation of TCII bound to 1. TCII structure: α-domain (red); β-domain (dark blue); unstructured linkage between domains (green). Structure of 1: B$_{12}$ (yellow ball and stick) with Co (III) (blue sphere); Insulin B-chain (yellow ribbon); A-chain (light blue ribbon). The structure was produced using the GROMACS software package with GROMOS96 (53a6) united-atom force field by collaborator Dr. Damien Allis at Syracuse University.
Figure 31. A representative structure of 1 bound within TCII. The structure was produced from a molecular dynamic simulation performed using the same software as previously described in Figure 30 by collaborator Dr. Damien Allis at Syracuse University.

Figure 32. GLP-1 (7-36) amide was purchased with N-terminal FMOC protection and K34R amino acid substitution (purple) which will be referred to as FMOC-K34R-GLP-1. These modifications allow for site directed B$_{12}$ conjugation at the Lys26 (blue).

Figure 33. RP-HPLC chromatogram showing separation of unconjugated B$_{12}$ and K34R-GLP-1 (7-36) amide from B$_{12}$ conjugated K34R-GLP-1 (7-36) amide with detection at 254 nm (solid line) and 360 nm (dash line). Absence of B$_{12}$ absorption at 360 nm indicated the peak at retention time (Tr) ~ 24 mins is unconjugated K34R-GLP-1 (7-36) amide now with two FMOC units, while the peak at Tr ~ 20 mins is B$_{12}$ conjugated K34R-GLP-1 (7-36) amide with one FMOC.

Figure 34. MALDI-ToF MS spectrum of 3 after FMOC removal showing the presence of a prominent peak at 4682 m/z representing a 1:1 B$_{12}$-K34R-GLP (7-36) amide.

Figure 35. A basic depiction of the intercellular mechanisms of the GLP-1R, which were monitored for in vitro assays described in section 3.2.5. RYR: ryanodine receptor; ER: endoplasmic reticulum; CRE: cAMP response element. RIP-1-CRE-Luc: rat insulin 1 promoter-cAMP response element-luciferase reporter.
The fold increase of luciferase activity was determined for 3 and K34R-GLP-1 (7-36) amide after HEK-GLP-1 cells were transfected with a RIP1-CRE-Luc reporter and were exposure to 3 for 4 hrs. 3 produced a concentration dependent response with an EC50 of 4.5 nM, while K34R-GLP-1 (7-36) amide had an EC50 of 4.1 nM.

Fura-2 determinations of intercellular [Ca2+] were obtained in HEK-GLP-1R cells that express endogenous P2Y purinergic receptors. Note that both forms of GLP-1 potentiated the action of ADP to mobilize intracellular [Ca2+] and that native GLP-1 was slightly more potent than 3. Data are averaged from 12 wells for each data point and similar results were obtained in four independent plates with three different batches of 3.

Pure human pancreatic islets isolated from digested pancreatic tissues, as seen under a microscope. Image reprinted from reference 31 with permission of Baishideng Publishing Group Co.

Static incubation assays using human pancreatic islets. Islets were exposed to KRB containing 2.8 mM glucose for 30 min and were then exposed to KRB containing 16.7 mM glucose with GLP-1 or 3 or without test substance. GSIS resulted in a 1.4 fold insulin secretion while a 2.1 fold increase was observed for 3, which is comparable to GLP-1 (2.8 fold).
Figure 40. Potential sites available for carboxylic acid formation that also maintain, at least in part, B_{12} uptake protein recognition and binding are indicated in red. The atom numbering scheme used for future NMR discussion (see section 4.2.2) is also indicated.

Figure 41. The structure of 2-iodoxybenzoic acid (IBX).

Figure 42. Purposed mechanism for the formation of 4 utilizing IBX and 2-hydroxypyridine (HYP).

Figure 43. The RP-HPLC spectra of 4 (T_r = 13.5 mins.) monitored at 360 nm. Unreacted B_{12} and 5' aldehyde co-elute at ~ T_r = 6 mins as indicated by MALDI-ToF MS.

Figure 44. MALDI-ToF MS spectrum of HPLC peak at T_r = 14 mins indicating 4 ([M-CN]^+ 1343.5 m/z) with angiotensin as an internal reference.

Figure 45. Absorbance spectrum of 4 in 35 % HPLC mobile phase B.

Figure 46. 500 MHz \textsuperscript{1}H NMR spectrum of 4.

Figure 47. 500 MHz \textsuperscript{1}H-\textsuperscript{13}C HSQC NMR spectrum of 4.

Figure 48. 500 MHz \textsuperscript{1}H-\textsuperscript{13}C HMBC NMR spectrum of 4.

Figure 49. 500 MHz \textsuperscript{1}H NMR spectrum of B_{12} (CNCbl).

Figure 50. Portion of HSQC of 4 (right) and B_{12} (left) highlighting the loss of protons (\delta_H = 3.75 and 3.92) due to oxidation at the R5 carbon (\delta_C = 63.3) in the 4.

Figure 51. Synthesis of 5.
Figure 52. Purification of the conjugate was established by RP-HPLC using an SAX column with a gradient from 100% water to 35% MeCN: phosphate buffer pH 2.

Figure 53. MALDI-ToF MS spectrum of 5 ([M-CN+H]+ 1432.5). Angiotensin was used as an internal reference (not shown).

Figure 54. 500 MHz $^1$H NMR spectrum of 5.

Figure 55. 500 MHz $^1$H-$^{13}$C HSQC NMR spectrum of 5.

Figure 56. 500 MHz $^1$H-$^{13}$C HMBC of 5.

Figure 57. Analysis of the 5 by 2D heteronuclear NMR spectroscopy. A. The atom numbering scheme for the benzylamine attached to the R5 of B$_{12}$ by an amide bond. B. Signal assignments made from HSQC. C. Signal assignments made from HMBC (Both $^1$H and $^{13}$C assignment for C$_A$ were made by HSQC).

Figure 58. Rendering of B$_{12}$-insulin (yellow-blue) bound by TCII (red and dark blue) produced by collaborator Dr. Damien Allis at Syracuse University. It is possible for such a structure to bind either the TCII receptor CD320 or the insulin receptor CD220 as indicated. The image was altered from 2 with permission of The Royal Society of Chemistry (RSC) on behalf of the Centre National de la Recherche Scientifique (CNRS).

Figure 59. Structure of B$_{12}$ with side chain propionamides indicated. Conversion of the propionamide at the b position to a carboxylic acid is cited as reducing TCII protein binding of B$_{12}$.
Figure 60. RP-HPLC of B$_{12}$ b carboxylic acid derivative.

Figure 61. Competitive binding study of human TCII vs. mouse TCII with B$_{12}$-Insulin (AKP1), B$_{12}$ b carboxylic acid-insulin (AKP2), B$_{12}$ b carboxylic acid (Acid) and control B$_{12}$ from two sources (B$_{12}$ USA and B$_{12}$ DK).
List of Tables

Table 1.  $\text{B}_{12}$-small molecule conjugates  

Table 2.  Optimization of the Reaction Conditions for 4.  

Table 3.  $^1\text{H}$ and $^{13}\text{C}$ chemical shifts for the $\text{B}_{12}$ 5'-carboxylic acid ($\text{B}_{12}\text{CA}$), $\text{B}_{12}$ 5'-carboxylic acid benzylamine ($\text{B}_{12}\text{CABA}$) and $\text{B}_{12}$ (CNCbl) in D$_2$O [500 MHz].
## List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Full Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>aa</td>
<td>Amino acid</td>
</tr>
<tr>
<td>AC</td>
<td>Adenyl cyclase</td>
</tr>
<tr>
<td>AdoCbl</td>
<td>Adensylcobalamin</td>
</tr>
<tr>
<td>ADP</td>
<td>Adenosine diphosphate</td>
</tr>
<tr>
<td>ANTIDE</td>
<td>Luteinizing hormone releasing hormone antagonists</td>
</tr>
<tr>
<td>ATP</td>
<td>Adenosine triphosphate</td>
</tr>
<tr>
<td>AUC</td>
<td>Area under the curve</td>
</tr>
<tr>
<td>B$_{12}$</td>
<td>Vitamin B$_{12}$, cobalamin</td>
</tr>
<tr>
<td>B12Act</td>
<td>B$_{12}$ activated by CDT</td>
</tr>
<tr>
<td>BOC</td>
<td>t-butyloxy carbonyl</td>
</tr>
<tr>
<td>cAMP</td>
<td>Cyclic adenosine monophosphate</td>
</tr>
<tr>
<td>CD</td>
<td>Circular dichroism</td>
</tr>
<tr>
<td>CD220</td>
<td>Insulin receptor</td>
</tr>
<tr>
<td>CDC</td>
<td>Center for Disease Control and Prevention</td>
</tr>
<tr>
<td>CDI</td>
<td>1, 1'-carbonyl diimidazole</td>
</tr>
<tr>
<td>cDNA</td>
<td>Complementary deoxyribonucleic acid</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>CDT</td>
<td>1, 1’-carbonyl-di-(1, 2, 4-triazole)</td>
</tr>
<tr>
<td>CH3CN</td>
<td>Acetonitrile</td>
</tr>
<tr>
<td>CHCA</td>
<td>α-cyano-4-hydroxycinnamic acid</td>
</tr>
<tr>
<td>CICR</td>
<td>Calcium induced calcium release</td>
</tr>
<tr>
<td>CNCbl</td>
<td>Cyanocobalamin</td>
</tr>
<tr>
<td>CNRS</td>
<td>Centre National de la Recherche Scientifique</td>
</tr>
<tr>
<td>CoA</td>
<td>Coenzyme A</td>
</tr>
<tr>
<td>CPE</td>
<td>Carboxypeptidase E</td>
</tr>
<tr>
<td>CR</td>
<td>Cysteine-rich region</td>
</tr>
<tr>
<td>CRE</td>
<td>cAMP response element</td>
</tr>
<tr>
<td>CT</td>
<td>C-terminal domain</td>
</tr>
<tr>
<td>CUB</td>
<td>Components C1r/C1s, the sea urchin protein Uegf, and bone morphogenic protein-1</td>
</tr>
<tr>
<td>DAG</td>
<td>Diacylglycerol</td>
</tr>
<tr>
<td>DCC</td>
<td>N,N’- dicyclohexylcarbodiimide</td>
</tr>
<tr>
<td>DM</td>
<td>Diabetes mellitus</td>
</tr>
<tr>
<td>DMB</td>
<td>5,6-dimethylbenzimidazole</td>
</tr>
<tr>
<td>DMF</td>
<td>Dimethylformamide</td>
</tr>
</tbody>
</table>
DMSO  Dimethyl sulfoxide

DPP-IV  Dipeptidyl peptidase IV

DTPA  Diethylenetriamine-N,N,N',N"'-pentaacetic acid

DTT  Dithiothreitol

EAS  Electronic absorption spectrometry

ECD  Extracellular domain

EC_{50}  Half maximal effective concentration

EDAC  1-ethyl-3-(3-dimethylaminopropyl)carbodiimide

EDC  Ethyl-3-(3-dimethylaminopropyl)carbodiimide

EGF  Epidermal growth factor

Enz  Methionine synthase

EPO  Erythropoietin

ER  Endoplasmic reticulum

Et2O  Diethyl ether

FBS  Fetal bovine serum

FMOC-K34R-GLP-1  Glucagon-like peptide (7-36) amide with a K34R amino acid substitution and N-terminal FMOC protection

FMOC-OSu  N-(9-Fluorenlymethoxycarbonyloxy)succinimide
Fn0-1-2  Fibronectin domains
Fura-2 AM  Fura-2 acetoxyethyl ester
G-CSF  Granulocyte-colony stimulating factor
GIIS  Glucose induced insulin secretion
GIP  Glucose-dependent insulinotropic polypeptide
GIT  Gastrointestinal tract
GLP-1  Glucagon like peptide-1
GLP-1R  Glucagon-like-peptide-1 receptor
GPCR  G protein-coupled receptor
GRP  Gastrin-releasing peptide
Gs  Gs alpha subunit
GSIS  Glucose stimulated insulin secretion
HC  Haptocorrin
HEK-293  Human embryonic kidney
HEK-GLP-1R  HEK-293 cells stably expressing the human GLP-1 receptor
HOBt  1-hydroxybenzotriazole
HPMA  Lysine-modified-hydroxypropyl-methacrylamide
hr  
Hour

HYP  
2-hydroxypyridine

IBX  
2-iodoxybenzoic acid

ICR  
Islet Cell Resource

IF  
Intrinsic factor

Ins  
Insulin insert domain

IP3  
Inositol 1,4,5-triphosphate

JM  
Juxtamembrane regions

K34R-GLP-1  
Glucagon-like peptide-1 (7-36) amide with a K34R amino acid substitution

KRB  
Krebs-ringer buffer

L1  
First leucine-rich-repeat domains

L2  
Second leucine-rich-repeat domains

LDL  
Low density lipoprotein

LMBD1  
Lipocalin receptor like protein

Luc  
Luciferase

MALDI-ToF  
Matrix assisted laser desorption ionization time of flight

MeCbl  
Methylcobalamin
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>min</td>
<td>Minutes</td>
</tr>
<tr>
<td>MRP1</td>
<td>Multidrug resistance protein 1</td>
</tr>
<tr>
<td>MS</td>
<td>Mass spectrometry</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular weight cutoff</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear magnetic resonance</td>
</tr>
<tr>
<td>NO</td>
<td>Nitric oxide</td>
</tr>
<tr>
<td>NOCbl</td>
<td>Nitrosylcobalamin</td>
</tr>
<tr>
<td>PC 1/3</td>
<td>Prohormone convertases 1/3</td>
</tr>
<tr>
<td>PC 2</td>
<td>Prohormone convertases 2</td>
</tr>
<tr>
<td>PEG</td>
<td>Poly(ethylene glycol)</td>
</tr>
<tr>
<td>RER</td>
<td>Rough endoplasmic reticulum</td>
</tr>
<tr>
<td>RIP1-CRE</td>
<td>Promoter region of the rat insulin 1 gene a cAMP response element</td>
</tr>
<tr>
<td>RIP1-CRE-Luc</td>
<td>Rat insulin 1 promoter-cAMP response element-luciferase reporter</td>
</tr>
<tr>
<td>RNR</td>
<td>Ribonucleotides reductase</td>
</tr>
<tr>
<td>RP-HPLC</td>
<td>Reverse-phase high pressure liquid chromatography</td>
</tr>
<tr>
<td>RSC</td>
<td>The Royal Society of Chemistry</td>
</tr>
<tr>
<td>RYR</td>
<td>Ryanodine receptor</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Name</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>STZ</td>
<td>Streptozocin</td>
</tr>
<tr>
<td>T1D</td>
<td>Type I diabetes mellitus</td>
</tr>
<tr>
<td>T2D</td>
<td>Type II diabetes mellitus</td>
</tr>
<tr>
<td>TBTU</td>
<td>O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate</td>
</tr>
<tr>
<td>TCII</td>
<td>Transcobalamin II</td>
</tr>
<tr>
<td>TEA</td>
<td>Triethylamine</td>
</tr>
<tr>
<td>TFA</td>
<td>Trifluoroacetic acid</td>
</tr>
<tr>
<td>TK</td>
<td>Tyrosine kinase domain</td>
</tr>
<tr>
<td>TM</td>
<td>Transmembrane regions</td>
</tr>
<tr>
<td>Tr</td>
<td>Retention time</td>
</tr>
<tr>
<td>TSTU</td>
<td>((N,N,N',N'-Tetramethyl-O-(N-succinimidyl) uronium tetrafluoroborate</td>
</tr>
<tr>
<td>TTHA</td>
<td>triethylenetetramine-N,N,N',N&quot;,N&quot;&quot;,N&quot;&quot;-hexacetic acid</td>
</tr>
<tr>
<td>TZD</td>
<td>Thiazolidinedione</td>
</tr>
</tbody>
</table>
Chapter 1 General Introduction

1.1 Diabetes Mellitus

Diabetes mellitus (DM) is a series of diseases characterized by the body’s inability to produce (type I, T1D) or properly use (type II, T2D) the hormone insulin, resulting in abnormally high blood glucose levels.\(^1\) T1D accounts for 5-10\% of all newly diagnosed cases and is a result of the body’s immune system destroying pancreatic \(\beta\) cells,\(^2,3\) which are the only cells in the body capable of producing insulin. In addition to genetic factors, environmental conditions have also been described as risk factors for T1D.\(^2,4\) There is no known prevention for T1D and it is usually diagnosed in people younger than 30.\(^2\) In T2D, there is an inability to produce enough insulin to meet the body’s insulin sensitivity level. As the disease progresses, pancreatic \(\beta\) cell mass declines and the insulin deficiency increases. The onset of T2D is usually later in life, although it has recently been diagnosed in children.\(^1,5,6\) T2D is much more prevalent then T1D, accounting for 90-95\% of all new cases. Research suggests physical inactivity and excess body weight increases your risk for T2D.\(^7-11\)

According to a study conducted in 2011 by the Center for Disease Control and Prevention (CDC), 8.3\% of the United States’ population is affected by DM.\(^1\) The number of DM cases is expected to increase significantly in the coming years, if current trends continue.\(^1\) DM is currently the third leading cause of death in the United States, just behind cancer and heart disease.\(^1,12\) The high morbidity rate associated with DM has been attributed to the microvascular (retinopathy, nephropathy, and neuropathy) and macrovascular (cardiovascular disease) complications that result due to prolonged hyperglycemia.\(^13\) The annual diabetes-related spending in the United States is also expected to increase from $113 billion to $336 billion by
The increasing prevalence and growing costs associated with DM has made it a major healthcare concern in the United States today.\textsuperscript{11, 12, 14}

1.2 Injectable Peptides and Proteins for the Treatment of DM

One of the primary goals for the treatment of DM is achieving and maintaining metabolic norms in regards to blood glucose levels (see Figure 1).\textsuperscript{15} Fluctuation in blood glucose levels often leads to serious health complications including retinopathy, nephropathy, and neuropathy, as previously mentioned.\textsuperscript{13}

\textbf{Figure 1.} The primary treatment goal for DM is to maintain glycemic control (indicated as normal glucose level). Glycemic control in a patient with diabetes can be challenging since the proper amount of insulin must be administered after meals. Not injecting enough or injecting too much insulin will result in hyperglycemia or hypoglycemia, respectively. While persistent hyperglycemia causes microvascular complications, hypoglycemia can cause diabetic coma or death.\textsuperscript{13, 16}
Patients with T1D require an exogenous source of insulin, often administered by subcutaneous injection, to maintain normal glycemic levels. In T2D the body still produces insulin, but the cells are unable to use it efficiently. The initial treatment for T2D is diet, exercise and oral anti-diabetic medications such as metformin, sulphonylureas, and thiazolidinedione (TZD) in an attempt to slow or reverse the disease. These oral treatments lose efficacy over time as the disease progresses and a transition to injectable medications is then required. Injectable insulin has been the primary medication for DM since its discovery in 1922. Recent research has focused on the importance of incretin peptides like glucagon like peptide-1 (GLP-1) for the co-management of DM.

1.2.1 Insulin

1.2.1.1 History and Discovery

A number of important scientific observations led to the ultimate discovery of insulin in 1922 by Canadian scientists Frederick Banting, Charles Best, John Macleod and Bertram Collip. In 1869, Paul Langerhans, a German medical student, observed a cluster of cells within the pancreatic tissues that did not have a reported function. This cluster would later be named the ‘islets of Langerhans’, and include insulin secreting β cells. Twenty years later, Oskar Minkowski and Joseph van Mering determined removing the pancreas of a dog resulted in diabetes, but ligating the duct through which pancreatic juices flow to the intestines only resulted in minor digestive problems, but not diabetes. This suggested the pancreas had at least two functions and one of them involved the regulation of glucose.

It was Dr. Frederick Banting that hypothesized that pancreatic digestive juice may be harmful to the pancreatic secretion produced by the islets of Langerhans. Banting convinced
Professor John Macleod at the University of Toronto to fund his idea and he began work in 1921. As a surgeon, Banting was able to ligate the pancreatic ducts prior to extraction of the pancreas. His work led to the successful isolation of insulin. The insulin extract was first tested in diabetic dogs with great success.\textsuperscript{28} The first attempt to inject the extract in humans, however, failed. With the help of biochemist Bertram Collip, the extract was further purified and used successfully in the treatment of a 14 year old diabetic patient in 1922.\textsuperscript{21,28} The news of the successful treatment of diabetes with insulin rapidly spread and in 1923 Banting and Macleod won the Nobel Prize in Physiology or Medicine. Banting decided to share his cash award with Best and Macleod shared his cash award with Collip.

\textbf{1.2.1.2 Biosynthesis and Secretion of Insulin}

Insulin is produced in the pancreatic $\beta$ cells of the islets of Langerhans. The process involves multiple steps beginning with the prohormones ‘preproinsulin’ and ‘proinsulin’ (see Figure 2).\textsuperscript{29} Initially, insulin mRNA is translated into the single chain precursor preproinsulin, which contains a 24 amino acid (aa), hydrophobic signal peptide directing the chain to the rough endoplasmic reticulum (RER).\textsuperscript{28,30,31} Shortly after entering the RER, a signal peptidase removes the signal peptide, forming proinsulin.
Figure 2. The subcellular organization of the insulin secretory pathway.

Proinsulin consists of a carboxyl-terminal A-chain, an amino-terminal B-chain, and a connecting peptide as the C-peptide (see Figure 3). In the RER lumen, proinsulin folds and three intrachain disulfide bonds form, linking residues 25-54 (B-chain) and residues 90-110 (A-chain) and giving insulin its native structure. The folding process is believed to be aided by a connecting segment at residue 55-89 (C-peptide), which is subsequently removed. Once folded, proinsulin is shuttled to the Golgi apparatus and is packaged into secretory vesicles. In the vesicles, proinsulin is cleaved by the prohormone convertases 1/3 (PC1/3), 2 (PC 2) and carboxypeptidase E (CPE), finally generating active insulin and C-peptide. PC1/3 and PC2
cleave between residue 32 and 33 at the C-peptide/B-chain junction and residue 65 and 66 at the C-peptide/A-chain. The removal of extra basic residues at the C-terminal end of the B-chain and C-terminal end of C-peptide occurs by CPE.

**Figure 3.** The post-translational modifications of preproinsulin, which results in active monomeric insulin and C-peptide.

Pancreatic β cells secrete insulin in a tightly regulated fashion to ensure glucose homeostasis. The mechanism for insulin secretion is a complex process involving the integration of intracellular and extracellular components and is still not completely understood (see Figure 4). Nutrients, hormones, and neurotransmitters have all been shown to directly
Glucose is the most important physiological regulator of insulin secretion. Calcium, adenosine triphosphate (ATP), cyclic adenosine monophosphate (cAMP) and phospholipid-derived molecules such as diacylglycerol (DAG) and inositol 1,4,5-triphosphate (IP3) have all been identified as key intracellular components as well. Glucose induced insulin secretion (GIIS) occurs by three different pathways: the triggering pathway ($K_{\text{ATP}}$ channel-dependent pathway), the metabolic amplifying pathway ($K_{\text{ATP}}$ channel-independent pathway) and the neurohormonal amplifying pathway. During the triggering pathway, intracellular transported glucose is metabolized causing an increase in ATP concentration and subsequent closure of ATP sensitive $K^+$ channels, which causes depolarization of the cell membrane and opens calcium channels, allowing calcium influx. The rise in intracellular $[Ca^{2+}]$ triggers the fusion of insulin granules to the plasma membrane. The mechanism of the metabolic amplifying pathway is less understood and believed to be more complex. It is known that the triggering pathway must be operational and there must be an increase in intercellular $[Ca^{2+}]$ levels for the metabolic amplifying pathway to occur. In addition to glucose stimulation, hormonal inputs are important for normal regulation of insulin and occur by the neurohormonal amplifying pathway, which has recently been studied and is summarized in Figure 4.
**Figure 4.** Glucose induced insulin secretion (GSIS) is regulated by three pathways: triggering pathway, metabolic amplify pathway and neurohormonal amplify pathway.\(^3^8\) Gs: Gs α subunit, AC: adenyl cyclase.

**1.2.1.3 Structure and Sequence of Bovine Insulin**

In 1955, insulin became the first protein to be fully sequenced, winning Frederick Sanger the 1959 Nobel Prize for Chemistry.\(^4^0\) Insulin’s primary structure consists of two polypeptide chains, a 30 aa B-chain and a 21 aa A-chain (see Figure 5).\(^4^1\) The amino acids are referred to by the letter of the chain (A or B) and the position within the chain. The aa sequence of insulin is highly conserved among different species, which allows the cross species use of insulin, for the most part.\(^4^2\) For research purposes the use of cross species insulin as a screen for biological
purpose is appropriate and well-studied. However, a transition to clinical use requires the use of human insulin to prevent immunological issues.

Bovine insulin, which was the source of insulin used for the work presented here, differs from human insulin by three aa: alanine for threonine at A8, valine for isoleucine at A10, and alanine for threonine at B30. The difference in aa does not alter the highly conserved secondary structure of the insulin. The insulin structure contains one intra-chain disulfide bond connecting residues 6 to 11 of the A-chain (A6-A11) and two inter-chain disulfide bonds linking the A- and B-chain (A7-B7 and A20-B19) (see Figure 5).

The secondary structure of insulin includes a helical section between residues B9 and B19, as well as a N-terminal helix (A1-A8) linked to an antiparallel C-terminal helix (A12-20).

![Insulin Primary Structure Diagram](image_url)

**Figure 5.** A representation of the primary structure of bovine insulin including the disulfide bonds formed between the A- and B-chains and the A-chain intrachain disulfide bond.

Insulin self-associates into dimers at micromolar concentrations, which is the concentration found in β cells during expression, suggesting dimers may be present in the
endoplasmic reticulum. The residues involved in dimerization are all located in the B-chain and have been identified as B8, B9, B12, B13, B16, B23-B29. Insulin further associates into hexamers at millimolar concentration in the presences of zinc ions. Insulin is stored as a hexamer in vacuoles of the β cells when blood glucose levels are low. When blood glucose is elevated, the hexamers are released into the blood where they dissociate into dimers and monomers. The dimer and hexamer formation is important for stabilization during storage. The cellular actions of insulin are initiated by monomeric insulin binding to its plasma membrane receptor.

1.2.1.4 The Insulin Receptor

The insulin receptor (CD220) is part of the receptor tyrosine kinase superfamily and is present on virtually all mammalian cells, although the concentration of CD220 varies across different cells. CD220 is a heterotetrameric glycoprotein consisting of two extracellular α-domains and two membrane spanning β-domains. The domains are connected by disulfide bonds forming a β-α-α-β structure. In simplest terms, insulin binds to an extracellular α-domain and causes phosphorylation of an intercellular β-domain.

The crystal structure of the CD220 ectodomain was published in 2006 and provided valuable information about the structural organization of the receptor (summarized in Figure 6) and domains involved in insulin binding. The crystal structure of CD220-insulin complex has not yet been acquired, but the CD220 crystal structure combined with various experiments involving insulin analogues has provided a basis for a reasonable CD220-insulin binding model. The insulin residues involved in receptor binding include residues A1, A2, A3, A5,
A19, A21, B12, B16, and B23-26.\textsuperscript{51-53} It has also been suggested that residues A13 and B17 are involved in receptor binding.\textsuperscript{29}

\textbf{Figure 6.} A carton representation of the domain organization within the homodimeric insulin receptor. L1, L2: first and second leucine-rich-repeat domains, CR: cysteine-rich region, Fn\textsubscript{0-1-2}: fibronectin domains, Ins: Insulin insert domain, TM/JM: trans- and juxta-membrane regions, TK: tyrosine kinase domain and CT: C-terminal domain.
1.2.2 Glucagon-Like Peptide-1

Glucagon-like-peptide-1 (GLP-1) is a gut hormone, produced in intestinal L cells, that increases the amount of insulin secreted from pancreatic β cells in response to ingested nutrients, also known as an incretin (see section 1.2.2.1). Orally delivered glucose produces a dramatic increase in the secretion of insulin compared to injected glucose. This response to oral glucose is referred to as the incretin effect and accounts for 70% of the total postprandial insulin secretion.

The role of GLP-1 in the postprandial homeostasis of glucose was established in the late 1980s and early 1990s. Recent preclinical and clinical research has established GLP-1’s role in increasing satiety, promoting modest weight loss, improving β cell function, and preserving β cell mass by inducing proliferation, which may help to slow or reverse the progression of T2D. The insulinotropic activity of GLP-1 is glucose dependent, making it an ideal treatment for hyperglycemia, since hypoglycemia is unlikely to occur as a result of treatment.

1.2.2.1 Discovery of the Incretin Effect

It was first speculated in the early 1900s that there was a component in the gut that was released after oral ingestion of nutrients, which caused insulin secretion. Zunz and Labarre were able to extract an intestinal component that produced a lowering of glucose levels in dogs. They introduced the term “incretin” to describe the extract. Several other groups pursued the gut component with some reports of the presence of glucose lower abilities. However, Loew and colleagues reported the inability to lower blood glucose levels in dogs with the extract and...
Interest in the extract quickly declined. It was later determined that Loew’s experiments involved fasting animals, which would not respond to the glucose dependent GLP-1.

Interest in the extract rekindled in 1960s with the development of a reliable radioimmunoassay for insulin, by Yalow and Berson. The ability to measure the circulating levels of insulin allowed researchers to determine the action of glucose alone could not account completely for the insulin secretion. In 1969, Creutzfeldt defined the criteria for an incretin as 1) nutrients like carbohydrates must cause incretin secretion and 2) at physiological levels, elevated blood glucose levels must cause it to stimulate insulin secretion. The first incretin, glucose-dependent insulinoactive polypeptide (GIP) was identified in the 1970s. In the 1980s, with the help of complementary deoxyribonucleic acid (cDNA) cloning, GLP-1 was identified.

1.2.2.2 Biosynthesis and Secretion

GLP-1 is produced in the intestinal L cells, which are located in the ileum and colon. As with insulin, the synthesis of GLP-1 involves a larger precursor, which is then posttranslationally modified to the functional form. The process begins with the prohormones “preproglucagon” and “proglucagon”. The proproglucagon has a signal peptide directing it to the secretory pathway, which is quickly removed to produce proglucagon. Posttranslational cleavage of proglucagon is organ specific with PC2 producing glucagon in pancreatic α cells and PC1/3 producing GLP-1 in intestinal L cells.

GLP-1 was originally synthesized and tested as 36 and 37 aa peptides, GLP-1 (1-36) amide and GLP-1 (1-37), respectively. The sequences were based on a predicted posttranslational modification based on the understanding that PCs cleave a prohormone to the active hormone at a dibasic aa sequence. Therefore, the cleavage was assumed to occur.
between the Lys and Arg resulting in a 37 or 36 aa peptide, depending on the presence or absence of the Gly residue at the C-terminus. C-terminal amidation of peptide hormones is common and involves the oxidation of a C-terminal Gly, resulting in the conversion of a carboxyl to an amide on the preceding aa. The conversion to the amide does not occur completely, resulting in non-amidated and amidated peptide. Amidation of the GLP-1 enhances the survival of GLP-1 in plasma.

After early tests determined the GLP-1 (1-36) amide and (1-37) isoforms had no insulinotropic effects, reanalysis of the aa sequence determined His7 should be the N-terminal amino acid. The reanalysis was based on alignment of other membranes of the glucagon superfamily of peptides. Testing of GLP-1 (7-36) amide and GLP-1 (7-37) confirmed the truncated forms were indeed the insulinotropic forms of the peptide. The functions of the GLP-1 isoforms GLP-1 (1-36) amide and (1-37) are still unknown. There are two biologically active forms of GLP-1: GLP-1 (7-37) and GLP-1 (7-36) amide, and both are found in the circulation of humans. GLP-1 (7-36) amide is the predominant form, accounting for typically 80% of circulating GLP-1.

GLP-1 secretion is stimulated by ingested nutrients and occurs in a biphasic pattern. Initially there is a rapid increase in the circulating GLP-1 level at 15-30 min after nutrient ingestion, followed by a second minor increase of GLP-1 at 90-120 min. The second increase of GLP-1 is believed to be a result of nutrients interacting directly with the L cell to cause GLP-1 secretion. Absence of L cells in the small intestines suggests the initial GLP-1 secretion is mediated indirectly, rather than by direct L cell contact with nutrients. Brubaker et al. proposed a mechanism involving the neuro/endocrine pathway for the indirect mediation of GLP-1 secretion. It was proposed that nutrients in the duodenum activate a proximal-distal
neuro/endocrine loop, which stimulates GLP-1 secretion from L cells in the ileum and colon. The incretin hormone GIP, the neurotransmitter acetylcholine and neuropeptide gastrin-releasing peptide (GRP) have all been identified as key regulators of GLP-1 secretion involved in the proximal-distal neuro/endocrine loop.  

1.2.2.3 Structure and Sequence of Human GLP-1

GLP-1 is a 3.3 kDa peptide containing 30 aa residues (see Figure 7). The primary structure is highly conserved in mammalian species. The conservation of GLP-1 reflects the fact that basically the entire aa sequence is required for full biological activity. Removal or addition of amino acids at the N-terminus of the GLP-1 (7-36) amide or GLP-1 (7-37) results in a dramatic loss of receptor binding and biological activity. Addition of an aa residue at the C-terminal also results in a significant reduction in biological activity.

![Figure 7. The primary structure of human GLP-1 (7-36) amide.](image)
The C-terminal amidation has been shown to increase plasma stability.
Site specific mutation studies identified the following amino acids as important for receptor binding and activity: 1 (His), 4 (Gly), 6 (Phe), 7 (Thr), 9 (Asp), 22 (Phe) and 23 (Ile).\(^{82}\) NMR studies of the structure of GLP-1 identified the N-terminus as a random coil from residues 1-7, two helical segments at residues 7-14 and 18-29 and a linker region at 15-17 (see Figure 8).\(^{84}\)

![Figure 8](image)

**Figure 8.** The secondary structure of GLP-1 (7-36) amide determined by solution NMR. Lys 26 (indicated as ball and stick) has been identified as the best position for conjugation.\(^ {85}\) Protein data bank ID: 1D0R. Modified using RCSB PDB protein workshop software 3.9.

### 1.2.2.4 The Glucagon-like Peptide Receptor

The therapeutic potential of GLP-1 receptor (GLP-1R) targeting for the treatment of T2D has been validated through extensive basic, clinical and translational research.\(^ {63,67,68,79,86}\) GLP-1 mediates its insulinotropic effects by binding to GLP-1Rs on pancreatic β cells. The GLP-1R is a seven-transmembrane class B heterotrimeric G protein-coupled receptor (GPCR). The B class is a subfamily of GPCR distinguishable by a large N-terminal extracellular domain (ECD),
which are typically 120 residues long and contain six highly conserved cysteine residues.\textsuperscript{87} The ECD is responsible for peptide recognition and selective binding. GLP-1R’s ECD is composed of an N-terminal α-helix, two anti-parallel β strands, and a final lobe composed of two additional anti-parallel β sheets and a short α-helical region.\textsuperscript{88}

The X-ray crystal structure (2.1 Å resolution) of GLP-1 bound to the ECD of GLP-1R identified GLP-1 (7-36) amide as a continuous α helix from amino acid residues 13 to 33, with a kink around residue 22.\textsuperscript{76} Interaction between the ECD and GLP-1 were noted between residues 24 and 33. Specifically, the hydrophobic face of GLP-1’s α-helix interacts with ECD at residues 24, 25, 28, 29, 32 and 33.\textsuperscript{76}

1.3 A Pursuit of a Noninvasive Peptide/Protein Treatment for Diabetes Mellitus

The oral delivery of peptides and protein is a major goal for therapeutics.\textsuperscript{89} Early work to develop an oral route focused primarily on a desire to increase patient adherence to medication and quality of life.\textsuperscript{90} Recent research has evolved to place new emphasis on the physiological importance of an orally administered peptide or protein,\textsuperscript{90,91} with such delivery postulated to bring a patient closer to metabolic norms through direct effects on hepatic glucose production.\textsuperscript{89} The major problem facing oral delivery of peptides and proteins is gastrointestinal tract (GIT) hydrolysis/proteolysis and the inability to passage the enterocyte.\textsuperscript{92}

The initial barrier for oral delivery of peptide and proteins is the enzymatic barrier.\textsuperscript{89} Pepsin is the major enzyme found in the stomach and is a key element in digestion. In the intestines there are three major proteases responsible for further digestion: trypsin, chymotrypsin and carboxypeptidase. Degradation of the peptide/protein by proteolysis will result in a reduced bioavailability.\textsuperscript{89}
If the peptide/protein makes it through the proteolytic barrier, the next major barrier is a physical barrier produced by the mucin lining of the intestines. Mucus plays an important role in determining the absorption and bioavailability of orally administered drugs.\textsuperscript{93} The intestinal wall is the next barrier for drug delivery and is also a physical barrier. Drug delivery is focused on the mucous layer of the intestinal wall, which is comprised of villi that extend out from the epithelial lining and are responsible for absorption of the majority of nutrients.\textsuperscript{94} Adsorption of nutrients from the intestine can occur by two major pathways: transcellular and paracellular (see Figure 9).

![Figure 9](image)

**Figure 9.** An illustration of the physical barriers and available pathways for oral drug delivery in the intestines.

The transcellular pathway involves passive or active diffusion through the intestinal epithelial cells, while the paracellular pathway takes place in the interstitial space between the epithelial cells. In transcellular passive diffusion, small nutrients can cross through nonspecific
permeability pathways. Larger nutrients, like vitamin B$_{12}$, must use a receptor mediated uptake pathway to utilize the transcellular route.$^{89}$ The paracellular route is regulated by epithelial tight junction proteins and is only open to small solutes, ions and water.$^{95}$ The tight junctions are important for keeping the cells tightly bound to one another.

1.3.1 Strategies for Oral Delivery of Peptides and Proteins

Numerous strategies have been attempted to overcome the barriers associated with oral delivery of peptides/proteins. The methods most commonly used include the use of enzyme inhibitors,$^{93,96-98}$ encapsulation,$^{99,100}$ chemical modification$^{101}$ and/or absorption enhancers.$^{102-104}$ While there have been great strides towards the delivery of oral peptides/proteins because of these strategies, all of the methods have negative aspects that have kept them from reaching a clinical level.$^{102}$

Early research focused on the use of enzyme inhibitors to reduce the rate of proteolytic degradation of orally delivered peptides/proteins.$^{98}$ The potential for this technique was demonstrated by the use of a duck ovomucoid, which provided insulin 100% protection from trypsin and chymotrypsin.$^{93}$ The major concern with the use of enzyme inhibitors is the potential side effects associated with continuous use including the potential for alteration to the body’s metabolism.$^{102}$

Absorption enhancers work by altering the permeability of the cell membrane and include surfactants, chelating agents, zonular occludens toxin, and fatty acids.$^{102}$ Detergents and surfactants disrupt the lipid bilayer of the cell membrane, enhancing the transcellular pathway.$^{94,95}$ Chelating agents and zonular occludens toxin disrupts the integrity of the tight junctions enhancing the paracellular pathway.$^{103,104}$ A major drawback to the use of absorption enhancers
is that by altering the permeability of the cell membrane potential toxins and pathogens have access to the circulatory system.\textsuperscript{102}

Various encapsulation techniques, including emulsions, liposomes, microsphere, and nanoparticles, have been attempted to protect peptides/protein.\textsuperscript{99,101,105} Nanoparticle based peptide and protein drugs have gained a lot of attention in recent years. Nanoparticles are usually designed to deliver a large encapsulated load of peptide/protein to the intestines where a trigger (often pH based) releases the load for uptake at the intestinal wall.\textsuperscript{99,106} The difficulties with the use of nanoparticles include the lack of specific control of drug release and the potential for buildup of non-degradable particles in tissues.\textsuperscript{99,102}

Chemical modification of peptides/protein includes amino substitution, to eliminate recognition sites of proteases or the attachment of various chemical moieties that are known to protect the peptide or protein from proteolysis.\textsuperscript{101,106-109} One example of attachment of a chemical moiety is the addition of PEG, which has the ability to protect peptides and proteins from intestinal proteolysis as well as alter the interactions of the molecular with blood proteins, allowing for an increase in systemic circulation time.\textsuperscript{101,106-109}

1.4 Vitamin B\textsubscript{12}

1.4.1 History

The early history of vitamin B\textsubscript{12} (B\textsubscript{12}) belongs entirely to the field of medicine and the treatment of pernicious anemia. Pernicious anemia was first described in 1824 and was an incurable and usually fatal disease.\textsuperscript{110} In 1920, Whipple discovered a diet of raw liver could regenerate the red blood cells of anemic dogs and in 1926, Minot and Murphy reported the same
findings in anemic patients.\textsuperscript{111} This discovery won all three the 1934 Nobel Prize in Physiology and Medicine.\textsuperscript{110}

In 1948, B\textsubscript{12} was isolated from the liver and determined to be the compound responsible for the improvement in pernicious anemia. The isolation was performed by two independent groups, Folkers and coworkers at Merck Laboratories in the U.S. and Smith and Parker at Glaxo Laboratories in the U.K.\textsuperscript{112,113} The X-ray crystal structure of B\textsubscript{12} was obtained in 1955 by Dorothy Hodgkin.\textsuperscript{110,114} This complex crystal structure of B\textsubscript{12} won Hodgkin the 1964 Nobel Prize for Chemistry. In 1972, Woodward and Eschenmoser reported the total synthesis of B\textsubscript{12}.\textsuperscript{115} It took over 100 scientists eleven years to complete this feat.\textsuperscript{116} Since these early years, thousands of papers have been published on the chemistry and biochemistry B\textsubscript{12}.

\textbf{1.4.2 Structure and Chemistry of Vitamin B\textsubscript{12}}

B\textsubscript{12}, or cobalamin, is the most structurally complex cofactor existing in nature.\textsuperscript{116} B\textsubscript{12} is a member of a family of cob(III)alamin species all of which are comprised of a cobalt (III) ion equatorially coordinated to four pyrrolic nitrogens of a corrin ring (see Figure 10).\textsuperscript{117} The corrin ring is similar to the more commonly known porphyrin structure but with a greater degree of saturation and increased sp\textsuperscript{3} carbons, which confers greater flexibility.\textsuperscript{117} In addition, there is greater asymmetry to the corrin ring compared to a porphyrin, due to the fusion of the A and D rings resulting in a 15-carbon ring structure over the porphyrin 16-carbon structure.\textsuperscript{117}
Figure 10. Structure of B<sub>12</sub> with a variable R group shown to the left. Three of the more common forms of B<sub>12</sub> (cyanocobalman, methylcobalamin and adenosylcobalamin) are shown to the right.

In humans, the cobalt atom is further bound to a 5,6-dimethylbenzimidazole (DMB) base at the α-axial position. The identity of the base varies in different species. DMB is connected to the D-ring of the corrin ring by a ribose-phosphate-amide bond linkage formed through the f amide side chain. The functional group attached at the upper β position of the cobalt ion, depicted as R in Figure 10, is variable depending on the type of cob(III)alamin species. The most common cob(III)alamin used in research and nutriceuticals is cyanocobalamin (see Figure 10, R = CN, CNCbl). CNCbl is not the biological active form of cob(III)alamin.
but is readily converted to the biologically active forms, methylcobalamin (R= CH₃, MeCbl) and adenosylcobalamin (R= 5’-deoxyadenosyl, AdoCbl) in the human body.¹¹⁹

The use of B₁₂ as a cofactor is a result of the ability of the cobalt atom to change oxidations states.¹¹⁶ In AdoCbl, the cobalt atom is converted from Co³⁺ to Co²⁺ (cob(II)alamin) by homolytic cleavage of the Co-C bond. In MeCbl, heterolytic cleave of the Co-C bond results in a Co⁺ (cob(I)alamin) species.¹¹⁶

1.4.3 Physiological Roles of B₁₂

B₁₂ plays an important role in the normal functioning of the brain and nervous system (myelin production), the formation of blood (tetrahydrofolate production), DNA synthesis (ribonucleotides reductase (RNR) conversion of ribonucleotides to deoxyribonucleotides), and fatty acid synthesis and energy production.¹²⁰⁻¹²² Many of B₁₂ functions can be replaced by folic acid (vitamin B₉) but two enzymes absolutely require B₁₂ as a cofactor, methylmalonyl coenzyme A mutase and methionine synthase.¹²²

Cytosolic methionine synthase is a MeCbl-dependent methyl transferase responsible for the formation of methionine by methylation of homocysteine.¹²³ The process is a catalytic cycle involving two methyl transfer half-reactions, outlined in Figure 11. The cobalt-bound methyl group of MeCbl is transferred to homocysteine, forming methionine and an enzyme-bound cob(I)alamin.¹²⁴ This highly nucleophilic cob(I)alamin then acquires the methyl group of N⁵-methyltetrahydrofolate, forming tetrahydrofolate and MeCbl.¹²³ Tetrahydrofolate is essential for the production of purines and pyrimidines in DNA synthesis.¹²³
Figure 11. Outline of methionine synthase (Enz) catalytic cycle, which involves two methyl transfer half-reactions: homocysteine to methionine and N^5^-methyltetrahydrofolate to tetrahydrofolate.

Methylmalonyl Coenzyme A mutase is an AdoCbl-dependent enzyme, which catalyzes the rearrangement of methylmalonyl-CoA to succinyl-CoA. This rearrangement occurs by the exchange of a hydrogen atom from the methyl group with the carbonyl-CoA group on the adjacent carbon. This conversion from methylmalonyl-CoA to succinyl-CoA is an important step in the degradation of branched-chain amino acids, and odd-chain fatty acids.
1.4.4 The Dietary Uptake Pathway Overview

B₁₂ is vital for the survival of all living organisms and is only synthesized naturally by bacteria, so mammals must acquire the vitamin through their diet. As a result, the body uses a complex uptake mechanism involving three soluble transport proteins (haptocorrin (HC), intrinsic factor (IF) and transcobalamin II (TCII)) to ensure successful isolation of B₁₂. The uptake pathway also utilizes four receptors (cubilin, amnionless, megalin, and TCblR/CD320), which only recognize a B₁₂ transport protein when bound to B₁₂. Any disruption in the uptake pathway is known to result in a B₁₂ deficiency. In the most basic terms, the uptake pathway begins with the protection of B₁₂ from the acidic environment of the stomach by HC, followed by the receptor mediated uptake of IF bound to B₁₂ in the intestines and B₁₂ delivery to the cell by TCII (see Figure 12).
Figure 12. An overview of the B\textsubscript{12} uptake pathway as outlined by a recent 2012 review.\textsuperscript{120} The transfer of B\textsubscript{12} from IF to TCII is currently being debated as discussed in section 1.4.7 and shown in Figure 15. Image is reproduced from reference 120 with permission of Nature Publishing Group. MRP1: multidrug resistance protein 1.

A number of important publications have been received in recent years providing a clearer picture of the exact mechanism of B\textsubscript{12} transport.\textsuperscript{125-132} There are still areas of the pathway that are not completely understood and some areas of the pathway that are controversial in the B\textsubscript{12} field. What follows is a review of the current understanding of the B\textsubscript{12} uptake pathway:
1.4.5 Haptocorrin

HC is a 45.6 kDa glycoprotein primarily secreted in human saliva at concentration of 50 nM. The predominant function of HC is the transport of B₁₂ through the acidic environment of the stomach and into the duodenum. In the duodenum, an increase in pH causes the release of B₁₂ and pancreatic proteases cleave the HC. The release and break down of HC results in the transfer of ~80% of bound B₁₂ to IF.¹³⁵

In the blood plasma, HC is believed to be responsible for the removal of partially synthesized or degraded B₁₂. HC is the only known B₁₂ transport protein capable of binding corrinoids, which are a large group of compounds containing the corrin ring with coordinated cobalt. Approximately 40% of plasma corrinoids are a B₁₂ analogue unable to serve as a cofactor. Due to HC presence in breast milk, tears and saliva, it was suggested that haptocorrin may also have antibacterial functions but a recent paper contradicts this hypothesis.¹²⁸

HC is the least understood of the B₁₂ transport protein. It was recently established that HC is not present in all species, which was an extremely important finding as a number of B₁₂ studies are conducted in nonhuman subjects.¹²⁷ No crystal structure has been reported, but the structure of HC is predicted to be very similar to IF (see section 1.4.5).¹³⁷ The recombinant expression of human haptocorrin was recently reported in human embryonic kidney (HEK) cells.¹³⁸ This successful expression will allow researchers the ability to produce and more extensively study HC in the coming years.

1.4.6 Intrinsic Factor

The second B₁₂ transport protein in the series, IF is a highly glycosylated protein secreted in the parietal cells and is crucial for B₁₂ transport across the intestinal wall.¹²⁵ IF binds B₁₂ in
the large intestines after proteases degrade HC and the pancreas secretes a bicarbonate solution that neutralizes the acidic stomach content.\textsuperscript{138} IF can only bind B\textsubscript{12} and is reported to be the first filter in the B\textsubscript{12} uptake pathway.\textsuperscript{122} The IF-B\textsubscript{12} complex crosses the intestinal wall in the ileum by receptor mediated endocytosis via the cubam receptor complex.\textsuperscript{125,139} The cubam receptor complex consists of two receptors cubilin and amnionless (see section 1.4.6).\textsuperscript{139} Once in the lysosome, IF is degraded by cathepsin L and B\textsubscript{12} is released.\textsuperscript{140,120} B\textsubscript{12} is transported across the lysosomal membrane with the aid of the lipocalin membrane receptor like protein LMBD1.\textsuperscript{130} Once B\textsubscript{12} is in the cytoplasm, there is debate as to how B\textsubscript{12} leaves the cell, which will be discussed in section 1.4.7.

The crystal structure of IF bound to B\textsubscript{12}, at 2.6-Å resolution was published in 2007, providing a picture of how the 60 kDa protein binds B\textsubscript{12} (see Figure 13).\textsuperscript{129} The structure of IF is very similar to the structure of TCII, which was reported in 2006 and is discussed in section 1.4.7.\textsuperscript{132} IF contains 399 aa residues plus 15\% carbohydrate and is comprised of two domains, a 270 residues α domain and a 110 residue β domain.\textsuperscript{129} As the names suggest, the α domain is an intertwined α\textsubscript{6}/α\textsubscript{6} helical barrel and the β domain is a cluster of β strands (see Figure 13). The B\textsubscript{12} binding site is located between the α and β domains and about 19\% of the B\textsubscript{12} is solvent exposed when bound by IF.\textsuperscript{129} There are five residues from the α domain (His73, Tyr115, Asp153, Asp204, and Gln252) and four residues from the β domain (Ser347, Val352, Phe370, and Leu377) that form direct hydrogen bonds with B\textsubscript{12}.\textsuperscript{129} The B\textsubscript{12} DMB group is maintained in a hydrophobic environment made up of a β-hairpin molded by residues 343-352 of the β domain.\textsuperscript{129}
Figure 13. Ribbon diagram of IF-B₁₂ complex with the B₁₂ shown in ball and stick. Used with permission of National Academy of Science from reference 129.

1.4.7 Cubilin and Amnionless

The transport of B₁₂ across the intestinal wall occurs in the terminal ileum with the aid of IF and the cubam receptor complex.¹²⁵, ¹³⁹, ¹⁴¹ The cubam receptor complex consists of two co-expressed receptors in the apical membrane of the ileum: cubilin and amnionless.¹³⁹ The cubilin receptor is a 460 kDa glycoprotein and is comprised of eight epidermal growth factor (EGF)-like repeats followed by 27 complement CUB (components C1r/C1s, the sea urchin protein Uegf, and
bone morphogenic protein-1) domains.\textsuperscript{125} The CUB domains are the ligand binding region and accounts for 85\% of the receptor.\textsuperscript{125} Cubilin has domains for the binding of a number of ligands including vitamins, lipids and hormones.\textsuperscript{142}

The IF-B\textsubscript{12} complex binds to the CUB 5-8 domains and the interaction is calcium dependent.\textsuperscript{125,143} B\textsubscript{12} and IF are only recognized by the CUB domain as a complex.\textsuperscript{142} A recent crystal structure of the CUB 5-8 domains bound to IF-B\textsubscript{12} identified both α and β domains of IF interacting with the CUB domains (see Figure 14).\textsuperscript{125} Specifically residues 25-299 of the IF α domain interact directly with CUB 6 domain and residues 308-417 of IF β domain interact directly with the CUB 8 domain.\textsuperscript{125} The internalization of CUB bound to IF-B\textsubscript{12} is facilitated by amnionless, a 48 kDa transmembrane protein.\textsuperscript{139} Amnionless contains internalization signals within its cytosolic domain known to induce endocytosis by clathrin-coated buds.\textsuperscript{125}
Figure 14. The structure of CUB 5-8 (left, blue) interacting with IF bound to B$_{12}$ (right, green and pink). The interaction is calcium (red sphere) dependent. The image is reproduced from reference 125 with permission from Nature Publishing Group.

1.4.8 Transcobalamin

TCII is the primary B$_{12}$ transporter in the blood serum and is responsible for the delivery of B$_{12}$ to the cells by facilitating binding of B$_{12}$ to the TCblR/CD320 receptor.\textsuperscript{120} TCII is the only nonglycosylated B$_{12}$ transport protein, known to date and is believed to be secreted from the vascular endothelium.\textsuperscript{144} When and where TCII binds B$_{12}$ is still debated.\textsuperscript{126,145} In the early 2000s, Alpers and Seetharam proposed the transfer from IF to TCII occurs within the enterocyte (see Figure 15a).\textsuperscript{145,146} Just recently, research has demonstrated the ability of B$_{12}$, with the aid of MRP1, to leave the enterocyte free of a transport protein (see Figure 15b).\textsuperscript{126} Further work needs to be done to determine when TCII binds B$_{12}$. What is known for sure is TCII is bound to B$_{12}$ in the blood serum and transports B$_{12}$ to the TCblR/CD320 or megalin receptor (discussed in section 1.4.9).\textsuperscript{137}
Figure 15. There are currently two theories involving the export of B$_{12}$. a) Seetharam et al. present B$_{12}$ bound to TCII in the cell while more recently b) Nexo et al. have established an alternative export system involving MRP1. Images used from references 146 (a) and 120 (b) with permission from Cambridge University Press and Nature Publishing Group, respectively.

TCII was the first transport protein crystallized and the structure revealed a number of important details about points of modification on B$_{12}$ including the presence of hydrogen bond formation between various residues in TCII and B$_{12}$’s phosphate moiety and side chains of the corrin ring. The crystal structure also showed that TCII does not completely encompass B$_{12}$ upon binding, and leaves an exposed section of the vitamin accessible to solvent. The 1.4 nm solvent accessible pocket of B$_{12}$ bound to TCII shows the phosphate and ribose hydroxyl groups
are left protruding into the solvent, but only the 5’-hydroxyl of the ribose group is open to the environment enough that it can easily accommodate the conjugation of large molecules (see Figure 16).  

**Figure 16.** A representation of the TCII-B_{12} complex. The TCII α-domain is shown on the bottom (red) with the β-domain on top (blue). The B_{12} molecule is shown in ball and stick and found sandwiched between the two domains (orange). Reprinted from reference 132 with permission of National Academy of Science.
**1.4.9 Megalin and TCblR/CD320**

The TCII-B₁₂ complex is recognized by two cell-surface TCII receptors: TCblR/CD320 and megalin. The renal filtration and reuptake of B₁₂ occurs in the kidney with the aid of TCII and the receptor megalin.¹⁴²,¹⁴⁷ Megalin is a 600 kDa LDL-receptor family protein.¹⁴⁸ The TCII-B₁₂ complex binds to megalin with high affinity and internalizes the complex, preventing loss of B₁₂.¹⁴⁹

In the bloodstream, TCII transports B₁₂ to the cell, where the TCII-B₁₂ complex is recognized by a newly identified receptor, TCblR/CD320.¹⁵⁰ TCblR/CD320 is a 58 kDa transmembrane protein that belongs to the low density lipoprotein (LDL) receptor family. TCblR/CD320 is present on the cell surface of virtually all tissue and a soluble form of the receptor has recently been identified in the bloodstream.¹⁵¹ The expression of the TCblR/CD320 receptor has been linked to the proliferative and differentiation status of the cell. The levels of TCblR/CD320 expression increase when cells are dividing, which relates to the fact that B₁₂ is essential for DNA synthesis.¹²³ Endocytosis of the TCII-B₁₂ complex results in the degradation of TCII in the lysosome and the release of B₁₂. Once the B₁₂ is released it can be stored in the cell, used as a cofactor or leave the cell.¹²⁰

In contrast to TCII, haptocorrin seems to have a less important role in B₁₂ delivery to cells, although haptocorrin that lacks terminal sialic acid is known to be rapidly cleared by the asialoglycoprotein receptor in the liver.¹³⁸,¹⁵² The role of this receptor is believed to be involved in degradation of HC. A receptor for sialylated HC has not been identified.¹³⁸
1.5 Using the B\textsubscript{12} Pathway for Drug Delivery

The use of B\textsubscript{12} for drug delivery is focused in two main areas of research: targeted delivery of therapeutic and imaging agents and increased oral bioavailability of peptides and proteins. B\textsubscript{12} has been utilized for the delivery of imaging agents and therapeutic drugs for the treatment and diagnosis of cancer, due to the increased need for B\textsubscript{12} in rapidly dividing cells.\textsuperscript{153} The increase in bioavailability of peptides and proteins is based on the potential the B\textsubscript{12} uptake pathway has for overcoming the major barriers to oral delivery of peptides and proteins.\textsuperscript{154} The following section will be a review of the field of B\textsubscript{12} drug delivery.

1.5.1 Targeted delivery

Upon delivery to the blood stream, a B\textsubscript{12} conjugate will typically be acquired by TCII, as previously discussed in section 1.4.5.\textsuperscript{120} In cancer therapy/imaging, the hypothesis for B\textsubscript{12} based targeted delivery has historically been that increased TCblR/CD320 expression (as much as 3- to 26- fold in certain patients)\textsuperscript{155} in a variety of cancers such as testis, breast, ovarian, thyroid, uterine, and brain cancer would provide sufficient selectivity over healthy cells.\textsuperscript{155} The increase in TCblR/CD320 expression in cancer cells has led to the production of both small organic and inorganic B\textsubscript{12} conjugates with various applications in medicinal chemistry, which will be briefly discussed in this section and can be found in Table 1.
Table 1. B\textsubscript{12}-small molecule conjugates.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Imaging agents:</th>
<th>Size</th>
<th>Conjugation Site</th>
<th>Linker (coupling agent)</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^{99m}\text{Tc-DTPA})</td>
<td></td>
<td>452.0</td>
<td>(b)-acid</td>
<td>1,4-diaminobutane (EDAC)</td>
<td>1997</td>
<td>[144]</td>
</tr>
<tr>
<td>(^{111}\text{In-DTPA})</td>
<td></td>
<td>467.7</td>
<td>(b)-acid</td>
<td>1,4-diaminobutane (EDAC)</td>
<td>1997</td>
<td></td>
</tr>
<tr>
<td>Naphthofluorescein</td>
<td></td>
<td>332.3</td>
<td>cobalt atom</td>
<td>3-aminopropyl</td>
<td>2001</td>
<td></td>
</tr>
<tr>
<td>Cy5 dye</td>
<td></td>
<td>792.0</td>
<td>ribose-5'-OH</td>
<td>1,6-diaminohexane (CDI)</td>
<td>2003</td>
<td></td>
</tr>
<tr>
<td>([^{99m}\text{TeO}_4])^\textsuperscript{-}</td>
<td></td>
<td>162.0</td>
<td>cyano ligand</td>
<td>imidazolecarboxylic acid (NA)</td>
<td>2004</td>
<td>[159]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>psalmine acid (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,4-dipicolinic acid (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>serine (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N,N-dimethylglycine (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(^{131}\text{I})</td>
<td></td>
<td>126.9</td>
<td>Pt(II) center</td>
<td>cisplatin (NA)</td>
<td>2007</td>
<td></td>
</tr>
<tr>
<td>([^{99m}\text{Te(CO)}_3(OH)_3])^\textsuperscript{-}</td>
<td></td>
<td>236.0</td>
<td>(b)-acid</td>
<td>propyl-PAMA-OEt (EDAC or TBTU)</td>
<td>2008</td>
<td>[166]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ethyl-PAMA-OEt (EDAC or TBTU)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>pentyl-PAMA-OEt (EDAC or TBTU)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>hexyl-PAMA-OEt (EDAC or TBTU)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fluorescence</td>
<td></td>
<td>475.2</td>
<td>ribose-5'-OH</td>
<td>trans-1,4-diaminocyclohexane (CDT)</td>
<td>2009</td>
<td>[162]</td>
</tr>
<tr>
<td>Rhodamine 6G</td>
<td></td>
<td>479.0</td>
<td>ribose-5'-OH</td>
<td>trans-1,4-diaminocyclohexane (CDT)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ge\textsuperscript{6+}</td>
<td></td>
<td>157.3</td>
<td>ribose-5'-OH</td>
<td>anhydride of DTPA (NA)</td>
<td>2009</td>
<td>[172]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>anhydride of TTHA (NA)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Therapeutics:**

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Imaging agents:</th>
<th>Size</th>
<th>Conjugation Site</th>
<th>Linker (coupling agent)</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Colchicine</td>
<td></td>
<td>359.4</td>
<td>cobalt atom</td>
<td>4-chlorobutyric acid chloride (EDAC/NHS)</td>
<td>2004</td>
<td></td>
</tr>
<tr>
<td>([\text{Re}(\text{OH})(\text{CO})_3])</td>
<td></td>
<td>288.0</td>
<td>cyano ligand</td>
<td>imidazolecarboxylic acid (NA)</td>
<td>2004</td>
<td>[159]</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>2,4-dipicolinic acid (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>serine (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>N,N-dimethylglycine (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cisplatin</td>
<td></td>
<td>300.0</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td>2005</td>
<td>[173]</td>
</tr>
<tr>
<td>Cisplatin-2'-deoxyguanosine</td>
<td></td>
<td>585.3</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cisplatin-methylguanosine</td>
<td></td>
<td>598.2</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>trans-PtCl(_3)(\text{NH}_3)_2</td>
<td></td>
<td>263.7</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td>2008</td>
<td>[174]</td>
</tr>
<tr>
<td>trans-PtCl(_3)(\text{NH}_3)</td>
<td></td>
<td>281.8</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>cis-PtCl(_3)(\text{NH}_3)</td>
<td></td>
<td>281.8</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PtCl(_3)</td>
<td></td>
<td>300.4</td>
<td>cobalt atom</td>
<td>cyano ligand (NA)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\text{VO}_2(\text{OH}/\text{H}))</td>
<td></td>
<td>118.0</td>
<td>cobalt atom</td>
<td>3-hydroxy-2-methyl-1-propyl-1H-pyridin-4-one (NA)</td>
<td>2008</td>
<td>[179]</td>
</tr>
<tr>
<td>(\text{VO}_2)</td>
<td></td>
<td>82.0</td>
<td>cobalt atom</td>
<td>3-hydroxy-2-methyl-1-propyl-1H-pyridin-4-one (NA)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

NA = Not available, CDT = 1,1'-carbonyl di-(1,2,4-triazole), DTPA = diethylentriamine-N,N,N',N''-pentaaacetic acid, DCC = N,N'-dicyclohexylcarbodiimide, TTHA = triethylenetetramine-N,N,N',N''-hexacetic acid, TBTU = O-(Benzotriazol-1-yl)-N,N,N',N''-tetramethyluronium tetrafluoroborate, TSTU = ((N,N',N'''-Tetramethyl-O-(N-succinimidyl) uromum tetrafluoroborate, HPMA = Lysine-modified hydroxypipropyl-methacrylamide
1.5.1.1 The Use of B\textsubscript{12} for Imaging

Early investigators demonstrated that cobalt radionuclides ($^{57}$Co, $^{58}$Co, $^{60}$Co) could be used to label B\textsubscript{12} for imaging.\textsuperscript{144,156} The half-life of the radionuclide, however, required the dosage for humans to be too small for successful external images to be taken and the radioactive compounds were shown to accumulate in the liver, pancreas, and kidneys, leading to organ damage.\textsuperscript{157} In an attempt to overcome these issues, a number of additional radionuclides including $^{111}$In (2.8 days; 0.245 MeV-$\gamma$),\textsuperscript{157} $^{131}$I (8 hrs; 364 keV-$\gamma$, 606 keV-$\beta$)\textsuperscript{158} and $^{99m}$Tc (6.02 hrs; 140 keV-$\gamma$)\textsuperscript{159} have been conjugated to B\textsubscript{12} for labeling purposes. Unfortunately, these experiments still resulted in undesired organ accumulation, limiting the amount of drug delivered to cancerous cells and demonstrating high background uptake in healthy cells.\textsuperscript{157,159,160}

Several fluorescent B\textsubscript{12} imaging agents have also been synthesized.\textsuperscript{161,161-163} The first reported fluorescent B\textsubscript{12} analogues utilized the $\beta$-axial position of the Co (III) ion for conjugation of the fluorophores (Fluorescein, napthofluorescein and Oregon Green), but low Co-C bond dissociation energy (~37 kcal/mol) caused the molecules to suffer from photochemical instability.\textsuperscript{164} A shift in conjugation sites to 5’-hydroxyl group of the ribose moiety\textsuperscript{162} and the addition of a rigid linker provided more stable B\textsubscript{12} fluorescent moiety with an increase in fluorescent lifetime.\textsuperscript{163} An increase in the overall fluorescence quantum yield was observed with the use of a rigid linker over a flexible linker because the rigid linker orients the fluorophore away from the corrin ring of B\textsubscript{12}, minimizing through space intramolecular interactions.\textsuperscript{163} The fluorescent agents are expected to be suitable as tumor markers to aid surgeons during surgery to remove the tumors, but no data supporting this claim has been published to date.\textsuperscript{163} The imaging agents will most likely suffer from nonselective TCII uptake \textit{in vivo}.
In an attempt to overcome background uptake, HC has more recently been targeted since receptors for haptocorrin are also overexpressed in a number of cancer lines.\textsuperscript{165,166} To target HC over TCII, Alberto \textit{et al.} disrupted the interactions between B\textsubscript{12} and TCII by utilizing the b propionamide site after modification to the b monocarboxylic acid.\textsuperscript{167} This is a considerable breakthrough for B\textsubscript{12} based cancer therapeutics since it demonstrates one of the often stated problems of using B\textsubscript{12} systems, namely non-specific cell uptake, can be addressed, at least towards cancer cell lines with \textit{de novo} HC expression.

\subsection*{1.5.1.2 Vitamin B\textsubscript{12} Therapeutics}

Nitrosylcobalamin (NOCbl) is a B\textsubscript{12} derivative with nitric oxide (NO), a molecule known to induce apoptosis, as the $\beta$-axial ligand.\textsuperscript{168} Once in the cell, the NO detaches from B\textsubscript{12} and has the potential to causes direct damage to DNA and inhibits cellular metabolism, leading to necrosis and apoptosis.\textsuperscript{169} In 2002, Bauer \textit{et al.} analyzed 22 human tumor cell lines and two non-tumor cell lines and concluded NOCbl had a greater selectivity for tumor cells and the tumor cells were more sensitive to the NOCbl.\textsuperscript{169} The NOCbl was determined to inhibit tumor growth \textit{in vitro} and \textit{in vivo} by activation of the extrinsic apoptotic pathway.\textsuperscript{168,170}

Colchicine is a cytotoxic agent, which disrupts mitosis by binding to tubulin and inhibiting microtubule polymerization.\textsuperscript{171} The use of colchicine suffers from a lack of selectivity for cancer cell among a field of healthy cells, resulting in undesired side effects.\textsuperscript{171} In 2004, Grissom \textit{et al.} employed B\textsubscript{12} to deliver colchicine specifically to cancerous cells.\textsuperscript{171} Using an acid labile hydrazone linker, colchicine was attached to the $\beta$ axial ligand of B\textsubscript{12}. The linker allowed for a pH-dependent release of colchicine in acidic conditions.\textsuperscript{171} Once taken into the cell \textit{via} the B\textsubscript{12} receptors, colchicine was released to interact with microtubules but showed a ten-fold
decrease in toxicity. In vivo experimentation demonstrating selectivity for cancer cells was not performed.

In 2009, Siega et al. determined it was possible to deliver Gd$^{3+}$ to cancer cells by conjugation to B$_{12}$. By using the metal chelating agent DTPA, a B$_{12}$–Gd$^{3+}$ conjugate was constructed. The Gd$^{3+}$ moiety conjugated to the 5'-hydroxyl group of the ribose moiety of the B$_{12}$ did not affect the binding of the B$_{12}$ transport proteins. Viability tests on human myelogenous leukaemia K562 cells incubated with the conjugate showed a significant decrease in cell viability compared to those incubated with the B$_{12}$ parent compound and/or the Gd$^{3+}$ ion alone. Given that this system utilized conjugation at the 5'-hydroxyl group of the ribose moiety, allowing TCII recognition, it can be postulated that this system, if tested in vivo, will give poor tumor specificity as well.

B$_{12}$ can act as a ligand for cisplatin by formation of a cyanide-bridged species between the β-axial ligand of B$_{12}$ and Pt. B$_{12}$–cisplatin conjugate retain a labile chloride ligand that can be exchanged with ligands such as guanine, allowing the conjugate to behave in a similar way to cisplatin. In 2008, Alberto et al. synthesized a number of prodrugs around the {B$_{12}$–CN–Pt–R} moiety. Using an in vitro adenosylation assay from Salmonella enteric, the group was able to show that Co (III) was reduced to Co (II) and the Pt (II) complex was released. In 2011, the preliminary in vitro cytotoxicity of the {B$_{12}$–CN–Pt–R} conjugate suggested a lower activity (IC$_{50}$ between 8 and 88 µM) than for cisplatin alone. It is hypothesized that the limited B$_{12}$ uptake capacity is most likely to blame. Studies on the effects of the uptake capacity on concentration are underway.
1.5.2 Oral Delivery of Peptides and Proteins

The investigation of B$_{12}$ for the oral delivery of peptides and proteins began in the early 1990s with the work of Russell-Jones.$^{176,177}$ In an attempt to establish a route of oral delivery for peptide and proteins, in general, Russell-Jones and co-workers conjugated B$_{12}$ to a wide range of peptides and proteins between 1994 and 1999.$^{176-178}$ The initial conjugate involved the attachment of luteinizing hormone releasing hormone antagonists (ANTIDE) to the B$_{12}$ carboxylic acid derivative.$^{176}$ Exploration of the spacer units established the ability to attach two biological active molecules. B$_{12}$ was also determined to increase solubility of the protein.$^{155}$

Later the same year, work with granulocyte-colony stimulating factor (G-CSF) and erythropoietin (EPO) established the ability of B$_{12}$ to deliver a biological active protein orally in an IF-dependent fashion *in vitro* and *in vivo*.$^{177}$

A number of issues with the B$_{12}$ uptake pathway were determined with these early studies including the reduction of biological activity of the peptides or proteins, although sometime only slight.$^{176}$ It was also determined the B$_{12}$ uptake pathway had a limited uptake capacity (~1 nM per dose in humans) and so the system was restricted to highly active compounds.$^{92}$ It was also stated that the system had no way of protecting the protein from proteolysis.$^{92}$

1.6 Points of Modification on B$_{12}$

Knowledge of the binding between B$_{12}$ and its various transport proteins is critical if a B$_{12}$ conjugate is to be successfully translated from bench to bedside. In the last five years there has been an explosion of critical structural data related to the B$_{12}$ uptake pathway with the publication of the IF,$^{129}$ TCII$^{132}$ and cubilin$_{5-8}$-IF-B$_{12}$ structures.$^{125}$ Researchers have a better understanding of how B$_{12}$ interacts with its transport proteins, and how these transport proteins...
interact with their receptors. The implications this can have on drug delivery and sites of potential conjugation can then be better detailed, rationalized and hence optimized.

The B12 and the molecule of interest can be: (1) coupled directly together, (2) held apart by ‘spacer’ units to produce distance between the B12 and drug or (3) carriers can be conjugated to B12 with the desired drug contained, unconjugated, within this carrier. Several functional groups are readily available for modification on B12 including propionamides, ethylamides, hydroxyl groups, the cobalt (III) ion and a phosphate moiety. Only a few modifications are capable of maintaining the recognition of all three transport proteins needed to utilize the full B12 uptake pathway effectively in an oral manner.

The size of the drug is important when considering the location for conjugation as well as the transport proteins being utilized since all three transport proteins are not created equally. Conjugation of drugs resulting in the recognition of all three transport proteins has been successful with B12 at three major sites: (1) to the peripheral corrin ring e propionamide; (2) to the 5’-hydroxy of the ribose; and (3) to the cobalt metal (see Figure 17).
**Figure 17.** Conjugation can occur at any of the functional groups (indicated in shaded shapes) without loss of binding by the three uptake proteins.
1.7 Summary

An orally administered peptide/protein medication is important for increasing patient adherence to therapy as well as the physiological importance of mimicking the natural pathway of the peptide/protein.\(^9\) The uptake pathway of B\(_{12}\) makes for an excellent candidate for the oral delivery of molecules because it has the potential to overcome the major problems associated with oral delivery. Before the full potential of the B\(_{12}\) uptake pathway can be realized, however, concerns about the pathway that have been brought to light in recent years, including the limited uptake capacity (~1.5-2.0 μg per dose in humans), need to be addressed.

The work presented here is based on the hypothesis that the B\(_{12}\) uptake pathway can be used to orally deliver active peptides and proteins. To address this hypothesis, three important goals will be set (1) Gaining a better understanding of the previously designed B\(_{12}\)-insulin conjugation system through the reexamination of the reaction conditions, alternative conjugation sites on insulin and modification of the B\(_{12}\) structure (2) Understanding the impact B\(_{12}\) conjugation and the B\(_{12}\) uptake pathway has on peptide/protein function (3) Appreciate the limitation of the B\(_{12}\) uptake pathway and explore peptides suitable for B\(_{12}\) delivery.

1.8 References


(9) Colagiuri, S. *Diabetes Obes. Metab.* 2010, 12, 463-473.


(14) Huang, E. S.; Basu, A.; O'Grady, M.; Capretta, J. C. *Diabetes Care* 2009, 32, 2225-2229.


(31) Steiner, D. F.; Chan, S. J.; Rubenstein, A. H. In *Biosynthesis of Insulin*; Comprehensive Physiology; John Wiley & Sons, Inc.: 2010; .


(105) Russell-Jones, G.; McTavish, K.; McEwan, J. J. Drug Target. 2011, 19, 133-139.


(111) Minot, G. R.; Murphy, W. P. JAMA 1926, 87, 470-476.


Chapter 2 Investigating the B\textsubscript{12} Uptake Pathway for the Oral Delivery of Insulin

2.1 Introduction

The oral delivery of insulin at clinically significant and controllable levels is a major goal in the treatment of DM\textsuperscript{1, 2}. The removal of insulin injections is predicted to increase patient adherence to insulin therapy and increase patient’s quality of life\textsuperscript{2, 3}. Side effects of current insulin therapy include weight gain and hypoglycaemia, which have been correlated with the over-insulization of the periphery due to subcutaneous injections\textsuperscript{4}.

Insulin is naturally secreted from pancreatic \(\beta\) cells directly into the portal vein and is then shuttled to the liver\textsuperscript{4}. During first pass through the liver, up to 70\% of secreted insulin is extracted by insulin receptors located on the liver\textsuperscript{5-7}. The insulin extraction by the liver results in a higher concentration of insulin in the portal vein compared to the systemic circulation\textsuperscript{4}. Subcutaneous injections bypass this first pass through the liver and result in an unnaturally higher concentration of insulin in the systemic circulation than in the portal vein. The oral route of administration is postulated to mimic the natural route of insulin secretion with absorption from the GIT into the portal vein\textsuperscript{4}.

The major problems facing oral delivery of peptides/proteins are hydrolysis/proteolysis in the GIT and an inefficient uptake mechanism for large peptides or proteins from the GIT (see section 1.3 for a more detailed discussion)\textsuperscript{8}. We are interested in the use of the B\textsubscript{12} dietary uptake pathway to address these hurdles to oral delivery.

The B\textsubscript{12} dietary uptake pathway (discussed in detail in sections 1.4.4-1.4.9) utilizes a series of transport proteins and receptor to protect and shuttle the B\textsubscript{12} from the mouth, across the intestines, and to the cell where it is utilized as a cofactor (see Figure 18). For the oral delivery of peptides and proteins, we are primarily interested in the two transport proteins, HC and IF,
and the intestinal IF dependent cubam complex. HC is important for the protection of B$_{12}$ in the mouth/stomach and facilitates the transport of B$_{12}$ to the intestines, where the second transport protein IF transports the B$_{12}$ across the intestinal wall with the aid of the cubam complex (see section 1.4.7).$^9$

**Figure 18.** A basic representation of the B$_{12}$ dietary uptake pathway.

Previously, Petrus *et al.* reported on the development and characterization of an orally active B$_{12}$-insulin conjugate with hypoglycemic properties in the streptozotocin (STZ)-induced diabetic rat model.$^{10}$ The oral administration of the B$_{12}$-insulin conjugate resulted in a 4.7 fold greater decrease in blood glucose levels compared to orally administrated insulin in the STZ rat model. It was determined that the B$_{12}$-insulin’s glucose lowering ability was a result of the use of the B$_{12}$ uptake pathway. This determination was based on experimentation with co-
administration of excess B\textsubscript{12} and the conjugate, which resulted in a loss of glucose lowering ability.\textsuperscript{10}

Results from the initial work with B\textsubscript{12}-insulin,\textsuperscript{10} stimulated further study into the implications of the B\textsubscript{12} conjugation on insulin’s ability to bind to its receptor (CD220) using immune-electron microscopy and molecular dynamics simulation.\textsuperscript{11} What is clear from the results obtained to date is that (i) B\textsubscript{12} can deliver insulin to rat blood serum at levels significantly greater than the same administered concentration of ‘free’ insulin, and (ii) The B\textsubscript{12}-insulin conjugate can still dock to CD220, insulin receptor and thus trigger glucose uptake, even when bound to TCII.

The use of the B\textsubscript{12} pathway to deliver a clinically relevant bolus of insulin, or fast-acting insulin usually given with a meal,\textsuperscript{12} is not possible however, given both the limited uptake capacity (1.5-2.0 μg per dose)\textsuperscript{13} of the pathway and the relatively slow uptake time for serum delivery (~several hours).\textsuperscript{8} The possibility of developing an oral, long acting basal therapy is more feasible, where a low level of long acting insulin is required and an initial uptake delay is not critical.\textsuperscript{12}

The earlier results with insulin\textsuperscript{10,11} have been highly encouraging and we have become focused on investigating and improving \textit{in vivo} residency of B\textsubscript{12}-insulin conjugates. Before such an area can be fully explored however, knowledge of (i) the effects of B\textsubscript{12} conjugation on the two main sites suitable for such on insulin (\textit{vide infra}) and (ii) dose-response of B\textsubscript{12}-insulin conjugates needed to be addressed. These important factors are discussed in this report.

Insulin has two primary sites available for conjugation (the N\textsuperscript{α}-amino group of phenylalanine B1 (PheB1) and the N\textsuperscript{ε}-amino group of lysine B29 (LysB29)), which allow
biological function to be maintained (see Figure 19).\textsuperscript{14} It is not completely understood which position is more suitable for modification.\textsuperscript{15,16} Significant differences between both positions upon B\textsubscript{12} conjugation would be of great importance given the limited B\textsubscript{12} uptake capacity and the need therefore to optimize such a conjugate.

**Figure 19.** The three amine groups available for conjugation on the bovine insulin structure are indicated. Conjugation at the GlyA1 has previously been shown to eliminate biological activity, therefore only the PheB1 and LysB29 are available for conjugation.

The original conjugate utilized the LysB29 position for B\textsubscript{12} attachment based on the understanding that this amino acid is involved in the dimerization of insulin.\textsuperscript{16-18} Insulin is more stable as a hexamer, but the monomer is the only biological active form.\textsuperscript{17} Therefore, the suppression of dimerization and subsequent hexamerization would result in faster absorption of the conjugate.\textsuperscript{17} It has been reported that conjugation at the LysB29 position results in a reduction in biological activity due to the close proximity to the insulin receptor binding site.\textsuperscript{15} Therefore, moving B\textsubscript{12} further away from the binding site could result in an increase in biological activity.
We describe herein the synthesis, characterization, and purification of a new B$_{12}$-insulin conjugate (1), which is attached at insulin PheB1 (Figure 19). The use of PheB1 for attachment of B$_{12}$ is also predicted to eliminate self-association of the insulin molecule because PheB1 is involved in insulin hexamerization.\textsuperscript{15} We hypothesize that this new conjugate may result in an increase in insulin’s biological activity, while maintaining the B$_{12}$ uptake capabilities due to the position of B$_{12}$ attachment in relation to the insulin receptor binding region. The synthetic route of the previously investigated B$_{12}$-insulin (B$_{12}$ bound to insulin LysB29, 2)\textsuperscript{10} was updated to increase yield and mimic as close as possible the synthetic route for the new B$_{12}$-insulin (B$_{12}$ bound to insulin PheB1, 1).

Both 1 and 2 were coupled at the B$_{12}$-ribose hydroxyl group through a carbamate linker. Insulinotropic properties of 1 and 2 were evaluated \textit{in vivo} in STZ induced diabetic rats. Molecular dynamics simulation studies were used to help rationalize the results. We also report for the first time the ability of a B$_{12}$-insulin conjugate to exhibit dose dependence in STZ diabetic rats.

\textbf{2.2 Results and Discussion}

\textbf{2.2.1 Synthesis of B$_{12}$-B$_{1}$Insulin (1) and B$_{12}$-B$_{29}$Insulin (2)}

Reaction of B$_{12}$ with 1, 1’-carbonyl-di-(1, 2, 4-triazole) (CDT) furnished an activated ester at the 5’-hydroxyl group (Figure 20), as previously reported by Russell-Jones \textit{et al.}\textsuperscript{19}
Figure 20. The activation of B\textsubscript{12} occurs at the 5'-hydroxyl group of the ribose with the aid of CDT. B\textsubscript{12}\textsubscript{Act} = CDT activated B\textsubscript{12}.

The synthetic procedures for 1 and 2 were identical unless otherwise noted (see Figure 21). The activated B\textsubscript{12} was added to a slowly stirring solution of bovine insulin in dimethyl sulfoxide (DMSO) with 0.1% triethylamine (TEA). The site directed B\textsubscript{12} conjugation utilizes the established difference in reactivity of the amino groups (LysB29 > GlyA1 >> PheB1) of insulin at pH 9.5.\textsuperscript{16} For 1, the more reactive LysB29 and GlyA1 were initially protected with \textit{t}-butyloxycarbonyl (BOC) to allow for specific conjugation of B\textsubscript{12} at the unprotected PheB1 residue as previously described by Kim \textit{et al.}\textsuperscript{16} 2 was synthesized without need for a protecting group due to the high level of reactivity of LysB29 under the reaction conditions chosen.
Figure 21. The synthetic route of 1 and 2 are outlined.

Both reactions were extracted with the addition of diethyl ether and centrifugation. The products were isolated by analytical reverse-phase high pressure liquid chromatography (RP-HPLC). Purification proved challenging due to the similarity in hydrophobicity of unconjugated insulin and B_{12} conjugated insulin. For 1, the BOC protection was adequate to produce a separation between B_{12} conjugated and free insulin (see Figure 22). At a pressure of 71 bar, HPLC yielded a dark red fraction at rt = 1.4 min, pink fractions at 7.4 and 8.5 and clear fractions at rt = 7.6, 8.1, 9.0 and 9.7. The pink fraction at 7.4 is B_{12} attached at the GlyA1 position with BOC at the LysB29. However, baseline separation was not achieved and the sample contained
BOC-Insulin, which eluted at 7.6. The pink fraction at 8.4 is B_{12} attached at PheB1-BOC^{2}-Insulin and the clear fraction at 9.0 is BOC^{3}-Insulin.

**Figure 22.** The RP-HPLC spectra of a BOC protected insulin plus B_{12} reaction monitored at 254 nm.

For 2, addition of an excess of N-(9-Fluorenylmethoxycarbonyloxy) succinimide (FMOC-OSu) was used to aid in purification. The FMOC moiety increases the retention time of the unconjugated (hence, unprotected) insulin allowing for separation (see Figure 23).
Figure 23. The RP-HPLC spectra of an FMOC protected insulin plus B\textsubscript{12} reaction monitored at 360 nm (Dash, indicating B\textsubscript{12} is present) and 254 nm (Solid, indicating insulin is present).

Following HPLC purification, the protecting groups (BOC or FMOC) were removed with trifluoroacetic acid (TFA) (1) or piperidine (2) and dialyzed against 5 L of H\textsubscript{2}O overnight in 7000 MWCO dialysis tubing.

2.2.2 MALDI-ToF Mass Spectrometry

Matrix assisted laser desorption ionization time of flight (MALDI-ToF) mass spectrometry (MS) detects ions. Therefore, the neutrally charged B\textsubscript{12} must fragment in order to produce detectable species. Experimentation with cyanocobalamin demonstrated fragmentation of B\textsubscript{12} occurs with the initial loss of the cyanide group. This is believed to result from the hemolytic cleavage of the Co-CN bond and subsequent protonation of the negatively-charged phosphate producing a [B\textsubscript{12}-CN+H]\textsuperscript{+} species with a 1329 m/z.\textsuperscript{20}
The exact site of B<sub>12</sub> conjugation was verified by MALDI-ToF MS analysis of reduced and digested fragments of 1 and 2. The conjugates were first treated with dithiothreitol (DTT) to reduce the interchain disulfide bonds linking the A- and B-chain. MS analysis of the reduced conjugates displayed m/z values of 2339 and 4755 corresponding to unmodified A-chain and \([B_{12}-CN+H]^+\) bound to the B-chain, respectively (see Figure 24).

![MalDI-TOF MS spectrum](image)

**Figure 24.** A representative MALDI-Tof MS spectrum of DTT reduced 1 or 2 shows mass at 4754 m/z representing \([B_{12}-CN+H]^+\) attached to the B chain of insulin. Absent is B<sub>12</sub> attached to the A strand, 3694 m/z.

To verify the specific B chain site, reduced conjugates were treated with trypsin, which cleaves peptides at the carbonyl side of arginine or lysine, except when followed by proline.\textsuperscript{21}
Reduced and digested 1 displayed a m/z of 3844 corresponding to B1-B22 plus [B_{12}-CN+H]^+ indicating conjugation at the B1 site (see Figure 25 b). While, reduced and digested 2 displayed a m/z of 2286 corresponding to B23-B30 plus [B_{12}-CN+H]^+, indicating attachment at the B29 site (see Figure 25 a).

Figure 25. (a) MALDI-Tof MS spectrum of reduced and tryptically digested 1 produces a 2286 m/z fragment consistent with [B_{12}-CN+H]^+ attached to the LysB29 position. (b) Reduced and
tryptically digested 2 produces a 3844 m/z fragment consistent with \([B_{12}-CN+H]^+\) at the N-terminus of the B strand, as expected with PheB1 conjugation.

### 2.2.3 Electronic Absorption Spectrometry

\(B_{12}\) (CNCbl) has a distinguishable electronic absorption spectrum, which provides experimental data on the nature of the axial ligands, the cobalt oxidation state and the concentration of the \(B_{12}\) solution. The absorption spectrum, seen in Figure 26, is characterized by two prominent bands, the \(\alpha/\beta\) in the visible region and the \(\gamma\) band in the UV region. All the prominent bands above 300 nm have been attributed to \(\pi\) to \(\pi^*\) transition of the corrin ring.\(^\text{22}\) Electronic absorption spectroscopy is a useful technique to determine the initial effects of synthesis on the vitamin \(B_{12}\) structure and the overall yield of the reaction.

![Figure 26](image-url)

**Figure 26.** The “typical” absorption spectra of CNCbl. The predominate bands highlighted (\(\alpha, \beta, \text{and } \gamma\)) are due to spin allowed intraligand charge transfers of the corrin ring.
Representative absorption spectrum for the B\textsubscript{12} insulin conjugate is shown in Figure 27. From this spectrum we can determine the B\textsubscript{12} is a Co (III) “base on” species and using Beer-Lambert law the concentration of the solution and subsequent yield of the reaction can be determined using an extinction coefficient of 27500 M\textsuperscript{-1} cm\textsuperscript{-1} for the band at 361 nm.

**Figure 27.** The electronic absorption spectra of B\textsubscript{12}-insulin conjugates contain the characteristic B\textsubscript{12} peaks and insulin peak.

### 2.2.4 In vivo experimentation

The dose dependent effect of 1 on blood glucose was assessed in STZ-rats and the results are shown in Figure 28. 1 was administered at concentrations of $10^{-9}$ M (n=3), $10^{-7}$ M (n=3) and $10^{-5}$ M (n=2). Area under the curve (AUC) was calculated for each dose as $1387 \pm 297$, $795 \pm$
202 and 125 ± 15 mmol/l 240 min below pre-gavage blood glucose concentration for 10⁻⁹ M, 10⁻⁷ M and 10⁻⁵ M doses, respectively. AUC of raw data identified a significant difference between 10⁻⁹ M and 10⁻⁵ M conditions, but there were no significant differences between the other concentrations. The data presented in Figure 28 demonstrates the ability of the conjugate to produce a dose dependent reduction in blood glucose. An increase in concentration from 10⁻⁹ M to 10⁻⁵ M resulted in an increase of 10-20% drop in blood glucose.

![Graph](image)

**Figure 28.** Relative (% change from pre-gavage) blood glucose concentration following administration of I at concentrations of 10⁻⁹ M, 10⁻⁷ M and 10⁻⁵ M (n = 3).
Comparison of the conjugates was assessed by measuring blood glucose concentration following oral delivery of 1 and 2 (see Figure 29). AUC transformation showed a significant difference (p=0.03) between 1 and free insulin control, consistent with that observed in our earlier report.\textsuperscript{12} Of interest here is that no significant difference in glucose drop (p=0.53) was observed between 1 and 2. Both produced similar and greater drops over free insulin of the same orally administered concentration. This data differs from previous literature\textsuperscript{15,18} that suggests conjugation position alters the biological activity of insulin and was not what we expected. We believe our findings may be explained by components of the B\textsubscript{12} uptake pathway, making this result not in conflict with these earlier findings but rather suggesting a unique result tied to the use of B\textsubscript{12} in insulin conjugation.
Figure 29. Relative (% pre-gavage) blood glucose concentration following administration of free insulin (n=4), 1 (n=6) and 2 (n=6).

To help explain the biological similarities between conjugates, molecular dynamic (MD) simulations of TCII interacting with 1 was performed by collaborator Dr. Damian Allis of Syracuse University (see Figure 30). Previous MD simulations, performed by Dr. Damian Allis, of 2 interacting with TCII\(^1, 23\) indicated that the lysine side chain provided an adequate tether length to place the unstructured insulin B-chain region (residues B20-B30) beyond the steric congestion of the B\(_{12}\) bound TCII structure. Furthermore, the B20-B30 tail provides considerable length for placing the most structured region of the insulin molecule (the complete A-chain and B7-B19 B-chain region) far from the \(\alpha\) and \(\beta\) domains of TCII, thereby not
impacting the binding geometry of the B_{12} in its internal TCII pocket or the interaction of insulin with the insulin receptor.

Formation of 1 places the B_{12} several residues from the structured region of insulin (the complete A-chain and B7-B19 B-chain region), but the effective separation of the structured insulin region from the TCII binding area is not obvious from the components themselves given (i) the inability to predict whether an accommodation for the B_{12} binding within its pocket will occur by insulin unfolding of the B1-B7 region and (ii) the inability to predict if the phenylalanine side chain itself would serve as a source of steric congestion to disrupt B_{12} binding in its pocket (in 2, the lysine side chain is itself the tether region between B_{12} and insulin). Insights into the nature of the B_{12}-insulin/TCII interaction upon PheB1 conjugation come from a 15 ns simulation of 1 interacting with the TCII complex. A representative of the geometry of 1 during these simulations is shown and labeled in Figure 30 and 31.
Figure 30. MD simulation of TCII bound to 1. TCII structure: $\alpha$-domain (red); $\beta$-domain (dark blue); unstructured linkage between domains (green). Structure of 1: B$_{12}$ (yellow ball and stick) with Co (III) (blue sphere); Insulin B-chain (yellow ribbon); A-chain (light blue ribbon). The structure was produced using the GROMACS software package with GROMOS96 (53a6) united-atom force field by collaborator Dr. Damien Allis at Syracuse University.$^{24}$

From the MD simulation we are able to see that the PheB1-CysB7 (B1-B7) region of the B-chain is unfolded from its unmodified, native $\alpha$ helix geometry (see Figure 31). The
disruption of α helical structure extends slightly past the disulfide linkage (CysB7). The result of this unfolding is that the B1-B7 residues become the long tether that separates the insulin core from TCII upon binding of the B_{12}-insulin conjugate. This tether originates from an otherwise stable α helix and is slightly longer than half of the length of the B20-B30 tether in the previous study. The retention of the B_{12} within its binding pocket with no observable deformation of this pocket indicates that the B1-B7 tether is long enough to separate the two larger fragments. It is this uncoiling of the B1-B7 region, induced upon binding by B_{12} uptake proteins that possibly explains why both sites on insulin ultimately equate in function and why such differences were not observed before.
Figure 31. A representative structure of 1 bound within TCII. The structure was produced from a molecular dynamic simulation performed using the same software as previously described in Figure 30 by collaborator Dr. Damien Allis at Syracuse University.24

2.3 Conclusion

We report herein on the synthesis, purification and characterization of a new B₁₂ insulin conjugate attached at the insulin PheB₁ position. The hypoglycemic properties resulting from oral administration (gavage) of such a conjugate in STZ diabetic rats was similar to that noted in a B₁₂-insulin conjugate covalently linked at the insulin LysB₂⁹ residue,¹⁰ demonstrating the
availability of both positions of insulin for, at least, B_{12} attachment. A possible rationale for this result is put forward from MD simulations, which suggest there is forced un-coiling of the insulin molecule at either the N- or C-terminal B strand insulin regions, depending on B_{12} conjugation, upon B_{12} binding protein interaction with 1 or 2. We also conclude that there is a dose dependent response that can be observed for B_{12}-insulin conjugates with doses of conjugate greater than 10^{-9} M necessary to observe a drop in glucose.

In our earlier report on the oral activity of 2, it was demonstrated that the blood glucose reduction could be blocked by the addition of a large excess of free B_{12}.^{10} However, the administered amount of conjugate (13.56 μg B_{12}) was in great excess of the expected B_{12} uptake capacity of the rat (1.5-2.0 μg per dose),^{13} signifying an alternative or additional mechanism for the response may be possible. Given the presence of an insulin receptor in the intestine,^{25} it may be that there is an alternate receptor at play in the uptake of B_{12}-insulin. While the intestinal insulin receptor was not the intended target for our conjugates it is possible that the conjugation of B_{12} has resulted in predominantly monomeric insulin (necessary for CD220 binding) capable of interacting with the intestinal insulin receptor.

It is also possible that B_{12} conjugation is greatly improving insulin stability, presumably as IF-B_{12}-insulin formed in situ. This hypothesis is more consistent with the observed glucose drop along with observed dose-dependent response. The blocking of uptake by large excess of B_{12} would not then be a consequence of cubam binding competition but rather prevention of IF binding and any associated protection. It is also possible that the excess B_{12} altered the oligerimization of the conjugate (‘modifying the formulation’) disrupting any B_{12} uptake and/or insulin receptor binding in the intestine. If such a route through the insulin receptor is made
available by B$_{12}$ conjugation it would be a paradigm shift in our viewing of the exploiting of B$_{12}$ for oral protein delivery.

2.4 References


(13) A Report of the Standing Committee on the Scientific Evaluation of Dietary Reference Intakes and its Panel on Folate, Other B Vitamins, and Choline and Subcommittee on Upper Reference Levels of Nutrients; Food and Nutrition Board; Institute of Medicine **1998**.


Chapter 3 Exploring the effect of B₁₂ on glucagon-like peptide-1 biological activity

3.1 Introduction

DM is characterized by a deficiency in pancreatic β cell function and mass.¹ In T1D, an autoimmune disease has completely destroyed β cells and the body is no longer able to produce the hormone insulin.² T1D will not be discussed further in this chapter. In T2D, the pancreatic β cells are not producing enough insulin to meet the body’s insulin demand and as the disease progresses, the loss of β cell mass increases the insulin deficiency. A primary goal of research in T2D diabetes is increasing the amount of functional insulin secreting β cells and slowing or stopping the progression of the disease.¹

GLP-1 (7-36) amide is a 30 amino acid amidated peptide, often referred to as GLP-1, that has gained a lot of attention in recent years because of its physiological role in the growth of β cells and suppression of β cell apoptosis.³ GLP-1 also has a wide variety of physiological roles centered on the postprandial homeostasis of glucose, including the secretion and promotion of insulin and suppression of glucagon production.⁴⁻⁵ GLP-1’s insulinotropic functions are glucose dependent, making it an ideal treatment for hyperglycemia since hypoglycemia is unlikely to occur.⁴⁻⁶ Furthermore, GLP-1 has been associated with appetite suppression,⁷ which is an important observation considering T2D is highly associated with obesity.⁸

There are two biologically active forms of GLP-1: GLP-1 (7-36) amide and GLP-1 (7-37), which refers to the amino acid residues making up the structure and whether the C-terminal is amidated (see section 1.2.2.2). The C-terminal amidation is the prominent form of GLP-1 in circulation, with the amidation predicted to increase stability.⁹⁻¹⁰
The therapeutic potential of GLP-1 receptor (GLP-1R) targeting for the treatment of T2D has been validated through extensive basic, clinical and translational research. The therapeutic use of GLP-1 is limited by a short half-life (1-2 min), due primarily to degradation by the enzyme dipeptidyl peptidase IV (DPP-IV). There is however currently a GLP-1 based therapeutic available for the treatment of diabetes known as liraglutide. Liraglutide is an injectable medication that exhibits an increased resistance to DPP-4 degradation due to the addition of a fatty acid moiety at Lys26. Although effective for the treatment of T2D, the injectable liraglutide does not have the ability to mimic the natural secretion of GLP-1, which may have physiological benefits yet to be observed.

GLP-1 is secreted from intestinal L cells in response to orally ingested fats and carbohydrates and enters the systemic circulation via intestinal capillaries draining into the hepatic portal vein. GLP-1 is rapidly degraded by DPP-IV and only 10-15% of secreted GLP-1 makes it intact into the systemic circulation. GLP-1 is therefore believed to be most concentrated in the intestinal extracellular space and hepatic portal vein. While it has been established that the physiological effects of GLP-1 are a result of interaction with the GLP-1R, it is unclear the exact mechanism and population of GLP-1Rs responsible for the various physiological effects. GLP-1R is expressed on the vagal afferent fibers of gastrointestinal and hepatoportal origin, pancreatic β cells and within CNS neurons.

As the potential of GLP-1 therapeutics is realized, a drug delivery system that can mimic the physiologic route of endogenous secretion is essential in fully understanding the mechanism of the GLP-1. It has been reported that GLP-1 acts on the afferent receptors near the site of secretion. Since native GLP-1 is secreted by the intestinal L cells, an orally delivered GLP-1 would have the potential to mimic the full physiological effects of native GLP-1.
As a peptide, GLP-1 cannot be delivered orally due to degradation and proteolysis in the gastrointestinal tract. The B12 uptake pathway, previously discussed in chapter 1 (section 1.4.4), has the potential to overcome these barriers and orally deliver GLP-1. The full potential of the B12 uptake pathway as a delivery system has yet to be fully realized. A number of limitations to this delivery strategy have been identified in recent years, including a limited uptake capacity and slow delivery rate, which has discouraged research into this area. The critical component to the success of the B12 delivery approach is the use of a peptide that has a therapeutic window within the B12 uptake capacity.

Using the B12 pathway for the delivery of peptides has a number of major advantages over other delivery systems including oral delivery, no known toxicity and the ability to cross the blood-brain barrier. The limited uptake capacity of the B12 uptake pathway refers to the transfer of B12 from the ileum to the systemic circulation through IF-mediated endocytosis via the cubilin receptor (see section 1.4.7 for discussion on the cubilin). It is important to note that cubilin recycles every few hours, and as such, multiple dosing (e.g., morning and evening) may be adopted to increase the quantity of peptide absorbed throughout the day. This recycling may also be important when considering dietary B12 competition.

We have previously demonstrated the ability to use B12 to orally deliver clinically relevant amounts of the appetite suppressant peptide tyrosine tyrosine (PYY). To adapt this uptake pathway for the delivery of GLP-1, recognition and affinity for the various binding B12 transport proteins must be maintained and the biological potency of the GLP-1 must not be diminished. We are now exploring in depth the effect of B12 conjugation on the biological potency of GLP-1 in an attempt to demonstrate the true potential of the uptake pathway for the delivery potent peptides and proteins.
3.2 Results and Discussion

3.2.1 Synthesis and Purification of B12-K34R-GLP-1 (7-36) amide (3)

GLP-1 (7-36) amide has three amine groups (N-terminal, Lys26 and Lys34) available for conjugation. For our lead conjugate, the peptide was purchased with modifications (amino acid substitution K34R, and FMOC protection of the N-terminal amine (FMOC-K34R-GLP-1)) to aid in selective conjugation at the Lys26 (Figure 32). No significant reduction in biological activity has been shown for conjugation at the Lys26.\textsuperscript{23, 24} Commercially available Liraglutide has the same amino acid substitution\textsuperscript{24, 25} and the FMOC protection was removed after synthesis.

![GLP-1 (7-36) amide](image)

**Figure 32.** GLP-1 (7-36) amide was purchased with N-terminal FMOC protection and K34R amino acid substitution (34) which will be referred to as FMOC-K34R-GLP-1 or K34R-GLP-1 when the FMOC is removed. These modifications allow for site directed B12 conjugation at the Lys26 (26).

Modifications at the 5’-hydroxyl group of B12 have been repeatedly shown to maintain binding with the transport proteins vital for oral B12 uptake.\textsuperscript{22, 26, 27} Reaction of B12 with 1,1’-carbonyl-di-(1,2,4-triazole) in DMSO at room temperature for 5 minutes furnished an activated ester at the 5’-hydroxyl group. An excess of activated B12 was then added to a slowly stirring
solution of FMOC-K34R-GLP-1 in DMSO with 0.1% TEA. The reaction proceeded for 3 hrs before an excess of FMOC-OSu was added to help with purification, as described below.

The reaction was extracted by diethyl ether addition and centrifugation. The product was isolated on a C\textsubscript{18} analytical column using reverse phase HPLC. Unreacted B\textsubscript{12} was easily separated from FMOC-K34R-GLP-1, however purification of unconjugated FMOC-K34R-GLP-1 and B\textsubscript{12} conjugated FMOC-K34R-GLP-1 proved challenging due to the similarity in hydrophobicity. The addition of excess FMOC-OSu was used to aid in purification, as previously reported for a B\textsubscript{12} insulin conjugate.\textsuperscript{27} The FMOC moiety increases the retention time of the unconjugated (hence, unprotected) FMOC-K34R-GLP-1, allowing for separation (see Figure 33). Following HPLC purification, the protecting group was removed with piperidine and dialyzed against 5 L of H\textsubscript{2}O overnight in 3500 MWCO dialysis tubing. A 10\% yield of 3 was calculated using electronic EAS with an extinction coefficient of 27500 M\textsuperscript{-1} cm\textsuperscript{-1} at 361 nm, corresponding to the B\textsubscript{12} \textgamma band, as described in section 2.2.3.

![RP-HPLC Chromatogram](image)

**Figure 33.** RP-HPLC chromatogram showing separation of unconjugated B\textsubscript{12} and K34R-GLP-1 from B\textsubscript{12} conjugated K34R-GLP-1 with detection at 254 nm (solid line) and 360 nm (dash line).
Absence of B\textsubscript{12} absorption at 360 nm indicated the peak at retention time (Tr) \sim 24 mins is unconjugated K34R-GLP-1 now with two FMOC units, while the peak at Tr \sim 20 mins is B\textsubscript{12} conjugated K34R-GLP-1 with one FMOC.

3.2.2 MALDI-ToF Mass Spectrometry

B\textsubscript{12}-FMOC-K34R-GLP-1 was identified as the peak at a Tr of 20 by MALDI-Tof MS. The peak at 4682 m/z represents a 1:1 [B\textsubscript{12}-CN+H]+ K34R-GLP-1 conjugate (see Figure 34).
Figure 34. MALDI-ToF MS spectrum of 3 after FMOC removal showing the presence of a prominent peak at 4682 m/z representing a 1:1 B₁₂-K₃₄R-GLP (7-36) amide.

3.2.3 In vitro Experimentation

GLP-1 mediates its insulinotropic effects by binding to specific GLP-1 receptors on the pancreatic β cells and stimulating cAMP formation, calcium induced calcium release (CICR) and insulin secretion (see Figure 35). The effect of B₁₂ conjugation on GLP-1 physiological function (cAMP production, CICR and insulin secretion) was studied in HEK-293 cells stably expressing the human GLP-1 receptor (HEK-GLP-1R) and isolated human pancreatic islets.
Figure 35. A basic depiction of the intercellular mechanisms of the GLP-1R, which were monitored for in vitro assays described in section 3.2.5. RYR: ryanodine receptor; ER: endoplasmic reticulum; CRE: cAMP response element. RIP1-CRE-Luc: rat insulin 1 promoter-cAMP response element-luciferase reporter.

To determine the ability of the conjugate to perform as a cAMP elevating agent, HEK-GLP-1R cells were transfected with a reporter construct containing the coding sequence for the luciferase (Luc) enzyme and a promoter incorporating a cAMP response element originally
identified within the 5’ promoter region of the rat insulin 1 gene (RIP1-CRE). This cAMP-responsive Luc reporter (RIP1-CRE-Luc) can be used to determine the effect of agonist for G protein-coupled receptors positively linked to cAMP production, such as the GLP-1 receptor.

The level of Luc activity was determined in whole-cell lysates by Luc-catalyzed oxidation of luciferin, which generated photons detected by a photomultiplier tube. After exposure to the drug for 4 hrs, Luc activity was determined and shown to be concentration dependent with an EC\textsubscript{50} of 4.4 nM (see Figure 36). These findings were comparable to that found for K34R-GLP-1 with an EC\textsubscript{50} of 4.1 nM (see Figure 36). B\textsubscript{12} was not shown to produce any response in the luciferase assay. The EC\textsubscript{50} values suggest the B\textsubscript{12} has little negative effect on the potency of the GLP-1, an important and highly positive result.

Figure 36. The fold increase of luciferase activity was determined for 3 and K34R-GLP-1 after HEK-GLP-1 cells were transfected with a RIP1-CRE-Luc reporter and were exposed to 3 for 4
hrs. 3 produced a concentration dependent response with an EC$_{50}$ of 4.4 nM, while K34R-GLP-1 had an EC$_{50}$ of 4.1 nM.

The stimulatory effects of GLP-1 on pancreatic islet insulin secretion have been attributed to the capability to increase levels of cytosolic [Ca$^{2+}$] in β cells. GLP-1 produces a rise in cytosolic [Ca$^{2+}$] by promoting CICR from intracellular storages. The ability of 3 to facilitate an increase in cytosolic [Ca$^{2+}$] was assessed in HEK-GLP-1R cells that express endogenous P2Y purinergic receptors using a fura-2 assay. Fura-2 is a widely used ratiometric fluorescent dye which is used to determine intercellular levels of calcium. Fura-2 determinations of [Ca$^{2+}$] were obtained in a 2-step injection protocol were 3 (A) or native GLP-1 (B) was administered to each well after baseline recording for 100 s. After a further 200 s the P2Y receptor agonist adenosine diphosphate (ADP) (10 nM) was added and the ratio of Fura 2 was recorded. Both forms of GLP-1 potentiated the action of ADP to mobilize intracellular [Ca$^{2+}$] and native GLP-1 was slightly more potent than 3 (see Figure 37).
Figure 37. Fura-2 determinations of intercellular [Ca\textsuperscript{2+}] were obtained in HEK-GLP-1R cells that express endogenous P2Y purinergic receptors. Note that both forms of GLP-1 potentiated the action of ADP to mobilize intracellular [Ca\textsuperscript{2+}] and that native GLP-1 was slightly more potent than 3. Data are averaged from 12 wells for each data point and similar results were obtained in four independent plates with three different batches of 3.
Since its formation in 2001, the national Islet Cell Resource (ICR) Center Consortium has provided isolated human pancreatic islets (see Figure 38) for clinical and basic research while also working to develop methods to perfect the isolation of islets for human transplantations.\textsuperscript{30} Human pancreatic islets are a critical tool for diabetes researchers. Human islets have provided invaluable information about the cellular biology of the pancreas and the pathology of diabetes. Furthermore, the uniqueness of the human pancreatic biology is an invaluable tool for the development of therapeutics for the treatment of diabetes.\textsuperscript{30}

\textbf{Figure 38.} Pure human pancreatic islets isolated from digested pancreatic tissues, as seen under a microscope. Image reprinted from reference 31 with permission of Baishideng Publishing Group Co.

To evaluate potential insulin secretagogue properties of B\textsubscript{12}-GLP-1, human islet insulin secretion was monitored under standard conditions of static incubation in which human islets were equilibrated in Krebs-ringer buffer (KRB) containing 2.8 mM glucose, followed by a subsequent elevation of the glucose concentration to 16.7 mM with or without the test substance.
The rise in glucose concentration will cause glucose stimulated insulin secretion (GSIS) in healthy insulin producing islets and the addition of a GLP-1 based conjugate should increase the GSIS. The initial GSIS was used to determine the viability of the human islets being tested. Human islets used for basic research are not healthy enough for human transplantations, which means they are not always viable for research purposes either.

The average of three batches of isolated human islets produced a GSIS of 2.2-fold insulin secretion (see Figure 39). GSIS was potentiated an additional 45% by 15 nM 3 (total 3.2 fold stimulation). This effect of 3 was similar to that which was measured when islets were instead treated with 15 nM GLP-1 (total 3.3 fold stimulation).

**Figure 39.** Static incubation assays using human pancreatic islets. Islets were exposed to KRB containing 2.8 mM glucose for 30 min and were then exposed to KRB containing 16.7 mM
glucose with GLP-1 or 3 or without test substance. GSIS resulted in a 1.4 fold insulin secretion while a 3.2 fold increase was observed for 3, which is comparable to the GLP-1 (3.3 fold increase).

3.3. Conclusion

We have reported on the synthesis, purification and in vitro characterization of a novel B₁₂-K34R-GLP-1 conjugate. Experimentation with the lead B₁₂-GLP-1 conjugate in vitro demonstrated B₁₂ attachment has a minimal effect on the insulinotropic nature of K34R-GLP-1 (EC₅₀ 4.1 nM vs. 4.4 nM). We analyzed various steps in the physiological pathway of GLP-1 stimulated insulin secretion to come to this conclusion. Initially we used a luciferase based assay to indirectly determine the ability of 3 to produce an increase in intercellular cAMP levels. The ability of B₁₂-GLP-1 to stimulate CICR was than monitored by Fura-2 assay. Finally we used transplanted human islets to confirm the ability of B₁₂-GLP-1 to stimulate insulin secretion. This is the first reported study of how B₁₂ conjugation affects a peptide’s in vitro activity. The data collected for the B₁₂-GLP-1 conjugate in vitro are very encouraging. The next step is in vivo experimentation to determine the biological activity of the B₁₂-GLP-1 conjugate, which is scheduled to occur in October 2012 in the lab of collaborator Dr. Christian Roth in Seattle Children’s Hospital, WA, US.

3.4 References


Chapter 4 Synthesis of a B\textsubscript{12} Derivative Capable of Amide Bond Formation

4.1 Introduction

The increasing use of B\textsubscript{12} for the delivery of therapeutics has encouraged the re-examination of potential conjugate sites of B\textsubscript{12}.\textsuperscript{1-4} Several positions are available on B\textsubscript{12} for conjugation, including the 5’-hydroxyl group of the ribose, the cobalt (III) ion, the phosphate moiety and propionamides off the corrin ring (if modified to a carboxylic acid first) (see Figure 40).\textsuperscript{5-7} Numerous groups have explored modifications of B\textsubscript{12} that allow for more facile, higher yielding or stable conjugate formation with continued recognition by B\textsubscript{12} dietary transport proteins.\textsuperscript{1,3,8-23} All such modifications to date suffer from issues including reduced or complete loss of interaction with one or more of the B\textsubscript{12} binding proteins, stability issues that result from the production of ester or carbamate bonds, low yielding conjugation reaction due to the nature of the functional group, or require problematic purification of the modified B\textsubscript{12} for subsequent conjugation.\textsuperscript{1,3,8-23} These limitations have been well investigated\textsuperscript{24-26} and reviews of such sites has been reported,\textsuperscript{27-30} however, in brief for purposes of perspective for the work described herein:

Conjugation resulting in the recognition of all three transport proteins has been most successful with B\textsubscript{12} at two major sites: 1) the 5’-hydroxy group of the ribose group and 2) the peripheral corrin ring e propionamide (see Figure 40). The side chain e carboxylic acid derivative, produced from hydrolysis of the propionamide, allows for a variety of modifications,\textsuperscript{31-33} but the synthesis of the derivative is low yielding and requires a laborious purification, the result, in part, of the formation of b and d carboxylic acid isomers in addition to the desired e carboxylic acid derivative.\textsuperscript{34-38} The 5’-hydroxyl
group of the ribose is a versatile site according to structure-activity relationships, however conjugates produced have typically been synthesized, as mentioned above, through carbamate or ester bonds.

**Figure 40.** Potential sites available for carboxylic acid formation that also maintain, at least in part, B$_{12}$ uptake protein recognition and binding are indicated in red. The atom numbering scheme used for future NMR discussion (see section 4.2.2) is also indicated.
We report the selective oxidation of the 5'-hydroxyl group of B$_{12}$ to the corresponding carboxylic acid using a hypervalent iodine reagent, 2-iodoxybenzoic acid (IBX). IBX (see Figure 41) has never been utilized for the oxidation of such a complex molecule as B$_{12}$. The work reported here expands the range of successful use of the environmentally friendly oxidizing agent.$^{39}$

![IBX structure](image)

**Figure 41.** The structure of 2-iodoxybenzoic acid (IBX).

4.2 Results and Discussion

4.2.1 Synthesis of B$_{12}$ 5'-carboxylic acid (4)

The B$_{12}$ 5'-carboxylic acid (4) was synthesized utilizing a modified method of Giannis *et al.*, which incorporates the addition of an O-nucleophile to aid IBX in the transformation of a primary alcohol to a carboxylic acid (see Figure 42).$^{40,41}$
Figure 42. Purposed mechanism for the formation of 4 utilizing IBX and 2-hydroxypyridine (HYP).
The use of the O-nucleophile N-hydroxysuccinimide (NHS) was initially attempted, but resulted in an inability to isolate the carboxylic acid, likely due to the formation of the activated NHS ester that has been reported as a dominating product of this reaction.\textsuperscript{40, 41} 2-hydroxypyridine (HYP) was used as the O-nucleophile. To optimize the reaction conditions, a series of experiments were performed where co-oxidant, temperature and time were varied (see Table 2).

**Table 2.** Optimization of the Reaction Conditions for 4.

<table>
<thead>
<tr>
<th>Entry</th>
<th>Co-Oxidant</th>
<th>Temp (°C)</th>
<th>Time (hr)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>50</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>HYP</td>
<td>50</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>HYP</td>
<td>50</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>HYP</td>
<td>50</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>HYP</td>
<td>50</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>HYP</td>
<td>60</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>

The best result was obtained in the presence of 2.6 equivalents of IBX and 5 equivalents of HYP in DMSO at 60 °C. These conditions resulted in the oxidation of B\textsubscript{12} to the 4 in two hours with a 30% yield. It was determined that a 30% yield was optimal for the IBX reaction, in this case, as increase in temperature, IBX/HYP ratios or time resulted in an increase in byproducts and/or decomposition of the B\textsubscript{12} (especially at temperatures in excess of 60 °C), and which also co-eluted with the desired carboxylic derivative, complicating purification. The 30% yield is ~ three times that previously
reported for the preparation of the e-carboxylic acid derivative and with far easier separation \((\text{vide infra})\). \(^{34-38}\)

The crude reaction was precipitated from DMSO by diethyl ether addition and dissolved in \(\text{H}_2\text{O}\) prior to purification by ion exchange chromatography using a SAX column. After a 5 minute hold at 100% water, a gradient from 100% water to 35% MeCN:phosphate buffer pH 2 (50:50 v/v) over 5 minutes was necessary for the separation and elution of 4 (see Figure 43). Slight increases in pH above 2 resulted in a loss of pure product isolation. The HPLC purification allowed for the recovery and reuse of \(\text{B}_{12}\) starting material, albeit with the oxidized 5’-aldehyde also present (see Figure 43). 4 was identified as the peak at retention time \((T_r)\) 13.5 mins. Purity of 4 was at least 95% as assayed by RP-HPLC.

**Figure 43.** The RP-HPLC spectra of 4 \((T_r = 13.5\) mins.) monitored at 360 nm. Unreacted \(\text{B}_{12}\) and 5’aldehyde co-elute at \(T_r = 6\) mins as indicated by MALDI-ToF MS.

MALDI-Tof mass spectrometry analysis of 4 demonstrated the anticipated molecular ion peak \([\text{M-CN}]^+\) of 1343.5 \(m/z\) (see Figure 44).
Figure 44. MALDI-ToF MS spectrum of HPLC peak at $T_r = 14$ mins indicating 4 ($[M-CN]^+ 1343.5 \, m/z$) with angiotensin as an internal reference.

Electronic absorption spectroscopy (EAS) was used to determine the oxidation state of the metal center, nature of the axial ligands bond to the metal center and yield of CA ($\varepsilon_{361}^{CA} = 27500 \, M^{-1} \, cm^{-1}$) (see Figure 45).
Figure 45. Absorbance spectrum of 4 in 35 % HPLC mobile phase B.

4.2.2 1D and 2D NMR analysis of 4

The structure of 4 was determined by assignments obtained from $^1$H NMR (see Figure 46) and $^1$H-$^{13}$C heteronuclear (HSQC and HMBC) correlations (see Figures 47 and 48). The $^1$H NMR spectra of 4 revealed the loss of 2 protons ($\delta_H = 3.75$ and 3.92), when compared to the spectrum of B$_{12}$ (CNCbl) (see Figure 49). A comparison of the HSQC of B$_{12}$ and 4 indicates the protons lost in 4 were originally attached to the R5 carbon ($\delta_C = 63.3$) (see Figure 50). HMBC was then chosen to further establish the addition of a carbonyl at the R5 position (see Figure 48), however it was established that the R5 carbon shows no connectivity to any of the ribose protons and so was assigned by HSQC.\textsuperscript{42}
Figure 46. 500 MHz $^1$H NMR spectrum of 4.

Figure 47. 500 MHz $^1$H-$^{13}$C HSQC NMR spectrum of 4.
Figure 48. 500 MHz $^1$H-$^{13}$C HMBC NMR spectrum of 4.
Figure 49. 500 MHz $^1$H NMR spectrum of B$_{12}$ (CNCbl).
Figure 50. Portion of HSQC of 4 (right) and B_{12} (left) highlighting the loss of protons ($\delta_H = 3.75$ and $3.92$) due to oxidation at the R5 carbon ($\delta_C = 63.3$) in the 4.

4.2.3 Synthesis and Characterization of B_{12}-5'-carboxylic acid benzylamine (5)

A simple amine containing organic molecule, benzylamine, was conjugate to the derivative to help with the identification of the R5 carbon by NMR. In the presence of 10 equivalents of ethyl-3-(3-dimethylaminopropyl) carbodiimide (EDC) and 20 equivalents 1-hydroxybenzotriazole (HOBt) in DMF with 1% H$_2$O, the desired conjugate was produced within 1 hour at 40% yield (Figure 51).
**Figure 51.** Synthesis of 5.

5 was purified using the same purification method as established for the carboxylic acid (see Figure 52) and characterized by MALDI mass spectrometry (see Figure 53) as well as 1D (see Figure 54) and 2D NMR (see Figure 55 and 56).

**Figure 52.** Purification of the conjugate was established by RP-HPLC using an SAX column with a gradient from 100% water to 35% MeCN: phosphate buffer pH 2.
**Figure 53.** MALDI-ToF MS spectrum of 5 ([M-CN+H]$^+$ 1432.5). Angiotensin was used as an internal reference (not shown).
Figure 54. 500 MHz $^1$H NMR spectrum of 5.

Figure 55. 500 MHz $^1$H-$^{13}$C HSQC NMR spectrum of 5.
Figure 56. 500 MHz $^1$H-$^{13}$C HMBC of 5.

**A**

![Diagram A](image)

**B**

![Diagram B](image)

**C**

![Diagram C](image)

Figure 57. Analysis of 5 by 2D heteronuclear NMR spectroscopy. A. The atom numbering scheme for the benzylamine attached to the R5 of B$_{12}$ by an amide bond. B.
Signal assignments made from HSQC. C. Signal assignments made from HMBC (Both \(^1\)H and \(^{13}\)C assignment for \(C_A\) were made by HSQC).

Addition of the benzylamine connectivity of the R5 carbon was established by HMBC (see Figures 56 and 57). Specifically the protons of \(C_A\) (\(\delta_H = 4.42\) and 4.50 ppm) showed direct connectivity to the R5 carbon (\(\delta_C = 173.8\) ppm) as well as \(C_B\) (\(\delta_C = 140.8\) ppm) and \(C_C\) (\(\delta_C = 131.4\) ppm) of the benzylamine (see Figure 57). The complete assignment of 4 and 5 can be found in Table 3.

**Table 3.** \(^1\)H and \(^{13}\)C chemical shifts for the \(B_{12}\) 5'-carboxylic acid (\(B_{12}\)CA), \(B_{12}\) 5'-carboxylic acid benzylamine (\(B_{12}\)CABA) and \(B_{12}\) (CNCbl) in D\(_2\)O [500 MHz].

<table>
<thead>
<tr>
<th>Assignment</th>
<th>(\delta^{13})C [ppm]</th>
<th>(\delta^{1}H) [ppm]</th>
</tr>
</thead>
<tbody>
<tr>
<td>C3</td>
<td>18.2 18.1 18.2</td>
<td>2.57 2.53 2.53</td>
</tr>
<tr>
<td>C5</td>
<td>18.3 18.3 18.3</td>
<td>2.54 2.52 2.49</td>
</tr>
<tr>
<td>C4</td>
<td>18.8 18.6 18.6</td>
<td>1.38 1.38 1.38</td>
</tr>
<tr>
<td>C25</td>
<td>19.7 19.6 19.7</td>
<td>1.49 1.38 1.38</td>
</tr>
<tr>
<td>Pr3</td>
<td>21.9 21.7 21.7</td>
<td>1.27 1.21 1.16</td>
</tr>
<tr>
<td>C36</td>
<td>22.0 21.9 21.9</td>
<td>1.87 1.83 1.83</td>
</tr>
<tr>
<td>C47</td>
<td>22.1 22.0 22.2</td>
<td>1.45 1.41 1.39</td>
</tr>
<tr>
<td>C20</td>
<td>22.2 22.2 22.2</td>
<td>0.45 0.45 0.45</td>
</tr>
<tr>
<td>B11</td>
<td>22.6 22.6 22.6</td>
<td>2.24 2.23 2.23</td>
</tr>
<tr>
<td>B15</td>
<td>22.6 22.6 22.6</td>
<td>2.27 2.26 2.26</td>
</tr>
<tr>
<td>C10</td>
<td>20.9 20.8 20.8</td>
<td>1.98 1.98 1.96</td>
</tr>
<tr>
<td>C44</td>
<td>30.1 31.3 30.5</td>
<td>1.99 1.91 1.90</td>
</tr>
<tr>
<td>C46</td>
<td>34.2 34.1 34.1</td>
<td>1.19 1.15 1.13</td>
</tr>
<tr>
<td>C42</td>
<td>34.4 34.2 34.4</td>
<td>1.01 0.84 0.91 0.80</td>
</tr>
<tr>
<td>C60</td>
<td>34.4 34.4 34.3</td>
<td>2.60 2.74 2.64 2.74</td>
</tr>
<tr>
<td>C56</td>
<td>35.2 34.9 35.5</td>
<td>2.10 2.66 2.09 2.61 2.56</td>
</tr>
<tr>
<td>C55</td>
<td>35.3 35.1 36.3</td>
<td>1.04 1.54 1.92 2.53 1.82 2.13</td>
</tr>
<tr>
<td>C49</td>
<td>37.5 37.6 37.6</td>
<td>2.61 2.65 2.53</td>
</tr>
<tr>
<td>C31</td>
<td>37.8 37.6 37.7</td>
<td>2.49 2.53 2.49</td>
</tr>
<tr>
<td>C18</td>
<td>41.9 41.7 41.7</td>
<td>2.76 2.76</td>
</tr>
<tr>
<td>C26</td>
<td>45.8 45.6 45.6</td>
<td>2.40 2.38 2.39</td>
</tr>
<tr>
<td>C37</td>
<td>45.9 45.7 45.7</td>
<td>2.18 2.16 2.15</td>
</tr>
<tr>
<td>Pr1</td>
<td>48.2 47.8 47.9</td>
<td>2.95 3.01 3.06 3.06 3.00 3.00</td>
</tr>
<tr>
<td>C2</td>
<td>50.1 50.1 50.2</td>
<td>3.34 3.30 3.19</td>
</tr>
<tr>
<td>C12</td>
<td>51.2 51.6 51.0</td>
<td>3.42 3.38 3.38</td>
</tr>
<tr>
<td>C7</td>
<td>54.5 54.4 54.6</td>
<td>4.18 4.13 4.19</td>
</tr>
<tr>
<td>C13</td>
<td>56.6 56.7 55.9</td>
<td>3.34 3.30 3.19</td>
</tr>
<tr>
<td>C8</td>
<td>58.7 58.4 58.5</td>
<td>3.42 3.38 3.38</td>
</tr>
<tr>
<td>C3</td>
<td>59.2 59.1 59.1</td>
<td>4.18 4.13 4.19</td>
</tr>
</tbody>
</table>

\(a\). The R5 carbon cannot be assigned because it shows no connectivities to any of the above protons and is usually assigned by HSQC.
4.3 Conclusion

In conclusion, we have developed a convenient synthesis for the site-specific oxidation of B$_{12}$ using IBX and HYP. This new carboxylic acid derivative is highly suited to the production of bioavailable amide bound conjugates. The one step purification is simplistic and the synthesis results in a 30% yield within 2 hours, which is far in excess of any previously reported B$_{12}$ carboxylic acid derivative.\(^9\) The resulting derivative has the potential to form high yielding conjugates and we believe will be of considerable use for the field of B$_{12}$ conjugation.

4.4 References


Chapter 5 Experimental Procedure

5.1 Materials and Methods

Chemicals and solvents were purchased from Sigma–Aldrich or Fluka and were used without further purification. Powdered bovine insulin was purchased from Sigma–Aldrich. Glucagon-like peptide (7-36) amide with a K34R amino acid substitution and N-terminal FMOC protection (FMOC-K34R-GLP-1) was purchased from NEOBiolab (Cambridge, MA). DMSO was dried over 4 Å molecular sieves (200–400 mesh, Sigma) under dry dinitrogen. Dialysis tubing (2000, 3500, and 7000 Da cutoff) was purchased from Pierce. H\textsubscript{2}O was distilled and deionized to 18.2 mΩ using a Barnstead Nano Diamond ultra-purification machine. RP-HPLC purification was carried out on an Agilent 1100 system with a variety of Agilent columns discussed below.

MALDI-TOF MS was performed on a Bruker Autoflex III Smartbeam machine with laser intensity ranging from 50–73%. The matrix used for MALDI-ToF was 10 mg α-cyano-4-hydroxycinnamic acid (CHCA) dissolved in dH\textsubscript{2}O/CH\textsubscript{3}CN (1:1 v/v) containing 0.1% TFA. Angiotensin I was used as an internal control (1296 m/z). B\textsubscript{12}-insulin conjugates were reduced with 10 mM DTT to determine B\textsubscript{12} attachment to A-Chain or B-Chain. Trypsin digestion of B\textsubscript{12}-insulin conjugates was accomplished with the addition of 0.02 moles trypsin to a 1 mole of B\textsubscript{12}-insulin to determine what fragment of B-chain B\textsubscript{12} attachment occurred. EAS were obtained on a Varian Cary 50 Bio spectrometer in a 1 mL quartz cuvette (Sigma) between 200 nm and 600 nm. All centrifugation was done at 4000 rpm for 10 minutes using a Sorvall centrifuge with a swinging rotor (Sorvall Heraeus 75006441 K). 96-well plates were coated with rat tail
collagen from BD Biosciences prior to assays. CMRL-1066 modified culture medium was obtained from Mediatech. Fetal bovine serum (FBS), Dulbecco’s modified eagle media (DMEM), Pluronic F-127 and Fura-2 acetoxymethyl ester (Fura-2AM) were purchased from Invitrogen Life Technologies, Carlsbad, CA. Insulin secretion was determined using a Mercodia Ultrasensitive Rat Insulin ELISA Kit #10-1247-10. Measurements of intracellular [Ca^{2+}] were performed using a FlexStation 3 microplate reader using SoftMaxPro software from Molecular Devices. A Lipofectamine 2000 kit was purchased from Invitrogen for transfections. A luciferase assay kit was purchased from Promega, Madison, WI.

5.2 Experimental Procedures

5.2.1 Synthesis of B_{12}-B{\textsuperscript{1}}Insulin (1)

The activation of the 5’-hydroxyl group of B_{12} was achieved using a previous reported method.\textsuperscript{1} \(B_{12} (10 \text{ mg, } 0.007 \text{ mmol})\) and CDT (1.7 mg, 0.010 mol) were stirred in DMSO (1 mL, dried) under dry dinitrogen for 1 h. The activated \(B_{12}\) was used without further purification. A variation of the literature procedure reported by Kim et al. was used for the BOC protection of insulin.\textsuperscript{2} BOC (4.0 mg, 0.0187 mmol) and NHS (4.4 mg, 0.04 mmol) dissolved in DMSO (100 \(\mu\)L) containing 0.05% TEA (v/v) stirred for 0.5 h prior to being added to a stirring solution of insulin (5.0 mg, 0.00087 mmol) dissolved in DMSO (1 mL) containing 0.05% TEA (v/v). After 0.5 h, an excess of activated \(B_{12}\) (2.4 mg, 0.0017 mmol) was added to the BOC protected insulin solution. The reaction stirred for 4 h and was isolated using diethyl ether (\(\text{Et}_2\text{O}\)) and centrifugation. The red solid was dissolved in 25/75 (v/v) \(\text{CH}_3\text{CN/H}_2\text{O}\) containing 0.1% TFA and purified using analytical
RP-HPLC with an Eclipse XDB-C$_{18}$ column (5 μm, 9.4 X 250 mm); flow rate: 1 mL/min, solvent A: dH$_2$O/0.1% TFA; solvent B: CH$_3$CN/0.1% TFA; method 0-10 min 75/25% A/B to 47.5/52.5% A/B. Collected 1 was then spiked with 50% TFA (0.5 ml) and stirred for 0.2 h to remove the BOC protection. 1 was dialyzed against 5 L dH$_2$O in snakeskin tubing (7000 Da cutoff) to remove residual free insulin and BOC. The dialysis bag content was then concentrated by lyophilization overnight yielding a light pink powder (0.31 mg, 5%). HPLC T$_r$ = 8.5 mins; MALDI-ToF: calculated for C$_{317}$H$_{464}$CoN$_{79}$O$_{89}$PS$_6$ Insulin-[B$_{12}$-CN+H]$^+$ 7090.9 m/z, found 7090.5 m/z; MALDI-ToF of reduced of 1: calculated for insulin’s B-Chain + [B$_{12}$]$^+$ 4754.5 m/z, found 4754.5 m/z; MALDI-ToF of reduced and digested 1: calculated for Insulin’s B23-B30 + [B$_{12}$]$^+$ 2285.9 m/z, found 2285.8 m/z. Yield was calculated from EAS using $\varepsilon_{361}$ =27,500 M$^{-1}$cm$^{-1}$.

5.2.2 Synthesis of B$_{12}$-B$_{29}$Insulin (2)

Powder bovine insulin (5.0 mg, 0.00087 mmol) was dissolved in DMSO (1 mL) containing 0.05% TEA (v/v). An excess of activated B$_{12}$ (2.4 mg, 0.0017 mmol) was transferred to the unprotected insulin and reacted for 4 hours. FMOC-OSu (0.5 mg, 0.0015 mmol) in DMF (1 mL) was added to the reaction mixture. The reaction mixture stirred for 2 h and was isolated using Et$_2$O and centrifugation. The red solid was then dissolved in a 25/75 (v/v) CH$_3$CN/H$_2$O containing 0.1% TFA and purified using analytical HPLC. Piperidine (10% v/v) was added to the pink fraction. Dialysis (7000 Da cutoff) against H$_2$O (5L) removed residual free insulin and FMOC. The dialysis bag content was then concentrated by lyophilization overnight yielding a light pink powder (0.62 mg, 10%). HPLC T$_r$ = 8.5 mins; MALDI-ToF: calculated for C$_{317}$H$_{464}$CoN$_{79}$O$_{89}$PS$_6$ Insulin-[B$_{12}$-CN+H]$^+$ 7090.9 m/z, found 7090.5 m/z; MALDI-ToF of reduced of 2:
calculated for insulin’s B-Chain + [B_{12}]^+ 4754.5 m/z, found 4754.5 m/z; MALDI-ToF of reduced and digested 2: calculated for Insulin’s B1-B22 + [B_{12}-CN+H]^+ 3844.6 m/z, found 3844.6 m/z. Yield was calculated from EAS using $\varepsilon_{361} = 27,500$ M$^{-1}$cm$^{-1}$.

5.2.3 Synthesis of B_{12}-K34R-GLP-1(7-36) amide (3)

B_{12} (10 mg, 0.007 mmol) was activated at the 5'-hydroxyl group using CDT (1.7 mg, 0.010 mmol) in DMSO (1 mL, dried) under dry dinitrogen for 30 mins. The activated B_{12} solution (0.096 mL, 0.0007 mmol) was subsequently added to a stirring solution of FMOC-K34R-GLP-1 (1.0 mg, 0.0028 mmol) dissolved in DMSO (0.1 mL) with 0.6% TEA (v/v). The reaction stirred for 3 hrs at RT before H$_2$O w/ 0.1% TFA (20% v/v) was added to the reaction to lower the pH. A solution of FMOC-OSu (0.028 mmol) in DMF was added and allowed to proceed an additional hr. The reaction was extracted with the addition of diethyl ether. Purification of the B_{12}-FMOC-K34R-GLP-1 conjugate was accomplished by RP-HPLC using an Agilent C$_{18}$ column and HPLC method solvent A: dH$_2$O/0.1% TFA; solvent B: CH$_3$CN; 0-25 min 75/35% A/B to 35/65% A/B. The pH of the solution was adjusted to pH 8 with NH$_2$HCO$_3$ immediately following isolation by HPLC. The removal of the N-terminal FMOC protection was accomplished with the addition of 10% piperidine after 10 minutes. The solution was transferred to 3500 MWCO dialysis tubing and dialyzed against H$_2$O for 18 hours, changing solution every 4-6 hours. 3 was collected as a light pink solution (0.10 mg, 10%). HPLC $T_r$ = 19.5 mins; MALDI-ToF: calculated for K34R-GLP-1-[B_{12}-CN+H]^+ 4680 m/z, found 4682 m/z; Yield was calculated from EAS using $\varepsilon_{361} = 27,500$ M$^{-1}$cm$^{-1}$.

5.2.4 Synthesis of B_{12}-5'-carboxylic acid (4)
B$_{12}$ (10.0 mg, 0.007 mmol), 2-iodoxybenzoic acid (5.37 mg, 0.019 mmol) and 2-hydroxypyridine (3.51 mg, 0.037 mmol) were dissolved in DMSO (1 mL). The reaction temperature (50-60°C) and time (2-5 hrs) were varied as outlined in Table 1. The crude reaction was precipitated by diethyl ether addition and dissolved in H$_2$O prior to purification. The derivative was purified by ion exchange chromatography using a SAX column. After a 5 minute hold at 100% H$_2$O, a gradient from 100% H$_2$O to 35% MeCN: phosphate buffer pH 2 (50:50 v/v) over 5 minutes was necessary for the separation and elution of the 4. 4 was then neutralized and mixed with amberlite XAD4 resin overnight. The resin, now red suggesting 4 binding, was collected and washed with H$_2$O to remove any salt prior to MeOH extraction of the 4. The MeOH was collected and lyophilized yielding a red powder (3 mg, 30% for 2 hrs at 60°C). HPLC $T_r$ =14 mins; MALDI-TOF MS: calculated for C$_{62}$H$_{86}$CoN$_{13}$O$_{15}$P [M-CN]$^+$ 1343.5, found 1343.5; See chapter 4 table 3 for NMR assignments.

5.2.5 Synthesis of B$_{12}$ 5'-carboxylic acid benzylamine (5).

4 (10.0 mg, 0.007 mmol), benzylamine (7.8 mg, 0.073 mmol), 1-EDC (14.0 mg, 0.073 mmol) and HOBt (19.7 mg, 0.146 mmol) were dissolved in DMF (1 mL) with 1% H$_2$O at rt for 1 hr. Isolation and purification were achieved as described for the 4 (4.0 mg, 40%). HPLC $T_r$= 10 mins. MALDI-TOF MS: calculated for C$_{69}$H$_{93}$CoN$_{14}$O$_{14}$P [M-CN+H]$^+$ 1432.5, found 1432.5; $\delta_H$ = 7.42, 7.41, 7.41, 7.39, 7.38 corresponding to the aromatics of benzylamine and 4.46 corresponding to C$_A$. $\delta_C$ = 140.8, 131.4, and 46.4 corresponding to C$_B$, C$_C$, and C$_A$ respectively. See chapter 4 table 3 for B$_{12}$ related NMR assignments.
5.3 *In vitro* Investigation of 3

5.3.1 Cell Cultures

Human embryonic kidney (HEK-293) cells stably expressing the human GLP-1 receptor were a gift of Novo Nordisk A/S (Bagsvaerd, Denmark). HEK-GLP-1R cells were cultured in DMEM containing 10% FBS, penicillin/streptomycin, and G418. The cells were maintained at 37 °C in a humidified incubator gassed with 5% CO₂ and were passaged by trypsinization once per week.

Human islets of Langerhans, obtained posthumously from anonymous donors, were provided under the auspices of the Integrated Islet Distribute Program of National Institutes of Health, National Center for Research Resources, and National Islet Cell Resource Centers. Primary cultures of human islets were maintained in a humidified incubator containing 95% air and 5% CO₂ at 37 °C in CMRL-1066 modified culture medium containing 10% (v/v) FBS.

5.3.2 Luciferase Assay

For GLP-1 activity, HEK-GLP-1R cells were plated in a rat tail collagen coated 96 well plate at a density of 55,000-65,000 cells/well and incubated overnight at 37°C. The next day, the cells were transfected with RIP1-CRE-Luc using Lipofectamine 2000 according to the manufacturer’s instructions. The cells (HEK-GLP-1R-RIP1-CRE-Luc) were cultured in DMEM containing 10% FBS overnight, then exposed for 4 h to serum-free culture medium containing 1% BSA and the test substance. Cells were then lysed and assayed for luciferase-catalyzed photoemissions using a luciferase assay kit and luminometer.
5.3.3 Fura-2 Assay

The experiment was performed using a monolayer of fura-2 loaded HEK-GLP-1R cells grown on rat tail collagen coated Costar 3904 plate. Fura-2 was loaded in a standard extracellular solution (SES) containing: 138 NaCl mM, 5.6 mM KCl, 2.6 mM CaCl₂, 1.2 mM MgCl₂, 10 mM HEPES pH 7.4 and 11.1 mM glucose. The SES was also supplemented with 20 μl per ml FBS, 1 μl per ml Pluronic F-127 and 1 μM fura-2 AM. Spectrofluorimetry was performed using excitation light at 355/9 and 375/9 nm (center/bandpass) delivered using a 455 nm dichroism mirror. Emitted light was detected at 505/15 nm and the ratio of emission light intensities due to excitation at 355 and 375 was calculated.

5.3.4 Insulin Secretion in Human Pancreatic Islets

Insulin secretion was determined by static incubation. After overnight culture in CMRL-1066 medium, human islets were transferred to MilliCell-PCF culture plate filter inserts at a density of 40 islets/insert. The inserts were mounted within individual wells of a 24-well cell culture plate and each well of the plate was filled with 1 ml of Krebs-ringer buffer (KRB) containing 24 mM sodium bicarbonate and the indicated concentrations of glucose (pH 7.4). Culture plates containing islets were placed within a cell culture incubator gassed with 95% air and 5% CO₂ at 37 °C. Prior to the start of an experiment, islets were exposed to KRB with 2.8 mM glucose for 30 min. The inserts were then transferred between adjacent wells on the plate to measure basal and stimulated insulin secretion. For each test solution, the duration of exposure was 30 min. A 200 μl
fraction of each solution was then collected and assayed for insulin content by an enzyme-linked immunosorbent assay after appropriate dilution of samples.

5.4 In vivo Experimentation of 1 and 2

Male Sprague Dawley rats (279 ± 31; ± SD; n=12) were rendered insulin-deficient via subcutaneous injection of streptozotocin (STZ; 80 mg/kg BW) dissolved in citrate buffer (100 mM; pH 4.2). Blood glucose concentration was assessed after four days to ensure fasting levels of greater than 15 mmol/L. Maintenance of animals and experimental protocols were conducted in accordance with federal regulations and approved by the Murdoch University Animal Ethics Committee. Animals were fasted overnight prior to the first blood sample being taken from the lateral tail vein and assessed for glucose concentration (One-touch Verio). Compounds were then administered by oral gavage using a flexible feeding tube and blood samples collected 60, 120, 150, 180, 210 and 240 min post gavage from the lateral tail vein. AUC transformations were conducted using the trapezoidal method in SigmaPlot, which was then subtracted from pre-gavage[mmol] x 240min; higher numbers represent a greater reduction in glucose concentration following oral gavage. Statistical significance was then assessed on this data using a one-way ANOVA with repeated measures (condition) after testing for normality (Shapiro-Wilk) and equal variance. Tukey post Hoc testing was used to assess differences between condition.

5.5 References

Chapter 6  Future Work

6.1 Vitamin B_{12}

Major strides towards establishing the role of B_{12} as a carrier for oral delivery of protein therapeutics have been achieved. The full potential of the B_{12} dietary uptake pathway, however, has not been completely realized. Further exploration into the transport proteins in terms of their role in protecting peptides/proteins in the stomach and intestines, and the potential loss of the peptide/protein to TCII in the blood stream needs to be done.

One of the major barriers facing oral drug delivery of peptides and proteins is the acidic/proteolytic environment of the GIT.\textsuperscript{1} Previous reports have discussed the inability of the transport proteins in aiding in the protection of peptides/proteins from these elements.\textsuperscript{1} The previous authors, however, had no data to support such claims and our \textit{in vivo} work with B_{12}-insulin suggests there \textit{is} a protecting effect from the transport proteins.\textsuperscript{2} \textit{In situ} experimentation with the HC-B_{12} peptide/protein complex or IF-B_{12} peptide/protein complex in the presence of simulated stomach and intestinal conditions would help to establish this important ability of the transport proteins. The production of simulated stomach and intestinal conditions has been well established and can be followed from literature.\textsuperscript{3}

Knowledge of whether there is loss of insulin activity through TCII-B_{12} based binding to the TCII receptor CD320, as opposed to the desired binding to the insulin receptor CD220 (see Figure 58), needs to be investigated. If, upon oral delivery to the circulatory system, a significant proportion of the conjugate binds to the CD320 and the
rest goes to the target CD220, then modification of the conjugate to minimize CD320 binding would immediately improve the response of the oral B_{12}-insulin conjugate and ensure an optimal system had been obtained.

**Figure 58.** Rendering of B_{12}-insulin (yellow-blue) bound by TCII (red and dark blue) produced by collaborator Dr. Damien Allis at Syracuse University. It is possible for such a structure to bind either the TCII receptor CD320 or the insulin receptor CD220 as indicated. The image was altered from \textsuperscript{2} with permission of The Royal Society of Chemistry (RSC) on behalf of the Centre National de la Recherche Scientifique (CNRS).

The effect of different B_{12} modifications on the binding affinity of the three transport proteins has been extensively documented,\textsuperscript{46} with studies showing conversion of the b propionamide group into a carboxylic acid (see Figure 59) reducing binding of
B₁₂ by human TCII.⁶⁻⁸ To investigate the binding of TCII, a B₁₂ b carboxylic acid based insulin conjugate (B₁₂ b acid-insulin) was synthesized with the intent of comparing its activity to that of previously reported, orally active B₁₂-insulin in a rat model. Recognition by IF is not completely removed for such a modification, making the B₁₂ b acid-insulin an intriguing and worthwhile study of CD320 vs. CD220, since if positive results were obtained, the final conjugate could still be used in oral form.⁶

**Figure 59.** Structure of B₁₂ with side chain propionamides indicated. Conversion of the propionamide at the b position to a carboxylic acid is cited as reducing TCII protein binding of B₁₂.⁷⁻⁸
Previous reports of the synthesis and purification of the b monocarboxylic acid show it is a very time consuming processes.\textsuperscript{7, 9-11} By utilizing a microwave reactor, the b monocarboxylic acid was synthesized in the Doyle lab in three min. Subsequently a one-step RP-HPLC purification method was established using a C\textsubscript{8} column with a step gradient of 2\% B for 3 min, 10\% B 3.1 to 8 min, 15\% B 8.1-13.0 min and 20-80\% B 13.1 to 20 min with A: H\textsubscript{2}O and B: MeOH (See Figure 60). The difficulty with purification is a result of the number of byproducts in the reaction. Carboxylic acid formation can occur at one (mono), two (di), or three (tri) propionamide positions (b, d, or e). The tricarboxylic acid elute first, followed by the dicarboxylic acids and the individual monocarboxylic acid. B\textsubscript{12} fragments are also byproducts which elute from T\textsubscript{r} 15-31 min. The peak at 13 minutes corresponds to the b monocarboxylic acids as determined by literature comparison and 1D and 2D NMR.

\textbf{Figure 60.} RP-HPLC of B\textsubscript{12} b carboxylic acid derivative.
The ability of TCII to bind our synthesized b-carboxylic acid derivative was explored by collaborators at Arrhus University in Denmark. The Nexo group had recently demonstrated that the B₁₂ transport proteins are species specific and that some species have only two transport proteins (IF and TCII) with altered TCII binding properties. Experimentation with human and mouse TCII established the b monocarboxylic acid derivative had reduced binding with human TCII, as literature would suggest, but was still bound by mouse TCII. Furthermore, the binding studies with the B₁₂ b carboxylic acid-insulin conjugate (AKP2 in Figure 61) vs. the B₁₂-insulin conjugate (AKP1 in Figure 61) were shown to bind at comparable levels in mouse and human TCII. This result was very interesting and suggested an interaction between insulin and TCII that negated the ability of TCII to bind the b carboxylic acid, which should be further studied. These studies also demonstrated the need to evaluate the animal model used for further exploration into the potential loss of protein/peptide to TCII binding.

![Graphs showing competitive binding study of human and mouse TCII with B₁₂-insulin conjugates](image)

**Figure 61.** Competitive binding study of human TCII vs. mouse TCII with B₁₂-Insulin conjugates.
(AKP1), B_{12} b carboxylic acid-insulin (AKP2), B_{12} b carboxylic acid (Acid) and control B_{12} from two sources (B_{12} USA and B_{12} DK).

6.2 **Glucagon-like Peptide-1**

The next step for the B_{12}-GLP-1 conjugate will be to assess its *in vivo* bioavailability and uptake upon oral administration. Bioavailability will be assessed through food intake and body weight gain following subcutaneous injection in male Sprague dawley rats (SD). The uptake and bioavailability of orally administered B_{12}-GLP-1 will be determined in an obese diabetic model, Zucker diabetic fatty (fa/fa) rats (ZDF). In these experiments, bioavailability of the conjugate will be assessed by: changes in food intake and weight gain relative to baseline and controls; and amelioration of the rodent’s insulin response to glucose stimulation as compared to GLP-1 and vehicle treated controls.

The therapeutic use of GLP-1 is limited by a short half-life (1-2 min), due primarily to degradation by the enzyme DPP-IV.\textsuperscript{12} Exendin-4 (Ex-4) is a DPP-IV resistant GLP-1 agonist currently approved for the treatment of T2D (as exenatide, brand name Byetta).\textsuperscript{13} A drawback of Ex-4 based therapy is that, as a peptide, it needs to be administered subcutaneously (s.c).\textsuperscript{1} An orally active Ex-4 may offer advantages in areas of patient acceptance/adherence to a regimen.\textsuperscript{14} Since native GLP-1 is secreted by the intestinal L cells, an orally delivered GLP-1 agonist would also have the highly significant possibility of mimicking the natural physiological effects of native GLP-1 more closely, in particular by interacting with the vagal afferent fibers of the gastrointestinal (GIT) tract (not affected by the s.c administered form).\textsuperscript{15,16}
A B$_{12}$-exendin-4 conjugate would be synthesized using CDT coupling chemistry between the B$_{12}$ ribose 5’-hydroxyl group and lysine K12 residue of Ex-4. No significant reduction in peptide activity has been shown when conjugation occurs at K12 of Ex-4.$^{17}$ Ex-4 will be purchased with FMOC protection at the N-terminus to aid in selective conjugation and the FMOC protection will be removed prior to biological testing. Purification will be achieved by RP HPLC using a C$_{18}$ column. B$_{12}$-exendin-4 will be characterized by electronic absorption and circular dichroism (CD) spectroscopy, MALDI-ToF mass spectrometry and B$_{12}$ binding studies with IF and TCII. These experiments will produce data confirming the conjugate is pure, coupled correctly, has remained folded through the chemical coupling procedure and the B$_{12}$ is still recognized by the transport proteins needed for oral delivery. To confirm the biological function of B$_{12}$-exendin-4 and to gauge the effects of B$_{12}$ conjugation on Ex-4, in vitro bioassays (CRE-luciferase in HEK 293-GLP-1R cells, insulin secretion in human islets monitored by ELISA, and Ca$^{2+}$ stimulation in HEK293-GLP1R cells) will be performed, as for B$_{12}$-GLP-1 previously investigated (see Chapter 3).

6.3 C-peptide

The B$_{12}$ uptake pathway is not without its pitfalls.$^{1}$ The limitations of the B$_{12}$ uptake pathway are known and include a limited uptake capacity. Therefore, the oral delivery pathway of B$_{12}$ may be better suited for a non life threatening peptide, such as C-peptide, which requires very little to produce a desired effect. Recent studies have shown the health benefits of C-peptide replacement therapy in patients with type I DM.$^{18}$ Recombinant protein expression of a C-peptide mutant will allow for conjugation to B$_{12}$. 
An attempt to orally delivery C-peptide could be a great step forward in the treatment of complications associated with diabetes.

6.4 References


Department of Chemistry

Susan Clardy-James BS, MS(Chemistry)
Syracuse University
Department of Chemistry
111 College Place
Syracuse, NY 13244
smclardy@syr.edu

Education

Syracuse University

2004-2008
Millikin University
Bachelor of Science, Chemistry

Experience

Advisor: Robert P. Doyle, Department of Chemistry, Syracuse University

Explored the use of vitamin B$_{12}$ for the oral delivery of peptides and proteins. Worked extensively with insulin and GLP-1. Results from this work expanded the use of vitamin B$_{12}$ for the delivery of potent peptides and provided general criteria for design of future B$_{12}$ peptide conjugates. Basic techniques used include extensive bioconjugate synthesis, recombinant expression in E.coli (including vector construction and transformation), mammalian cell growth, assay using HEK and human islets, and transient transfection. Purification techniques for modified substrates, protein conjugates and recombinant protein include dialysis, HPLC/FPLC (C$_{18}$/C$_{8}$/IEC/SEC/Affinity) chromatography, and gel electrophoresis. Characterization techniques for substrates, conjugates and purified protein include 1D and 2D NMR spectroscopy, MALDI and electrospray mass spectrometry, circular dichroism spectroscopy, and bioassays (CRE-luciferase, Insulin secretion monitored by ELISA, Fura-2 Ca stimulation). (August 2008-Present)

Meda Pharmaceuticals
QC/QA Division
Decatur, IL, USA

Tested raw materials and established product stability and composition prior to distribution utilizing standard operating procedures outlined by the FDA. (June 2007- July 2008)

Tate & Lyle
Intern, Physical Chemistry Division
Decatur, IL, USA

Used Rapid-Visco Analzyer (RVA) and Lumisizer to determine the viscosity and stability of various substances. (June 2006- June 2007)
Teaching Experience
Syracuse University, Department of Chemistry
Syracuse, NY, USA

Honor General Chemistry Lab (Spring 2010, Spring 2011, January 2012-Present)
Undergraduate Student (August 2010-Present)
Introductory Chemistry (August 2011-December 2011)
High School Student (May 2011-August 2011)

Presentations


"The Path to an Orally Administered Protein Therapeutic for the Treatment of Diabetes Mellitus." S. Clardy-James. 2012 ACS GOLD lecturer at Millikin University. Decatur, IL. April 2012. (Invited)


Publications


References
Professor Robert P. Doyle
Syracuse University
Departments of Chemistry and Biology, 111 College Place
Syracuse, NY 13244
rpdoyle@syr.edu 315-443-3584

Professor George G. Holz
Upstate Medical University
Departments of Medicine and Pharmacology, 4310 Institute For Human Performance
Syracuse, NY 13210
holgz@upstate.edu 315-464-9841

Dr. Timothy J. Fairchild
Murdoch University
Department of Exercise Physiology
Murdoch, W. Australia, 6150
T.fairchild@murdoch.edu.au (61) 08 9360 2959
Vitamin B₁₂ in drug delivery: breaking through the barriers to a B₁₂ bioconjugate pharmaceutical

Susan M Clardy, Damian G Allis, Timothy J Fairchild & Robert P Doyle†
†Syracuse University, Syracuse, Department of Chemistry, NY 13244-4100, USA

Importance of the field: Vitamin B₁₂ (B₁₂) is a rare and vital micronutrient for which mammals have developed a complex and highly efficient dietary uptake system. This uptake pathway consists of a series of proteins and receptors, and has been utilized to deliver several bioactive and/or imaging molecules from ⁹⁹mTc to insulin.

Areas covered in this review: The current field of B₁₂-based drug delivery is reviewed, including recent highlights surrounding the very pathway itself.

What the reader will gain: Despite over 30 years of work, no B₁₂-based drug delivery conjugate has reached the marketplace, hampered by issues such as limited uptake capacity, gastrointestinal degradation of the conjugate or high background uptake by healthy tissues. Variability in dose response among individuals, especially across ageing populations and slow oral uptake (several hours), has also slowed development and interest.

Take home message: This review is intended to stress again the great potential, as yet not fully realized, for B₁₂-based therapeutics, tumor imaging and oral drug delivery. This review discusses recent reports that demonstrate that the issues noted above can be overcome and need not be seen as negating the great potential of B₁₂ in the drug delivery field.

Keywords: B₁₂, cobalamin, haptocorrin, intrinsic factor, peptides, transcobalamin

1. Introduction

The consumption of vitamin B₁₂ (B₁₂; cobalamin) is essential for the survival of all living cells. B₁₂ is produced naturally by bacteria and all other species must acquire the vitamin through their diet. There are two major biologically active forms of B₁₂: methylcobalamin and adenosylcobalamin. Methionine synthase uses methylcobalamin to produce the amino acid methionine from homocysteine, and methylmalonyl-CoA mutase uses adenosylcobalamin as a cofactor to produce succinyl CoA, an important molecule in the TCA cycle [1]. Mammals have developed a complex uptake pathway for B₁₂ involving a series of transport proteins [1-3]. For the purposes of this review, the uptake pathway is discussed only superficially. For a more in-depth discussion of the transport proteins, the reader is referred to recent reviews by Banerjee et al. [1] and Randaccio et al. [4].

B₁₂ is a water-soluble vitamin (molar mass 1355.38 g/mol as cyanocobalamin) with a highly complex structure, comprising a midplanar corrin ring composed of four pyrrole elements linked to a central cobalt(III) atom (see Figure 2 later). The corrin ring is similar to the more commonly known porphyrin structure but with key differences. Corrin rings have a greater degree of saturation compared with porphyrins and the increased number of sp³ carbons confers greater flexibility to the corrin. In addition, there is greater asymmetry to the corrin ring, the latter the result of a 15-carbon ring over the porphyrin 16-carbon structure [1].
Vitamin B<sub>12</sub> in drug delivery

**Article highlights.**

- An overview is provided of B<sub>12</sub> structure, chemistry, and uptake proteins and receptors.
- An overview is given of the use of the B<sub>12</sub> dietary pathway in drug delivery.
- Methods and highlights in the use of B<sub>12</sub> conjugates to utilize/target specific B<sub>12</sub> proteins and associated receptors.
- In the section on general uses of B<sub>12</sub> in drug delivery, oral delivery is discussed, especially of proteins and peptides, and general small molecule delivery by oral and invasive means is discussed.

This box summarizes key points contained in the article.

Several functional groups are readily available for modification on B<sub>12</sub>. Only a few modification sites, however, maintain the recognition needed to utilize the full B<sub>12</sub> uptake pathway necessary for oral delivery (see Figure 1). Other modifications can be used to target specific proteins while reducing affinity for others, a fact recently exploited to target haptocorrin (HC)-positive tumors [5]. An in-depth discussion of modification sites can be found in Section 2.

B<sub>12</sub> is initially released from food by the action of peptic enzymes and the acidic environment of the gastrointestinal system [6]. It is then bound and transported by two glycoproteins, HC and intrinsic factor (IF) [2]. Haptocorrin is secreted by salivary glands and released also by the gastric mucosa. Haptocorrin has a high affinity for B<sub>12</sub> under acidic conditions (pH < 3) and so protects B<sub>12</sub> from acid hydrolysis. The HC:B<sub>12</sub> complex travels from the stomach to the duodenum, where the increased pH (> 5) decreases the affinity of HC for B<sub>12</sub> [7]. Haptocorrin is also enzymatically digested here. Release from HC, B<sub>12</sub> binds to the second of the two gastric transport glycoproteins, IF.

Intrinsic factor is a 43.4 kDa glycosylated protein that is secreted from the gastric mucosa and the pancreas [2]. The IF protein facilitates transport across the intestinal enterocyte, which occurs by receptor-mediated endocytosis at the apically expressed IF-B<sub>12</sub> receptor (cubilin) [8]. Cubilin works to transport B<sub>12</sub> in concert with an anchoring protein amnionless (Am) [6]. Following transcytosis, and between 2.5 and 4 h after initial ingestion, B<sub>12</sub> appears in blood plasma bound to the third trafficking protein, transthyretin (TCII) [3]. The holo-TCII is cellularly internalized by the TCII receptor (TCII-R), B<sub>12</sub> is released by the degradation of TCII by lysozyme. Another receptor, megalin (MG), can reabsorb filtered holo-TCII from primary urine. The TCII-R and megalin receptors then ensure widespread delivery and maximal use of B<sub>12</sub> [1].

Knowledge of the binding between B<sub>12</sub> and its various transport proteins is critical if the system is to be successfully translated from bench to bedside. In the last 5 years there has been an explosion of critical structural data related to the B<sub>12</sub> uptake pathway, with the publication of the IF [9], TCII [10] and cubilin/5,8)-IF-B<sub>12</sub> [8] structures. One missing piece in the B<sub>12</sub> puzzle is the crystal structure of HC. Indeed, a real understanding of the functions of HC remains elusive, with suggestions it may play a bacteriostatic role in the mouth and bloodstream [11]. What is true now is that researchers have a better understanding of how B<sub>12</sub> interacts with its transport proteins, and how these transport proteins interact with their receptors. The implications this can have on drug delivery and sites of potential conjugation can then be better detailed, rationalized and hence optimized.

Along with the wealth of knowledge from the experimental field is the recent use of modern theoretical approaches to B<sub>12</sub>-based drug design. The foundational work for B<sub>12</sub> molecular modeling studies has all but been completed with the publication of several recent papers providing both force field parameters and charge density information at levels of theory comparable to, or exceeding, the quality of the force fields within which the parameters have been used [12-17]. With very little modification, a family of B<sub>12</sub> structures can be turned into topologies for force field parameter assignments. This has opened the door to a wider range of B<sub>12</sub> molecular modeling studies, including studies of the periplasmic binding protein BtuF [16] and the TonB-dependent transporter BtuB [17]. Computational approaches have shown themselves to be a powerful interpretive tool for explaining the various structures involved in B<sub>12</sub> transport and providing a useful tool for rational drug design.

Based on the results of recent years, it is clear there is great hope for a B<sub>12</sub>-based bioconjugate to reach the marketplace sooner rather than later. This review details and expands on the points made in this introduction and discusses the future of the field.

2. Review of the B<sub>12</sub> dietary pathway in drug delivery

2.1 Modifiable points – solvent-accessible pocket

The solvent-accessible surface of B<sub>12</sub> is critical when considering B<sub>12</sub>-based bioconjugates. The key point to be made upfront is that there is solvent accessibility. For TCII this solvent exposure is ~ 6.5% (~ 80 Å<sup>2</sup>). For IF this exposure is twice as high at ~ 13% (~ 163 Å<sup>2</sup>), with HC the least accessible at 3.2% (~ 40 Å<sup>2</sup>) [18]. This exposure allows for select sites to be utilized when designing a B<sub>12</sub> conjugate, for general pathway acceptance or for selecting specific parts thereof.

B<sub>12</sub> and the molecule of interest can be: i) coupled directly together; ii) held apart by ‘spacer’ units to produce distance between the B<sub>12</sub> and drug; or iii) carriers can be conjugated to B<sub>12</sub> with the desired drug contained, unconjugated, within this carrier. Several functional groups are readily available for modification on B<sub>12</sub>, including propionamides, acetamides, hydroxyl groups, the cobalt(III) ion and the phosphate moiety. Only a few modifications are capable of maintaining the recognition of all three transport proteins needed to utilize
the full $\text{B}_{12}$ uptake pathway effectively in an oral manner. This limitation is a result of the manner by which the $\text{B}_{12}$ is bound by the transport proteins and a strong indicator even in the absence of structural information on the similar solvent accessibility of $\text{B}_{12}$ positions on binding in all cases. The size of the drug is important when considering the location for conjugation, as are the transport proteins being utilized, because the three transport proteins are not created equally. Conjugation of drugs resulting in the recognition of all three transport proteins has been successful with $\text{B}_{12}$ at four major sites: i) to the peripheral corrin ring $\text{e}^{-}$-propionamide [19]; ii) through the 5'-hydroxy group of the ribose unit of the $\alpha$-tail [20]; iii) through the 2'-hydroxy group of the ribose unit of the $\alpha$-tail [21]; and iv) to the cobalt cation (see Figure 2) [22].

Examination of the recently published crystal structures of $\text{B}_{12}$ bound with TCII [10] and IF [9] provides a better understanding of why such positions are available for conjugation without complete loss of recognition. TCII was the first transport protein crystallized [10]. The structure revealed several important details about points of modification on $\text{B}_{12}$, including the presence of hydrogen bond formation between various residues in TCII and $\text{B}_{12}$’s phosphate moiety and the side chains of the corrin ring. The further stability found to occur with the replacement of the Co’s axial ligand with a histidine residue from TCII [10] was also rationalized. However, this latter substitution is not necessary for the binding of TCII and will not occur if a stable Co–C bond has been formed, such as the Co-CN in cyanocobalamin. The crystal structure [10] and molecular dynamics studies [14] have also shown that TCII does not completely encompass $\text{B}_{12}$ on binding, and leaves an exposed section of the vitamin accessible to solvent. The 1.4 nm solvent-accessible pocket of $\text{B}_{12}$ bound to TCII shows that the phosphate and ribose hydroxyl groups are left protruding into the solvent, but only the ribose 5'-hydroxyl tail is open to the environment enough such that it can easily accommodate the conjugation of large molecules (see Figure 3). Depending on the size of the conjugate, the three other sites have been shown to accommodate molecules without interference with TCII binding. It has been established that recognition and uptake of the vitamin still occur on modification of the $\beta$-axial position of the $\text{B}_{12}$’s cobalt, the $\alpha$-axial ribose tail and several exterior carboxylic acid groups (produced by hydrolysis of side-chain propionamides) [10,19,22].
The crystal structure of IF was published in 2007 [9]. This ~60 kDa protein is highly glycosylated and proved difficult to crystallize. The IF-B12 complex has several similarities to the previously discussed TCII-B12 complex, including several of the same hydrogen bond formations between the protein and the vitamin. The IF-B12 complex does not incorporate a histidine residue in the β-axial position, however, and ~13% of the B12 is solvent-accessible compared with only 6.5% of the B12 in the TCII-B12 complex [18]. The crystal structure revealed a possible mechanism, involving pH, for the transfer of B12 from IF to TCII, allowing researchers a better understanding of the last few steps of the B12 pathway [9].

While attention has concentrated on maintaining the binding between the transport proteins and B12 bioconjugates, little is known about what impact, if any, modifications to the B12-transport protein complex might have on subsequent receptor binding. Until recently the structure of a holoprotein complex bound to a B12 receptor was unknown. In 2010, Andersen et al. published the structure of cubilin(5–8)-IF-B12 (see Figure 4) [8]. This publication provided an outline of how the holo-complex interacts with the receptor, allowing researchers a new tool to determine the implications of conjugation on receptor binding.

2.2 Targeting specific transport proteins

On delivery to the bloodstream, a B12 conjugate typically will be acquired by the TCII protein. In cancer therapy/imaging, the hypothesis in B12-based targeted delivery has historically been that increased TCII-R expression (as much as 3- to 26-fold in certain patients [6]) in a variety of cancers such as testis, breast, ovarian, thyroid, uterine, and brain cancer would provide sufficient selectivity over healthy tissue. This characteristic of B12 led to the production of both small molecule organic and inorganic B12 bioconjugates with various applications in medicinal chemistry. Although the bioconjugates are discussed in more detail in the next section, it should be noted here that the major problem associated with these bioconjugates has been background uptake by healthy tissue.

In an attempt to overcome background uptake, HC has been targeted more recently. Cytoplasmic or membrane-associated HC has been suggested to be de novo expressed [5] or overexpressed [23] in certain cancer lines. To target HC over TCII, Waibel et al. disrupted the interactions between

---

Figure 2. Structure of vitamin B₁₂. Molecular structure of XCbls: the central cobalt(III) atom is six-coordinated, with the equatorial positions filled by the nitrogen atoms of the corrin macrocycle. The (conventionally) ‘lower’, ‘α’-axial site is occupied by an imidazole nitrogen atom from a 5’,6’-dimethylbenzimidazole base whereas the ‘upper’, ‘β’-axial site can be occupied by various X groups (e.g., CN-, CH₃-, Ado-, NO₃-, SCN-, SeCN-, SO₃⁻ and thiourea). The corrin ring incorporates seven amide side chains, three acetamides (a, c, g) and four propionamides (b, d, e, f). The four pyrrole rings are usually indicated as A, B, C and D, as shown. Areas lightly shaded in pink have been commonly used for conjugates that maintain HC, IF and TCII binding. The area shaded in yellow, namely the β-propionamide, has been used, on conversion to the carboxylic acid, for HC-specific targeting. This HC-specific targeting is important where background TCII-R uptake prevents selectivity of cancerous tissue over healthy tissue.

Haptocorrin; IF: Intrinsic factor; TCII: Transcobalamin II; TCII-R: Transcobalamin II receptor.
B12 and TCII by utilizing the b-propionamide site after modification to the b-monocarboxylic acid (the yellow shading seen in Figure 2) [5]. This method negates the possibility of oral uptake but demonstrated greater tumor targeting by reducing prevalent TCII-based uptake of the conjugates into healthy cells. This is a considerable breakthrough for B12-based cancer therapeutics because it demonstrates that one of the oft-stated problems of using B12 systems, namely non-specific cell uptake, can be addressed, at least towards cancer cell lines with de novo HC expression. Images obtained with single-photon-emission computed tomography (SPECT) agents targeting HC and TCII are shown in Figure 5.

2.3 General uses of B12 in drug delivery

2.3.1 Increased solubility
The solubility behavior of a drug is one of the key physico-chemical properties that can determine clinical efficacy and is often the cause of the failure of a drug to reach the market [24]. For orally administered drugs, the compound needs to be highly soluble or become immersed in the intestinal fluid to allow for absorption (assuming enterocyte passage is also feasible). For intravenously administered agents, sufficiently high solubility in the plasma is vital to diminish undesirable precipitation in the systematic circulation. In its native form, B12 is a highly soluble vitamin [10.2 mg/ml] that can be used to increase the solubility of compounds that would otherwise have little ability to be used as a drug.

2.3.2 Oral delivery
Few peptide/protein-based drugs have the ability to survive the gastrointestinal tract and/or cross the intestinal wall to make it to the bloodstream. Therefore, delivery of a pharmaceutically viable amount of peptide/protein is unachievable through simple oral dosing. At present, administration of these compounds is done through subcutaneous injections. This approach is time-consuming, painful, inconvenient, and can lead to allergic reactions near the sight of injection, as well as lower patient compliance. The improved ease of administration associated with the oral-enteric pathway provides an attractive means for the delivery of many pharmaceutical drugs because higher patient compliance is likely [25,26]. As a highly soluble, non-toxic vitamin with an extremely effective uptake pathway, B12 makes for an attractive vehicle for the oral delivery of drugs [26]. Although great strides have been taken in the field, the use of the uptake pathway is not without disadvantages, including a limited oral uptake capacity (1–2 µg of B12 per day [27]) and the inability to protect peptides/proteins from proteolytic degradation [28]. However, recent results, both experimental and theoretical, offer new strategies to overcome these issues, and these are discussed below.

2.3.3 Delivery of small molecules
For years, B12 has been utilized for the delivery of imaging agents and therapeutic drugs for the treatment and diagnosis of rapidly proliferating cells owing to the increased need for B12 in such cells (see Table 1). Early investigators demonstrated that cobalt radionuclides (57Co, 58Co, 60Co) could be used to label B12 for imaging [29-31]. The half-life of the radionuclides, however, required the dosage for humans to be too small for successful external images and the radioactive compounds were shown to accumulate in the liver, pancreas and kidneys, leading to organ damage [32]. In an attempt to overcome these issues, some extra radionuclides, including 111In (2.8 days; 0.245 MeV-γ) [32], 131I (8 h; 364 keV-β), 606 keV-β) [33] and 99mTc (6.02 h; 140 keV-γ) [34], have been conjugated to B12 for labeling purposes since these early days. Unfortunately, these experiments still resulted in
Vitamin B₁₂ in drug delivery

Figure 4. Cubilin$_{5-8}$-IF-B₁₂ structure. This structure, published in 2010 (see Protein Data Bank accession code 3KQ4), shows for the first time the interaction between holointrinsic factor (pink and green) and CUB$_{5-8}$ domains (blue with calcium as red spheres). The structure provides a road map to study B₁₂ conjugates bound to IF to explore the implications of B₁₂ modification on IF binding and subsequent receptor interactions. The structure also offers up potential mutagenesis sites on IF for use of IF as a delivery agent itself. Neither of these two latter points has been studied extensively so far.

Image modified with permission from [8].

For personal use only.

Undesired organ accumulation, limiting the amount of drug delivered to cancerous cells and demonstrating high background uptake in healthy cells [31–34].

In 2009, Siega et al. determined it was possible to deliver Gd³⁺ to cancer cells by conjugation to B₁₂. By using the metal chelating agent DTPA, a B₁₂–Gd³⁺ conjugate was constructed. The Gd³⁺ moiety conjugated to the 5′ ribose of the B₁₂ did not affect the binding of the B₁₂ transport proteins. Viability tests on human myelogenous leukemia K562 cells incubated with the conjugate showed a significant decrease in cell viability compared with those incubated with the B₁₂ parent compound and/or the Gd³⁺ ion alone [35]. Given that this system utilized conjugation at the 5′-ribose position, allowing TCII recognition, it can be postulated that this system, if tested in vivo, will give poor tumor specificity.

Nitric oxide (NO) has been demonstrated as a potential candidate for antitumor therapy owing to its ability to cause both necrosis and apoptosis [36]. In 2002, Bauer et al. set out to demonstrate that conjugation of NO to B₁₂ would result in a greater antineoplastic effect against malignant cells compared with normal cells. Nitrosylcobalamin (NO–B₁₂) is a B₁₂ derivative with NO as the β-axial ligand. Once the conjugate was in the cell NO was detached from B₁₂, causing inhibition of cellular metabolism and directly damaging DNA, leading to apoptosis. The antiproliferative efficacy of the conjugate was determined in 22 human tumor cell lines and two non-cancer cell lines. It was determined that tumor cell lines were more sensitive to the conjugate than the normal cell lines. The conjugate was determined to inhibit tumor growth in vitro and in vivo by activation of the extrinsic apoptotic pathway [36].

Colchicine is a highly toxic compound that has been used as a therapeutic agent for a wide variety of diseases since the sixteenth century [37]. In recent years, it has been investigated as a potential anticancer drug. It has a similar mode of action as taxanes but is far more water-soluble, making it more accessible to biological applications. Unfortunately, colchicine does not have the ability to distinguish healthy from cancerous cells and colchicine chemotherapy results in overwhelming systemic toxicity [37]. In 2004, Grissom and co-workers used B₁₂ to deliver colchicine to cancerous cells. To make the B₁₂ conjugate, the colchicine analogue 2-(4-acetylphenoxy)-N-acetophenocolchicine was synthesized. To make the B₁₂ conjugate, the colchicine analogue 2-(4-acetylphenoxy)-N-acetophenocolchicine was synthesized. Compared with the unmodified colchicine, the analogue retained activity with only a minimal loss of toxicity. Using an acid labile hydrazine linker, the analogue was then attached to the β-axial ligand of B₁₂. The linker allowed for a pH-dependent release of colchicine in acidic conditions. Once taken into the cell by means of the B₁₂ receptors, colchicine was released to interact with microtubules but showed a 10-fold decrease in toxicity [38]. As noted before, this system in vivo would again be predicted to suffer from TCII-based non-selective uptake/accumulation.

A progression of papers has recently appeared in the use of B₁₂ to deliver Pt drugs, focusing on cisplatin and analogues thereof. In this instance it was shown that B₁₂ can act as a ligand for cisplatin by the formation of a cyanide-bridged species between the β-axial ligand of B₁₂ and Pt [39]. Alberto and co-workers showed that the B₁₂–cisplatin conjugate retains a labile chloride ligand that can be exchanged with ligands such as guanine, allowing the conjugate to behave in a similar way to cisplatin [40]. In 2008, Alberto and co-workers continued their initial work with the cisplatin derivatives and synthesized several prodrugs around the {B₁₂-CN-Pt-R} moiety (see Figure 6) [22].

Using an in vitro adenosylation assay from Salmonella enterica, the group was able to show that Co III was reduced to Co II and the Pt II complex was released [22]. As the Pt II complexes were shown to be released, B₁₂-mediated delivery of Pt complexes showed promise. In 2010, the in vitro cytotoxicity of the {B₁₂-CN-Pt-R} conjugates was published [41]. The preliminary results suggested a lower activity (IC₅₀ between 8 and 88 µM) than for cisplatin alone. It is hypothesized that the limited B₁₂ uptake capacity, discussed earlier, is most probably to blame. Studies on the effects of the uptake capacity on concentration are underway [41].
2.3.4 Delivery of peptides/proteins

Table 2 lists the peptide/protein B12 conjugates reported so far. A comprehensive review of this area has been reported recently and the reader is referred there [42]. Instead, herein is discussed the use of molecular dynamics in the development and understanding of B12-based drug delivery systems using a B12-insulin conjugate system (see Table 2) as an example of the power of this approach.

3. Molecular modeling considerations

Molecular modeling serves invaluably in both predictive and interpretive capacities in modern drug design. The field is limited in application only by either the availability of proper theoretical descriptions of biologically relevant molecules (largely in the form of molecular force field parameters and topologies, limitations addressable by parameter generation and validation) or the absence of experimental data on which to base proper analyses. For molecular dynamics studies, the limitation is largely in the form of absent crystallographic or solution-phase structural information against which to identify structural candidates and test new designs. As applied to the study of bioconjugates, the atom-level description of molecular phenomena afforded synthetic chemists by even modest computational efforts reveals a great deal about the binding environment of a molecule. The potential difficulties of designs based on the placements of steric bulk (that prohibit binding of the native molecule) or physical inaccessibility of targeted functionalization (that changes the binding of the native molecule) can be visualized. Insights into possible molecular designs not readily obvious from either crystallography or macroscale sample characterization without synthetic efforts that survey possible candidate positions can also be highlighted.

As a complement to many of the experimental studies of functionalization positions [43,44], molecular modeling studies reveal the extent to which the modification potential of the B12 framework itself is greatly limited for new bioconjugate designs by the nature of its binding within its transport proteins. Using the binding of B12 to TCII (the complex for which the best structural information is available [1,3,10]) as a basis for all subsequent bioconjugate design work (as a disruption at any step in the transport pathway for B12 is probably deleterious for any bioconjugate design), several design considerations were revealed. These molecular modeling studies provide both a reinforcement of previous experimental work and insights into approaches for potential new B12-based delivery designs based on functionalization positions deemed accessible, themselves based on the geometry of the B12–TCII binding interaction.

Figure 5. SPECT/CT scans of mice 24 h after intravenous injection of B12-99mTc conjugates designed to be non-binders to TCII (HC binder) and binders to TCII. A, C. HC targeting compound. B, D. TCII binding compound. Note increased tumor specificity obtained with HC targeting. Color scale relates activity.

Reproduced with permission from [5].
HC: Haptocorrin; IF: TCII: Transcobalamin II.
<table>
<thead>
<tr>
<th>Molecule</th>
<th>Size</th>
<th>Conjugation site</th>
<th>Linker (coupling agent)</th>
<th>Use</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small molecules</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$^{99m}$Tc-DTPA</td>
<td>452.0</td>
<td>$b$-acid</td>
<td>1,4-diaminobutane (EDAC)</td>
<td>Imaging of TCII receptors</td>
<td>1997</td>
<td>[31]</td>
</tr>
<tr>
<td>$^{111}$In-DTPA</td>
<td>467.7</td>
<td>$b$-acid</td>
<td>1,4-diaminobutane (EDAC)</td>
<td>Imaging of TCII receptors</td>
<td>1997</td>
<td>[31]</td>
</tr>
<tr>
<td>b-nido-carborane</td>
<td>144.2</td>
<td>$b$-acid</td>
<td>Diaminobutane (HOBt/EDAC)</td>
<td>Antitumor therapy</td>
<td>2000</td>
<td>[46]</td>
</tr>
<tr>
<td>d-nido-carborane</td>
<td>144.2</td>
<td>$d$-acid</td>
<td>Diaminobutane (HOBt/EDAC)</td>
<td>Antitumor therapy</td>
<td>2000</td>
<td>[46]</td>
</tr>
<tr>
<td>bis-nido-carborane</td>
<td>288.5</td>
<td>$b$- and $d$-acid</td>
<td>Diaminobutane (HOBt/EDAC)</td>
<td>Antitumor therapy</td>
<td>2000</td>
<td>[46]</td>
</tr>
<tr>
<td>NO</td>
<td>30.0</td>
<td>Cobalt atom</td>
<td>Direct</td>
<td>Antitumor therapy</td>
<td>2002</td>
<td>[36]</td>
</tr>
<tr>
<td>Cy5 dye</td>
<td>792.0</td>
<td>Ribose-5'-OH</td>
<td>1,6-diaminoheptane (CDI)</td>
<td>Imaging</td>
<td>2003</td>
<td>[47]</td>
</tr>
<tr>
<td>Colchicine</td>
<td>399.4</td>
<td>Cobalt atom</td>
<td>4-chlorobutyric acid chloride (EDAC/NH5)</td>
<td>Delivery</td>
<td>2004</td>
<td>[38]</td>
</tr>
<tr>
<td>$^{99m}$TcO$_{4}^-$</td>
<td>236.0</td>
<td>N/A</td>
<td>N/V-dimethylglycine (NA)</td>
<td>Radiodiagnosis</td>
<td>2004</td>
<td>[48]</td>
</tr>
<tr>
<td>[Re(OH$_2$)$(CO)_3$]</td>
<td>288</td>
<td>Cyano ligand</td>
<td>N/V-dimethylglycine (NA)</td>
<td>Radiodiagnosis</td>
<td>2004</td>
<td>[48]</td>
</tr>
<tr>
<td>Cisplatin</td>
<td>300.0</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2005</td>
<td>[39]</td>
</tr>
<tr>
<td>Cisplatin-2'-deoxyguanosine</td>
<td>585.3</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>Cisplatin-methylguanosine</td>
<td>598.2</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>trans-PtCl($NH$_2$)$_2$</td>
<td>263.7</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>trans-PtCl$_2$(NH$_3$)</td>
<td>281.8</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>Cys-PtCl$_2$(NH$_3$)</td>
<td>281.8</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>PtCl$_2$</td>
<td>300.4</td>
<td>Cobalt atom</td>
<td>Cytosine ligand (NA)</td>
<td>Linker for biomolecules</td>
<td>2006</td>
<td>[40]</td>
</tr>
<tr>
<td>$^{99m}$TcO$_{4}^-$</td>
<td>236.0</td>
<td>$b$-acid</td>
<td>N/V-dimethylglycine (NA)</td>
<td>Radiodiagnosis</td>
<td>2004</td>
<td>[48]</td>
</tr>
<tr>
<td>VO$_2$(OH)</td>
<td>118.0</td>
<td>Cobalt atom</td>
<td>Propyl-PAMA-OEt (EDAC+TBTU)</td>
<td>Imaging</td>
<td>2008</td>
<td>[5]</td>
</tr>
<tr>
<td>VO$_2$</td>
<td>82.0</td>
<td>Cobalt atom</td>
<td>Ethyl-PAMA-OEt (EDAC or TBTU)</td>
<td>Imaging</td>
<td>2008</td>
<td>[5]</td>
</tr>
<tr>
<td>Fluorescence</td>
<td>475.2</td>
<td>Ribose-5'-OH</td>
<td>Butyl-PAMA-OEt (EDAC or TBTU)</td>
<td>Imaging</td>
<td>2008</td>
<td>[5]</td>
</tr>
<tr>
<td>Khodamine 6G</td>
<td>479.0</td>
<td>Ribose-5'-OH</td>
<td>Pentyl-PAMA-OEt (EDAC or TBTU)</td>
<td>Imaging</td>
<td>2008</td>
<td>[5]</td>
</tr>
<tr>
<td>Gd$_3^+$</td>
<td>157.3</td>
<td>Ribose-5'-OH</td>
<td>Hexyl-PAMA-OEt (EDAC or TBTU)</td>
<td>Imaging</td>
<td>2008</td>
<td>[5]</td>
</tr>
<tr>
<td>Rhodamine isothiocyanate</td>
<td>536.08</td>
<td>NA</td>
<td>3-hydroxy-2-methyl-1-propyl-L-pyridin-4-one(NA)</td>
<td>Imaging</td>
<td>2008</td>
<td>[49]</td>
</tr>
</tbody>
</table>

Any design of B₁₂ bioconjugates must ultimately consider all of the steps in the B₁₂ transport pathway. In these considerations, it is known that B₁₂ binding to its transport proteins (HC, IF and TCII) must be conserved. The availability of partial or complete molecular geometries for TCII[10] and IF [9] reveals that these two proteins are largely similar in geometry and binding mode, and indicates that HC is likely to bind B₁₂ in a similar manner. As the initial computational study of B₁₂-insulin [44] bioconjugate began with preliminary experimental results that indicated successful transport (as determined by cellular glucose uptake), the molecular modeling study focused on TCII. As part of any subsequent design based on B₁₂ delivery, the characterization of the binding mode of B₁₂ and the flexibility of the α/β domain ‘clamshell’ mechanism of binding by TCII (see Figure 7) produce an elegant picture of one set of readily accessible ‘best-practice’ approaches for bioconjugate design. This is a very straightforward model of how adhering to this set of approaches is likely to produce viable candidates for a wide variety of delivery targets.

In the previously reported molecular dynamics studies of TCII with B₁₂ [15], B₁₂-insulin [44] and a B₁₂ bioconjugate modified with a short peptide tether [14], the simulations revealed the extent to which the B₁₂ is itself the hydrogen-bonding glue that binds the TCII into the encapsulation structure observed in diffraction studies (Figure 7). This work indicated that the interaction between side chains at the surface-surface interface between the α- and β-domains is not a major contributor to the stability of the complex. It also indicated that TCII in the absence of B₁₂ probably consists of two, poorly interacting fragments linked by a short tether (obviously the design best suited to binding and protecting a large biomolecule that cannot easily migrate into the binding pocket of a less flexible or more inherently structured protein). As a tool for directing design, the result of the molecular dynamics simulation showed that modifications to the B₁₂ are best performed at the positions on the B₁₂ molecule at the α/β domain interface. Appropriate modifications at this interfacial region are not expected to interfere with the individual B₁₂-α domain and B₁₂-β domain interactions that produce the encapsulated complex and, further, are made viable because the ‘clamshell’ design of this complex results in a α/β breathing motion in the MD simulations that makes this interfacial region accessible under ambient conditions and amenable to accommodating small tethering fragments to couple B₁₂ to some arbitrary molecular beyond the solvent-accessible surface of TCII (see Figure 8).

The MD simulations provided atom-level explanations for all of the bioconjugate design work on B₁₂ performed so far. This study specifically shows why certain chemical modifications produce viable candidates and some modifications produce B₁₂ bioconjugates that, by their disruption of the B₁₂-transport protein binding interaction, do not. Modification at the cobalt (Co) is shown to be a reasonable approach for bioconjugate design because the Co coordination site lies at the α/β interface and small changes to side chains on the α-domain reduce steric congestion for coordinated ligands. This is obvious for small molecules, given the known binding of B₁₂-CN and B₁₂-CH₃. Whereas the cobalt is more embedded than the ribose 5′-hydroxyl position used as the point of conjugation in the B₁₂-insulin bioconjugate, the use of a small, flexible tether that is linked to one molecule and is Co-coordinated should serve as a reasonable mechanism for new Co-tether designs, a result supported by previous experimental work [44].

The MD simulation work also predicts that, owing to the mechanism of encapsulation and the solvent accessibility of
Table 2. B<sub>12</sub>-protein/peptide bioconjugates.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>Size</th>
<th>Conjugation site</th>
<th>Linker (coupling agent)&lt;sup&gt;§&lt;/sup&gt;</th>
<th>Use</th>
<th>Year</th>
<th>Ref.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Directly conjugated and encapsulate insulin delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>BS Albumin</td>
<td>66.0</td>
<td>Phosphate</td>
<td>Phospho-amine (EDAC)</td>
<td>NA</td>
<td>1971</td>
<td>[52]</td>
</tr>
<tr>
<td>YG-globulin</td>
<td>150.0</td>
<td>Phosphate</td>
<td>Phospho-amine (EDAC)</td>
<td>NA</td>
<td>1971</td>
<td>[52]</td>
</tr>
<tr>
<td>HS albumin</td>
<td>66.0</td>
<td>e-acid</td>
<td>GABA (EDAC)</td>
<td>antibody response</td>
<td>1979</td>
<td>[53]</td>
</tr>
<tr>
<td>IFN-con</td>
<td>22.0</td>
<td>ribose-5'-OH</td>
<td>glutaroyl (CDI)</td>
<td>24 – 28% activity&lt;sup&gt;*&lt;/sup&gt;</td>
<td>1994</td>
<td>[54]</td>
</tr>
<tr>
<td>G-CSF</td>
<td>19.6</td>
<td>e-acid</td>
<td>disulfide (SPDP)</td>
<td>61 – 66% activity&lt;sup&gt;†&lt;/sup&gt;</td>
<td>1995</td>
<td>[55]</td>
</tr>
<tr>
<td>EPO</td>
<td>34.0</td>
<td>e-acid</td>
<td>Amide (EDAC)</td>
<td>ND-100% activity&lt;sup&gt;‡&lt;/sup&gt;</td>
<td>1995</td>
<td>[55]</td>
</tr>
<tr>
<td>ANTIDE-1</td>
<td>1.6</td>
<td>e-acid</td>
<td>EGS (EDAC)</td>
<td>ND</td>
<td>1995</td>
<td>[56]</td>
</tr>
<tr>
<td>ANTIDE-3</td>
<td>1.6</td>
<td>e-acid</td>
<td>EGS (EDAC)</td>
<td>ND</td>
<td>1995</td>
<td>[56]</td>
</tr>
<tr>
<td>LHRH</td>
<td>1.2</td>
<td>e-acid</td>
<td>Amide (DC/NHS)</td>
<td>45% absorbed</td>
<td>2000</td>
<td>[57]</td>
</tr>
<tr>
<td>DP3</td>
<td>0.9</td>
<td>e-acid</td>
<td>Amide (EDAC) hexyl (EDAC)</td>
<td>23% absorbed 42% absorbed</td>
<td>2000</td>
<td>[57]</td>
</tr>
<tr>
<td>Insulin</td>
<td>5.7</td>
<td>ribose-5'-OH</td>
<td>amide (CDI, CDT)</td>
<td>26% drop in glucose</td>
<td>2007</td>
<td>[20]</td>
</tr>
<tr>
<td>Encapsulated insulin delivery</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>B&lt;sub&gt;12&lt;/sub&gt; coated-dextran nanoparticles</td>
<td>5.7</td>
<td>ribose-5'-OH</td>
<td>amide (CDI)</td>
<td>70 – 75% drop in plasma glucose</td>
<td>2007</td>
<td>[58]</td>
</tr>
</tbody>
</table>

<sup>*</sup>compared with native IFN-con.

<sup>†</sup>compared with unconjugated G-CSF and EPO.

<sup>§</sup>linker chosen on basis of greatest yield and/or activity. ND = not determined.

BS, HS: Bovine and human serum; IFN-con: Consensus interferon; G-CSF: Granulocyte colony stimulating factor; EPO: Erythropoietin; ANTIDE: N-Ac-D-Nal(2)D, D-Phe (pCl), D-Pal(3),ser, Lys (Nic), D-Lys(Nic), Leu, Tyr(Ph), Pro, D-Ala-NH2; LHRH: Luteinizing hormone-releasing hormone; DP3: Octapeptide (Glu-Ala-Ser-Ala-Ser-Tyr-Ser-Ala); GABA: y-aminobutyric acid; EDAC: l-ethyl-3-(3-dimethylaminopropyl)carbodiimide; CDI: 1,1’carbonyl-diimidazole; SPDP: N-succinimidyl 3-(2-pyridyldithio)-propionate; SMPT: 4-[(Succinimidyl)oxy]-carbonyl]-a-methyl-a-2-pyridyldithiotoluene; NHS: N-hydroxysuccinimide; DCC: N,N-dicyclohexylcarbodiimide; CDT: U’-carbonyl-di-(1,2,4-triazole); EGS: Ethylene glycol bis(succinimidylsuccinate); EDAC: l-ethyl-3-(3-dimethylamino)propyl)carbodiimide.
several regions of B12, the cobalt and ribose 5′-hydroxyl positions need not be mutually exclusive as targets for bioconjugate design. Instead, with the accessibility of both positions attributed to the breathing motion of the α/β groups constrained largely by their strong binding to the B12 molecule, the possibility exists for three-component bioconjugates. This result was hinted at, but was not obvious from, the static model of the B12–TCII binding interaction in the crystallographic study. The success of such designs is dependent on the modes of delivery and the differences in where and when the degradation of the combined product occurs. This route does offer a promising path for new work in B12 delivery approaches that are strongly supported from the synergy of experimental precedent and supportive computational work.

With the availability of force field parameters and the reported success of B12 MD simulations in several instances, one of the few issues in B12 bioconjugate design hampering continued and more thorough progress is the lack of structures for proteins within the B12 transport pathway. Although the success of the B12-insulin bioconjugate indicates that, even in the presence of a large tethered molecule, the solvent-accessible binding positions on the

---

**Figure 7. The binding of B12 within TCII.** Substructures (domains) are differentiated by color. The α- and β-domains at the B12 binding position interact only through hydrogen bonding (and by way of a 10-residue covalent linkage), while numerous interactions between both domains and the B12 have been characterized experimentally and through previous molecular dynamics simulations. TCII structure based on [10].

**Figure 8. The time-averaged structure of B12-TCII from a 50 ns MD simulation.** The accessibility of the B12 for chemical functionalization but potentially limited B12 binding inhibition is shown by way of solvent-accessible surface representations at both the cobalt (left) and ribose (right) positions.
B<sub>12</sub> are conserved in the members of the transport protein family, greater subtlety in the modes of chemical modification and targeted B<sub>12</sub> design will come most easily through MD simulations that follow a B<sub>12</sub> bioconjugate design through each step in the pathway, for which a modeling protocol is only as complete as the experimental characterization of all components. Within even the narrow selection of B<sub>12</sub> modification positions hinted at by the solvent-accessible crystal structure maps and MD simulations of the native TCII-B<sub>12</sub> complex, a wealth of bioconjugate design can be envisioned. Completeness in all-computational approaches to B<sub>12</sub> bioconjugate design simply awaits adequate models of HC and IF, the members of the transport pathway either partially characterized (IF) or still unavailable (HC).

4. Expert opinion

B<sub>12</sub> and the B<sub>12</sub> uptake pathway play a fundamental role in biology and have been explored extensively from a structural and mechanistic viewpoint. The unique structure of B<sub>12</sub> and the exquisite relationship with its binding proteins have driven researchers for several decades and continue to this day to offer up exciting new questions and new avenues of research. One of these avenues is the use of B<sub>12</sub> in drug delivery, from small molecule in vivo imaging agents to the oral delivery of large proteins. Inherent restrictions of the B<sub>12</sub> pathway, namely dose limitations and the demand across cells, healthy or otherwise, for the vitamin, have so far prevented the realization of a B<sub>12</sub> bioconjugate pharmaceutical reaching the market. Despite this, several groups are working to address the challenges of the field, with recent successes including HC-specific targeting overcoming TCII background and the use of polymeric carriers attached to B<sub>12</sub> to increase drug payload.

The future of the field then lies in expanding on these successes. The authors envisage building B<sub>12</sub> bioconjugates with multiple payloads, using multiple sites on the B<sub>12</sub> molecule simultaneously. The use of B<sub>12</sub> to deliver protein vaccine components also has great potential. By targeting the immune response, potentially low doses need be delivered that require concentrations within the boundaries of the B<sub>12</sub> uptake pathway capacity. The use of the B<sub>12</sub> uptake proteins themselves has also not been explored. Coating a viral particle that requires an enteric IgA response with HC, for example, may offer gastrointestinal protection and ultimately provide for an oral vaccine for such a virus. The very investigation of the role of HC in blood serum and its structure will also be a highlight of the coming years in the B<sub>12</sub> field. The take home message of this review is that this is an exciting time for B<sub>12</sub>-based drug development, with many of the barriers thought to be road blocks to successful clinical trials being overcome.

Declaration of interest

The authors declare no conflict of interest and have received no payment in preparation of this manuscript.
Bibliography

Papers of special note have been highlighted as either of interest (*) or of considerable interest (**) to readers.


15. Allis D, Fairchild T, Doyle R. The binding of vitamin B12 to transcobalamin (II); structural considerations for biocjugate design – a molecular dynamics study. Mol BioSyst 2010;6:1611-18


30. Wooley K, Collins D, Morton K. Transcobalamin II receptor imaging via radiolabeled...
Vitamin B12 in drug delivery


47. McGreavy J, Cannon M, Grissom C. Minimally invasive lymphatic mapping through the cobalamin pathway in the rat intestine. Pharmaceutical research 2007;103:118-34


Affiliation
Susan M Clardy1, Damian G Allis1, Timothy J Fairchild2 & Robert P Doyle1†
1Author for correspondence
2Syracuse University, Department of Chemistry, Syracuse, NY 13244-4100, USA
E-mail: rpdoyle@syr.edu
2Murdoch University, Chiropractic and Sports Science, 90 South Street, Perth, WA 6150, Australia
Examining the effects of vitamin B$_{12}$ conjugation on the biological activity of insulin: a molecular dynamic and in vivo oral uptake investigation†

Susan Clardy-James,*a Damian G. Allis,*a Timothy J. Fairchild*abh and Robert P. Doyle*aa

Received 10th February 2012, Accepted 20th March 2012
DOI: 10.1039/c2md20040f

The practical use of the vitamin B$_{12}$ uptake pathway to orally deliver peptides and proteins is much debated. To understand the full potential of the pathway however, a deeper understanding of the impact B$_{12}$ conjugation has on peptides and proteins is needed. We previously reported an orally active B$_{12}$ based insulin conjugate attached at LysB29 with hypoglycemic properties in STZ diabetic rats. We are exploring an alternative attachment for B$_{12}$ on insulin in an attempt to determine the effect B$_{12}$ has on the protein biological activity. We describe herein the synthesis, characterization, and purification of a new B$_{12}$-insulin conjugate, which is attached between the B$_{12}$ ribose hydroxyl group and insulin on the protein biological activity. We describe herein the synthesis, characterization, and purification of a new B$_{12}$-insulin conjugate, which is attached between the B$_{12}$ ribose hydroxyl group and insulin PheB1. The hypoglycemic properties resulting from oral administration (gavage) of such a conjugate in STZ diabetic rats was similar to that noted in a conjugate covalently linked at insulin LysB29, demonstrating the availability of both positions on insulin for B$_{12}$ attachment. A possible rationale for this result is put forward from MD simulations. We also conclude that there is a dose dependent response that can be observed for B$_{12}$-insulin conjugates, with doses of conjugate greater than $10^{-9}$ M necessary to observe even low levels of glucose drop.

Introduction

Oral delivery of therapeutic peptides and proteins at clinically significant and controllable levels is a major challenge in current medicinal chemistry. The oral delivery route aims to open up new areas of peptide/protein therapeutics associated with the removal for a need for injection and the concomitant promise of greater patient quality of life and regimen adherence. The major problems facing oral delivery of peptides/proteins is hydrolysis/proteolysis in the gastrointestinal tract (GIT) and an inefficient uptake mechanism for peptides/proteins from the GIT. We are interested in the use of the vitamin B$_{12}$ (B$_{12}$) dietary uptake pathway to address these hurdles and recently demonstrated clinically relevant plasma concentration of a peptide hormone (hPYY3-36) achieved through the B$_{12}$ system in a rat model.

B$_{12}$ is vital for the survival of virtually all living cells. B$_{12}$ is synthesized naturally by certain bacteria and all other species must acquire the vitamin exogenously through diet. Mammals have a series of transport proteins in the GIT for the absorption and cellular uptake of B$_{12}$. The first transport protein, haptocorrin (HC), binds to B$_{12}$ upon release from food, either in the mouth or in the stomach and transfers it into the duodenum.

Once in the intestine, proteases cause the partial degradation of HC, while a rise in pH promotes the interaction of B$_{12}$ with a second transport protein, intrinsic factor (IF). This results in the transfer of ~80% of bound B$_{12}$ to IF. The IF–B$_{12}$ complex crosses the ileum via the cubam complex. IF digestion occurs during transcellular passage allowing release of B$_{12}$ to the serum transport protein transcobalamin II (TCII). TCII is responsible for cellular entry through the TCIIR/CD320 receptor and is postulated to facilitate blood–brain barrier transport.

Previously, we reported on the development and characterization of an orally active B$_{12}$-insulin conjugate with hypoglycemic properties in the streptozotocin (STZ)-induced diabetic rat model. Results of this work stimulated further study into the implications of the B$_{12}$ conjugation on insulin’s ability to bind to its receptor (INSR/CD220) using immune-electron microscopy and molecular dynamics simulation. What is clear from the results obtained to date is that (i) B$_{12}$ can aid in delivering insulin to rat blood serum at levels significantly greater than the same administered concentration of ‘free’ insulin, and (ii) the B$_{12}$--insulin conjugate can still dock to CD220 and thus trigger glucose uptake.

The use of the B$_{12}$ pathway to deliver a clinically relevant bolus of insulin is not possible however, given both the limited uptake capacity of the pathway and the relatively slow uptake time for serum delivery (~3 to 4 hours) and plasma maximum (~7 hours). The possibility of developing an oral, long acting basal therapy is more feasible, where a low level of long acting insulin is required and an initial uptake delay is not critical.

† Electronic supplementary information (ESI) available: Experimental details including in vivo studies, molecular dynamics details, LC and mass spectra. See DOI: 10.1039/c2md20040f

*aDepartment of Chemistry, Syracuse University, Syracuse, NY, USA. E-mail: rpdoyle@syr.edu; Tel: +1 315 443 3584
bSchool of Chiropractic and Sports Sciences, Murdoch University, Perth, Australia. E-mail: T.Fairchild@murdoch.edu.au

This journal is © The Royal Society of Chemistry 2012
The earlier results with insulin\textsuperscript{12} and PYY\textsuperscript{4} have been highly encouraging and we have become focused on investigating and improving in vivo residency of B\textsubscript{12}–insulin conjugates. Before such an area can be fully explored however, knowledge of (i) the effects of B\textsubscript{12} conjugation on the two main sites suitable for such on insulin (\textit{vide infra}) and (ii) dose–response of B\textsubscript{12}–insulin conjugates need to be addressed. These important factors are discussed in this report.

Insulin has two primary sites available for conjugation (the N\textsuperscript{\alpha}-amino group of phenylalanine B1 (PheB1) and the N\textsuperscript{\alpha}-amino group of lysine B29 (LysB29)), which allow biological function to be maintained.\textsuperscript{14} It is not completely understood which position is more suitable for modification.\textsuperscript{16} Significant differences between both positions upon B\textsubscript{12} conjugation would be of great importance given the limited B\textsubscript{12} uptake capacity and the need therefore to optimize such a conjugate. It is believed that conjugation at the LysB29 position results in a reduction in biological activity due to the close proximity to the insulin receptor binding site.\textsuperscript{15–18} Therefore, moving B\textsubscript{12} further away from the binding site could result in an increase in biological activity.\textsuperscript{15}

We describe herein the synthesis, characterization, and purification of a new B\textsubscript{12}–insulin conjugate (1), which is attached at insulin PheB1 (Fig. 1). We hypothesized that 1 may result in an increase in insulin’s biological activity, while maintaining the B\textsubscript{12} uptake capabilities due to the position of B\textsubscript{12} attachment in relation to the insulin receptor binding region. Our previously investigated B\textsubscript{12}–insulin conjugate (B\textsubscript{12} bound to insulin LysB29, 2)\textsuperscript{12} was used for comparison purposes (along with free insulin controls). Both 1 and 2 were coupled at the B\textsubscript{12}–ribose hydroxyl group through a carbamate linker. Hypoglycemic properties of 1 and 2 were evaluated \textit{in vivo} in STZ induced diabetic rats. Molecular dynamics simulation studies were used to help rationalize the results. We also report dose dependence in STZ diabetic rats for a B\textsubscript{12}–insulin conjugate.

Results and discussion

Chemistry

The synthetic procedures for 1 and 2 were identical unless otherwise noted. Reaction of B\textsubscript{12} with 1,1\textsuperscript{-carbonyl-di-(1,2,4-triazole)} furnished an activated ester at the 5\textsuperscript{\textprime}-hydroxyl group. The activated B\textsubscript{12} was then added to a slowly stirring solution of bovine insulin. The site directed B\textsubscript{12} conjugation utilizes the established difference in reactivity of the amino groups (LysB29 > GlyA1 \textsuperscript{\geq} PheB1) of insulin at pH 9.5.\textsuperscript{17} For 1, the more reactive LysB29 and GlyA1 were initially protected with tert-butylxycarbonyl (BOC) to allow for specific conjugation of B\textsubscript{12} at the unprotected PheB1 residue as previously described by Kim et al.\textsuperscript{17} 2 was synthesized without need for a protecting group due to the high level of reactivity of LysB29 under the reaction conditions chosen.

Both reactions were extracted by ether precipitation and centrifugation. The products were isolated by analytical reverse-phase HPLC. Purification proved challenging due to the similarity in hydrophobicity of unconjugated insulin and B\textsubscript{12} conjugated insulin. For 1, the BOC protection was adequate to produce a separation between B\textsubscript{12} conjugated and free insulin (see Fig. S2†). For 2, addition of an excess of FMOC-OSu was used to aid in purification. The FMOC moiety increases the retention time of the unconjugated (hence, unprotected) insulin allowing for separation (Fig. 2). Following HPLC purification, the protecting groups (BOC or FMOC) were removed with TFA (1) or piperidine (2) and dialysed against 5 L of H\textsubscript{2}O overnight in 7000 MWCO dialysis tubing.

The exact site of conjugation was verified by matrix assisted laser desorption ionization time of flight (MALDI-ToF) mass spectrometry (MS) analysis of reduced and digested fragments of 1 and 2. The conjugates were first treated with dithiothreitol to reduce the interchain disulfide bonds linking the A- and B-chain. MS analysis of the reduced conjugates displayed m/z values of 2339.0 and 4755.4 corresponding to unmodified A-chain and B\textsubscript{12}–ribose hydroxyl group (after B\textsubscript{12} attachment) bond to B\textsubscript{12}–ribose hydroxyl group (after B\textsubscript{12} attachment) bond, respectively (Fig. 3). To verify the specific B-chain site, reduced conjugates were treated with trypsin, which cleaves peptides at the carbonyl side of arginine or lysine, except when followed by a proline. Reduced and digested 1 displayed a m/z of 3844.9 corresponding to B1–B22 plus B\textsubscript{12} indicating conjugation at the B1 site (Fig. 3). While, reduced and digested 2 displayed a m/z of 2286.0 corresponding to B23–B30 plus B\textsubscript{12}, indicating attachment at the B29 site (see Fig. S2†).

![Fig 1 Molecular dynamic stimulation snapshot of TCII (red = A-chain, dark blue = B-chain, green = unstructured linkage) bound to 1 [B12 (yellow with light blue Co(m)] conjugated to the first amino acid (PheB1) on the insulin B-chain (also yellow).](image1)

![Fig 2 The RP-HPLC spectra of an FMOC protected insulin plus B12 reaction monitored at 360 nm (dash) indicating B12 is present and 254 nm (solid) indicating insulin is present.](image2)
In vivo testing and molecular dynamics (MD)

Comparison between both conjugates was assessed by measuring blood glucose concentration following oral delivery of 1 and 2 (Fig. 4). Area under the curve (AUC) transformation showed a significant difference (p = 0.03) between 1 and free insulin control, consistent with that observed in our earlier report.12 Of interest here is that no significant difference in glucose drop (p = 0.53) was observed between 1 and 2. Both produced similar and greater drops over free insulin of the same orally administered concentration. This data differs from previous literature15–18 that demonstrated conjugation position alters the biological activity of insulin and was not what we expected. We believe our findings may be explained by components of the B12 uptake pathway, making this result not in conflict with these earlier findings but rather suggesting a unique result tied to the use of B12 in insulin conjugation.

To help explain this result we conducted MD simulations on the binding of 1 with the B12 transport protein TCII. Previous MD simulations of 2 interacting with TCII13,14 indicated that the lysine side chain alone provided an adequate tether length to place the unstructured insulin B-chain region (residues B20–B30) beyond the steric congestion of the B12 bound TCII structure. Furthermore, the B20–B30 tail provides considerable length for placing the most structured region of the insulin molecule (the complete A-chain and B7–B19 B-chain region) far from the α- and β-Domains of TCII, thereby not impacting the binding geometry of the B12 in its internal TCII pocket or the interaction of insulin with the insulin receptor.13

Formation of 1 does place the B12 several residues from this structured insulin region, but the effective separation of the structured insulin region from the TCII binding area is not obvious from the components themselves given (i) the inability to predict insulin unfolding of the B1–B7 region (the residues prior to the first disulfide linkage to the A-chain) to accommodate the B12 binding within its pocket and (ii) the inability to predict if the phenylalanine side chain itself would serve as a source of steric congestion to disrupt B12 binding in its pocket (in 2, the lysine side chain is itself the tether region between B12 and insulin). Given the structural similarities between the B12 transport proteins TCII and IF (and presumed structural similarities between both and haptocorrin),19 the TCII simulations serve generally as a guide for predicting how the entire family of B12 transport proteins might behave upon B12 conjugation, at least when tethered at the B12 ribose hydroxyl group. Insights into the nature of the B12–insulin/TCII interaction upon PheB1 conjugation come from a 15 ns simulation of 1 interacting with the TCII complex. A representative of the geometry of 1 during these simulations is shown and labeled in Fig. 5.

The unstructured B20–B30 region of the insulin B-chain was removed from such calculations. This unstructured loop points away from the TCII when the B12 is linked at the B1 position.
requiring a considerable extension of the solvent box, which adds to total atom count without providing useful information about the TCII interactions with I. The representative geometry in Fig. 6 shows that the PheB1–CysB7 (B1–B7) region of the B-chain is unfolded from its unmodified, native α-helix geometry. In the snapshot shown, the disruption of α-helical structure extends slightly past the disulfide linkage (CysB7). This unfolding of the α-helix enables several local hydrogen-bonding interactions to form transiently throughout the 15 ns simulation. One of the more persistent interactions is the newly formed hydrogen bond between B4 and B13 (Fig. 5). The result of this unfolding is that the B1–B7 residues become the long tether that separates the insulin core from TCII upon binding of the B12–insulin conjugate. This tether originates from an otherwise stable α-helix and is slightly longer than half of the length of the B20–B30 tether in the previous study. The retention of the B12 within its binding pocket with no observable deformation of this pocket indicates that the B1–B7 tether is long enough to separate the two larger fragments. As an MD simulation of the B12–insulin molecule already bound in TCII, this simulation indicates only that, once bound, there are no expected steric issues preventing retention of the B12 in its binding pocket. Given the clamsHELL-like binding mechanism proposed for the B12–TCII complex, it is presumed that the B12 can bind with at least one domain of TCII first with the insulin bound, after which the insulin B1–B7 unfolding is a subsequent step in the enclosure of the complete B12 by TCII. The 5.0 ns timesteps shown in Fig. 6 (shown without the TCII for clarity) reveals the extent to which the PheB1–CysB7 is completely uncoiled upon stable B12 binding in the TCII pocket. It is this uncoiling induced upon binding by B12 uptake proteins that possibly explains why both sites on insulin ultimately equate in function and why such differences were not observed before.

Fig. 6 5.0 ns snapshots of the I along the B-chain α-helix axis (V1) and perpendicular to the α-helix axis (V2), with both views showing the gradual uncoiling of the B-chain (PheB1 through LeuB11) over the course of the simulation. Light blue = insulin A-chain; yellow = insulin B-chain and B12. The TCII protein is not shown for clarity.

Fig. 7 5.0 ns snapshots of the PheB1-linked B12–insulin/TCII complex. Light blue = insulin A-chain; yellow = insulin B-chain and CN–Cbl; red = TCII α-Domain; green = TCII α/β-Domain linker; dark blue = TCII β-Domain.

Fig. 8 Backbone RMSD plots for the PheB1-linked B12–insulin/TCII complex. The removal of the α-Domain α-helix-linking unstructured loops in the TCII α-Domain RMSD analysis (light red) reveals the persistence of the α-Domain B12 binding region compared to the overall α-Domain RMSD analysis (dark red).
However, the administered amount of conjugate (100 nM) was in gates with doses of conjugate greater than 10^{-9} M acting with the intestinal insulin receptor, as suggested by depending on B12 conjugation site, upon B12 binding protein molecule at either the N- or C-terminal B strand insulin regions, possible rationale for this result is put forward from MD simulation of a new B12 insulin conjugate attached at the insulin LysB29 residue, which suggests there is forced un-coiling of the insulin structure. This hypothesis is consistent with the observed glucose drop along with observed dose-dependent response. Under this latter scenario, the blocking of uptake by a large excess of B12 (>10^{-5} M) as observed in our earlier work would still be attributed to loss of IF binding, but rather than loss of cubam mediated uptake, the effect of such loss would be reduced gastric protection.

If such a route to oral insulin through the insulin receptor is made available by B12 conjugation it would be a paradigm shift in our view of exploiting the B12 uptake pathway for oral protein delivery.

Acknowledgements

R.P.D wishes to acknowledge funding from the Office of the Vice President of Research at Syracuse University. In vivo studies were funded by a grant from the McCluster Charitable Foundation to Murdoch University (T.J.F.), Perth, Australia.

Notes and references

1. R. Eldor, M. Kidron and E. Arbit, Diabetes, Obes. Metab., 2010, 12, 179.
Site-Selective Oxidation of Vitamin B$_{12}$ Using 2-Iodoxybenzoic Acid

Susan Clardy-James, Jaime L. Bernstein, Deborah J. Kerwood, Robert P. Doyle*
Department of Chemistry, Center for Science and Technology, Syracuse University, Syracuse, NY 13244-4100, USA
Fax +1(315)4434070; E-mail: rpdoyle@syr.edu
Received: 18.06.2012; Accepted after revision: 31.07.2012

Abstract: Reaction of vitamin B$_{12}$ (B$_{12}$) with excess 2-iodoxybenzoic acid and 2-hydroxypyridine leads to selective oxidation of the 5′-hydroxyl group of the ribose tail of B$_{12}$ in a 30% isolated yield. The acid derivative was purified in one step by HPLC chromatography and characterized by MALDI-TOF mass spectrometry, 1H NMR and 2D (HSQC and HMBC) NMR. The new carboxylic acid derivative is perfectly suited to make stable amide-based B$_{12}$ bioconjugates.

Key words: vitamins, oxidation, 2-iodoxybenzoic acid, carboxylic acids, cobalamin

Interest in using vitamin B$_{12}$ (B$_{12}$) for the oral delivery of peptides and proteins$^1$ has encouraged us to further explore the available conjugate sites of B$_{12}$. Modifications of B$_{12}$ that allow for more facile, higher yielding or stable bioconjugate formation with continued recognition by B$_{12}$ dietary uptake and transport proteins have been explored by numerous groups.$^2$-$^4$ Several functional groups are available on B$_{12}$ for conjugation, including the 5′-hydroxyl group of the ribose tail, the cobalt(III) ion, the phosphate moiety, and suitably modified propionamides off the corrin ring.$^5$ All such modifications suffer from issues including reduced or complete loss of interaction with one or more of the B$_{12}$ binding proteins, stability issues that result from the production of ester or carbamate bonds, low yielding conjugation reaction due to the nature of the functional group, or require problematic purification of the modified B$_{12}$ for subsequent conjugation.$^2$-$^4$ These limitations have been well investigated$^6$ and a review of such sites has been reported,$^7$ however, in brief for purposes of perspective for the work described herein: Conjugation resulting in the recognition of all three transport proteins has been most successful with B$_{12}$ at two major sites: (i) the peripheral corrin ring e-propionamide, and (ii) the 5′-hydroxyl group of the ribose unit of the a'-tail'. The side chain e-carboxylic acid derivative (Figure 1) allows for a variety of modifications,$^8$ but the synthesis of the derivative is low yielding and requires laborious purification, the result, in part, of the formation of b- and d-carboxylic acid isomers in addition to the desired e-isomer.$^9$ The 5′-hydroxyl group has been cited as being a more versatile site according to structure–activity relationships, however, the conjugates have typically been synthesized as mentioned above, through carbamate or ester bonds.

The goal of this work was to create a simple route for the production of a 5′-carboxylic acid derivative of B$_{12}$ with facile purification to ultimately produce stable amide-linked bioconjugates at this optimal position for conjugation, especially in regard to intrinsic factor binding (critical for oral uptake of B$_{12}$ bioconjugates).$^{10}$ Herein, we report the use of the hypervalent iodine reagent, 2-iodoxybenzoic acid (IBX) for the selective oxidation of the 5′-hydroxyl group of B$_{12}$ to the corresponding carboxylic acid. This is the first time IBX has been utilized for the oxidation of such a molecule as B$_{12}$, and expands the range of use of this environmentally friendly oxidizing agent.$^{11}$

The B$_{12}$ derivative was synthesized through a modification of the method developed by Giannis et al., which incorporates the addition of an O-nucleophile to aid IBX in the transformation of a primary alcohol into a carboxylic acid (Scheme 1).$^{12}$ The use of N-hydroxysuccinimide (NHS) was initially attempted but resulted in the formation of a product mixture from which it was not possible to isolate the carboxylic acid, likely due to the formation of the activated NHS ester, which has been reported as a dominating product of this reaction.$^{12}$ Therefore, 2-hydroxypyridine (HYP) was used as the O-nucleophile. To optimize the reaction conditions, a series of experiments were performed in which the co-oxidant, temperature, and time were varied (Table 1). The best result was obtained in the
presence of 2.6 equivalents of IBX and 5 equivalents of HYP in dimethyl sulfoxide (DMSO) at 60 °C.

These conditions resulted in the oxidation of B12 to the 5′-carboxylic acid derivative B12CA in two hours with a 30% yield (see the Supporting Information). It was determined that a 30% yield was optimal for the IBX reaction in this case. An increase in temperature, IBX/HYP ratios, or time resulted in an increase in the amount of byproducts and/or decomposition of the B12 (especially at temperatures in excess of 60 °C), reducing the yield and complicating purification. The 30% yield is approximately three times that previously reported for the preparation of the e-carboxylic acid derivative, with far easier separation (see below). 9

The crude reaction mixture was precipitated from DMSO by ether addition and the precipitate was redissolved in water prior to purification by ion exchange chromatography using a SAX column. After a 5 minute hold at 100% water, a gradient from 100% water to 35% MeCN–phosphate buffer (pH 2; 50:50 v/v) over 5 minutes was necessary for the separation and elution of B12CA (Figure 2). Slight increases in pH above pH 2 resulted in a reduction in the amount of pure isolated product. HPLC purification allowed the recovery and reuse of B12 starting material, albeit with the oxidized 5′-aldehyde also present (Figure 2). B12CA was identified as the peak at 13.6 minutes retention time. The purity of B12CA was at least 95%, as assayed by RP-HPLC. MALDI-TOF mass spectrometry analysis of B12CA revealed the anticipated molecular ion peak (m/z = 1343.5 [M – CN]+; Figure 3).

Electronic absorption spectroscopy (EAS) was used to establish the oxidation state of the metal center as Co(III), with no indication of derivation of the axial ligands bound to the metal center (ε$_{CA}$361 = 27500 M–1cm–1; see Figure 4).

The structure of B12CA was determined on the basis of 1H NMR (see Figure S2) and 1H–13C heteronuclear (HSQC and HMBC) correlations (see Figures S3 and S4). The 1H NMR spectra of B12CA revealed the loss of two protons (δ$_{H}$ = 3.75 and 3.92 ppm), when compared to the spectrum of B12 (CNCbl; see Figures S2 and S5). A comparison of the HSQC of B12 and B12CA indicates the protons that are absent in B12CA were originally attached to the R5 carbon (δ$_{C}$ = 63.5 ppm; see Figure S6). HMBC was attempted to further establish the presence of a carbonyl at the R5 position (see Figure S4), however, it was not possible to assign the R5 carbon from the HMBC spectra because it

Table 1 Optimization of the Reaction Conditions

<table>
<thead>
<tr>
<th>Entry</th>
<th>Co-oxidant</th>
<th>Temp (°C)</th>
<th>Time (h)</th>
<th>Yield (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>None</td>
<td>50</td>
<td>2</td>
<td>17</td>
</tr>
<tr>
<td>2</td>
<td>HYP</td>
<td>50</td>
<td>2</td>
<td>23</td>
</tr>
<tr>
<td>3</td>
<td>HYP</td>
<td>50</td>
<td>3</td>
<td>26</td>
</tr>
<tr>
<td>4</td>
<td>HYP</td>
<td>50</td>
<td>4</td>
<td>28</td>
</tr>
<tr>
<td>5</td>
<td>HYP</td>
<td>50</td>
<td>5</td>
<td>22</td>
</tr>
<tr>
<td>6</td>
<td>HYP</td>
<td>60</td>
<td>2</td>
<td>30</td>
</tr>
</tbody>
</table>
showed no connectivities to any of the ribose protons, and so was assigned by HSQC.\textsuperscript{13}

A simple amine-containing organic molecule, benzylamine, was conjugated to the derivative to help with the identification of the R5 carbon by NMR analysis. In the presence of 10 equivalents of ethyl-3-(3-dimethylaminopropyl)-carbodiimide (EDC) and 20 equivalents 1-hydroxybenzotriazole (HOBt) in N,N-dimethylformamide (DMF) with 1% water, the desired conjugate was produced at 40% yield over the first hour (Scheme 2), thus establishing proof-of-concept for the new B12CA.

The conjugate was purified by using the same purification method established for the carboxylic acid (Figure 5) and characterized by MALDI mass spectrometry (Figure S7) as well as 1D (Figure S8) and 2D NMR (Figures S9 and S10) spectroscopy.

With the addition of the benzylamine, connectivity to the R5 carbon was established by HMBC (Figure 6). Specifically, the protons of CA (δ\textsubscript{H} = 4.42 and 4.50 ppm) showed direct connectivity to the R5 carbon (δ\textsubscript{C} = 173.8 ppm) as well as to C\textsubscript{B} (δ\textsubscript{C} = 140.8 ppm) and C\textsubscript{C} (δ\textsubscript{C} = 131.4 ppm) of the benzylamine. The complete assignment of B\textsubscript{12}CA and B\textsubscript{12}CABA can be found in Table S1.

In conclusion, we have developed a convenient site-specific oxidation of B\textsubscript{12} using IBX and HYP. This new carboxylic acid derivative is highly suited to the production of vitamin B\textsubscript{12}.
of bioavailable amide bound bioconjugates. The one-step purification is simple and the synthesis results in a 30% yield within two hours, which is far in excess of any previously reported B12 carboxylic acid derivative. The resulting derivative has the potential to form high-yielding conjugates and we believe will be of considerable use in the field of B12 conjugation.

Acknowledgment

R.P.D. gratefully acknowledges the W. M. Wrigley Jr. company (Chicago, IL) for funding this work and iEARN program at Syracuse University for funding J.L.B. We would also like to acknowledge Professor John Chisholm (Department of Chemistry, Syracuse University) for his insight into IBX.

Supporting Information

For this article is available online at http://www.thieme-connect.com/ejournals/toe/synlett.

References and Notes


