UNIVERSITY OF CALIFORNIA

Santa Barbara

Addressing challenges in Type 1 diabetes:

non-invasive insulin delivery and advanced diagnosis

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Chemical Engineering

by

Kelly Noblit Ibsen

Committee in charge:

Professor Samir Mitragotri, Chair

Professor Patrick Daugherty

Professor Glenn Fredrickson

Professor Scott Shell

Professor Adele Doyle

June 2018

The dissertation of **Kelly Noblit Ibsen** is approved.

Patrick Daugherty

Adele Doyle

Glenn Fredrickson

Scott Shell

Samir Mitragotri, Committee Chair

June 2018

Addressing challenges in Type 1 diabetes:

non-invasive insulin delivery and advanced diagnosis

By

Kelly Noblit Ibsen

DEDICATION

For my daughter Nikki, and all those with type 1 diabetes, who face each day with bravery, humor and optimism and ask only in return for some semblance of normalcy. I am humbled by your fortitude and strength, and so very proud to be a part of the fight.

ACKNOWLEDGEMENTS

First and foremost, thank you to my advisors, Samir Mitragotri and Patrick Daugherty. Both of you allowed me the freedom to explore topics that were personally important to me and helped me find the scientific importance in them. Your guidance, support and critical appraisals helped me refine my skills and become confident enough to pursue ever more challenging projects. The spirit of collaboration you fostered in your labs was a key aspect to my success, and one that is unique, so thank you.

I have been privileged to work with so many amazing people throughout grad school, both at UCSB and Harvard. Everyone in my labs freely provided insight, ideas and expertise, of which I partook deeply and often. Special thanks to Max, Tyler, Amrita, Michael Evans and Michael Zakrewsky from the Mitragotri lab and Joel, Michael, Sera, Jen and Bob from the Daugherty lab.

From the very start, I have been surrounded by family and friends that never wavered in their support, and I cannot overstate how much that helped me complete this chapter of my life. Special thanks go to Gregg Beckham, my longtime friend and sounding board, for not falling off his chair laughing when I told him I wanted to go to grad school at 50. Much love to my personal cheerleaders Nikki, Cory, Christine and Timmy – you guys are all amazingly giving, and whenever I needed a shot of confidence, all I had to do was talk to one of you. To the next generation - Ally, Arthur and Fin, find a star and reach for it. And to Tim, who has unfailingly been there every step of the way, especially the rough ones, I give my most ardent gratitude and love. How you know exactly what I need (loud country music, coffee and salt), and why you so unstintingly give of your time, energy and intellect to help me succeed every day is one of the great mysteries of life. I know I will never meet another human like you.

VITA OF KELLY NOBLIT IBSEN

EDUCATION

2013-2018	**Ph.D. in Chemical Engineering (Expected)** **Bioengineering Emphasis** University of California, Santa Barbara, CA
1981-1985	**B.S. in Chemical Engineering with Distinction** Colorado State University, Ft. Collins, CO

PROFESSIONAL EXPERIENCE

2017-2018	**Visiting Fellow, School of Engineering & Applied Science** Harvard University, Cambridge, MA
2013-2017	**Graduate Researcher and Teaching Assistant** University of California, Santa Barbara, CA
2009-2013	**Owner/Principal Engineer** Lynx Engineering LLC, Denver, CO
2010-2013	**Chief Operations Officer** National Advanced Biofuels Consortium
2007-2009	**Director of Tech Transfer** Myriant Corporation, Boston, MA
2004-2007	**Biochemical Platform Lead**
2002-2004	**Group Manager, Analysis & Partnership Development**
1997-2002	**Process Engineering & Analysis Team Lead**
1992-1997	**Pilot Plant Team Lead/Process Engineer** National Renewable Energy Laboratory, Golden, CO
1989	Licensed Professional Engineer certification
1986-1992	**Senior Project Engineer/Plant Operations Engineer** Nestle Foods Research and Development, Van Nuys, CA

CERTIFICATIONS, APPOINTMENTS & AWARDS

2014-2018	National Science Foundation Graduate Research Fellowship
2017	Best poster, Controlled Release Society Annual Meeting
2017	UCSB Grad Slam finalist
2016	Best poster, Clorox-Amgen Graduate Student Symposium
2016	Best poster, Mellichamp Lecture in Systems Biology

PUBLICATIONS

- **Ibsen, K.**, Ma, H., Banerjee, A., Tanner, E., Nangia, S., Mitragotri, S. The mechanism of antimicrobial activity of choline-based ionic liquids (CAGE). *ACS Biomaterials Science & Engineering* (2018).
- **Ibsen, K.,** Banerjee, A., Iwao, Y., Zakrewsky, M., Mitragotri, S. Transdermal protein delivery using choline and geranate (CAGE) deep eutectic solvent. *Advanced Healthcare Materials* (2017).
- **Ibsen, K**., Daugherty, P. Prediction of antibody structural epitopes via random peptide library screening and next generation sequencing. *Journal of Immunological Methods* (2017).
- Banerjee, A., **Ibsen, K.,** Brown, T., Chen, R., Agatemor, C., Mitragotri, S. Ionic liquids for oral insulin delivery. *PNAS* (2018).
- Tanner, E., **Ibsen, K,** Mitragotri, S. Transdermal insulin delivery using choline-based ionic liquids (CAGE). Submitted to *J Contr Rel* (2018).
- Agatemor, C., **Ibsen, K**., Tanner, E., Mitragotri, S. Ionic liquids for addressing unmet needs in healthcare. *Bioengineering & Translational Medicine* (2018).
- Evans, M., Huang, P., Iwamoto, Y., **Ibsen, K.**, Chan, E., Garcia, J., Hitomi, Y., Ford, P., Mitragotri, S. Macrophage-Mediated Delivery of Light Activated Nitric Oxide Prodrugs with Spatial, Temporal and Concentration Control, *Chemical Science* (2018).
- Paull, M., Johnston, T., **Ibsen, K.,** Daugherty, P. A general approach for identifying protein epitopes targeted by antibody repertoires using whole proteomes. In preparation (2018).
- Pantazes, R., Reifert, J., Bozekowski, J., **Ibsen, K**., Murray J., Daugherty P. Identification of disease-specific motifs in the antibody specificity repertoire via next-generation sequencing. *Scientific Reports* (2016).

PRESENTATIONS

- Transdermal insulin delivery via a choline-based deep eutectic solvent. Controlled Release Society annual meeting, 2017; Boston, MA. *Poster*.
- Transdermal protein delivery via a choline-based deep eutectic solvent. American Institute of Chemical Engineers annual meeting, 2017; Minneapolis, MN. *Oral presentation*.
- Profiling the antibody response in Type 1 Diabetes via peptide display. Mellichamp Lecture in Systems Biology and Bioengineering. 2016, Santa Barbara, CA. *Poster.*
- 2016 Clorox-Amgen Graduate Student Symposium poster. “Mapping conformational epitopes via random peptide display and high-throughput sequencing”. *Poster*.
- Profiling the Humoral Antibody Response in Type 1 Diabetes Via Peptide Display. American Institute of Chemical Engineers annual meeting, 2016; San Francisco, CA. *Poster*.

Addressing challenges in Type 1 diabetes:

non-invasive insulin delivery and advanced diagnosis

by

Kelly Noblit Ibsen

ABSTRACT

The treatment of disease is unarguably one of the most important applications of the emerging field of bioengineering, which seeks to apply engineering principles to biological challenges. In the field of medicine, bioengineers have played a role in a multitude of areas including diagnostics, monitoring and therapy, working with scientists and physicians to develop imaging equipment, biocompatible implants, and targeted cancer therapies. These advances often used materials that a few decades ago would have been unheard of in healthcare applications, such as titanium, polymers and nanoparticles. In keeping with this pioneering spirit, my research sought to use bioengineering applications to address the areas of diagnosis and treatment of disease. I chose Type 1 diabetes (T1D) as a model disease for this work, as challenges exist in both areas. For T1D patients taking exogenous insulin, alternatives to injection therapy could reduce the stress of daily multiple injections, increase adherence to dosing regimens and lower the incidence of skin infections. In addition, early diagnosis of T1D can provide earlier intervention, providing an opportunity for the development of therapies to slow beta cell destruction and prolong endogenous insulin production.

In the first part of this work, we demonstrate the potential of ionic liquids as both antimicrobials and drug delivery vehicles. Probably best known as “green” solvents in the chemical industry, ionic liquids are sparking interest in healthcare due to their innate antimicrobial qualities and excellent drug solvation properties. These characteristics and properties can be both finely tuned and widely modulated by changes in ion type, ratio and ion structure, providing an almost unlimited design framework. To demonstrate this, we studied how the simplest tuning parameter, ion ratio, affects both antibacterial efficacy and transdermal drug delivery of a variety of proteins including insulin. We also sought to investigate the breadth of ionic liquids in drug delivery by showing their potential to effect oral delivery of insulin. Using a combination of *in silico, in vitro, ex vivo* and *in vivo* models, we showed that a choline and geranic acid (CAGE) ionic liquid is a highly effective bactericide as well as an enhancer of transdermal and intestinal protein delivery.

In the second portion of the work, we investigate the use of bacterial peptide display libraries and next generation sequencing (NGS) in two ways; to improve epitope mapping and to search for new humoral biomarkers in T1D. We were able to demonstrate that NGS data can improve the results from PepSurf, a web-based epitope mapping program.

Finally, by profiling the antibody repertoire of plasma from T1D and non-T1D matched controls, we discovered several peptide motifs with good (30-50%) sensitivity and 100% specificity. Challenging these motifs to a larger control population of 238 assumed non-T1D samples resulted in one motif with moderate (20%) sensitivity and >99% specificity. Further development and maturation of these motifs could provide new biomarkers for T1D.

Contents

Chapter 1

1. Dissertation Overview

This dissertation is loosely organized around Type I diabetes (T1D) and insulin as a model disease and therapeutic to showcase the application of bioengineering in the diagnosis and treatment of disease, specifically the identification of plasma biomarkers for T1D and non-invasive methods for the delivery of insulin. While T1D, a disease that affects over one million people in the U.S. currently with an expected 5 million by 2050 (JDRF, 2018), figures prominently in this work, the results discussed here have the potential for broader impact. The text is divided into two sections: I) medical applications of ionic liquids, specifically fighting infection and drug delivery, and II) epitope mapping and serum biomarker discovery using peptide libraries and NGS. Both topics are highly relevant to a goal of improving the health care industry. Non-invasive drug delivery can improve the lives of many patients living with chronic diseases, while more sensitive biomarkers can provide earlier intervention and targeted therapy.

I begin the first part of this dissertation by discussing the development of ionic liquids as antimicrobials and drug delivery vehicles (Chapters 2-5). In Chapter 2, I introduce ionic liquids and their properties related to biotechnology. In Chapter 3, in collaboration with Prof. Nangia's lab at Syracuse University, we determined the antibacterial properties of CAGE, a choline-based ionic liquid, and the mechanism by which CAGE attacks *E. coli*, a Gram-negative bacterium. In Chapters 4 and 5, I discuss the use of CAGE in transdermal and oral drug delivery. Non-invasive drug delivery is the holy grail for clinicians and patients alike, with the potential to provide a simpler, less painful and ultimately higher adherence treatment, but

increasing transport and bioavailability top the list of challenges in each area. My aim for this portion of the dissertation is to showcase how broadly applicable ionic liquids can be as part of therapies for both chronic and acute disease with specific examples of bacterial neutralization and transdermal and oral insulin delivery.

Part II begins in Chapter 8 with background information regarding the use of bacterial peptide display libraries coupled with next-generation sequencing (NGS) in the identification of antigenic peptides. Chapter 9 presents a new method to use NGS datasets for epitope mapping while Chapter 10 focuses on using the tools of peptide display and NGS to search for novel biomarkers for T1D that can serve as diagnostics to identify at-risk populations before clinical onset of symptoms. Throughout, I provide a summary of the important findings and potential future research directions for each area.

Permissions and Attributions

1. Some content in Chapter 3 was previously published in *ACS Biomaterials Science and Engineering*, and the material is adapted with permission from the American Chemical Society.
2. Some content in Chapter 4 was previously published in *Advanced Healthcare Materials*, and the material is adapted with permission from Wiley.
3. Some content in Chapter 5 was previously published in *Proceedings of the National Academy of Science*, and the material is adapted with permission from the National Academy of Sciences.
4. Some content in Chapter 9 was previously published in the *Journal of Immunological Methods*, and the material is adapted with permission from Elsevier.

Chapter 2

2. Introduction to ionic liquids

2.1 A brief history

The first reported ionic liquid (IL) was discovered in 1914 by Paul Walden, a Latvian chemist, who is credited with the discovery of ethylammonium nitrate, or EAN (Walden, 1914). Most simply, ionic liquids are cation/anion pairs with low melting points. Walden synthesized his "molten salt" by neutralizing ethylamine with concentrated nitric acid. Fast forward to a century later, and one can shop for ILs at company websites like Solvionic.com - just select your cation, anion, or even a choose desired property to see which ILs fit your need. Let's take a closer look at the development path of this important class of liquids.

EAN was highly reactive (Emel'yanenko, et al., 2014), which likely contributed to the lack of interest, and ILs' next appearance wasn't until 1931, in a patent describing the use of nitrogen-containing salts such as 1-benzylpyridinium chloride to dissolve cellulose, creating highly reactive cellulose derivatives (Graenacher, 1931), but this IL also failed to ignite the hearts of chemists. The discovery of aluminum chloride/ethylpyridinium bromide salts for aluminum electrodeposition finally convinced electrochemists to rigorously study, characterize and ultimately expand this group of ILs, adding the imidazolium cation family (Plechkova & Seddon, 2008). One drawback of the chloroaluminate (III) ILs was extreme water reactivity, requiring all work be done in an inert-atmosphere box. In 1992, Wilkes and Zaworotko reported the preparation and characterization of ILs that contained imidazolium cations, but with a range of anions ($[CH_3CO_2]$, $[NO_3]$, $[BF_4]$) that were less corrosive, and

stable in air and water, allowing much broader use (Wilkes & Zaworotko, 1992). Other classes of cations based on pyrrolidinium and phosphonium followed.

The realization that over one million ILs could be made by simple combinations of the available cations and anions (**Figure 2.1**) (Seddon, 1999) (Stark & Seddon, 2007) gave rise to

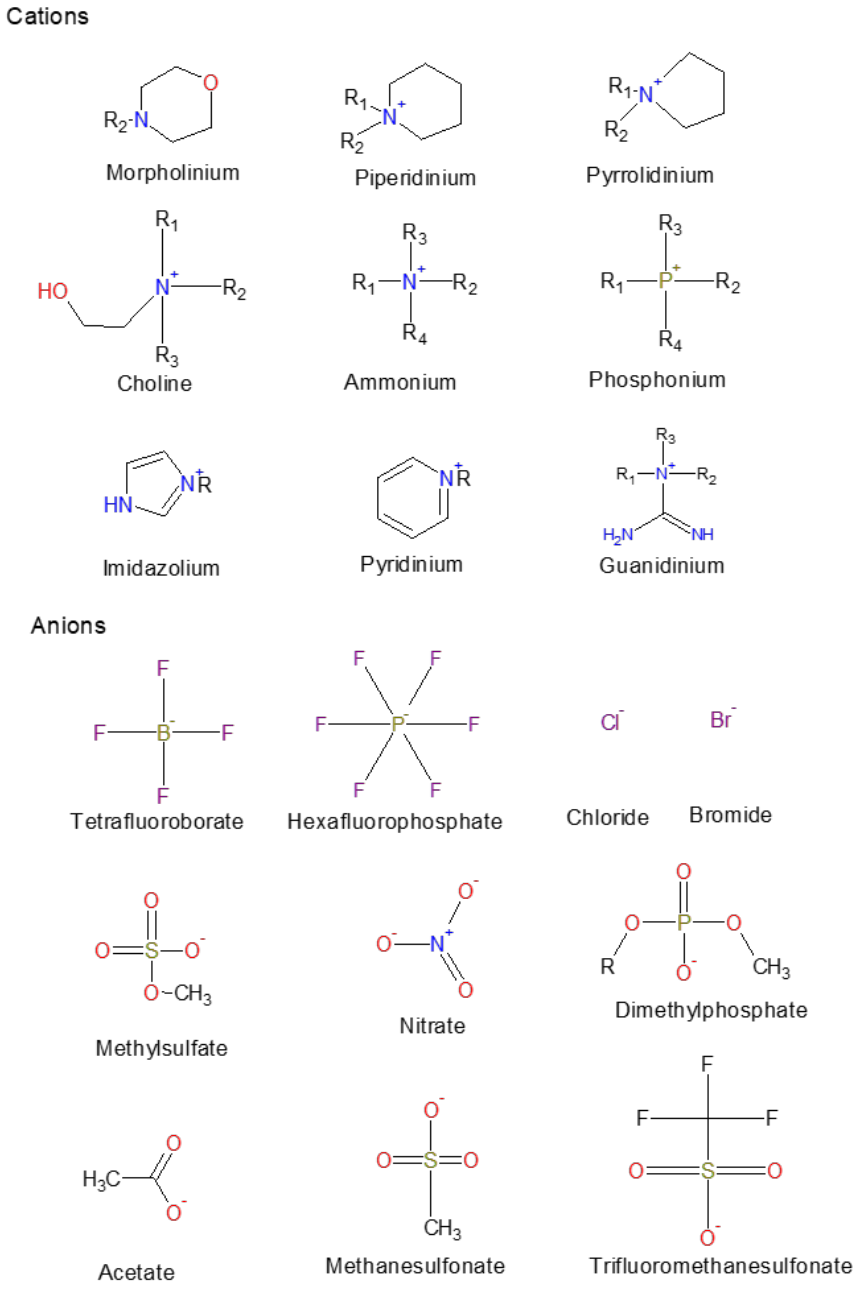


Figure 2.1 Some common ionic liquid cations and anions.

the tantalizing idea that ILs could be synthesized for almost any function given an understanding of how their composition, substitution and pairing affected their properties, and prompted several chemical companies to file patents outlining the use of ILs as solvents for organic synthesis (Fields, et al., 1998) (Abdul-Sada, et al., 1995) (Abdul-Sada, et al., 1995). The greening of the chemistry industry was underway, and due to their negligible volatilities and superior solvating properties, ILs were considered the front runners to replace volatile, toxic solvents like benzene and toluene. By 2007, companies like BASF, Eastman Chemical Company, and Degussa were developing IL-based technologies for a variety of chemical processes (Plechkova & Seddon, 2008).

Today, ionic liquids are commonly divided into protic (containing hydrogen atoms that can be donated, such as in a hydroxyl group) and aprotic (no H donors) subclasses. Another important class of ionic solvents are deep eutectic solvents (DES), which are systems formed from a eutectic mix of acids and bases with anionic or cationic species (e.g. choline chloride + urea) (Ashworth, et al., 2016). Both ILs and DESs are typically described by their negligible vapor pressure, low flammability and defined as having melting points at or below 100 °C (a number courtesy of Walden that survives today).

2.2 ILs in biotechnology

In 2007, Hough elegantly traced the evolution of ionic liquids, from Generation 1 (ILs with unique tunable physical properties) to Gen 2 (ILs with targeted chemical properties combined with chosen properties), finally proposing what was then a nascent Gen 3 – ILs with targeted biological properties combined with chosen physical and chemical properties to create active pharmaceutical ingredients (Hough, et al., 2007). Understanding how to harness the inherent

characteristics of ILs – low melting point, superior solvation, good thermal and chemical stability and innate antimicrobial activity might make them compelling candidates to address some of the most pressing challenges in healthcare including drug delivery, development of robust antimicrobials, and biosensing. There have been several reviews published that outline the potential role of ILs in biotechnology and specifically medicine; recent ones include Agatemor (Agatemor, et al., 2017), Egorova (Egorova, et al., 2017) and Adawiyah (Adawiyah, et al., 2016). The focus of this work was the application of ILs in two important areas: 1) the development of antimicrobial therapies to treat infection and 2) non-invasive drug delivery.

ILs have innate antiseptic qualities. As early as the 1930s, the broad antimicrobial activity of quaternary ammonium compounds was recognized and harnessed for use in disinfectants and antiseptics (Tischer, et al., 2012) (Rajkowska, et al., 2016). Additionally, ecology studies from the chemistry industry have provided a wealth of information on the effect of ILs on a variety of species and their environments (Adawiyah, et al., 2016). This aspect of ILs is perhaps the timeliest, given the urgent need to develop antibiotics to fight an increasing number of resistant microbial strains. It has been postulated that the mechanism of microbial neutralization lies in the ability of ILs to disrupt the cell membrane, potentially by alkyl chain insertion, which is supported by the recognized effect of alkyl chain length. Pernak and others found that increasing side chain length reduced the minimum bactericidal concentration (Pernak & Skrzypczak, 1996) (Pernak, et al., 2003) (Docherty & Kulpa Jr, 2005) (Pernak, et al., 2007) (Yu & Nie, 2011) (Jeong, et al., 2012). Zakrewsky also demonstrated that ILs can neutralize biofilm-forming pathogens (Zakrewsky, et al., 2016), an important quality for antimicrobials used to treat infections in chronic wounds such as diabetic foot ulcers (Zhao, et al., 2013) (Percival, et al., 2014).

Understanding the toxicity of ILs is important not only to harness it for the treatment of infectious or invasive organisms, but also to minimize ILs' effect on healthy organs, tissue and cells. The toxicity of ILs is primarily determined by its molecular structure including the nature of the cation, the alkyl chain length, and the functionalization of both ions. Amde (Amde, et al., 2015) reported the relative toxicity for cationic head groups as:

Choline<piperidinium<pyrrolidinium<pyridinium=imidazolium<ammonium<phosphonium

ILs have also been shown to increase the solubility of drug molecules that are sparingly soluble in water and most pharmaceutically accepted organic solvents such as DMSO (Kumar, et al., 2007) (Moniruzzaman, et al., 2010) (Mizuuchi, et al., 2008) (Williams, et al., 2014) (Azevedo, et al., 2014). The solvation ability of an IL is dependent on its properties, and those of the anion appear to be most important followed by the cation side chain, owing to the formation of hydrogen bonds between the IL anions and the drug molecules. Morrison showed that a urea-choline chloride eutectic increased the solubility of benzoic acid, danazol and itraconazole by 20-, 16- and 22,000-fold respectively compared to water, and that the components alone (aqueous choline chloride or urea) did not confer the same solubility (Morrison, et al., 2009). Additional information on drug solvation and IL composition can be found in the 2016 review by Adawiyah.

In addition to increasing solubility of proteins, ILs can preserve the catalytic activity of a variety of enzymes; for example, lipases are reported to have increased selectivity and operational stability in ILs compared to traditional media (Kragl, et al., 2002) (Rantwijk, et al., 2003). The thermal stability of 3 mM cytochrome c was increased (up to 110 °C) in ILs using a dihydrogen phosphate anion with varying cations including choline (Fujita, et al., 2005).

From the proven effect of ILs on microbial cells, and the likely mechanism of membrane disruption, it seems plausible to consider ILs as permeation enhancers for transdermal drug delivery, which requires the disruption of a lipid bilayer in the stratum corneum, similar to that found in a bacterial cell membrane. There several classes of chemical penetration enhancers already in use to increase skin permeability, including alcohols, surfactants and fatty acids, many of which share similar properties to ILs. Monti increased permeability of diltiazem, a calcium channel blocker for hypertension, in excised rat skin using acyclic octane-based ILs (Monti, et al., 2017) and Zakrewsky used a variety of ILs including choline and tetraalkylphosphonium cations paired with carboxylic acid anions to deliver a small molecule (mannitol) and cefadroxil, an antibiotic, *ex vivo* in porcine skin (Zakrewsky, et al., 2014). In addition, fatty acids have been shown to be excellent permeation enhancers for intestinal absorption of drugs (Brayden, et al., 2014) (Bruno, et al., 2013) (Muheem, et al., 2016). With their similarity to fatty acids, we hypothesized that ILs with anions of organic acids might also demonstrate permeation enhancement in the gut.

Combining these three aspects of drug delivery – solubility, stability and permeation – seems a daunting task to ask a single system to provide. The unique aspect of ILs, their tuneability, makes them ideal candidates to take up the challenge (**Figure 2.2**). Adding to ILs' potential benefits are the availability of cations with low toxicity to human cells. With this challenge in mind, our lab began investigating the use of ILs for drug delivery.

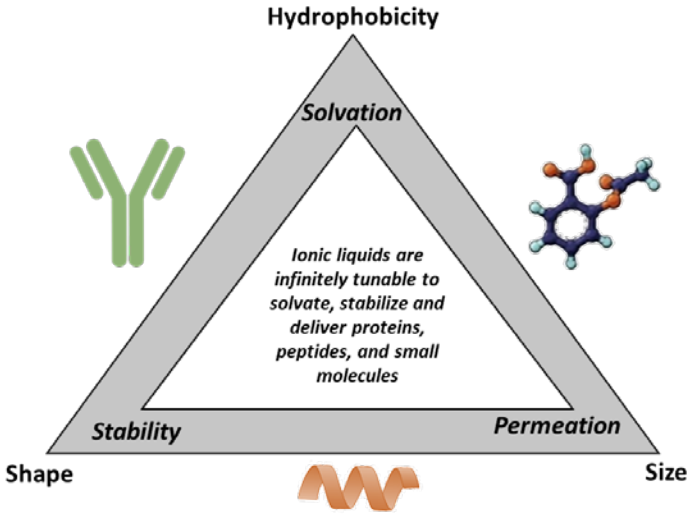
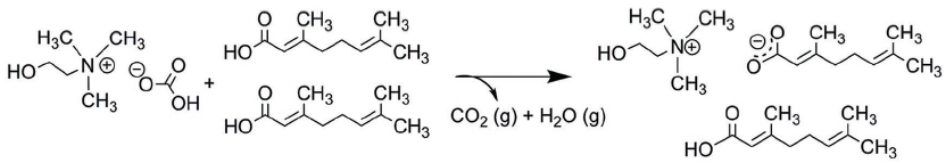


Figure 2.1. ILs' tuneability make them ideal candidates for solving major challenges in drug delivery. Low solubility in water makes many hydrophobic drugs hard to deliver in efficacious concentrations. Additionally, peptides and proteins can be denatured or degraded by components in the host system, such as proteases in the gut. Finally, many molecules of interest are too large to passively transport through biological barriers in the body. ILs have the potential to address all three areas with low toxicity to human cells and can be infinitely tuned via ion selection and further functionalization via side chain substitutions.

2.3 Development of CAGE

In 2013, the Mitragotri group undertook a large study of ILs and DESs to identify candidates suitable for development in medical applications (Zakrewsky, et al., 2014). The study examined several biologically relevant characteristics of ILs, including viscosity, conductivity, density and lipophilicity. In addition, they investigated each IL for its ability to transport small molecules like mannitol across the skin, neutralize pathogens, and show minimal cytotoxic action against mammalian cell lines. The best candidate was CAGE, a choline and geranic acid deep eutectic (1:2 choline:geranic acid) that exhibited good skin transport, low cytotoxicity and high antibacterial activity against *Pseudomonas aeruginosa* and *Salmonella enterica*. Choline is a water-soluble essential nutrient, made in the liver, and present in phospholipids that are abundant in cell membranes. Geranic acid, commonly used as a flavoring agent, is naturally occurring in lemongrass, which has reported antimicrobial activity itself (Friedman,

et al., 2004). CAGE is easily produced via a salt metathesis reaction of choline bicarbonate and geranic acid, which are both FDA-listed as GRAS (Generally Recognized As Safe), producing water and CO_2 as easily removed byproducts:



Choline is in the family of quaternary ammonium compounds (QACs) which have long been utilized as surfactants (Richmond, 1990) (Holland, 1991), disinfectants (Petrocci, 1983), and preservatives (Cross, 1994). In drug delivery, QACs have potential as penetration enhancers for transnasal and transbuccal delivery of vaccines (Klinguer, et al., 2001) and as liposomal materials (Liu & Huang, 2002).

To expand our understanding of how the composition of CAGE affects its properties, which in turn allows the use of the tunability aspect of ILs and DESs, a panel of CAGE variants was developed by altering the ion ratios, namely 2:1, 1:1, 1:2 (the initial CAGE formulation) and 1:4 choline:geranic acid. This panel was utilized in several studies discussed below to help elucidate the mechanisms of CAGE in both transdermal and intestinal permeation and bacterial neutralization. Only CAGE 1:1 is a true IL; the others are DESs, but hereafter, for simplicity, all will be identified as ILs.

2.4 CAGE characterization

All CAGE variants were analyzed via nuclear magnetic resonance spectroscopy, differential scanning calorimetry and thermos-gravimetric analysis. The density, conductivity and viscosity of CAGE 1:2, used most widely in the following studies, was also measured.

2.4.1 Nuclear magnetic resonance (NMR) spectroscopy

^{1}H NMR was consistent with choline:geranic acid ratios, with the expected number of protons appearing for each variant. For those variants comprised of higher ratios of geranic acid than choline, both geranic acid and geranate will be present. For the 2:1 variant, the carbon NMR contained a signal at 146 ppm that was not present in other CAGE variants, indicating the signature of a carbon attached to a carbonyl group, which corresponds to the residual bicarbonate present. NMR spectra can be found in section 7.1.

2.4.2 Thermo-gravimetric analysis (TGA)

From thermo-gravimetric analysis, it is seen that the short-term stability of the CAGE variants decreases in the order CAGE 1:2 > 1:4 > 1:1 > 2:1. (**Figure 2.2**). The range of decomposition temperatures is small (< 42 °C), indicating the ion ratio does not significantly change thermal stability.

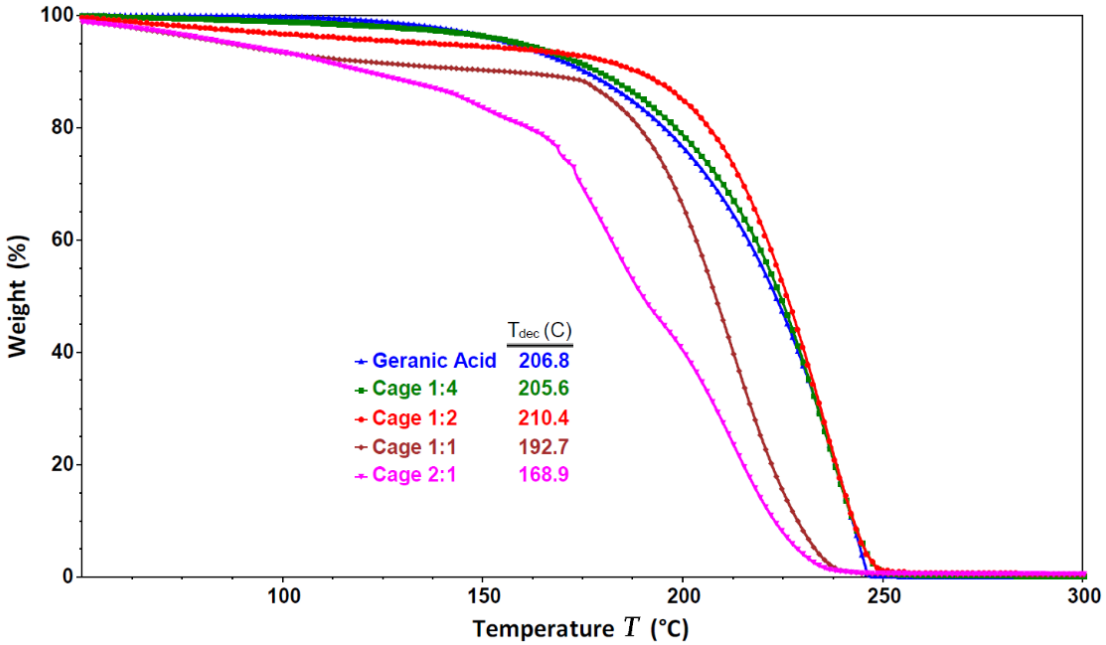


Figure 2.2. Dynamic TGA curves at a heating rate of 10 °C/min for the CAGE variants and pure geranic acid. The onset temperature (T_{dec}) was determined for each sample.

2.4.3 Differential scanning calorimetry (DSC)

The material was heated from either -150 °C or -125 °C to 125 °C and subsequently cooled in the second part of the cycle. A total of 3 complete heat/cool cycles were performed. All 4 variants exhibit a characteristic step change in heat capacity corresponding to glass transition between -50 and -150 °C (**Figure 2.3**). The T_g increase order is CAGE 2:1 < 1:2 < 1:4 < 1:1 and was much lower for CAGE 2:1. The increasing T_g trend follows the order of increasing geranic acid/geranate content for the three eutectics, while the 1:1 true ionic liquid has the highest T_g. This trend could provide some insight into how the ions in different CAGE variants interact; for example, higher T_g has been linked to alkyl chain length on cations (Zheng, et al., 2011), and more rigid structures due to cyclic aromatics (Cowie, 1991). Two variants, CAGE

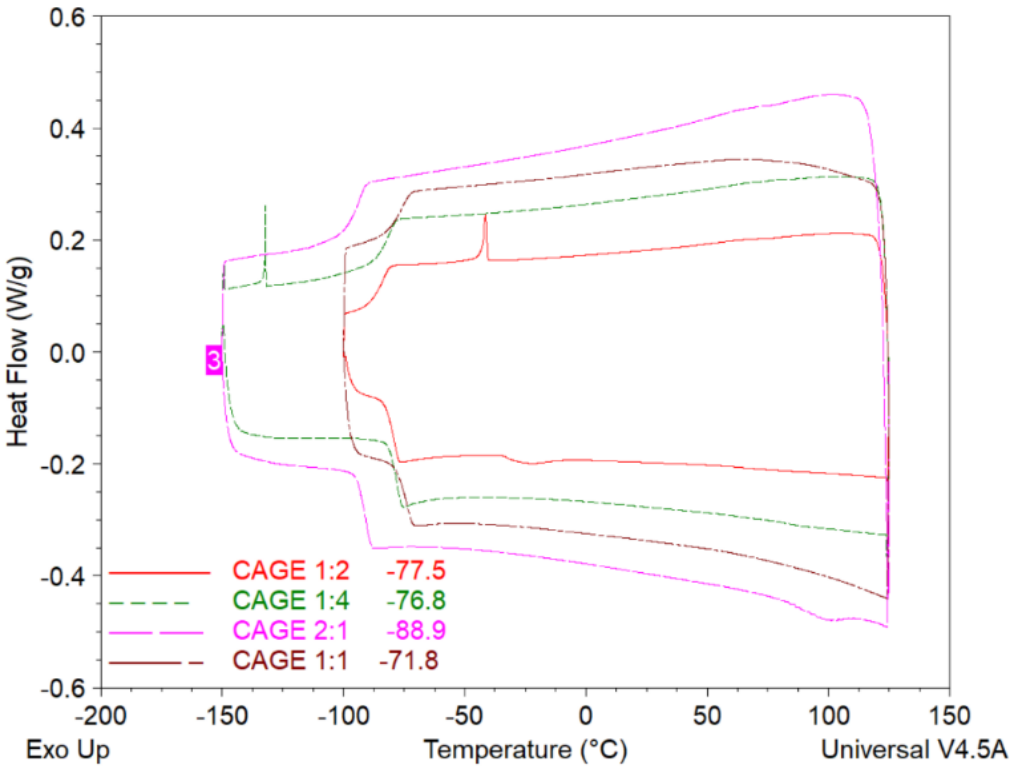


Figure 2.3. DSC heating and cooling scans at 10 °C/min for CAGE variants. All variants exhibit glass transition behavior, and the two with free geranic acid, 1:2 and 1:4, have a crystallization event upon cooling. T_g are listed.

1:2 and 1:4, exhibited cold crystallization events upon cooling at -41.7 °C and -132.2 °C respectively. These events could be attributed to the free geranic acid in the 1:2 and 1:4 variants; pure geranic acid crystallizes at -34.5 °C (**Figure 2.4**).

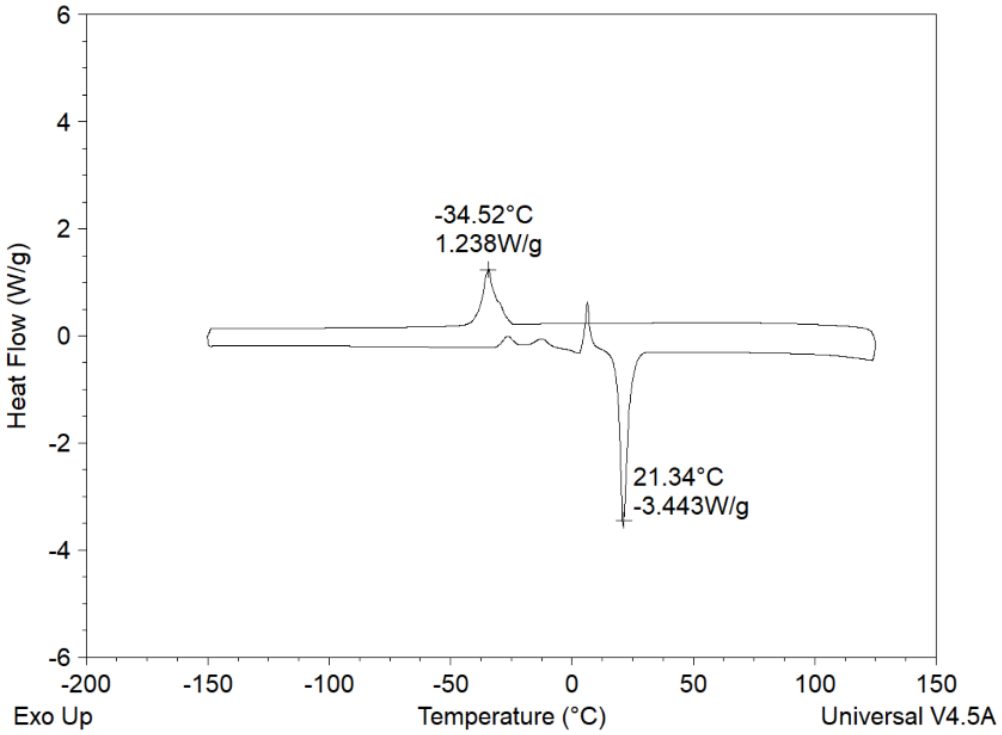


Figure 2.4. DSC heating and cooling scans at 10 °C/min for pure geranic acid. Geranic acid freezes at 21.34 °C and exhibits a crystallization event at -34.5 °C.

2.4.4 Density

Density was measured at 1.02 g/mL using an ARES-LS1 rheometer with a 25 mm diameter parallel plate to calculate the volume of a known weight of CAGE 1:2.

2.4.5 Viscosity

The viscosity of CAGE 1:2 was measured using an ARG2 Rheometer with a 40 mm diameter aluminum 2° cone over a shear range of 10 to 0.1 1/s. CAGE 1:2 exhibited a viscosity of 729

± 5 mPa/s over this shear range. For comparison, the viscosity of glycerin is 950 mPa/s (EngineeringToolbox, 2018).

2.4.6 Conductivity

Conductivity of CAGE 1:2 was measured at 31.6 mS/m with a hand-held conductivity probe and at 23 mS/m using through-plan dielectric relaxation spectroscopy (Bartels, et al., 2017) conducted by Dr. Josh Bartels of UCSB on a sample provided by the author.

CAGE variant properties are summarized in **Table 2.1**. Viscosity, conductivity and density of the CAGE variants were determined by Dr. Eden Tanner (Tanner, et al., 2018). These values are within an order of magnitude for those initially determined by the author.

Table 2.1. CAGE variant properties

Variant	*Viscosity (mP/s)*	*Conductivity(mS/m)*	*Density (g/mL)*	T_{dec} *(°C)*	T_g *(°C)*
2:1	737 ± 32	57.55 ± 0.07	1.17	168.9	-88.9
1:1	492 ± 18	16.28 ± 0.76	1.23	192.7	-71.8
1:2	568 ± 19	13.79 ± 0.28	1.23	210.4	-77.5
1:4	223 ± 26	5.28 ± 0.07	1.12	206.8	-76.8

There is an opportunity to expand our understanding of CAGE and associated ILs with the identification and development of additional characterization methods to more fully describe their thermal and chemical properties. For example, the use of vibrational spectroscopy can help to better understand intermolecular interactions (Paschoal, et al., 2017) and small/wide angle X-ray scattering (SAXS/WAXS) can identify structural changes as a result of temperature (Rengstl, et al., 2014). This data will provide important information to inform studies in biologically-based applications of CAGE and are currently being developed for addition to the characterization protocol.

2.5 Assessment of insulin stability in CAGE

To ensure that CAGE could be used to solvate proteins such as insulin without damage to either structure or function, two studies were performed. Insulin in neat CAGE was stored at 37, 25 and 4 °C overnight. In all cases, the insulin was dispersed, rather than dissolved in the CAGE. This observation is similar to what others report, that proteins are usually insoluble or sparingly soluble in ILs, but instead are in a highly dispersed state (Fujita, et al., 2005). Following that study, insulin in CAGE was stored at 25 and 4 °C for several months, with aliquots tested each month to confirm both the presence of the secondary structure and *in vivo* activity.

2.5.1 Secondary structure

Circular dichroism (CD) spectrophotometry was performed in order to determine the effect of CAGE on the structure of insulin. Secondary protein structures (alpha helices, beta sheets and random coils) give rise to characteristic shapes in a CD spectrum. Since insulin is made up of alpha helices only, it yields characteristic double negative troughs at 208 nm and 222 nm in a CD spectrum (Hua, et al., 2002) (Huang, et al., 2004). The alpha helical conformation of insulin was retained after storing in CAGE overnight at 4, 25 (room temperature) and 37 °C indicating that insulin remains in its physiologically active conformation when dispersed in CAGE **(Figure 2.5)**. However, it should be noted that CAGE is a poor solvent for CD, exhibiting low transparency in the wavelength range of interest; insulin was recovered from CAGE by repeated washings with phosphate buffered saline (PBS).

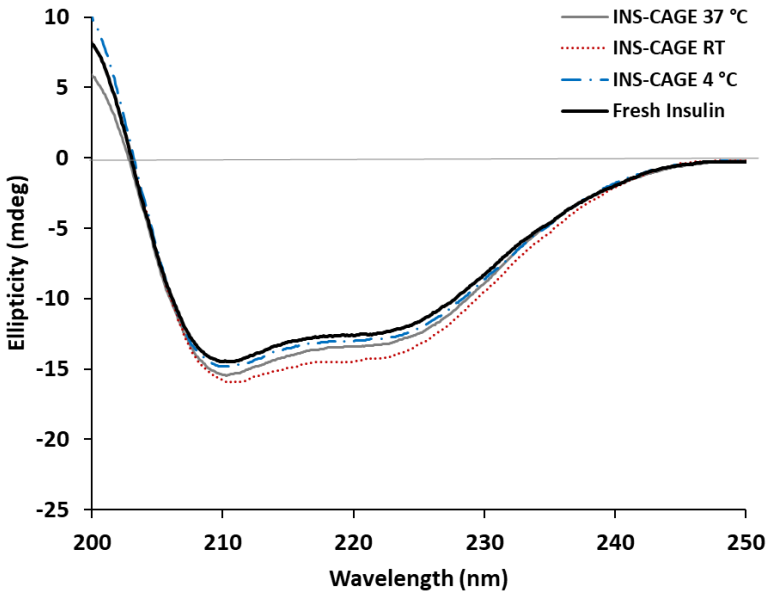


Figure 2.5. Insulin retains structure overnight in CAGE. Human insulin (100 U) was suspended in 1 mL of CAGE and incubated either under refrigeration (4 °C, blue dash-dot line), room temperature (RT, 25 °C, red dotted line) or physiologically relevant temperature (37 ° C, grey solid line). Circular dichroism spectra in the far-UV region of each sample compared to a fresh insulin-PBS control (black solid line) showed no change in insulin secondary structure.

We then undertook an extended study, storing insulin in CAGE for several months at room temperature and under refrigeration and testing aliquots every 30 days to look for changes in structure. No difference in the shape or degree of ellipticity was noticed between freshly prepared insulin solution and insulin stored in CAGE at RT or 4 °C for up to 3 and 4 months respectively (**Figure 2.6)**, suggesting CAGE helps retain insulin structure for prolonged periods.

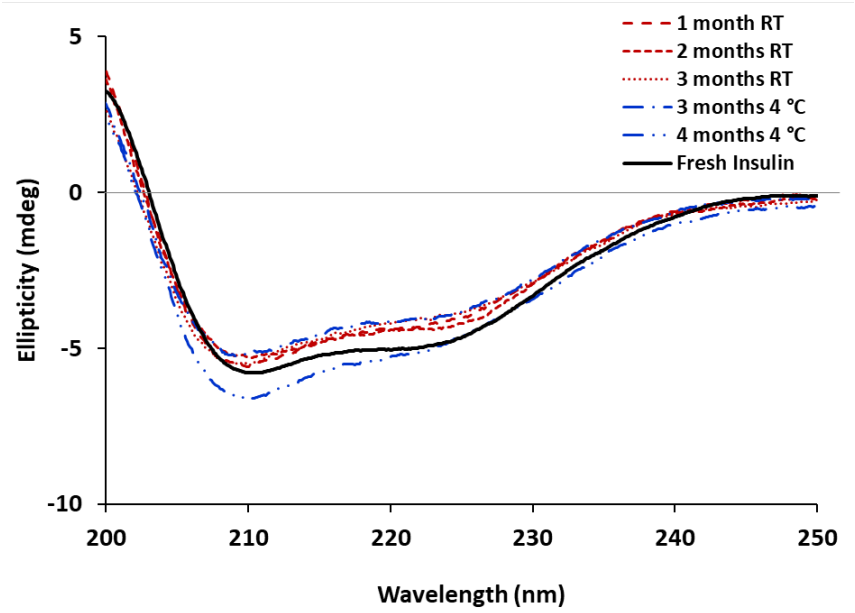


Figure 2.6. Insulin retains structure after extended storage in CAGE. Human insulin (100 U) was suspended in 1 ml CAGE and incubated under refrigeration (4 °C, blue dash-dot lines)) or at room temperature (25 °C, red dashed or dotted lines) for up to 4 months. Circular dichroism spectra in the far-UV region of each sample compared to a fresh insulin-PBS control (black solid line) showed no change in insulin secondary structure up to 3 months at RT (red dotted line) and 4 months at 4 °C (blue dash-dot-dot line).

2.5.2 In vivo activity

To confirm that insulin stored in CAGE for long periods not only remained structurally stable, but also retained its biological activity, we recovered insulin from CAGE via centrifugation and washing. Non-diabetic male Wistar rats were dosed via injection with 1 International Unit (U) per kg of the recovered insulin in saline. Compared to freshly prepared insulin solution, insulin stored in CAGE for up to 2 months at room temperature and 4 months under refrigeration showed a similar pattern in blood glucose change (**Figure 2.7**).

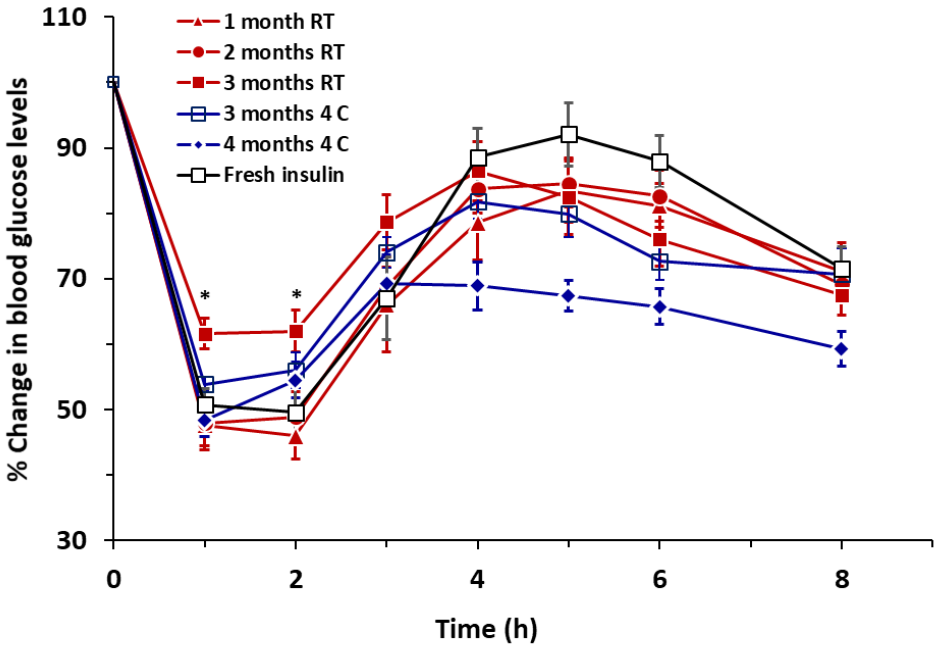


Figure 2.7. In vivo activity of insulin after long-term storage in CAGE. Insulin was mixed into CAGE and stored at room temperature (RT) or under refrigeration (4 °C). Aliquots were sampled every month. RT samples; 1 month (red triangles), 2 months (red circles), 3 months (red squares). Refrigerated samples; 3 months (hashed blue squares), 4 months (blue diamonds). All samples were compared to fresh insulin in PBS (open black squares). *p<0.05 for insulin at 3 months RT compared to fresh insulin.

Loss of insulin potency can result from aggregation, hydrolysis, or intermolecular transformations (Brange & Langkjoer, 1993). Protic ionic liquids are reported to stabilize insulin's conformation and mitigate its tendency to aggregate (Micaelo & Soares, 2008) (Kumar, 2013). Ionic liquids can also reduce water-protein interactions, thereby increasing structural stability by reducing hydrolytic activity. CAGE may also play a role in preventing aggregation, which enables efficacy by maintaining insulin in its monomeric, bioactive form.

2.6 Summary

The physical characterization of ILs is important from two standpoints; first, to provide a relevant set of descriptors to compare them among each other and with those described in

literature, and second, to use in quantifying how different environments (water, air, pH) change the properties of ILs. This second objective becomes a priority when using ILs in applications like microbial neutralization and protein stability/delivery, which we discuss in the following chapters.

Chapter 3

3. CAGE as an antimicrobial agent

3.1 Motivation

The development of robust, broad-spectrum antibiotics that do not trigger microbial resistance continues to be one of the major challenges facing the healthcare community today. ILs, with their intrinsic antimicrobial qualities and highly tunable nature, have the potential to address this need. As described earlier, ILs are a broad class of compounds consisting of ions and most commonly described by their low melting points (<100 °C) and non-volatility. IL cations (e.g. imidazolium, pyridinium, quaternary ammonium and phosphonium) can be combined with a variety of anions to create a diverse set of compounds with properties that can be further tuned via functionalization of either ion. While the activity of ILs against Gram-negative and Gram-positive bacteria, mycobacteria and fungi is well documented, the mechanism of action remains unknown. ILs are structurally analogous, and therefore likely have similar antimicrobial mechanisms to established cationic biocides and surfactants such as benzalkonium chloride, a quaternary ammonium compound (Pendleton & Gilmore, 2015). Of particular interest to this study, several groups have reported on the antibacterial properties of ILs containing functionalized choline, another quaternary ammonium cation, combined with a variety of counterions (Pernak, et al., 2007) (Petkovic, et al., 2010) (Zakrewsky, et al., 2014) (Siopa, et al., 2016) (Zhao, et al., 2015).

The systematic evaluation of several different cation/anion pairings that resulted in the development of CAGE showed it possessed superior antimicrobial activity and low cytotoxicity to human cells (Zakrewsky, et al., 2014). The envisioned application for CAGE is

as an antimicrobial agent to treat infections in humans. In earlier work, CAGE was tested for antimicrobial efficacy against 47 different strains of Gram-negative and Gram-positive bacteria, fungi and viruses. In all cases, low concentrations of CAGE provided complete pathogen neutralization, including against *E. coli* (5% CAGE), methicillin-resistant *S. aureus* (<1%), *C. albicans* (<1%) and Herpes Simplex Virus Type-1 (1%). In addition, the concentration of CAGE required to kill bacteria (*E. coli* and *S. epidermis*) was lower than the LC_{50} for human keratinocytes (Zakrewsky, et al., 2016).

These results are highly encouraging, and to leverage CAGE to its full advantage, we must fully understand the mechanism it uses to neutralize organisms. In 1996 Pernak and Skrzypczak reported a correlation between the concentration of an imidazolium chloride IL and its minimum inhibitory concentration against bacteria (Pernak & Skrzypczak, 1996). Other studies followed to confirm this relationship and provide a second mechanistic hypothesis, namely that an IL's antibacterial activity is correlated with the length of its alkyl chain (Pernak, et al., 2003) (Docherty & Kulpa Jr, 2005) (Pernak, et al., 2007) (Yu & Nie, 2011) (Jeong, et al., 2012). Some postulated that aliphatic chains of ILs insert into the bacterial membrane with a mechanism similar to that used by surfactants or pesticides. Others attributed the functionality of ILs to the inhibition of acetylcholinesterase because of the cation (Arning, et al., 2008) (Torrecilla, et al., 2009).

Molecular dynamics (MD) simulations have been used to gain insights into the action of select imidazolium ILs on model lipid bilayers (Bingham & Ballone, 2012) (Yoo, et al., 2016) (Klähn & Zacharias, 2013). These simulations showed that imidazolium cations interact with the polar head groups of the lipids and insert their hydrophobic tails into the membrane. The

interactions, however, are highly dependent on the charge and structure of the cation, the counter anion, as well as the complexity of the membrane lipids. Literature studies on ILs have focused on simple phospholipid bilayer models as surrogates for bacterial membranes due to the lack of available force fields for the Gram-negative bacterial outer membranes (Bingham & Ballone, 2012) (Yoo, et al., 2016) (Klähn & Zacharias, 2013). Moreover, the atomistic MD simulations have been limited to short timescales (hundreds of nanoseconds) that are unable to provide adsorption kinetics of an IL cation on bacterial membranes. In recent years, there have been advances in coarse-grained force field libraries for bacterial membranes (Ma, et al., 2015) (Hsu, et al., 2016) (Van Oosten & Harroun, 2016) (Ma, et al., 2017) (Ma, et al., 2017) that can aid in elucidating the IL-induced morphological reorganization of the bacterial membranes.

Despite these experimental and computational investigations into the interaction of ILs with lipid membranes, the exact molecular mechanism remains unknown, including the secondary effects of membrane disruption on cellular signaling and other cellular functions (Pendleton & Gilmore, 2015) (Bhattacharya, et al., 2017) (Bhattacharya, et al., 2018). The lack of a complete mechanistic description hampers the effective development of antimicrobial ILs for the treatment of infections, especially as it pertains to avoiding imparting resistance. Combining a full mechanistic knowledge with the fact that IL properties (hydrophilicity, hydrophobicity, density, viscosity, conductivity, and polarity) can be widely and readily tuned could provide a wealth of new IL-based antimicrobials with maximum efficacy and minimum toxicity.

3.2 Study design

We choose Gram-negative bacteria *Escherichia coli (E. coli)* as the starting point for our CAGE mechanistic studies. Using a combination of experimental and simulation techniques, we investigated the antibacterial activity and mechanism of the panel of CAGE variants on *E. coli* cells. Each variant was tested to determine the minimum bactericidal concentration (MBC) required to kill *E. coli*, and MD simulations by collaborators at Syracuse University were performed to compute the interfacial properties of CAGE variants with *E. coli*, as well as choline bicarbonate and pure geranic acid.

Using fluorescent flow cytometry, we determined that the loss of cell membrane integrity was an integral part of the mechanism and qualitatively analyzed the membrane damage with scanning electron microscopy (SEM). Fourier Transform Infrared spectroscopy (FTIR) was employed to investigate changes in the membrane lipid profile which provide evidence of membrane disruption. One variant, CAGE 1:2, was further tested for the potential to impart antibacterial resistance.

3.3 Antibacterial activity of CAGE against *E. coli*

The results of MBC testing showed a clear trend in bactericidal activity among the four CAGE variants, with increasing geranic acid/geranate content resulting in lower minimum bactericidal concentrations. While the MBCs of the variants were not significantly different from their nearest neighbors when ordered by geranic acid content, the extremes (1:4 and 2:1) are significantly different ($p<0.05$). All variants had significantly lower MBCs than either choline bicarbonate or geranic acid, suggesting a combination of the two ions is necessary for effective bacterial inactivation (**Figure 3.1**). Increasing geranic acid content resulted in a lower MBC.

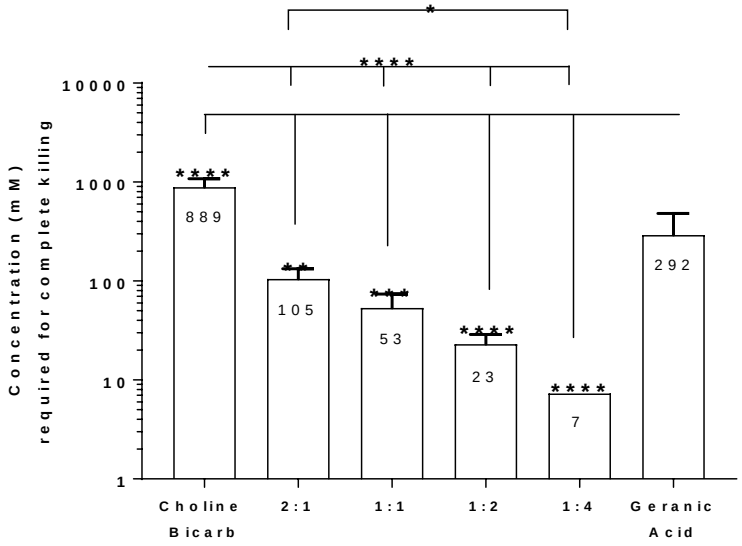


Figure 3.1. Determination of MBC of CAGE variants. *E. coli* cells were incubated in CAGE reactants or variants for 2 h at 37 °C. The MBC was determined as the lowest concentration with no growth on the plates. The MBC for the CAGE variants was lower than either of the starting choline bicarbonate or geranic acid, indicating a synergistic effect. Mean MBCs ± S.E. n=5, shown for each variant or reactant. *p<0.05, **p<0.01, ***p<0.001, ****p<0.0001.

The kinetics of bactericidal activity were studied in detail using one variant, CAGE 1:2, at concentrations below the MBC, specifically 13 (50% of a lethal dose) and 6.5 mM (25%), measuring optical density (OD_{600}, where 1 unit of absorbance at 600 nm is equivalent to 8×10^8 colony forming units (CFU) per mL of *E. coli*) over 14 h. *E. coli* did not grow in 13 mM as indicated by an unchanged cell density (**Figure 3.2 left)** but the culture was positive via overnight plating, indicating an immediate bacteriostatic effect. Bacteria treated at 6.5 mM exhibited slight growth at early time points, but the cell count plateaued at a number significantly lower than untreated cells, between 15 – 22% of the untreated culture.

To further evaluate the difference between bacteriostatic and bactericidal doses, we stained CAGE-treated *E. coli* cells with propidium iodide and analyzed the cells using flow cytometry (**Figure 3.2, right**). Untreated, live cells showed low PI staining, while heat-killed cells exhibited the expected high PI fluorescence. Cells treated with a bacteriostatic dose of 8

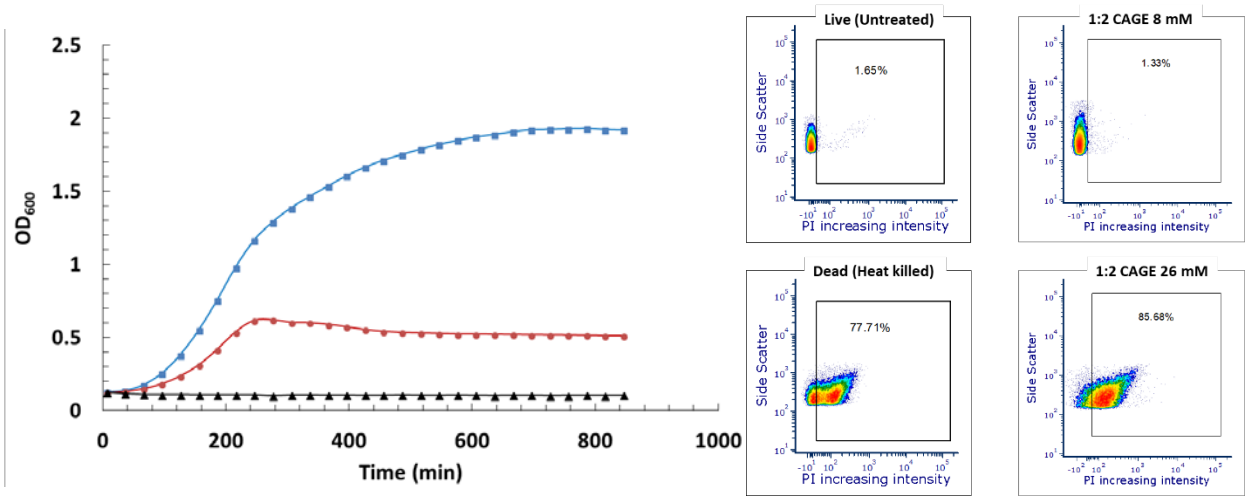


Figure 3.2. Lower concentrations of CAGE are bacteriostatic. While 26 mM 1:2 CAGE completely kills E. coli, lower doses are bacteriostatic. Left: Cells treated with 13 mM (black triangles) inhibit growth immediately, while those treated with 6.5 mM (red circles) exhibit abnormal growth up compared to untreated cells (blue squared) up to about 4 h, after which growth is halted. Starting cell loading was 3.4×10^7 CFU/mL and cell viability was confirmed via overnight plating. Right: *E. coli* cells were incubated for 2 h at 37 °C at either a lethal (26 mM) or sub-lethal (8 mM) dose. Untreated cells and heat-killed cells served as negative and positive controls, respectively. Cells were then separated and washed with buffer followed by staining with propidium iodide (PI), and flow cytometry was used to quantify the PI-stained populations. Population gates were set using unstained samples (not shown).

mM for 2 h showed low PI intensity indicating cell viability, while cells at 26 mM CAGE were significantly stained by the PI. Cell viability was confirmed via overnight culturing in media. Flow cytometry results support the kinetic results. A 2 h treatment of 26 mM CAGE resulted in a PI intensity (85.68%) similar to heat-killed cells (77.71%), indicating significant cell membrane disruption. In contrast, a sub-lethal, or bacteriostatic dose of 8 mM for 2 h resulted in few PI-positive cells (1.33%) similar to the untreated control (1.65%). This is consistent with the results of the kinetic study, which showed that cells were viable and able to replicate for up to 3-4 h in a bacteriostatic dose of 6.5 mM 1:2 CAGE.

There are no previous studies on cholinium-geranic acid salts, however some groups have investigated other choline-based ILs for their bacterial activity, and our results are generally consistent with these studies. Petkovic (Petkovic, et al., 2010) synthesized a group of ILs using

a choline cation paired with a range of linear alkanoate anions ($[C_nH_{2n+1}CO_2]^-$, where n=1-9) and found that the longer anion chains resulted in lower minimum fungicidal concentrations. Choline chloride, tested as a proxy for the choline cation alone, showed the lowest toxicity. Zhao (Zhao, et al., 2015) synthesized a variety of choline-based DESs using choline chloride and several different types of hydrogen-bond donors including organic acids, amines, alcohols and sugars; only the organic acid-containing DES showed bacterial inhibition. The CAGE mechanism of membrane attraction and insertion, while similar to those suggested for cation-substituted ILs, has a unique feature – the hydrophobic long chain can dissociate from the more bulky, hydrophilic cation and penetrate deeper into the membrane. This dissociation ability may prove useful in developing highly effective antimicrobial ILs. It is also interesting to note that geranic acid is structurally similar to free fatty acids, which have demonstrated bioactivities related to chain length and degree of saturation (Desbois. & Smith, 2010), but poor solubility. Combining choline, a hydrophilic molecule, with hydrophobic geranic acid may improve its ability to contact cells in aqueous environments like wounds.

3.4 MD simulation of CAGE-cell membrane interactions

Molecular Dynamic (MD) simulations were performed by Huilin Ma in Prof. Nangia's lab at Syracuse University. The simulations show that the choline geranate pair exhibits a cooperative profile against the *E. coli* membrane. The choline cation, with its short side chain is small enough to penetrate the lipopolysaccharide (LPS) domain and form stable ionic interactions with the negatively charged membrane. This facilitates the penetration of geranate, which inserts its long chain, acting like a short chain fatty acid, into the lipid A tails while the charged head group remains above. As the concentration of geranic acid/geranate increases, the amount

in the membrane increases (**Figure 3.3 b through e**). In the case of pure geranic acid (**Figure 3.3 f),** the penetration is significantly lower than the CAGE variants, suggesting that choline is necessary.

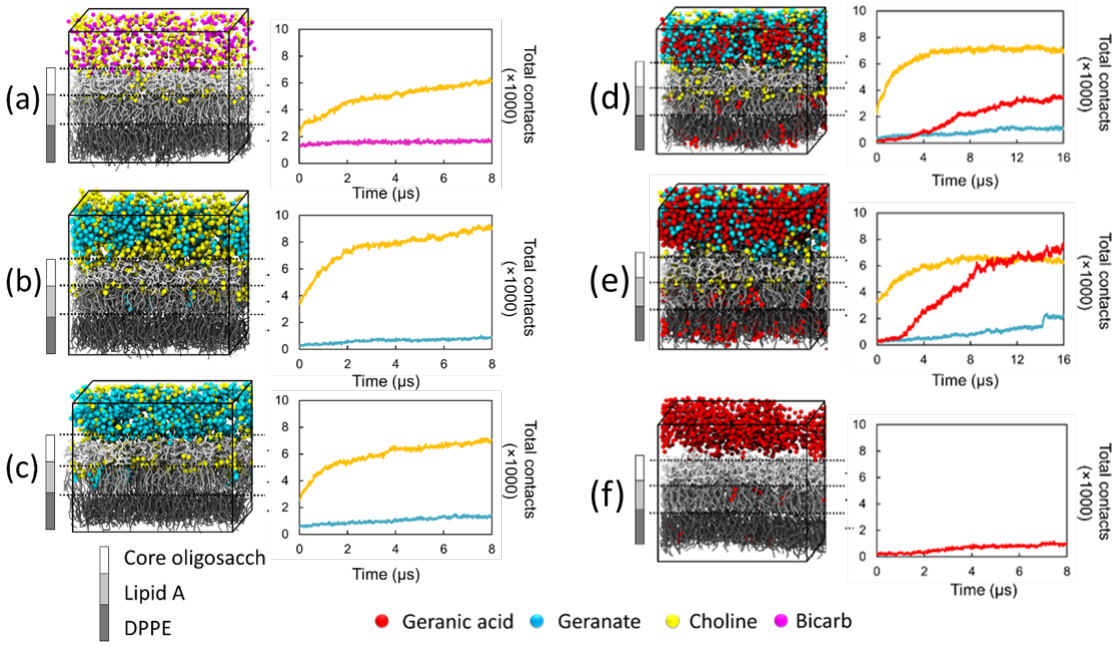


Figure 3.3. MD simulation snapshots and contact plots for CAGE interactions with the E. coli cell membrane. (a) Choline bicarbonate, (b) CAGE 2:1, (c) CAGE 1:1, (d) CAGE 1:2, (e) CAGE 1:4, (f) pure geranic acid.

3.5 Visualization of membrane disruption via SEM

To visualize the cell membranes after a lethal (MBC concentration) CAGE treatment, cells were imaged using SEM. The images show a significant change in morphology in cells incubated for 2 h in all CAGE variants (**Figure 3.4**). Cell surfaces had a roughened, flaked appearance compared to the smooth surface of untreated cells. Only a few cells in each sample were completely ruptured. The 2:1 CAGE treatment resulted in cells that appeared more bubbled or swollen compared to other variants.

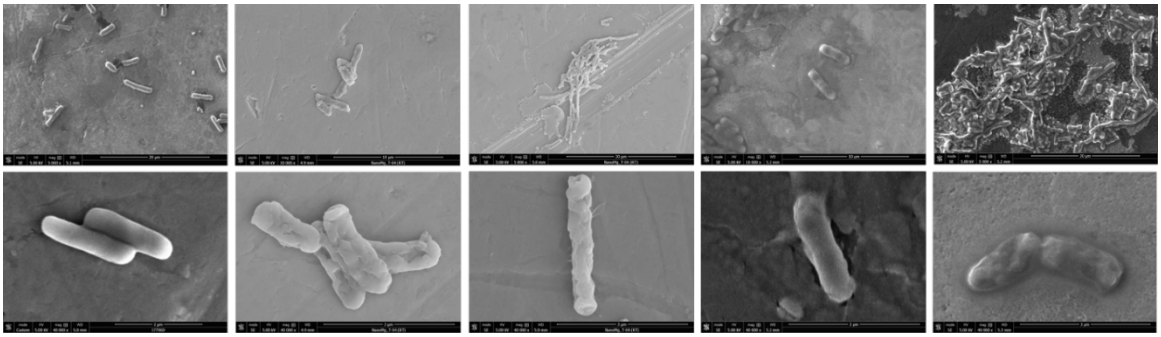


Figure 3.4. Investigation of membrane disruption via SEM. *E. coli* cells were incubated in CAGE variants for 2h at 37 °C. Cells were fixed, washed and ethanol dehydrated followed by ambient drying prior to SEM. Images show surface disruption on CAGE-treated cells compared to an untreated control. Magnification of top row images is 5,000 – 10,000× and bottom row images is 40,000×. From left to right: untreated cells, CAGE 1:4, 1:2, 1:1 and 2:1 variant.

3.6 Effect of CAGE on the cell membrane lipid profile

FTIR has been shown to be a powerful tool in analyzing biological systems, and of interest for this study, lipid content in the cell membrane (Zoumpopoulou, et al., 2010) (Papadimitriou, et al., 2008) (Lin, et al., 2004). To look for alterations in the lipid content of the cell membrane due to CAGE treatment, we used FTIR with an attenuated total reflection (ATR) crystal to collect the spectra of cells grown for 2 or 24 h in a sub-lethal dose (13 mM) of the 1:2 CAGE variant. In a sub-lethal dose, the cells remain viable, but under stress, and sub-cellular responses, e.g. increased lipid content, can be investigated in addition to the effect of geranic acid insertion. Both the 2 and 24 h treated samples showed increased lipid peak heights compared to untreated cells incubated for the same time in medium only (**Figure 3.5**). While the 2 h peak increases were modest, the 24 h data showed a significant increase in peak heights associated with lipid bond vibrations, namely 2,853 (symmetric CH_2), 2,874 (anti-symmetric CH_3), 2,924 (anti-symmetric CH_2) and 2,959 cm^{-1} (symmetric CH_3).

Additionally, the frequency of the symmetric CH_2 stretching increased in the 2 and 24 h challenged cells (2,855 cm^{-1}) compared to the untreated cells (2,853 cm^{-1}). A slight upward

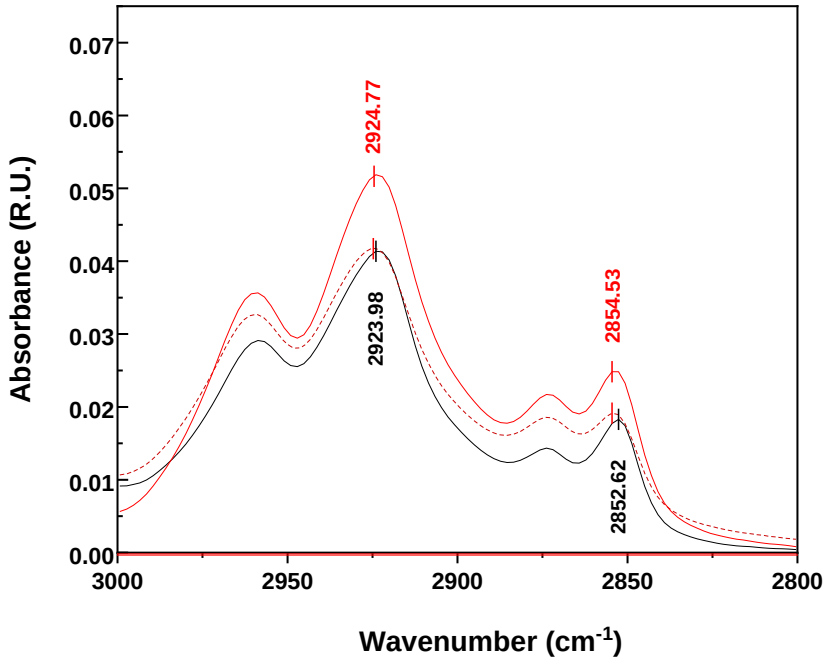


Figure 3.5. FTIR-ATR spectra of CAGE-treated *E. coli* suggest a change in membrane content and order. Cells incubated in sub-lethal doses of 1:2 CAGE for 2 (dashed red line) or 24 h (solid red line) exhibited both an increase in peak heights indicative of lipid content compared to an untreated control (solid black line), and an increase in vibrational frequency of CH_2 stretching was observed at 2,853 and to a lesser extent at 2,924 cm^{-1}. Cell count was kept constant between all samples.

shift was observed in the frequency of anti-symmetrical CH_2 stretching around 2,924 cm^{-1}. The frequencies of these band shifts are known to correspond to a change in the conformational order of lipid acyl chain systems, *e.g.* from an ordered gel to a disordered liquid-crystalline phase upon melting (Tamm & Tatulian, 1997). The vibrational modes originating from the terminal methyl groups at 2,959 (anti-symmetric) and 2,874 (symmetric) cm^{-1} are also conformation sensitive, although to a lesser degree (Mantsch & McElhaney, 1991). No shifts were observed in the methyl group peak frequencies in this study.

In the FTIR spectra of cells treated with sub-lethal doses of CAGE for 2 and 24 h, there was a clear and reproducible (seen in all four 24 h samples) increase in the vibrational frequency of symmetric CH_2 stretching in the CAGE-treated cells. A shift in the anti-

symmetric CH_2 stretching frequency was seen in one out of four 24 h samples. An increase in the symmetric and anti-symmetric CH_2 vibrational frequencies is seen upon the transition of a phospholipid system from an ordered gel to a disordered crystalline phase (Cameron, et al., 1980) (Cortijo, et al., 1982). The symmetric band is usually monitored because it is freer from overlap by other vibrational modes and shifts about 2 cm^{-1} as a result of the chain melt process (Mantsch & McElhaney, 1991). Using model lipid systems, two early studies showed that incorporation of cholesterol into dipalmitoylphosphatidylcholine (DPPC) bilayers decreased the conformation order of the lipids and increased symmetric CH_2 vibrational frequencies by about 2 cm^{-1} (Cortijo, et al., 1982) (Reis, et al., 1996), while Choi (Choi, et al., 1991) saw an increase of about 1 cm^{-1} of the anti-symmetric CH_2 stretching frequency with the addition of cholesterol to phosphatidylserine bilayers that introduced disorder in the acyl chain packing. Tsai (Tsai, et al., 1989) applied polar anesthetics to phospholipid membranes and saw a similar increase in the CH_2 anti-symmetric stretching frequency. Hence, the 1.9 cm^{-1} symmetric and 0.8 cm^{-1} anti-symmetric frequency shifts seen in this study indicate a change in the conformational order of the lipid layers in the *E. coli* membrane.

The FTIR spectra for the CAGE-treated cells also showed increased lipid peak heights for CH_2 and CH_3 stretching frequencies; the increases at 2,855 (symmetric CH_2), 2,874 (symmetric CH_3), 2,924 (anti-symmetric CH_2) and 2,959 cm^{-1} (anti-symmetric CH_3) are slight in the 2 h sample but signficant in the 24 h incubation sample. While the MD simulations show the extraction of lipids from the membrane as a result of CAGE, there may be a cellular response to stress manifested as changes to the lipid content that is not represented in the simulation. Corte (Corte, et al., 2015) reported an increase in the height of CH_2 symmetric and anti-symmetric peaks in surfactant-treated *E. coli*. Other groups have found that low pH

environments can cause changes in phospholipid fatty acid type in bacterial and fungal membranes, presumably as a means to maintain membrane integrity and functionality (Brown, et al., 1997) (Chen & Ganzle, 2015) (Suchodolski, et al., 2017) (Hartmann & Silva Pereira, 2013).

3.7 Resistance challenge

The potential for the *E. coli* to develop resistance to 1:2 CAGE via vertical evolution (i.e. mutations passed on to progeny) was investigated using two different techniques: 1) short, repeated sequential challenges, and 2) with extended one-time challenges. Based on the kinetic study, we knew that the cells did not undergo many doublings, so the effect would likely be to exert pressure on the current population to create mutations that would carry to their progeny. In 6 sequential 2 h challenges, the MBC for 1:2 CAGE did not change from 26 mM (test range was from 6.5 – 52 mM) (**Figure 3.6 left**). Here, the cells used in each successive challenge were taken from the plate with growth from the highest CAGE concentration in the previous challenge. When cells from the ultimate sequential challenge were cultured in regular media and analyzed via FTIR, there was no difference in lipid profile from the wild type culture (**Figure 3.6 right**).

To test whether growing bacteria for extended periods under selection pressure imparted resistance, cells were grown in media and then subcultured into 13 mM 1:2 CAGE and incubated for either 24 or 48 h. The final OD_{600} of the 24 and 48 h samples were 1.4 and 1.5, respectively, from a starting OD_{600} of 0.05, indicating at best 5 doublings. Following incubation, the cultures were plated to confirm viability, and colonies from those plates were

grown and the MBC determined. Both the 24 and 48 h cultures had viable cells, but those cells did not exhibit a higher MBC (**Figure 3.6 left**).

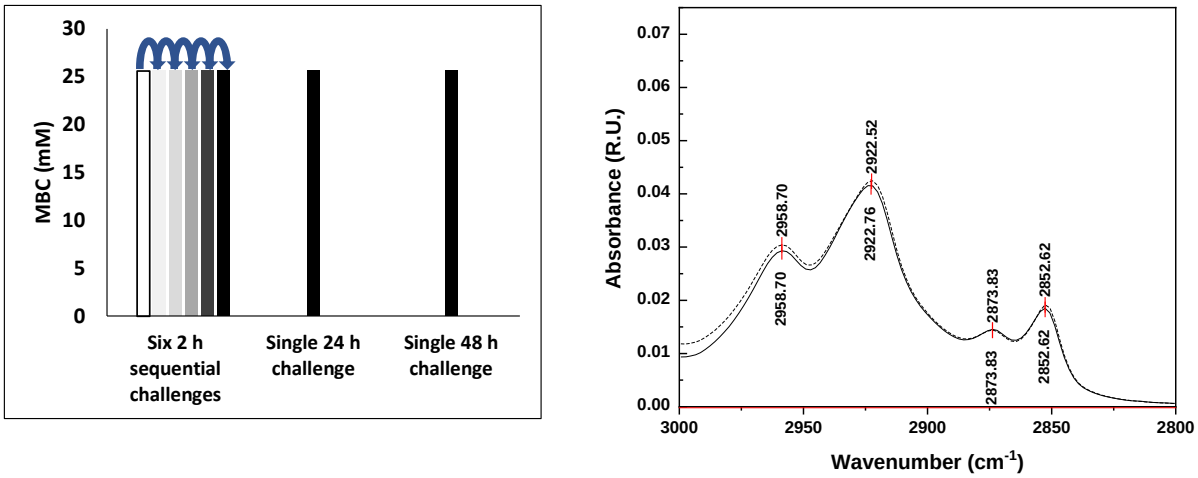


Figure 3.6. MBC and lipid profile did not change when cells were exposed to sub-lethal doses of 1:2 CAGE. Left: 2 h sequential challenges: *E. coli* cells were exposed to varying concentrations of 1:2 CAGE (6.5 – 52 mM) and then plated overnight. A colony from the highest concentration plate with growth was cultured and challenged again, for a total of 6 challenges. Light to dark bars represent the first to sixth challenge. 24 and 48 h challenges: Cells were grown in sub-lethal 1:2 CAGE for 24 or 48 h and the MBC was determined. The MBC did not change as a result of either type of challenge. Right: Cells from the final sequential challenge cultured in standard media (dashed line) exhibited the same lipid profile as untreated cells (solid line).

It has been suggested that sub-lethal doses of antibiotics lead to multidrug resistance due to the production of reactive oxygen species leading to an increased mutagenesis rate (Kohanski, et al., 2010). Allowing subsequent generations to recover in regular media after growing under selection pressure arguably increases the chance that changes to MBC values will be due to genomic mutations. We did not see evidence of developing resistance, as demonstrated by a constant MBC concentration across 6 sequential 2 h challenges and two longer (24 and 48 h) challenges. In addition, the changes in the lipid profile seen in cells grown in media spiked with CAGE (**Figure 3.5**) were not present in CAGE-challenged cells that were subsequently grown in regular media (**Figure 3.6 right**), suggesting that the lipid profile changes were due to environmental pressure rather than genetic alterations. This is not

surprising, since CAGE's target is the membrane; mutations to several genes would have to occur to produce a mutated functional membrane. It should be noted that vertical evolution is only one means of resistance acquisition. This was recently addressed by Sommer, who argues that the current methods for testing are overly focused on the *in vitro* acquisition rate of spontaneous mutations (via vertical transfer) that confer resistance and undervalue the role of horizontal gene transfer in resistance acquisition. Sommer also suggests that vertical evolution testing should include fitness cost evaluation, since some mutations, while imparting good resistance, may be too costly for the organism (Sommer, et al., 2017).

3.8 Summary

A full understanding of how a candidate antibiotic acts against a pathogen is of upmost importance to develop antibiotics with high efficacy and low potential to impart resistance. Since the properties of ILs can be finely tuned, a full mechanistic knowledge of their action of cell disruption can allow us to the design a wealth of IL-based antimicrobials.

Mechanistic hypotheses for ILs' antibacterial activity most commonly include cell membrane disruption as a result of interaction with the ionic species (Zakrewsky, et al., 2014) (Pernak & Skrzypczak, 1996) (Pernak, et al., 2003) (Hsu, et al., 2016) (Bhattacharya, et al., 2018). Some studies suggest additional signal interruptions as a result, but what actually causes cell death remains unknown. The presence of a negatively charged outer leaflet comprised of LPS makes the bacterial membrane unique compared to phospholipid bilayers and makes it susceptible to penetration by CAGE (**Figure 3.7**). In MD simulations, the density profiles and

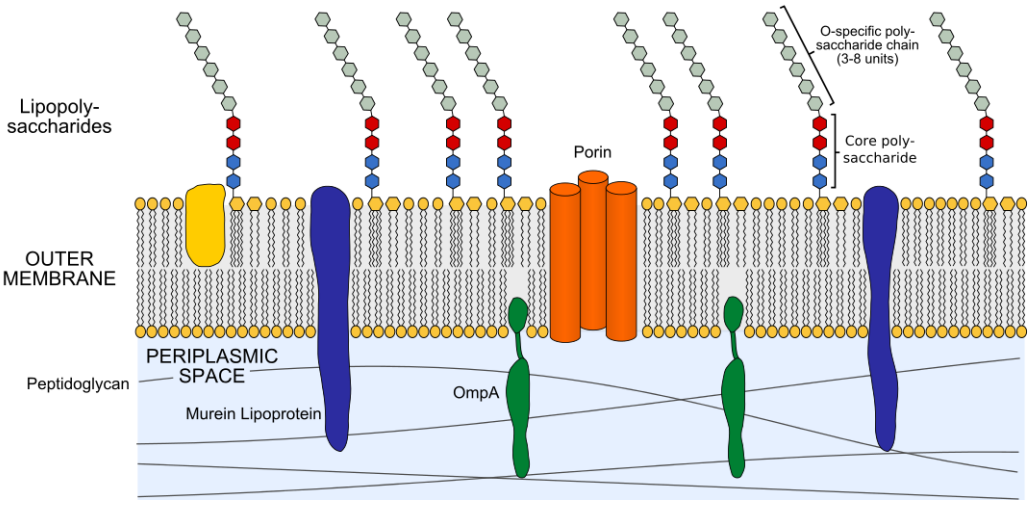


Figure 3.7. Structure of the outer wall of a Gram negative bacterial cell. Cross-section of the cell wall of a Gram negative bacterial cell showing the LPS, which is negatively charged, providing an attractive surface for positively-charged choline. Image credit: Jeff Dahl.

the contact analysis demonstrate that the negatively charged LPS core forms a barrier for geranate and geranic acid. In the presence of choline-containing CAGE variants, however, the LPS negative charge is effectively screened and choline is able to facilitate the geranate and geranic acid penetration into the membrane.

Among the six compounds simulated, 1:4 CAGE has the highest penetration, which explains the high toxicity observed in the experiments. Using the CAGE component penetration as a measure of their efficacy, the simulation results show the following order of CAGE variants toxicity: 1:4 > 1:2 > 1:1 > 2:1> choline bicarbonate > pure geranic acid. The order corroborates with the experimentally observed CAGE toxicity (**Figure 3.8**).

By varying the ion ratios in CAGE, we were able to show that increasing the geranic acid content increases the biocidal activity. Through MD simulations we identified cell membrane disruption via choline attraction to the negatively-charged cell membrane and geranic acid insertion as a disrupting mechanism, and qualitatively showed such disruption via SEM

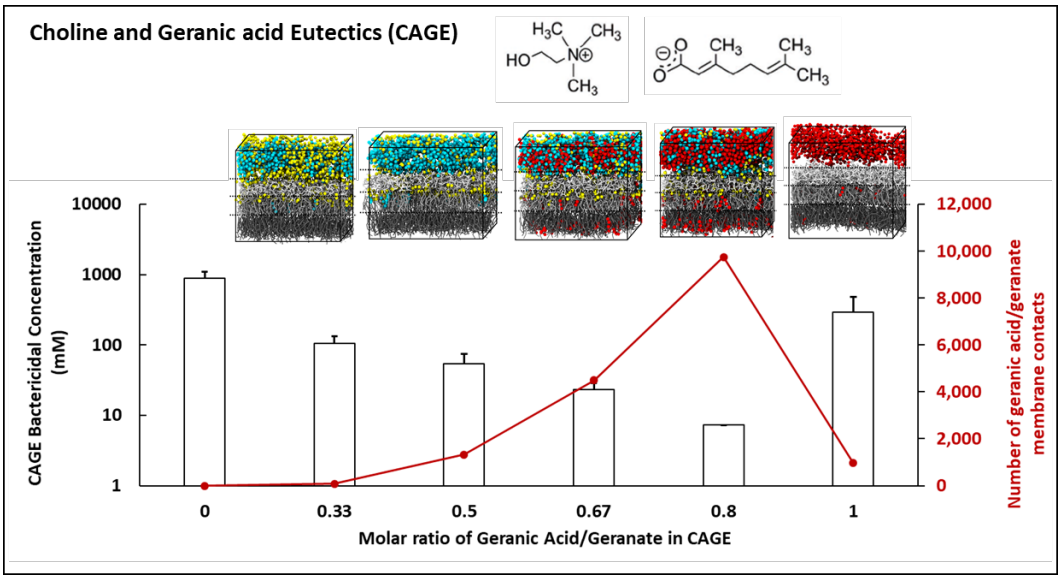


Figure 3.8. Increasing geranic acid/geranate content in CAGE results in greater toxicity to E. coli. While pure geranic acid does not effectively kill bacterial cells, a combination of choline and geranic does. The mechanism identified requires both ions to interact with and disrupt the membrane via geranic acid insertion, shown in the MD simulation snapshots for each CAGE molar ratio.

imaging. Overall, this study provides the basic mechanism for choline-based IL activity on the cell membrane of Gram-negative *E. coli*. Through kinetic studies and flow cytometry, we confirmed that low (6.5 -13 mM) concentrations of CAGE are bacteriostatic, and that cells can propagate for a few hours at delayed growth rates at the lowest concentration, while modestly higher concentrations (≥ 26 mM) are bactericidal. FTIR spectra confirm that treatment of *E. coli* cells with CAGE resulted in a more disordered membrane lipid organization. Finally, CAGE exhibited no short-term tendency to impart vertical evolution-based resistance to *E. coli.*

Chapter 4

4. CAGE as a transdermal drug delivery agent

4.1 Motivation

With its large area and easy accessibility, skin is an ideal candidate for delivery of therapeutic molecules. Drugs delivered transdermally are exposed to vasculature in the dermis for systemic uptake, avoiding first-pass metabolism (Brown, et al., 2006) (Prausnitz, et al., 2004) (Thomas & Finnin, 2004) and the epidermal layer contains a rich immunological community for vaccine activation (Kupper & Fuhlbrigge, 2004). While some molecules such as nicotine exhibit passive transport due to their small size (<500 Da) and lipophilicity and are currently available as patches, (Kalia, et al., 1998) (Prausnitz & Langer, 2008) (Bos & Meinardi, 2000) (Yano, et al., 1986) the majority of drugs of interest for dermal delivery, including peptides and proteins require methods to enhance their transport at efficacious concentrations. This is due to the extremely efficient barrier created by the highly organized and dense "brick and mortar"

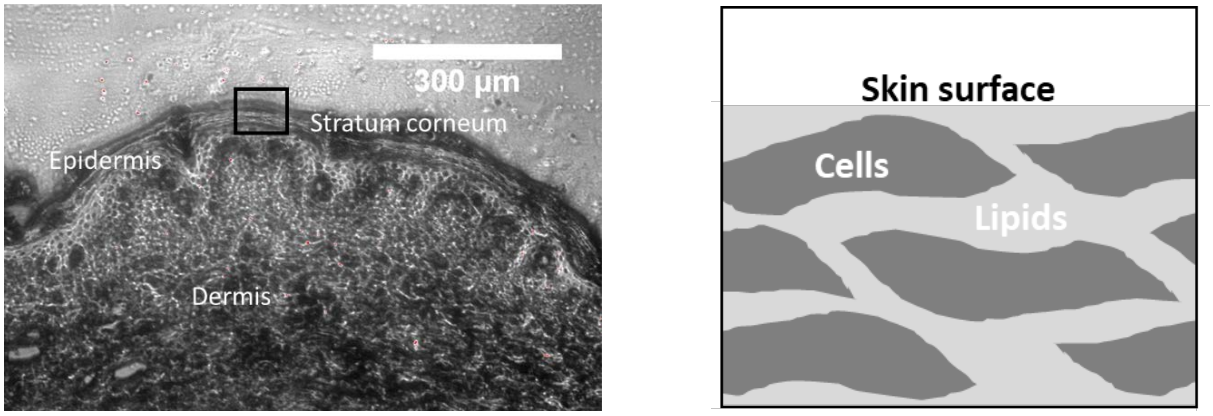


Figure 4.1. Porcine skin structure. Porcine skin structure is remarkably like human skin, making it a good *ex vivo* model for transdermal drug delivery. The primary skin layers are, from the top, the stratum corneum, epidermis and dermis (left). An enlarged section of the SC (black box, left) would look similar in structure to the cartoon on the right, showing the brick and mortar structure made up of corneocytes (dead cells) and lipid bilayers. Magnification of left image: 40X.

structure of the stratum corneum (SC), the topmost layer of skin (**Figure 4.1**) (Bouwstra & Ponec, 2006) (Barry, 2004) (Lee, et al., 2009). The SC consists of dead cells called corneocytes made up of dense keratin fibrils (bricks) about 1 μm in width and 25 μm long and spaced about 0.05 μm apart, with multilaminar lipid layers in between (mortar). The lipid layers are made up of ceramides, fatty acids and cholesterol. Several physical and chemical penetration enhancers (CPE) have been developed to disrupt the SC barrier in different ways, improving permeability of large or hydrophilic molecules (**Table 4.1**) (Kalia, et al., 2004) (Prausnitz, et al., 1993) (Mitragotri, et al., 1995) (Mikszta, et al., 2002) (Prausnitz, 2004) (Lee, et al., 2002) (Karande, et al., 2004) (Williams & Barry, 2004) (Kim, et al., 2012) (Lopes, et al., 2015).

Table 4.1 Permeation enhancers for dermal delivery

Physical (device-based)	***Chemical (formulation-based)***
Iontophoresis/electroporation	Glycols
Sonophoresis (ultrasound-based)	Alcohols
Microneedles	Surfactants
Laser ablation	Fatty acids

Limitations exist for both types of enhancers; device-based systems can typically only be used on a small fraction of available skin area, and formulation-based systems can exhibit toxicity. Both systems can result in chronic disruption of the SC, leaving the body vulnerable to infection (Karande & Mitragotri, 2009) (Pathan & Setty, 2009) (U.S. Food and Drug Administration Report, 2016) (Zakrewsky, et al., 2016) (Zakrewsky, et al., 2014). Previous investigations of CAGE for cytotoxicity suggest it is benign to primary keratinocytes, making it an ideal permeation enhancer for dermal delivery (Zakrewsky, et al., 2016).

Dermal delivery of insulin has been studied by several groups. Ling & Chen used dissolving polymer microneedle patches to deliver insulin to diabetic rats (Ling & Chen, 2013), while Pettis et al. used a microneedle-based intradermal system to deliver Lispro® insulin, showing a faster insulin absorption and action than a subcutaneous injection (Pettis, et al., 2011). Hadebe studied transdermal delivery of insulin by an amidated pectin hydrogel matrix patch with up to 0.144 U per patch (Hadebe, et al., 2014).

4.2 Study design

Previous studies showed the ability of CAGE to deliver small molecules mannitol and cefadroxil (Zakrewsky, et al., 2014); this study sought to expand the range of molecules able to permeate across skin with CAGE to larger molecules, with a wide range of hydrophobicities. To further determine the ability of CAGE to enhance permeation of macromolecules across porcine skin *ex vivo*, we selected three biomolecules of varying size and hydropathicities; insulin (5.8 kDa, 0.278 GRAVY score), bovine serum albumin (~66 kDa, -0.429) and ovalbumin (~45 kDa, -0.001) and qualitatively investigated the ability of CAGE 1:2 to enhance their transport across the SC via confocal imaging. Transport of insulin and BSA were further quantified by measuring the transport of fluorescein isothiocyanate (FITC)-Insulin or tritium-labelled (^{3}H-BSA). The efficacy of CAGE 1:2 was compared with two well-known CPEs, ethanol and diethylene glycol monoethyl ether (DGME) (Karande & Mitragotri, 2009) (Trommer & Neubert, 2006) (Javadzadeh, et al., 2015) (Mura, et al., 2000). *In vivo* studies were conducted to examine the effect of CAGE 1:2 and insulin topically applied to rats on blood glucose levels.

The four CAGE variants, along with choline and geranic acid were used to qualitatively determine the effect of increasing geranic acid content on permeation of insulin to, and the lipid profile of, porcine skin *ex vivo*. Confocal imaging revealed an obvious dependence of permeation with geranic acid content in the variants. FTIR analysis of SC samples was conducted to determine if a similar trend existed in lipid content after incubation in CAGE.

4.3 Delivery of proteins into porcine skin *ex vivo*

Using confocal microscopy, penetration of fluorescein isothiocyanate (FITC)-labeled BSA, OVA and insulin into the epidermis and dermis was evaluated and found to be significantly higher in the presence of CAGE compared to PBS (**Figure 4.2**), suggesting CAGE facilitates

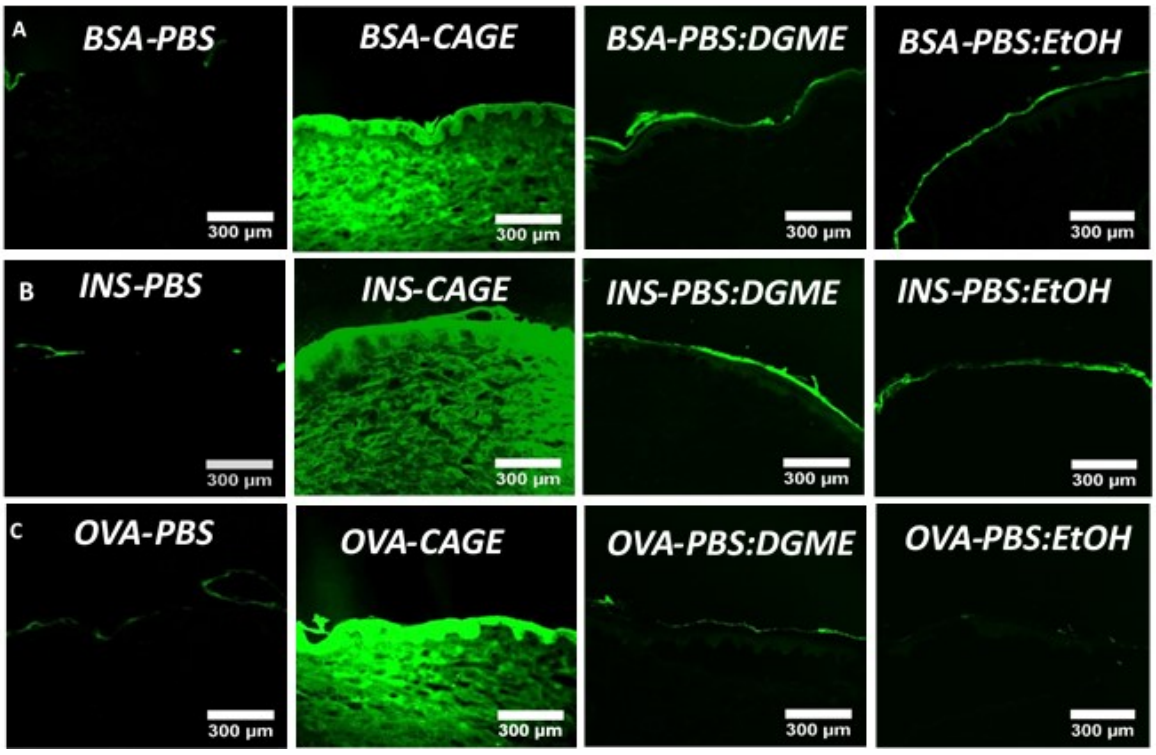


Figure 4.2. Confocal images of skin penetration of fluorescently labeled proteins in PBS, CAGE 1:2, 50:50 v/v PBS:DGME or 50:50 v/v PBS:Ethanol. A) FITC-BSA, B) FITC-Insulin, or C) FITC-OVA. FITC-proteins were suspended in 1 mL of PBS, CAGE, PBS:DGME or PBS:Ethanol and applied to porcine skin in a Franz diffusion cell. After 24 h at 37 °C, confocal imaging of 20 µm sections at 10X magnification showed FITC-protein penetration into the dermis layer for the CAGE samples but not in the PBS or CPE controls. Scale bar: 300

the delivery of large proteins into deep skin layers. The penetration of BSA, OVA and insulin from CAGE was compared with that from DGME (Transcutol®) and ethanol (EtOH). In all cases, penetration of BSA, OVA and insulin was significantly higher in the presence of CAGE than the CPEs, implying that CAGE is more effective in enhancing permeation of proteins across the skin compared to these widely used CPEs.

The ability of CAGE to enhance macromolecule skin penetration was quantitatively assessed using tritium (^{3}H) labeled BSA (**Figure 4.3**) and FITC-labeled insulin (**Figure 4.4**). CAGE significantly enhanced ^{3}H-BSA penetration into and through the deep layers of porcine skin (measured as ^{3}H-BSA in the acceptor fluid of the Franz diffusion cell). Specifically, ^{3}H-BSA content in the epidermis (2.84 ± 0.59 μg/cm^2) and dermis (1.61 ± 0.09μg/cm^2) was significant after 12 h compared to PBS controls (epidermis: 0.080 ± 0.01 μg/cm^2 and dermis: 0.04 ± 0.04 μg/cm^2). The ^{3}H-BSA content in 24 h in the dermis of CAGE-BSA treated skin was similar to earlier time points at 1.178 ± 0.01 μg/cm^2 but was significantly higher in acceptor compartment (3.497 ± 0.36 μg/cm^2). After 48 h, ^{3}H-BSA content in the dermis (1.72 ± 0.17 μg/cm^2) and acceptor was increased (6.02 ± 1.04 μg/cm^2), indicating that CAGE-mediated transport through the dermis may be time dependent (**Figure 4.3A**). Total amounts of ^{3}H-BSA transported into the skin after 6 and 12 h were distributed approximately evenly between the SC and epidermis/dermis/acceptor (E/D/A) (6 h: 45% in SC, 55% in epidermis/dermis/acceptor, 12 h: 51% in SC, 49% in epidermis/dermis/acceptor). At 24 and 48 h, the distribution was significantly skewed towards the deeper skin layers, with at least 80% present in the epidermis, dermis or acceptor fluid (**Figure 4.3B**). The transport rate of BSA was found to be 0.19 μg/cm^2/h with a total transport of 9.34 μg/cm^2 after 48 h. Skin

penetration of ^{3}H-BSA in PBS was negligible at all time points; testing was suspended after 24 h (**Figure 4.3C**). Compared to CPEs such as ethanol and DGME at 50% concentrations,

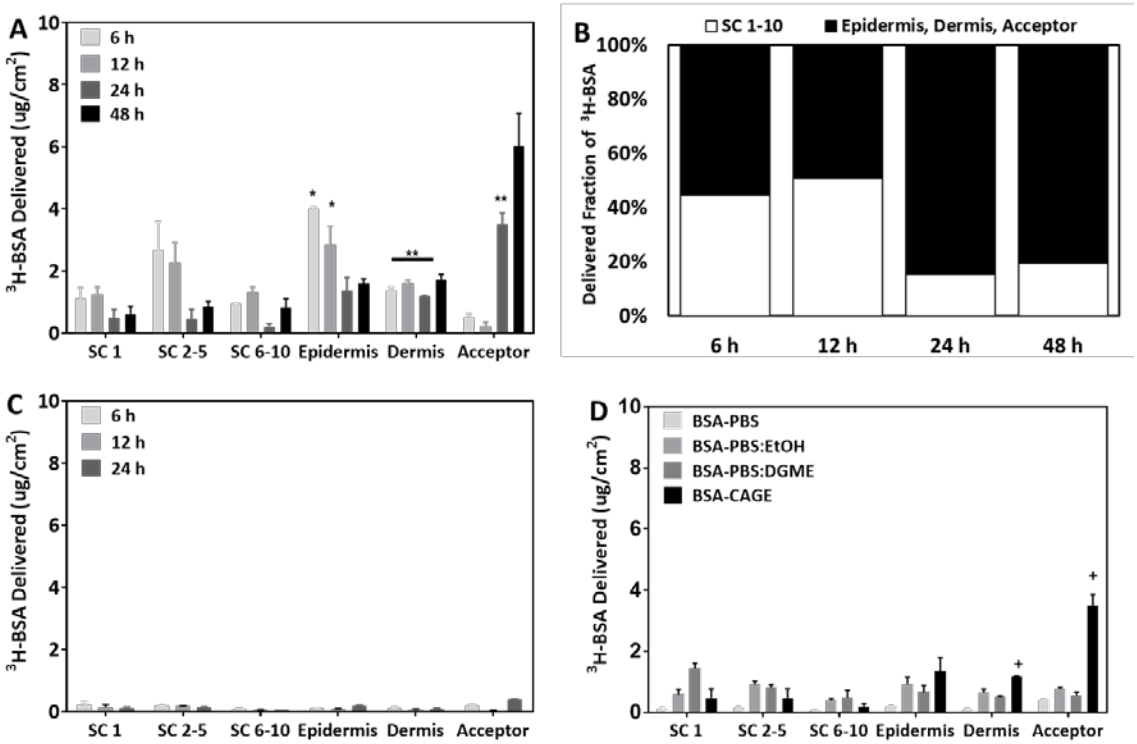


Figure 4.3. CAGE improves skin penetration of high molecular weight BSA. (A) CAGE enhances BSA penetration compared to PBS control into and across the dermis. Bars represent 6, 12, 24 and 48 h penetration. *$p < 0.04$, **$p < 0.005$. (B) Total penetration amount of BSA into skin after CAGE treatment. White and black bars represent the total of SC1-10 and combination of the epidermis, dermis and acceptor, respectively. (C) BSA penetration into individual layers of porcine skin *ex vivo* is negligible in PBS control. Skin depth increases from left to right where sample SC 1 is the first tape-stripped layer of the SC, SC 2-5 is the combination of the next four layers, and SC 6-10 is the deepest layers. (D) Enhancement of BSA penetration into and across porcine dermis is significantly improved with CAGE compared to 50:50 v/v PBS:Ethanol ($p < 0.05$ for dermis, $p < 0.02$ for acceptor) or PBS:DGME ($p < 0.001$ for dermis, $p < 0.01$ for acceptor) after 24 h. Data represented as mean ± S.E. (n=3).

neat CAGE exhibited significantly higher penetration of BSA into and across the dermis (**Figure 4.3D**).

CAGE also significantly enhanced FITC-insulin penetration into porcine skin, with significant concentrations in the dermis at 24 h: 5.46 ± 0.04 μg/cm^2 (**Figure 4.4A**) and 48 h:

17.22 ± 1.43 μg/cm^2 (**Figure 4.4B**). A larger portion of the insulin fraction delivered was present in the epidermis/dermis/acceptor fluid for insulin compared to BSA (93% at 24 and 96% at 48 h). The difference in molecular weights between BSA and insulin may have contributed to this effect (**Figure 4.4C**). For insulin, the rate after 48h was 0.57 μg/cm^2/h with a total transport of 28.35 μg/cm^2.

4.4 Transdermal delivery and increasing geranic acid content

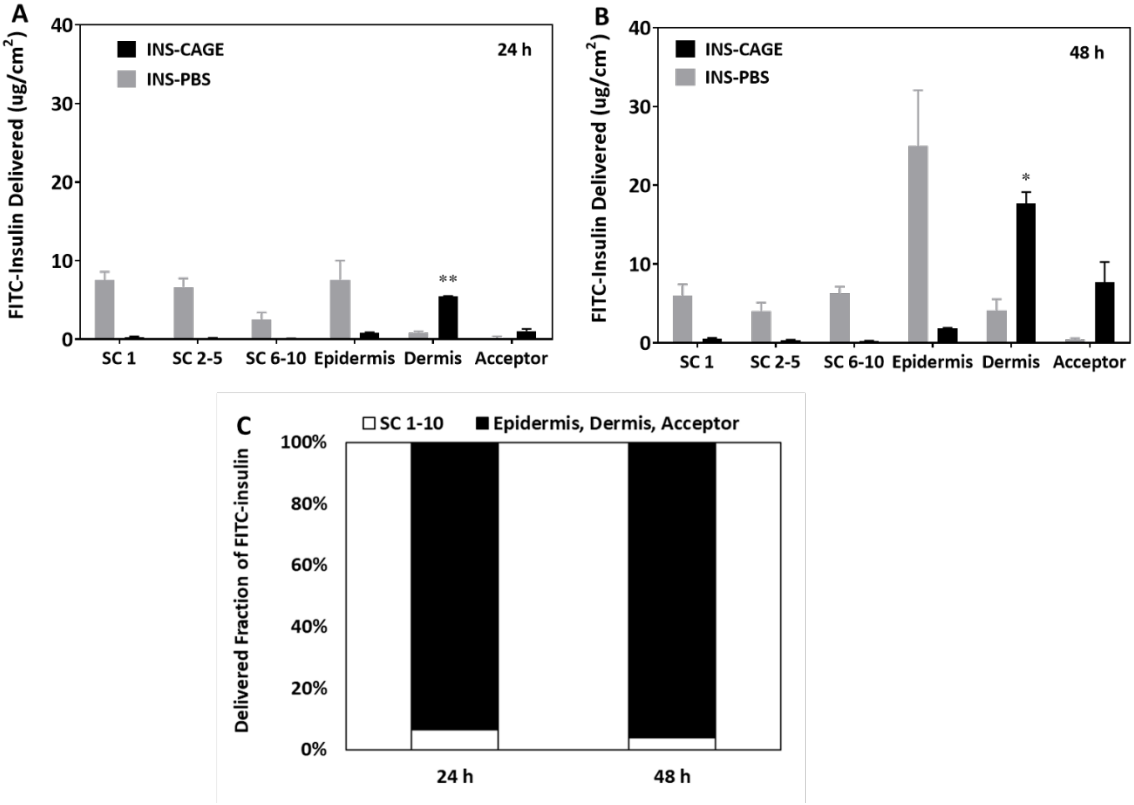


Figure 4.4. CAGE improves skin penetration of insulin. CAGE significantly enhances insulin penetration to the dermis compared to PBS control. (A) 24 h (**p<0.005); (B) 48 h (*p < 0.04); and (C) Total penetration amount of insulin into skin after CAGE treatment. White and black bars represent the total of SC 1-10 and combination of the epidermis, dermis and acceptor, respectively. Data represented as mean ± S.E. (n=3).

To identify the role of geranic acid in enhancing the permeation effect of CAGE, all 4 CAGE variants plus their constituents, choline bicarbonate and geranic acid, were applied to porcine

skin *ex vivo* in Franz diffusion cells. The CAGE 1:1 and 2:1 samples showed no transport across the SC (**Figure 4.5 a and b),** while the 1:2 and 1:4 samples showed FITC-insulin permeating into the dermis after 24 h (**Figure 4.5 c and d)**. Samples containing geranic acid or choline bicarbonate and FITC-insulin (**Figure 4.5 f and g**) showed that neither alone is responsible for the enhancement of insulin delivery seen using the 1:2 and 1:4 CAGE formulations. It is notable that there is less fluorescence in the geranic acid sample then either the buffer (**Figure 4.5 e**) or choline bicarbonate sample, suggesting that insulin may be less soluble in highly hydrophobic geranic acid than buffer or hydrophilic choline bicarbonate.

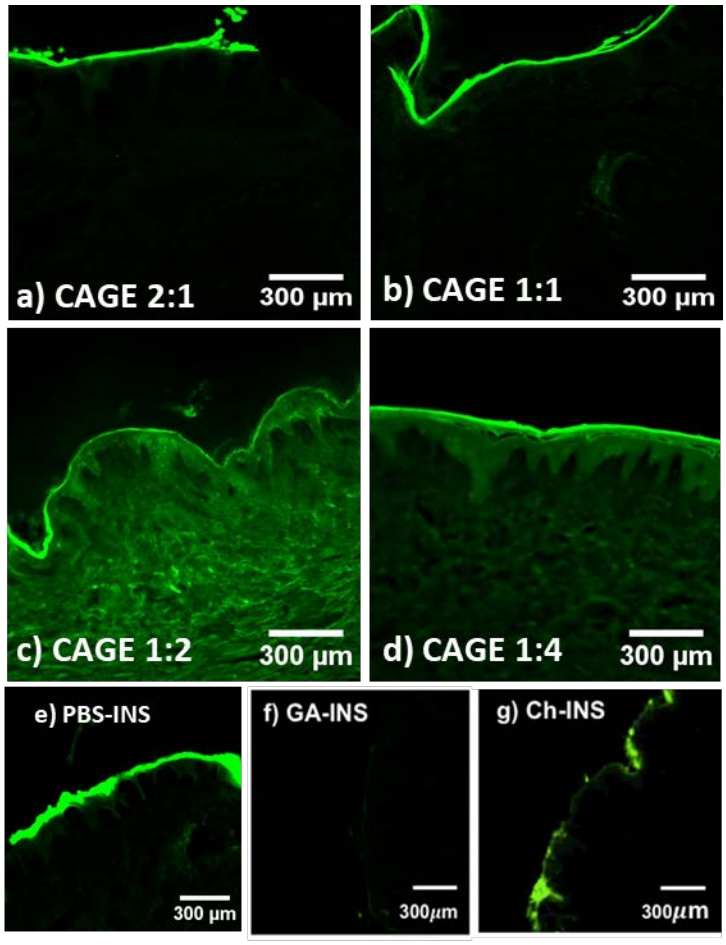


Figure 4.5. Confocal microscopy images of FITC-Insulin in porcine skin solvated in CAGE variants. FITC-insulin in CAGE 2:1 (a) and 1:1 (b) did not exhibit permeation across porcine SC while CAGE 1:2 (c) and 1:4 (d) did show transport of FITC-Insulin into the dermis. FITC-insulin in buffer (e), geranic acid (f) and choline (g) did not permeate into the skin.

4.5 SC permeabilization by CAGE

The four CAGE variants were evaluated for SC disruption in correlation with geranic acid content. SC samples were incubated in CAGE 2:1, 1:1, 1:4, 2.26M choline or neat geranic acid and the FTIR spectra recorded for each sample before and after incubation. FTIR spectra of SC samples incubated in CAGE variants showed a trend of reduced peak height at 2,850 cm^{-1} (symmetric CH_2 bond stretching) and 2,920 cm^{-1} (asymmetric CH_2 bond stretching) with

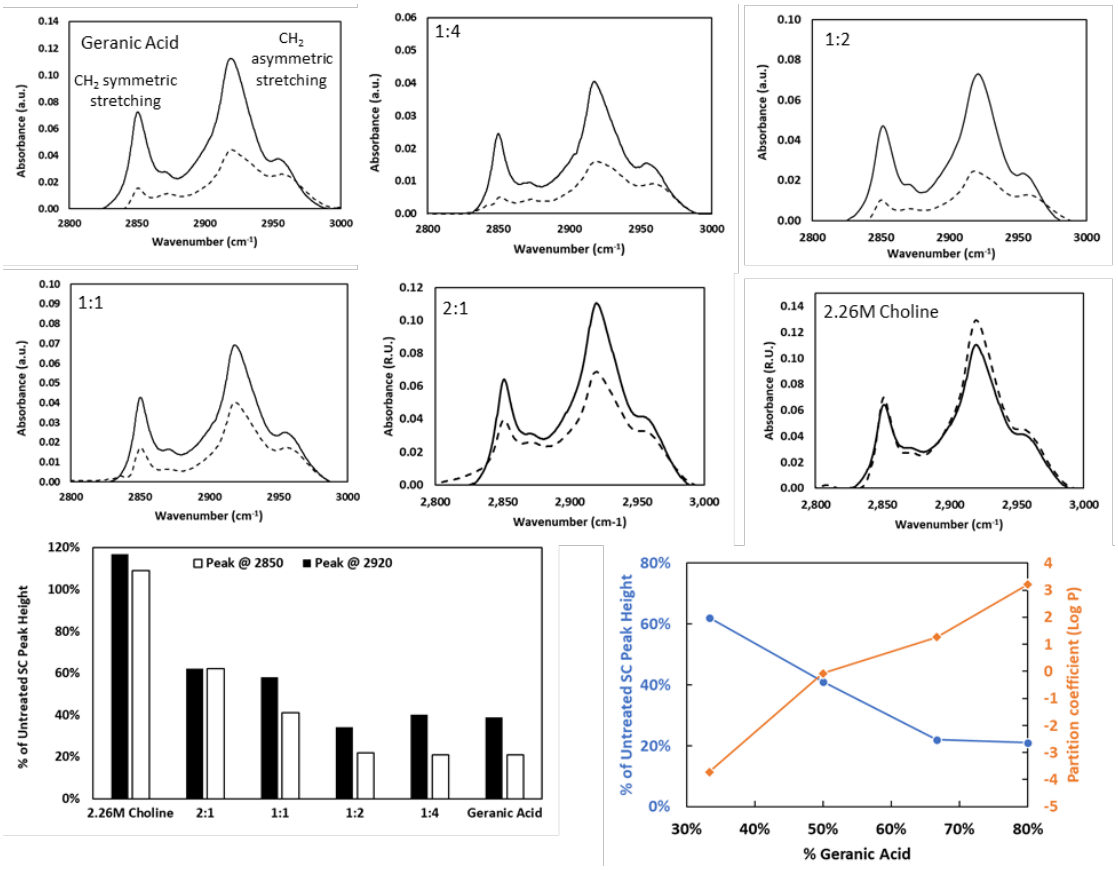


Figure 4.6. CAGE variants appear to reduce lipid content in SC. FT-IR spectra of the lipid region for SC samples before (solid line) and after (dashed line) 24h incubation in geranic acid, 2.26M choline or CAGE variants. Peak heights at 2850 cm^{-1} (white columns) and 2920 cm^{-1} (black columns) after CAGE treatment as % of untreated (starting) SC peak height). Comparison of peak height at 2,850 cm^{-1} (left axis, blue circles) with the calculated CAGE variant partition coefficient (right axis, orange diamonds).

increasing geranate/geranic acid content (**Figure 4.6**), suggesting lipid extraction may be part of the mechanism by which CAGE increases penetration of molecules into the skin.

FTIR studies were also performed to more fully evaluate the effect of CAGE 1:2 on the SC structure when combined with proteins **(Figure 4.7)**. It was observed that neat CAGE, BSA-CAGE and INS-CAGE reduced the area of the 2,850 and 2,920 cm^{-1} peaks which are indicators of the SC lipid content (Tsai, et al., 2004), while PBS, BSA-PBS and INS-PBS did not. A decrease in the peaks of CH_2 symmetric stretching at 2,850 cm^{-1} or asymmetric

stretching at 2,920 cm^{-1} are indicative of lipid extraction (Kaushik & Michniak-Kohn, 2010) (Kuma, et al., 2015). This suggests that CAGE acts as a lipid extractor to improve permeability of drugs across the skin. All spectra were deconvoluted into individual peaks (**Figure 4.8**) to facilitate comparison of the pre- and post- incubation spectrum. In the SC sample, before application of protein-CAGE, two to four peaks are observed to contribute to the 2,850 and 2,920 cm^{-1} absorbance peaks, respectively. Post-incubation, all peaks' absorbance was significantly reduced. Inspection of FTIR spectra of the individual proteins or neat CAGE show that they contain few peaks in the region of 2,800 to 3,000 cm^{-1}, however neither of the proteins (BSA or insulin), nor CAGE contain peaks at 2,850 or 2,920 cm^{-1}. A more detailed assignment of peaks is beyond the scope of this study, but it appears that a CAGE peak around 2,930 cm^{-1} may contribute to both post-application SC spectra. While absorbance spectra of residual protein or CAGE in the SC sample may contribute to the post-application SC spectra, overall there was a significant reduction in peak areas at 2,850 and 2,920 cm^{-1}.

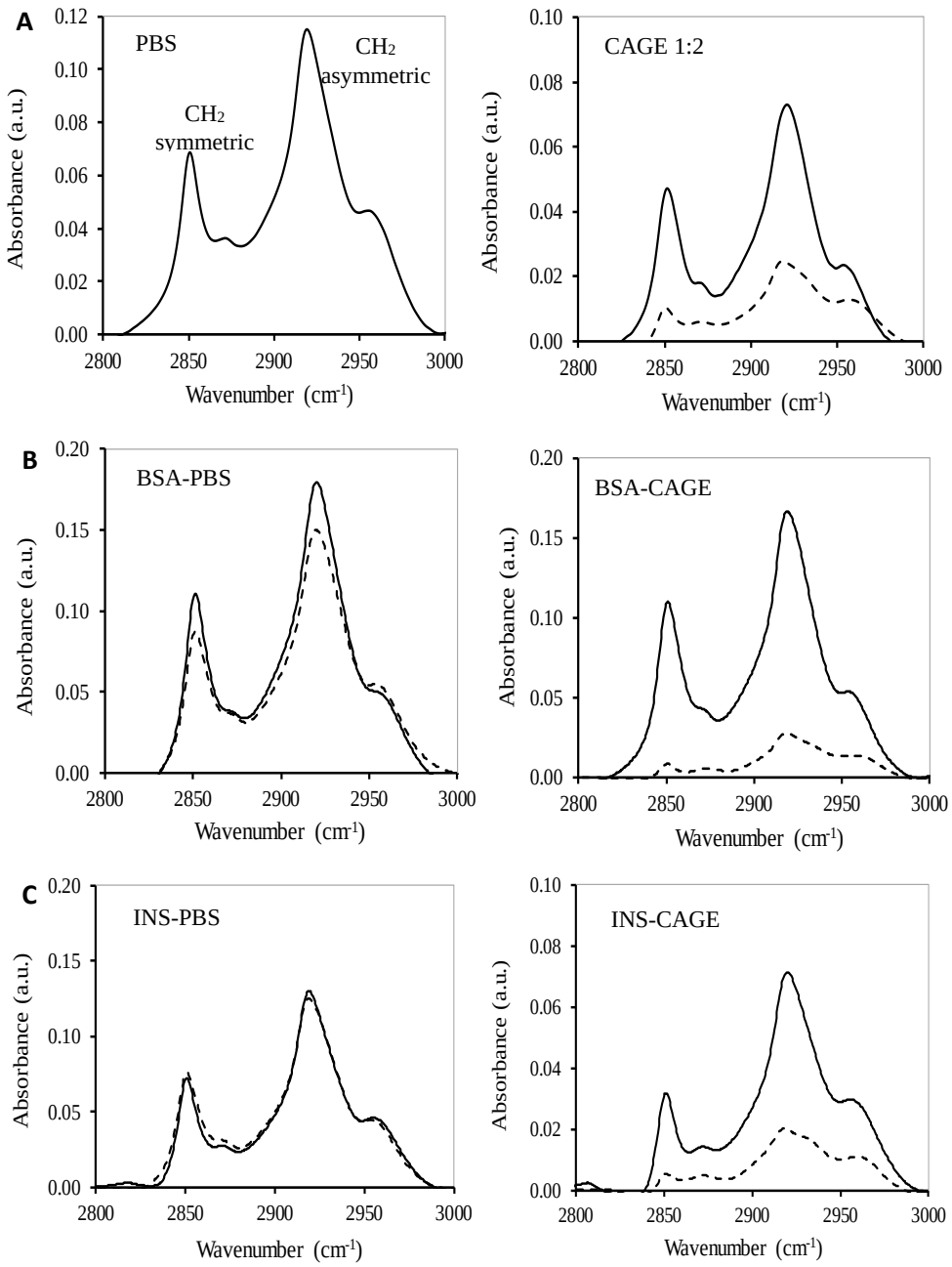


Figure 4.7. FTIR spectra corresponding to lipid content of the SC. Measured as by peak area of the CH_2 symmetric (2,850 cm^{-1}) and asymmetric (2,920 cm^{-1}) stretching bands. SC before (solid black line) and after (dashed black line) application of (A) PBS (dashed and solid lines are overlaid) or neat CAGE, (B) BSA-PBS or BSA-CAGE, and (C) INS-PBS or INS-CAGE for 24 h. Protein-CAGE application reduced peak areas in the spectral region indicative of lipid content, while protein alone (protein-PBS) did not.

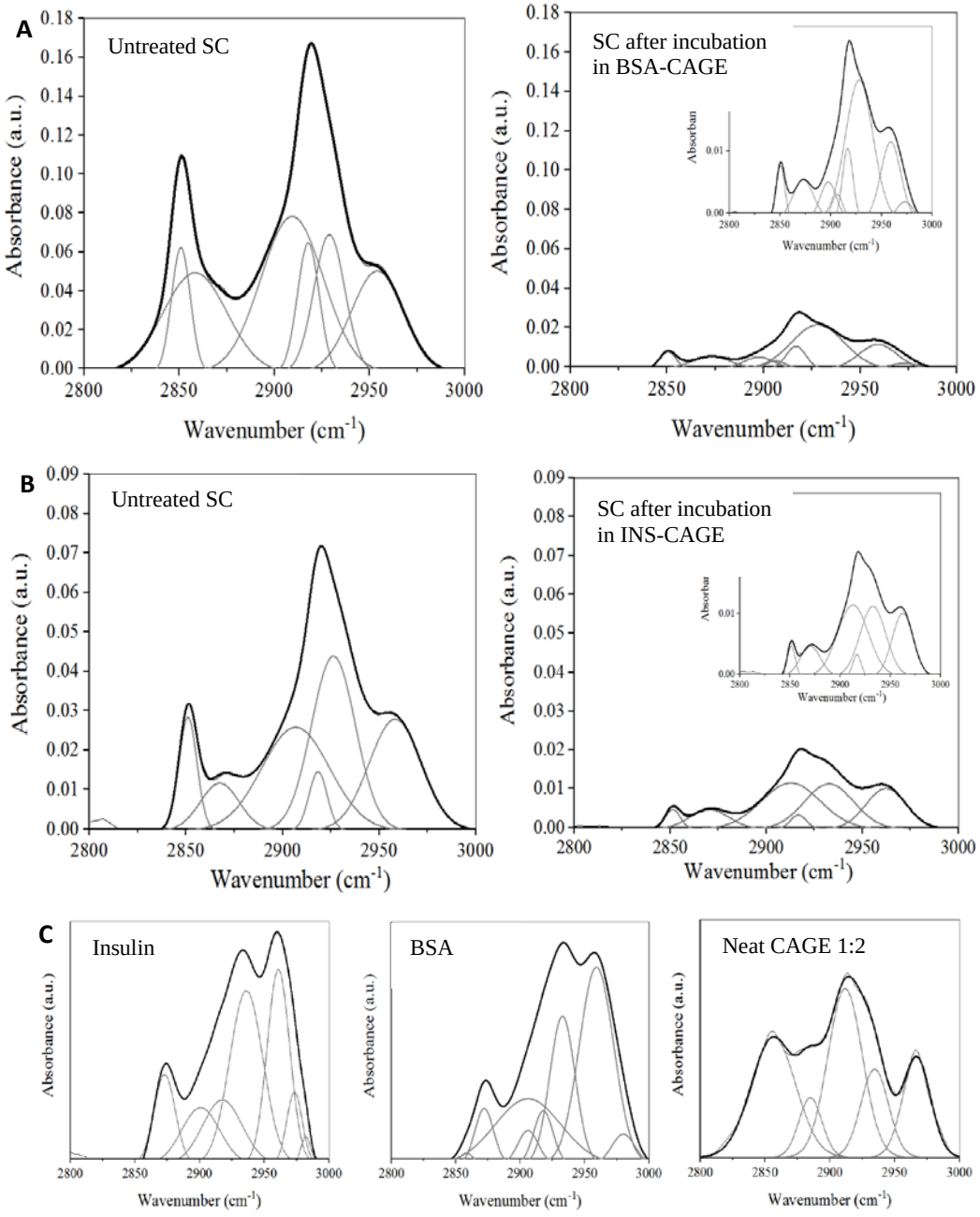


Figure 4.8. Deconvoluted FTIR spectra for the region indicative of stratum corneum (SC) lipid content. Spectra of SC before and after application of protein-CAGE shows a reduction in the deconvoluted peaks at 2,850 and 2,920 cm^{-1}. The same SC sample before and after incubation in (A) BSA-CAGE or (B) INS-CAGE for 24 hr followed by washing with PBS and drying. Inset graphs show enlarged peak deconvolution. (C) Reference FTIR spectra of insulin powder, BSA crystals and neat CAGE 1:2 (no stratum corneum).

4.6 Transdermal delivery of insulin via CAGE *in vivo*

In an *in vivo* study performed by A. Banerjee with CAGE-insulin supplied by the author, the efficacy of CAGE 1:2 in delivering macromolecules across the skin was evaluated. Rats were treated with a topical application of 25 U/kg CAGE-insulin and their blood glucose levels (BGL) were measured over a 12 h period (**Figure 4.9**). The efficacy of the formulation was compared to a subcutaneously injected dose of 1 U/kg insulin in saline and also to controls for formulation (CAGE alone) and passive permeation (insulin in PBS). Neither topically-applied CAGE or PBS-insulin resulted in a significant drop over the course of the study. The group that received the CAGE-insulin application experienced a 25% drop in BGL at 2 h, which continued to decrease until 4 h, culminating in a 40% reduction. After 4 h, BGL climbed slightly and remained at about 70% of initial BGL throughout the remainder of the study. There was a significant difference in efficacy for the CAGE-insulin topical treatment compared to the PBS-insulin control, indicating the CAGE improved permeation of insulin into the dermis.

In contrast, the group which received a subcutaneous injection experienced a rapid drop in BGL, 45% at 1.5 h followed by a recovery to initial BGL at 4 h. Following 4 h, BGL in the injected group dropped slowly due to the fasted state of the animals; this trend was seen in the other control groups. The quick response and subsequent recovery seen here is emblematic of a typical response in T1D patients to injected insulin. Known as the “valley and peak” phenomena (Bolli, 2002), patients can experience hypoglycemic symptoms as a result of an injection (the valley), which they tend to overtreat with glucose, which leads to high BGL (the peak). This can become a frustrating cycle of treatment which is stressful and time-consuming. It is notable, therefore, that the CAGE-insulin application resulted in a slower and sustained

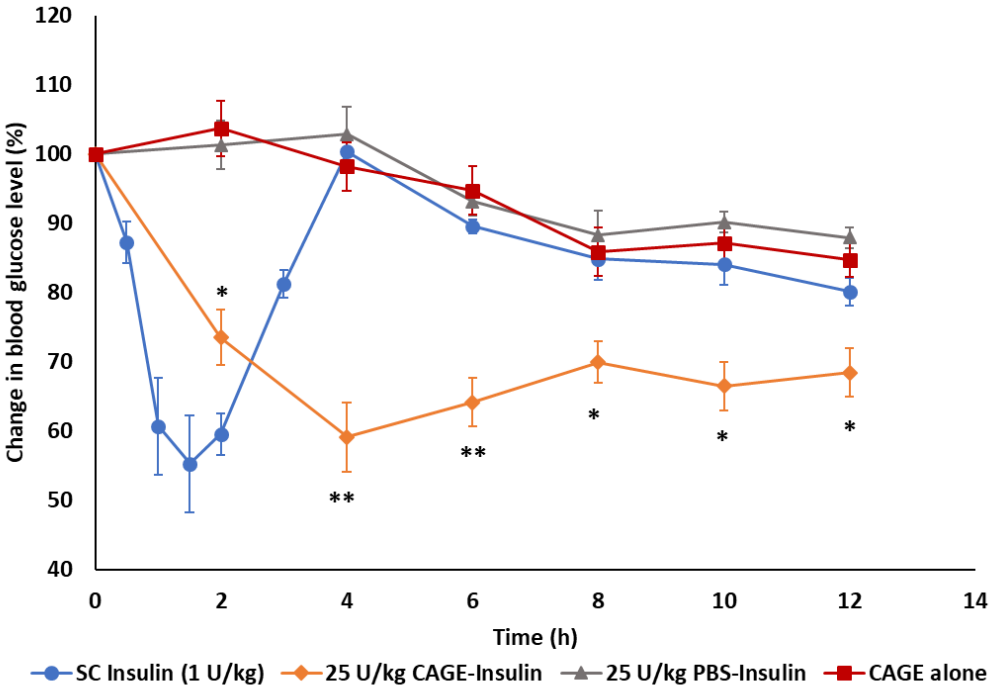


Figure 4.9. *In vivo* efficacy evaluation of CAGE as a transdermal insulin delivery vehicle. Change in blood glucose levels (BGL) were measured over time after administering different test treatments to non-diabetic rats. Topical application of CAGE alone (0 U/kg insulin, red squares), insulin in buffer (25 U/kg insulin, grey triangles), insulin in CAGE (25 U/kg, orange diamonds) or subcutaneous insulin in saline (1 U/kg, blue circles) were delivered and blood glucose levels were measured from the tail vein at least every 2 h for 12 h. BGL after CAGE-Insulin application were significantly different than PBS-Insulin (*$p<0.01$, **$p<0.001$). BGL are plotted as percent change compared to the initial (before treatment, time 0 h) value. Data represented as mean ± S.E. (n=6 for transdermal applications, n=3 for subcutaneous injection).

lowering of BGL compared to injection; this could be an important aspect of this formulation in helping to improve not only BG control in diabetes, but spike effects from other injected therapies.

4.7 Summary

Transdermal delivery of therapeutic proteins and peptides is challenging due to very low permeability of these macromolecules across the skin. Traditionally, various physical enhancement techniques have been used to address the delivery challenge. This study demonstrated the effectiveness of CAGE 1:2, a deep eutectic solvent as a potential transdermal

delivery vehicle for therapeutic proteins and peptides. CAGE was found to enhance permeability of large macromolecule drugs through the skin both *ex vivo* and *in vivo*, in part by appearing to extract lipids from the stratum corneum. In addition, topical administration of 25 U/kg insulin in CAGE demonstrated similar peak efficacy as 1 U/kg subcutaneous administration and efficacy was sustained for a prolonged period. The formulation is easy to prepare and can be stored at room temperature for several weeks without compromising protein secondary structure. Additional studies further investigating CAGE's mechanism of action, toxicity upon chronic use and investigation of insulin-CAGE efficacy in diabetic rats are required before consideration for further development. In conclusion, the study suggests that CAGE can be used as a transdermal alternative to injectable protein formulations that could significantly improve the quality of life for people currently using injections to treat chronic diseases.

It is to be noted that chemical enhancers have successfully been used mostly for small molecule delivery, as they do not provide similar degrees of penetration enhancement as physical enhancement methods (Paudel, et al., 2010). This notion was corroborated in our study where no improvement in penetration of FITC-BSA, FITC-OVA or FITC-insulin across the skin was observed in the presence of CPEs such as ethanol or DGME, while a significant enhancement in delivery of these proteins was noted in the presence of CAGE. Further quantitation of ^{3}H-BSA in different layers of the skin after 24 h treatment with different materials also demonstrated significantly higher accumulation of BSA in the dermis and acceptor chamber when the protein was delivered using CAGE compared to PBS or the CPEs. This accumulation of ^{3}H-BSA using CAGE was noted to be a time-dependent process with more accumulation in the stratum corneum up to 12 h, followed by higher accumulation in the

epidermis, dermis and acceptor chambers at longer time points (24 and 48 h). Additionally, dermis concentrations after 24 and 48 h were significantly higher for FITC-insulin delivered using CAGE compared to PBS.

For effective transdermal delivery, it is incumbent that proteins are not significantly modified to compromise bioactivity and permeation enhancers do not denature proteins or cause skin irritation. This study shows that CAGE 1:2 is a safe and effective transdermal delivery vehicle that retains protein secondary conformation and significantly improves insulin delivery in rats compared to formulation controls. Previous histopathological and toxicity studies in rats and mice showed no observable skin irritation or toxicity upon CAGE application (Zakrewsky, et al., 2016). Based on the FTIR studies we postulate that CAGE improves permeability through SC lipid extraction, a quintessential chemical enhancer mechanism of action as seen for DMSO, alcohols, azones and fatty acids (Karande & Mitragotri, 2009) (Karande, et al., 2005) (Benson, 2005). However, additional studies using small and wide-angle X-ray diffraction and other techniques are required to understand the exact mechanism of action of CAGE on skin components (Moghadam, et al., 2013).

Finally, this study showed that differing the amount of geranic acid affected permeation enhancement of insulin and lipid removal; the 1:4 variant resulted in a lipid peak reduction similar to geranic acid, but geranic acid alone did not show permeation of insulin. This suggests that choline with geranic acid is necessary to facilitate the transport. Quantitative studies are needed to confirm the relationship between geranic acid content and permeation of proteins.

.

Chapter 5

5. CAGE as an oral drug delivery agent

5.1 Motivation

Oral delivery of drugs is preferred over injections for several reasons, including ease of administration, lower manufacturing costs, and reduced risk of infection. For some patients, oral dosing also results in higher adherence to the therapy regime. However, many drugs are currently not suitable for delivery via the oral route, either due to a narrow therapeutic window or their susceptibility to degradation in the gastric system.

Insulin, a peptide used by all Type 1 and some Type 2 diabetics to control blood glucose due to lack of endogenous insulin production, suffers from both issues, but the significant benefits of oral dosing over subcutaneous injection still makes it an attractive target. Insulin, while highly effective in managing hyperglycemia to mitigate associated health risks including retinopathy and neuropathy, is only commercially available in injected form, and as many as 60% of patients intentionally miss injections due to pain or inconvenience (Fonte, et al., 2013) (Peyrot, et al., 2010), which can result in short-term complications requiring hospitalization and increased morbidity (Peyrot, et al., 2010). Another benefit of oral insulin delivery is the difference in glucose homeostasis. Injected insulin goes directly into systemic circulation, and only a small portion (20%) ends up in the portal vein. Endogenously produced insulin from the pancreas is first transported to the liver via the portal vein, where a large portion (80%) is retained. The lower insulin concentration in the portal vein from injection results in unsuppressed levels of glucagon, a hormone that promotes the liver to release glucose, which causes hyperglycemia (Arbit & Kidron, 2017) (Fonte, et al., 2016) (Sonaje, et al., 2010). The

rising prevalence and morbidity of diabetes worldwide highlights the need for a cheaper and easier to administer form of insulin. Diabetes incidence has increased in every country with low and middle-income countries experiencing the most rapid growth (Kharroubi & Darwish, 2015) (Olokoba, et al., 2012) (World Health Organization, 2016). In 2012, the World Health Organization reported 1.5 million deaths directly attributable to diabetes and another 3.2 million due to disease caused by chronic elevated blood glucose levels. (World Health Organization, 2016).

The barriers to oral delivery of drugs are many; the primary *in vivo* ones facing developers of biologics destined for the gastrointestinal tract are shown in **Figure 5.1** (Araujo, et al., 2017). In the stomach, the low pH environment is designed to digest proteins; those that do make it through the stomach intact encounter proteolytic enzymes in the intestine, whose job is to reduce proteins to small, easily absorbed amino acids that can traverse the mucus layer adjacent

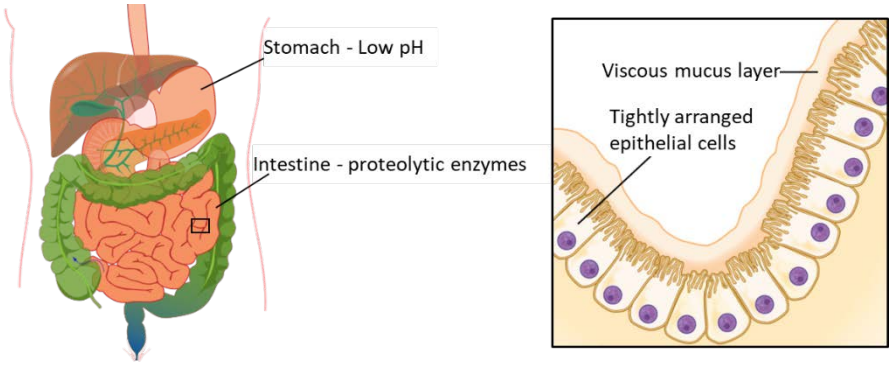


Figure 5.1. Barriers to oral delivery of biopharmaceuticals and requirement for a delivery system to overcome them. The digestive system (left) is designed to break proteins into easily absorbed amino acids using a multi-stage process catalyzed by acids in the stomach and enzymes in the intestine. To ensure that only desired molecules are absorbed, the intestinal wall (right) provides additional barriers in the form of a mucus layer and a tightly arranged epithelium. Image credit: OpenStax College.

to the intestinal wall and pass through the intestinal epithelium. Larger molecules encounter significant resistance to passage through this tight cell structure, designed in large part to keep microbes that commensally inhabit the gut from accessing tissue and circulatory pathways where they can cause illness. Finally, a high level of variability is exhibited across human subjects in terms of gastric emptying and intestinal transit time and the effect of food or water on these parameters (Soares, et al., 2017) (Steingoetter, et al., 2006) (Kwiatek, et al., 2009) (Karsdal, et al., 2008) (Tanko, et al., 2004). All of these factors contribute to extremely low oral bioavailabilites (<1%) for orally-dosed proteins (Shaji & Patole, 2008).

Strategies to improve drug bioavailability typically include encapsulation or coating to protect from stomach acids, use of protease inhibitors, protein modification, or addition of absorption enhancers to increase uptake. Capsules are commonly coated with Eudragit® to ensure their survival past the stomach. Once in the intestine, the use of soybean trypsin inhibitor or other enzyme inhibitors can protect from degradation. Absorption enhancers such as fatty acids, bile salts or surfactants act on the epithelial cell wall to increase transport, either by modulating the cell membrane to increase transcellular transport, or opening the tight junctions in between cells to enhance paracellular transport (Wong, et al., 2016). Hydrophobic molecules are thought to affect the transcellular route, while hydrophilic ones work on tight junction integrity.

Using a variety of these methods, reserachers have worked to improve oral insulin bioavailablities. Using liposomes with bile salts to encapsulate insulin, Niu et al reported an oral bioavailability of 8.5% with sodium glycocholate-liposomes in non-diabetic rats, attributed to reduced enzymatic degradation (Niu, et al., 2012). Sheng et al achieved an oral

bioavailability of 18% in rats by conjugating insulin with low molecular weight protamine, a cell penetrating peptide, and loading it into mucoadhesive nanoparticles (Sheng, et al., 2016). Sonaje et al achieved a 15% oral bioavailability in rats with a pH-responsive nanoparticle composed of chitosan and poly (gamma-glutamic acid) as the insulin carrier (Sonaje, et al., 2010). Insulin-loaded dextran sulfate/chitosan mucoadhesive, negatively charged nanoparticles orally delivered to rats resulted in a 5.25% bioavailability (Sarmento, et al., 2007).

CAGE has the potential to address many of these hurdles including cargo biocompatibility, protection from proteolytic enzymes and circumventing the intestinal epithelium barrier properties. We have to date demonstrated CAGE's biocompatibility with mammalian cells, its ability to protect cargo (insulin) from thermal degradation, and to increase permeation into another barrier, the stratum corneum in skin. Additionally, CAGE has the potential for intermolecular association with the cargo due to both hydrophilic and hydrophobic segments, which may further aid stability and transport of cargo.

5.2 Study design

Having seen the superior permeation enhancing ability of CAGE 1:2 in skin, we wanted to investigate the potential of CAGE to increase permeation of insulin across the intestinal wall. In addition, knowing that insulin in CAGE remains stable for several months, we posited that CAGE might afford protection against proteolytic enzyme degradation and thus be a candidate for an oral delivery route. Thus, the study objectives were to 1) determine the compatibility of CAGE with intestinal cells, 2) understand how CAGE might affect transport across the

intestinal epithelium, and 3) determine if CAGE-insulin could be orally delivered with good efficacy.

Cell viability was investigated with a standard MTT viability assay at two time points; immediately after a 5 h incubation with CAGE, and after 19 h of incubation with no CAGE, to determine the ability of the cells to recover. To confirm whether CAGE would enhance penetration, we conducted TEER studies, comparing CAGE with a known penetration enhancer, sodium caprate. From there, we conducted mechanistic studies using paracellular and transcellular markers to determine the route of transport for CAGE-drug complexes. To determine if CAGE might affect transport through the mucus layer next to the intestinal wall, we tested the effect of CAGE on simulated mucus viscosity. *In vivo* studies were conducted to determine the efficacy of CAGE in delivering insulin intrajejunally and orally. Finally, a small *in vivo* toxicity study was conducted to collect preliminary histology data on CAGE.

5.3 Compatibility of model intestinal cells with CAGE

To determine the compatibility of intestinal cells with 1:2 CAGE, varying concentrations of CAGE (0, 10, 25 and 50 mM) were applied for 5 h to a Caco-2 monolayer grown in a transwell insert. Inserts were either assayed immediately for viability (5 h) or incubated for an additional 19 h after CAGE was replaced with medium and assayed at 24 h. Good viability was seen at all concentrations (**Figure 5.2**), with the best viability at the lowest CAGE concentration. At 24 h, all concentrations had improved cell viability compared to 5 h.

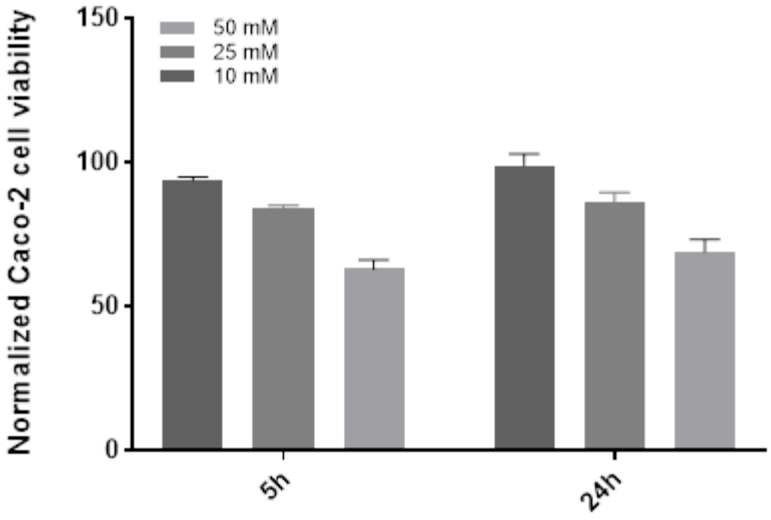


Figure 5.2. Caco-2 cell viability after incubation in varying concentrations of CAGE for 5 h. 24 h samples were analyzed after an additional 19 h incubation with no CAGE.

5.4 TEER evaluation

To determine if CAGE affected monolayer permeability, we measured trans epithelial electrical resistance (TEER) across Caco-2 seeded transwell inserts. A decrease in TEER correlated with increasing CAGE concentrations (**Figure 5.3**). Cells treated with 10 mM CAGE showed a significant decrease in TEER at 1 and 5 h with a complete recovery at 24 h, while those treated with 25 and 50 mM CAGE showed a significant (30-50%) decrease in TEER across all time points and a recovery to 90 and 58% of initial levels at 24 h respectively. Notably, the TEER drop due to 50 mM CAGE was similar to that from 10 mM sodium caprate (Krug, 2013), but cells treated with 50 mM CAGE recovered some monolayer resistance after 24 h, while those treated with sodium caprate did not.

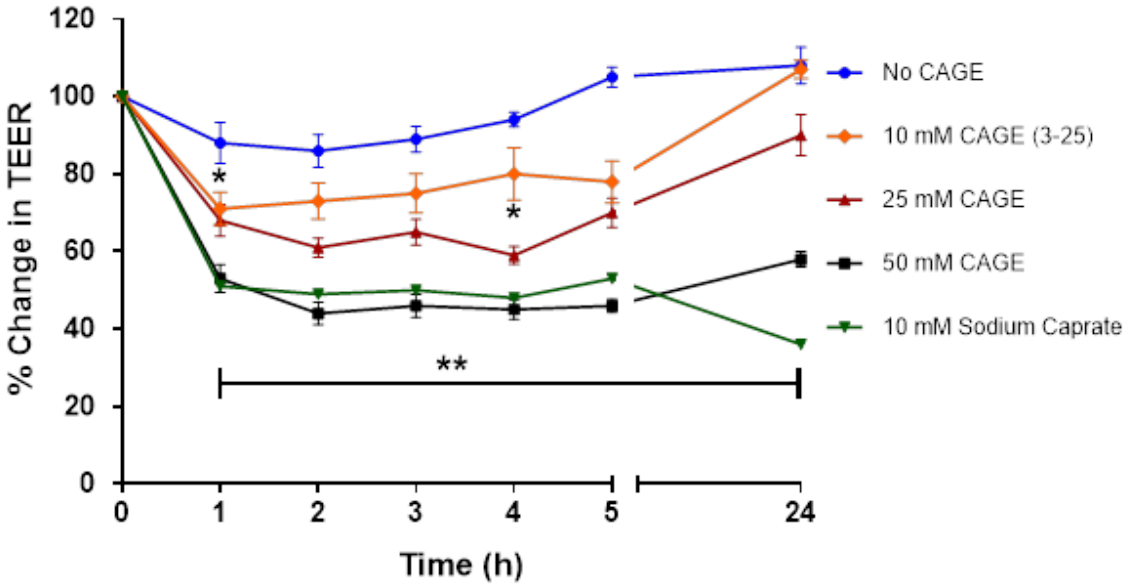


Figure 5.3. Effect of TEER in Caco-2 monolayer after treatment with varying CAGE concentrations or sodium caprate. Data represented as mean ± S.E. (n=4); *p<0.05, **p <0.0001; all treatments compared to no CAGE treatment.

5.5 CAGE-mediated transport across an epithelial model

Lucifer Yellow (LY) and FITC-dextran (4 kDa) are known paracellular transport markers (Konsoula & Barile, 2005). When combined with 0, 10, 25 or 50 mM CAGE, a concentration-dependent transport profile was seen for LY and FITC-dextran (**Figure 5.4**). LY transport was increased by 1.5 to 10x when mixed with 10 to 50 mM CAGE compared to no CAGE (LY in medium). For 4 kDA FITC-dextran, which is of similar size to insulin, transport increased 10 – 12x when combined with 50 mM CAGE compared to a no-CAGE control. From these results, it appears that CAGE affects the integrity of the tight junctions in epithelial cells, effecting increased paracellular transport.

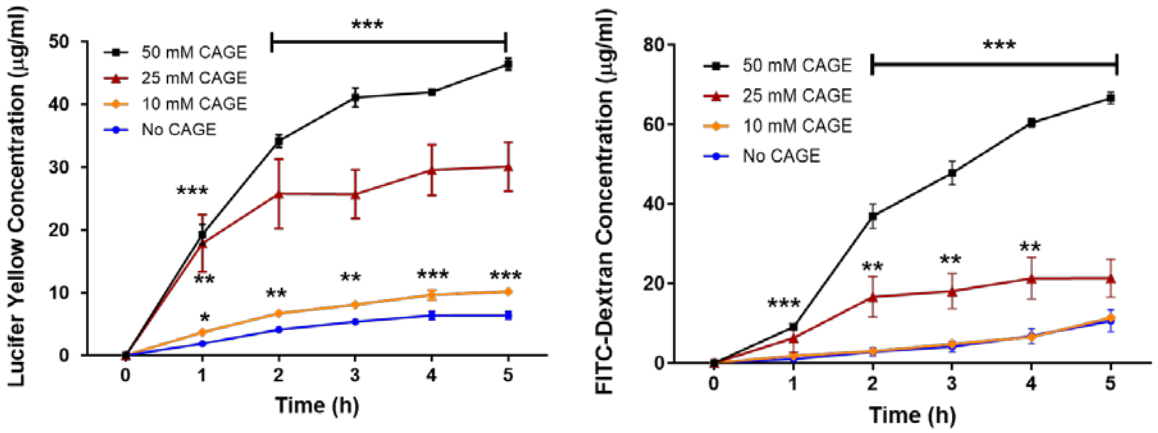


Figure 5.4. Basolateral chamber concentration of paracellular markers LY and 4 kDa FITC-dextran in CAGE. Starting apical chamber concentrations 500 µg/mL for both markers. Data represented as mean ± S.E. (n=4); *p<0.05, **p <0.001, ***p<0.0001; all treatments compared to no CAGE treatment.

Next, the transcellular route was investigated using coumarin-6 (Simovic, et al., 2015) and antipyrine (Siew, et al., 2011), both of which are known to passively transport across the cell membrane. Interestingly, all concentrations of CAGE significantly halted uptake of both coumarin-6 and antipyrine (**Figure 5.5**). This is convincing evidence that CAGE does not enhance transcellular uptake.

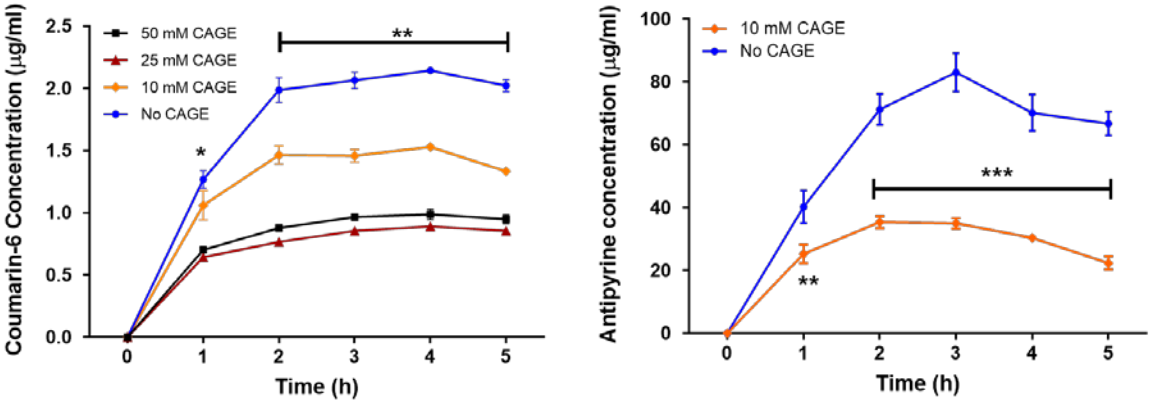


Figure 5.5. Basolateral chamber concentration of transcellular markers coumarin-6 and antipyrine. Starting apical chamber concentration 5 µg/mL for coumarin-6 and 500 µg/mL for antipyrine. Data represented as mean ± S.E. (n=4 for LY, n=6 for antipyrine); *p<0.05, **p <0.0001; all treatments compared to no CAGE treatment.

As a final transport study, we evaluated the movement of FITC-insulin in CAGE or dimethyl palmitoyl ammonio propanesulfonate, (PPS), an enhancer identified earlier in our lab (Gupta, et al., 2013). 10 and 25 mM CAGE performed about equally, while 50 mM CAGE significantly improved transport of insulin, by 10-15x ((**Figure 5.6)**. Moreover, the extent of FITC-insulin transport across the intestinal monolayer was similar to that from PPS.

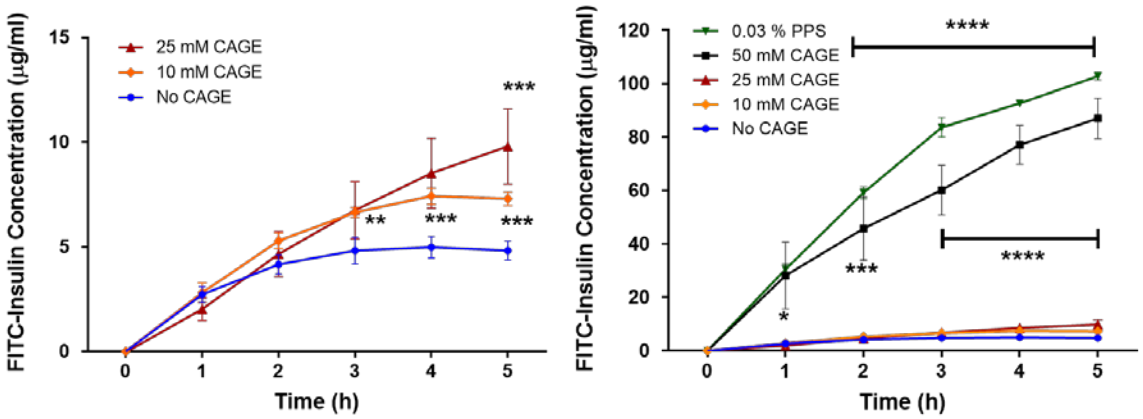


Figure 5.6. Basolateral chamber concentration of FITC-insulin. Starting apical chamber concentration 500 µg/mL. Data represented as mean ± S.E. (n=4 for CAGE, n=3 for PPS); *p<0.05, **p <0.001, ***p<0.0001, ****p<0.00001; all treatments compared to no CAGE treatment.

The apparent permeation coefficients, P_{app}, are tabulated in **Table 5.1**.

Table 5.1. Apparent permeation coefficients (10^6 cm/s)

CAGE concentration (mM)	*Lucifer Yellow*	*FITC-4 kDa Dextran*	*FITC-Insulin*	*Antipyrine*	*Coumarin-6*
0	1.42	1.67	1.17	7.41	456
10	2.2	1.77	1.67	2.49	320
25	6.99	4.81	1.93	-----	191
50	10.3	13.5	17.6	-----	213

5.6 Effect of CAGE on simulated mucus viscosity

Simulated Mucus (SM) made from 2% dried pig mucin dissolved in saline exhibits a similar

shear-thinning profile as reported for gastric mucus (Lai, et al., 2009) and the SM viscosity profile compared well to literature values for healthy human duodenal gastric mucus (Curt & Pringle, 1969). For example, at a shear rate of 46 1/s, the literature value was 12.3 cP compared to the measured SM mean value of 11.3 cP at 50.12 1/s. CAGE incubated with SM resulted in a very significant drop in viscosity (**Figure 5.7**). The addition of 1, 5 and 10% CAGE reduced

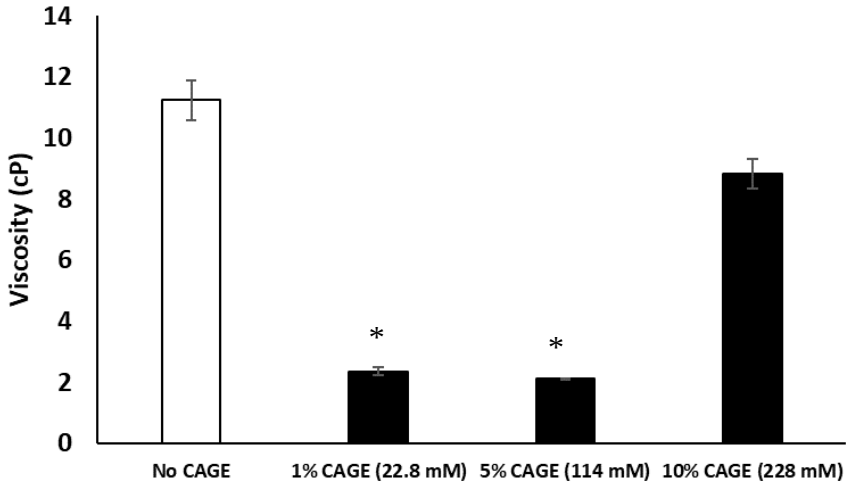


Figure 5.7. Reduction in mucus viscosity after addition of CAGE. Mean viscosity values in centipoise (cP) at a shear rate of 50.12 1/s are shown for simulated mucus with no CAGE (0%), 1, 5 and 10% w/v CAGE. Data represented as mean ± S.E (n=3); *p< 0.001, CAGE treatment compared to no CAGE treatment.

the viscosity throughout the entire measured shear range. The viscosity reduction from 10% CAGE was not statistically significant compared to SM alone (no CAGE). Viscosity reduction suggests that CAGE would assist in mucus penetration *in vivo*, thus facilitating delivery of insulin to the intestinal epithelia. At higher CAGE concentrations of 25 and 50%, the viscosity increased, likely due to contributions from CAGE's inherent viscosity.

5.7 Pharmacokinetic and pharmacodynamic studies

Studies were conducted by the author and Dr. A. Banerjee to determine CAGE's *in vivo* effects.

5.7.1 Efficacy and bioavailability of intrajejunally-delivered insulin-CAGE

Rats were intrajejunally injected with either insulin-CAGE or insulin-saline (5 U/kg) and blood glucose was measured over time (**Figure 5.8 top**). Plasma samples were collected to determine insulin concentration in the blood over time (**Figure 5.8 bottom**). Animals given the insulin-CAGE injections experienced a drop in blood glucose level (BGL), 70% within 2 h, following a similar but slightly delayed pattern of a subcutaneous (SQ) injection (2 U/kg), and recovering to 65% of initial levels at 5 h. Animals who received insulin-saline did not see an appreciable drop in BGL during the course of the study. Insulin concentration in the blood peaked at 1 h in rats receiving insulin-CAGE, which was comparable to a subcutaneous injection. There was no insulin measured in the blood samples of rats given insulin-saline. Additionally, insulin-CAGE appeared to have an increased half-life compared to injected insulin. Integration of each concentration curve gives a pharmacokinetic bioavailability of 51% for insulin-CAGE compared to injected insulin.

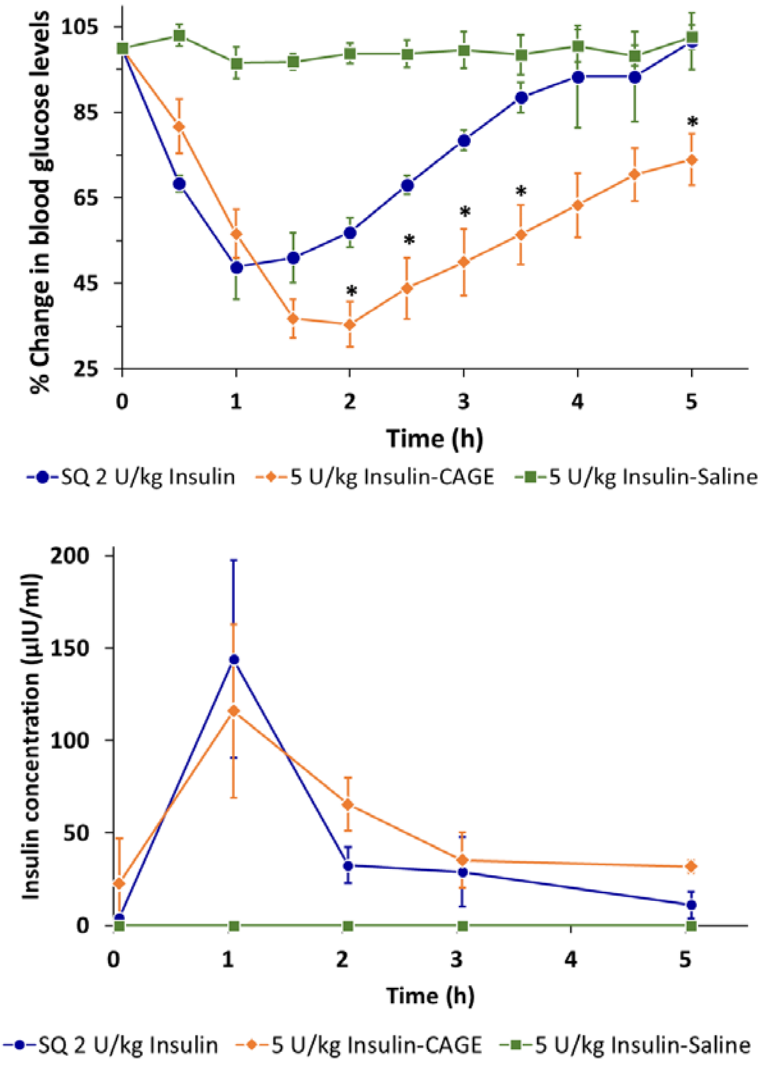


Figure 5.8. Intrajejunal injection of CAGE-insulin. Top: An IJ injection of insulin-CAGE (5 U/kg, orange diamonds) administered to rats had a significantly different BG profile than insulin-saline (5 U/kg, green squares). Bottom: The insulin-CAGE profile was not statistically different from subcutaneously injected insulin (2 U/kg, blue circles). Data represented as mean ± S.E. (n=4 for insulin, n=6 for BGL); *p<0.05, insulin-CAGE compared to SQ insulin.

5.7.2 Efficacy of orally-delivered insulin-CAGE

To investigate the efficiency of insulin-CAGE delivered orally in a capsule, rats were orally gavaged with either insulin-CAGE (10 U/kg), neat CAGE or empty capsules. In addition, one group was orally gavaged with an insulin-saline solution (10 U/kg). A positive control group

was subcutaneously injected with insulin-saline (2 U/kg). The insulin-CAGE dosed animals experienced a sustained drop in BGL with an initial decrease of 65% at 2 h and remaining between 55 – 70% for the duration of the study (**Figure 5.9**). The insulin-saline oral solution exhibited a gradual decrease over time, as did the empty capsule and neat CAGE dose; this is attributed to the fasting condition of the non-diabetic animals. Compared to SQ insulin, insulin-CAGE capsules provided a slower rate of BGL drop and a sustained lowered BGL; this is a potentially preferred BGL profile for patients, as discussed in the previous section.

5.7.3 Effect of delivery timing on efficacy

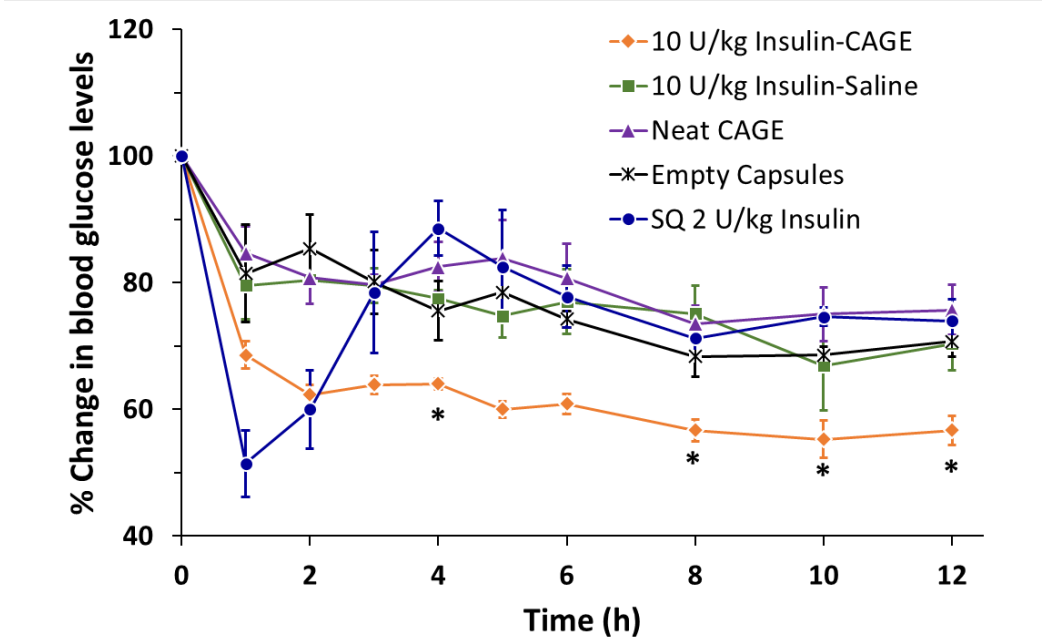


Figure 5.9. Efficacy of CAGE-insulin delivered orally. Capsules containing either insulin-CAGE (10 U/kg, orange diamonds) or neat CAGE (purple triangles) were orally administered to rats. Empty capsules (black x's) and insulin-saline solution (10 U/kg green squares) were dosed to negative control groups. A positive control of insulin (2 U/kg blue circles) was subcutaneously injected to a final control group. Data represented as mean ± S.E. (n=6); *p<0.05, insulin-CAGE compared to SQ insulin.

Investigation of delayed permeabilization by CAGE of the intestinal epithelium was done by dosing rats with neat CAGE capsules followed by capsules containing 10 U/kg insulin powder 30 minutes later. Oral dosing test articles included CAGE-insulin capsules, insulin only

capsules, and insulin-saline liquid. There was a significant efficacy difference observed at the early timepoints (1 and 2 h) between a CAGE-insulin combined capsule and the sequential dosing of CAGE and insulin capsules (**Figure 5.10**). Conversely, no significant difference was seen between insulin solution (10 U/kg) and sequential dosing of insulin and CAGE, suggesting that combined CAGE and insulin is required for effective permeation.

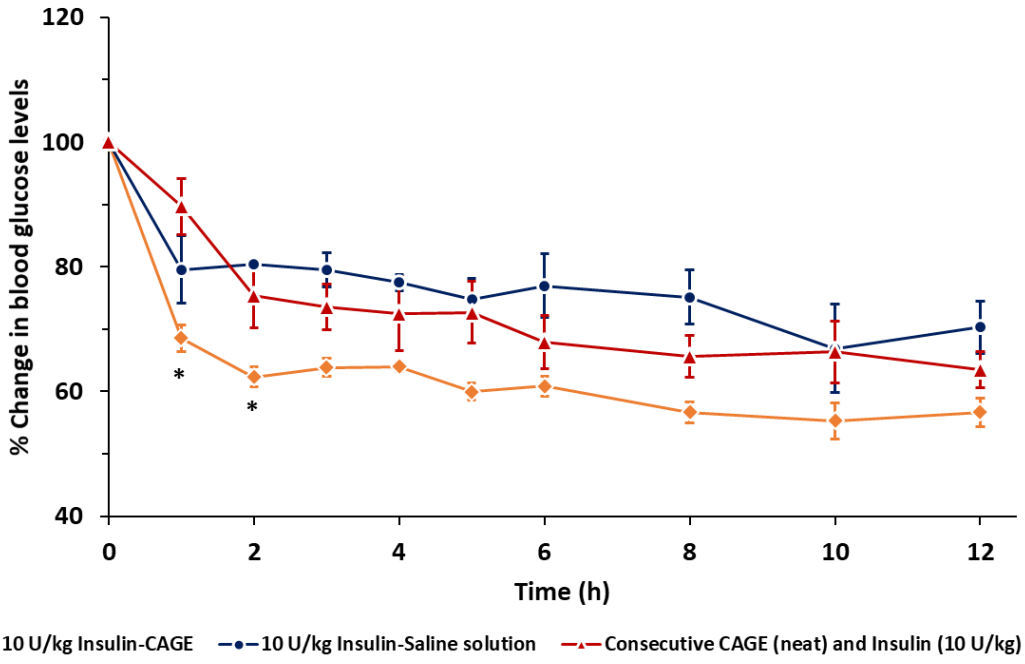


Figure 5.10. Effect of oral delivery of CAGE-insulin capsule compared to sequential CAGE and insulin capsule dosing (30 min delay). CAGE-insulin (10 U/kg, orange diamonds) administered to rats orally in a single capsule had a significantly different BG profile than sequentially dosed CAGE and insulin capsules (10 U/kg, red triangles), whose profile was not statistically different from insulin solution dosed alone (10 U/kg, blue circles). Data represented as mean ± S.E. (n=6).

5.8 Toxicity study

To determine *in vivo* safety, we dosed non-diabetic male Wistar rats with either neat CAGE, 10 U/kg insulin, or 10 U/kg CAGE-insulin capsules once a day for 7 days. Following the study, sections of the small intestines were examined for histological changes via hematoxylin and eosin (H&E) staining by Dr. Renwei Chen of UCSB, who saw no remarkable morphological

differences between the three groups (**Figure 5.11**). This finding suggests oral tolerability of CAGE, but additional studies must be conducted. It should be noted that the daily CAGE dose in the study was 80 mg (one capsule), containing 27 mg choline and 53 mg geranic acid. This is much lower than the reported LD_{50} of either component, 3,700 mg/kg and 3,400 mg/kg for rats, respectively (Davis, et al., 1978) (Anonymous, 1979), and lower than the recommended daily dose of choline (550 mg for men and 425 mg for women) (Zeisel & Costa, 2009).

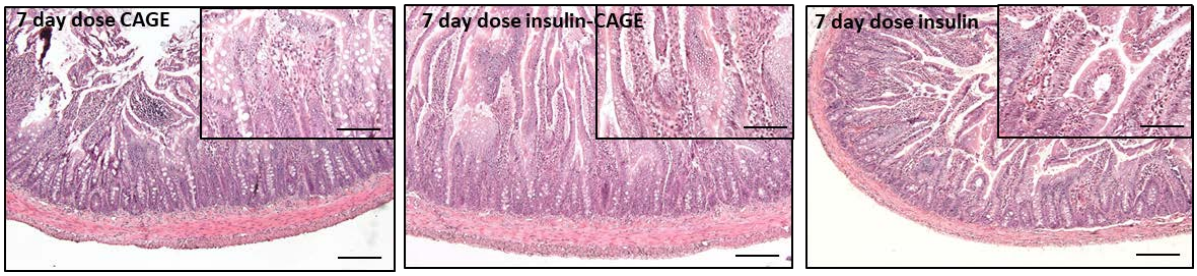


Figure 5.11. Intestinal cross-sections stained with Hematoxylin & Eosin. Animals were dosed once a day for 7 days with capsules containing either neat CAGE (left), insulin-CAGE (middle) or insulin powder (right). Scale bar: 200 µm; inserts are mucosal surface with scale bar of 50 µm.

5.9 Summary

Based on this set of studies, we believe that CAGE assists insulin transport across the mucus layer and mediates paracellular uptake through the opening of tight junctions. Additionally, it was shown that CAGE can reduce the degradation of insulin by trypsin (Banerjee, et al., 2018). CAGE appears to be biocompatible with intestinal tissue, a critical quality for oral delivery systems. Thus, CAGE fills the role of protease inhibitor, permeation enhancer, and mucal transport facilitator in a single compound, without completed modification, conjugation or encapsulation strategies.

While insulin was used in this study, the applicability of CAGE to deliver other peptides orally is a logical conclusion. GLP-1 agonists, used in the treatment of Type 2 diabetes, growth

hormone, and other peptides are all good targets for further study with CAGE. Many of these do not have the narrow therapeutic window of insulin. Taken as a rapid-acting form as a meal bolus, and long-acting form as a daily control of basal glucose levels released by the liver, the pursuit of an oral insulin formulation is likely to best succeed by first focusing on the transition of the basal insulin formulation from injection or oral.

Chapter 6

6. Future directions

The three studies described here combined with previous work in the area of transdermal drug delivery and antimicrobial action provide convincing proof that CAGE has multi-faceted capabilities to address several healthcare challenges, and we have identified three important future goals to enable the robust design of CAGE-based treatments.

6.1 Interaction of CAGE and biological systems

Elucidating how CAGE interacts with a variety of biological systems should continue to be a priority in any future work with the IL as a drug delivery system or antimicrobial. For the antimicrobial effort, continued use of a mix of MD simulation and experimental studies on Gram positive bacteria, fungi and viruses will provide mechanistic insights. Study of the interaction between components of biofilm matrices and CAGE should also be included. For transdermal work, the priority should be on describing the change in lipid profile in skin as a result of CAGE application. This can be accomplished through lipidomic studies which make use of mass spectroscopy and which have been successfully used to study variability in skin due to gender and aging (t'Kindt, et al., 2012) (Smeden, et al., 2014) (Sadowski, et al., 2017). For oral delivery, studies to better elucidate CAGE's effect on epithelial cells *in vivo* is key to improving confidence that CAGE reliably and repeatably performs in the difficult environment of the digestive tract.

6.2 CAGE characterization

Continuing to build out the characterization panel for CAGE and its variants will provide insight into how to tune CAGE for a specific task. Great strides have been made into understanding how CAGE acts in the presence of water (data not shown); expanding these basic studies to other co-solvents and cargo, like organics and proteins is an important next step.

6.3 Investigation of the CAGE landscape

Characterizing the larger landscape of CAGE variants is another key area that will inform all the potential use areas. While work is underway to build out from the simple ion ratio panel used here, a larger effort, undertaken within a thoughtful and logical framework of choline substitutions and different anion options will continue to provide insights as to how the constituents act. For example, fatty acids are well known penetration enhancers; one, caprate (decanoic acid) was used to make choline-caprate. Due to the melting point of caprate, the final compound was gelatinous, but other shorter chain fatty acids with lower melting points could change the form while still imparting good permeation enhancment. For example, octanoic and nonanoic acids have lower melting points and are reported to be less toxic to endothelial cells while still providing permeation effects (Brayden, et al., 2014).

Chapter 7

7. Experimental methods for CAGE

Materials

Geranic acid, choline bicarbonate, BSA, fluorescein isothiocyanate (FITC)-insulin, FITC-BSA, DMSO, DGME, FITC-dextran, dimethyl palmitoyl ammonio propanesulfonate (PPS), sodium caprate (98% purity), pig mucin, coumarin-6, human insulin and human insulin ELISA kit were purchased from Sigma-Aldrich (St. Louis, MO, USA). FITC-OVA was purchased from ThermoFisher (Rockford, IL, USA). Caco-2 human colorectal adenocarcinoma cells were bought from American Type Culture Collection (Manassas, VA, USA) while Dulbecco modified eagle medium (DMEM) with or without phenol red, fetal bovine serum (FBS), penicillin/streptomycin (P/S) solution, Dulbecco's Modified Eagle Medium (DMEM), fetal bovine serum (FBS), Penicillin/Streptomycin (P/S, 5,000 units/mL penicillin with 5,000 ug/mL streptomycin), MTT powder (3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide), Dulbecco's phosphate buffered saline (DPBS), 0.25% trypsin solution, propidium iodide and Remel^{TM} 5% sheep's blood agar plates were purchased from Thermo Fisher Scientific (Waltham, MA, USA). Intestinal epithelium growth medium comprising basal seeding medium (BSM), enterocyte differentiation medium (EDM) and MITO+ serum extender was purchased from Corning (Corning, NY, USA). Millicell®-PCF cell culture inserts (3.0 μm pore size, 12 mm diameter) and the TEER measuring device, Millicell®-ERS were obtained from Millipore Sigma (Burlington, MA, USA) while TEER measuring electrodes were obtained from World Precision Instruments, Inc (Sarasota, FL, USA). Tritium labeled BSA (^{3}H-BSA) was purchased from American Radiolabeled Chemicals Inc. (St. Louis,

MO, USA). Porcine skin was obtained from Lampire Biological Laboratories (Pipersville, PA, USA). Capsule oral gavage and size 9 elongated capsules were obtained from Torpac (Fairfield, MA, USA). Male Wistar rats weighing between 200 – 300 g were purchased from Charles River Laboratories (Wilmington, MA, USA). The blood glucose measuring meter (Aimstrip plus) and strips were bought from Fisher Scientific (Pittsburgh, PA, USA). BD Vacutainer® red top tubes were obtained from Becton, Dickinson and Company (Franklin Lanes, NJ, USA). Eudragit polymer was purchased from Evonik (Essen, Germany). *E. coli* BL21(DE3) was procured from New England BioLabs (Ipswich, MA, USA). Lucifer Yellow (LY) and Tryptic Soy Broth (TSB) was purchased from VWR (Radnor, PA, USA).

7.1 Characterization of CAGE

To prepare CAGE 1:2, two equivalents of neat geranic acid (20 g, 0.119 mol), recrystallized at least five times at −70 °C in acetone, were added to one equivalent of choline bicarbonate (80 wt % solution, 12.274 g, 0.059 mol) in a 500-mL round-bottom flask. The mixture was stirred at 40 °C until CO_2 production ceased. Remaining water was removed via rotary evaporation at 60°C for 20 min, and further dried in a vacuum oven for 48 h at 60°C. Choline bicarbonate and geranic acid molar ratios of 1:4, 1:2, 1:1 and 2:1 were used during the salt metathesis reaction step to create the 4 variants.

7.1.1 Nuclear Magnetic Resonance (NMR) Spectroscopy

Each CAGE formulation was characterized via NMR spectroscopy using an Agilent DD2 600 MHz spectrometer by placing dried, neat CAGE into an NMR tube with a co-axial insert filled with DMSO-d_6. NMR studies were conducted by Dr. Eden Tanner.

CAGE 1:1: ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.25 (d, *J* = 6.5 Hz, 6H, C(C<u>H</u>$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOOH), 1.68 (s, 3H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(C<u>H</u>$_3$)CHCOOH), 1.71 (m, 4H, C(CH$_3$)$_2$CHC<u>H</u>$_2$C<u>H</u>$_2$C(CH$_3$)CHCOOH), 2.94 (s, 9H, N(C<u>H</u>$_3$)$_3$CH$_2$CH$_2$OH), 3.25 (s, 2H, N(CH$_3$)$_3$C<u>H</u>$_2$CH$_2$OH), 3.65 (s, 2H, N(CH$_3$)$_3$CH$_2$C<u>H</u>$_2$OH), 4.75 (s, H, C(CH$_3$)$_2$C<u>H</u>CH$_2$CH$_2$C(CH$_3$)CHCOOH), 5.28 (s, H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)C<u>H</u>COOH), 5.85 (s, b, 2H, N(CH$_3$)$_3$CH$_2$CH$_2$O<u>H</u>: C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOO<u>H</u>). ^{13}C NMR (600 MHz, DMSO-d_6) 20, 28, 29, 43, 56, 58, 70, 127, 133, 148, 175 ppm.

CAGE 1:2: ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.31 (d, *J* = 6.5 Hz, 12H, C(C<u>H</u>$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOOH), 1.76 (s, 6H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(C<u>H</u>$_3$)CHCOOH), 1.78 (m, 8H, C(CH$_3$)$_2$CHC<u>H</u>$_2$C<u>H</u>$_2$C(CH$_3$)CHCOOH), 2.98 (s, 9H, N(C<u>H</u>$_3$)$_3$CH$_2$CH$_2$OH), 3.3 (s, 2H, N(CH$_3$)$_3$C<u>H</u>$_2$CH$_2$OH), 3.65 (s, 2H, N(CH$_3$)$_3$CH$_2$C<u>H</u>$_2$OH), 4.79 (s, 2H, C(CH$_3$)$_2$C<u>H</u>CH$_2$CH$_2$C(CH$_3$)CHCOOH), 5.41 (s, 2H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)C<u>H</u>COOH), 5.85 (s, b, 3H, N(CH$_3$)$_3$CH$_2$CH$_2$O<u>H</u>: C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOO<u>H</u>). ^{13}C NMR (600 MHz, DMSO-d_6) 20, 27, 29, 43, 56, 58, 70, 123, 126, 133, 154, 173 ppm.

The CAGE 1:2 variant NMR spectra was in good agreement with earlier characterizations.

CAGE 1:4: ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.32 (d, *J* = 6.5 Hz, 24H, C(C<u>H</u>$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOOH), 1.78 (m, 28H, C(CH$_3$)$_2$CHC<u>H</u>$_2$C<u>H</u>$_2$C(C<u>H</u>$_3$)CHCOOH), 3.0 (s, 9H, N(C<u>H</u>$_3$)$_3$CH$_2$CH$_2$OH), 3.3 (s, 2H, N(CH$_3$)$_3$C<u>H</u>$_2$CH$_2$OH), 3.8 (s, 2H, N(CH$_3$)$_3$CH$_2$C<u>H</u>$_2$OH), 4.81 (s, 4H, C(CH$_3$)$_2$C<u>H</u>CH$_2$CH$_2$C(CH$_3$)CHCOOH), 5.51 (s, 4H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)C<u>H</u>COOH), 10.4 (s, b, 5H, N(CH$_3$)$_3$CH$_2$CH$_2$O<u>H</u>: C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOO<u>H</u>). ^{13}C NMR (600 MHz, DMSO-d_6) δ 20, 28, 29, 43, 56, 58, 70, 121, 126, 134, 158, 173.

CAGE 2:1 ^{1}H NMR (600 MHz, DMSO-d_6) δ 1.18 (d, *J* = 6.5 Hz, 6H, C(C<u>H</u>$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOOH), 1.52 (s, 3H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(C<u>H</u>$_3$)CHCOOH), 1.54 (m, 4H, C(CH$_3$)$_2$CHC<u>H</u>$_2$C<u>H</u>$_2$C(CH$_3$)CHCOOH), 2.78 (s, 18H, N(C<u>H</u>$_3$)$_3$CH$_2$CH$_2$OH), 3.15 (s, 4H, N(CH$_3$)$_3$C<u>H</u>$_2$CH$_2$OH), 3.45 (s, 4H, N(CH$_3$)$_3$CH$_2$C<u>H</u>$_2$OH), 4.62 (s, H, C(CH$_3$)$_2$C<u>H</u>CH$_2$CH$_2$C(CH$_3$)CHCOOH), 5.18 (s, H, C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)C<u>H</u>COOH), 5.7 (s, b, 2H, N(CH$_3$)$_3$CH$_2$CH$_2$O<u>H</u>: C(CH$_3$)$_2$CHCH$_2$CH$_2$C(CH$_3$)CHCOO<u>H</u>). ^{13}C NMR (600 MHz, DMSO-d_6) δ 17, 25, 26, 40, 53, 55, 67, 123, 125, 130, 146, 160, 172.

7.1.2 Thermogravimetric Analysis (TGA)

The decomposition temperatures (T_{dec}) of the variants were measured using a TA Discovery Thermo-Gravimetric analyzer (TA Instruments, New Castle, DE, USA) with a ramp of 10 °C/min from 100 to 800 °C. Decomposition temperatures were determined from the onset of mass loss derived from the intersection of the baseline before thermal decomposition with the tangent during mass loss.

7.1.3 Differential Scanning Calorimetry (DSC)

Glass transition temperatures (T_g) and cold crystallization points (T_{cc}) of the variants were analyzed using a Modulated DSC Q2000 (TA Instruments, New Castle, DE, USA). A sample between 6 – 10 mg was sealed in an aluminum crucible and cooled to either -150 or -125 °C, then scanned to 125 °C with a heating rate of 10 °C /min for a total of 6 cycles.

7.1.4 Density

Density was measured using an ARES-LS1 rheometer with a 25 mm diameter parallel plate to calculate the volume of a known weight of CAGE 1:2 at a known temperature.

7.1.5 Viscosity

Viscosity was measured on a TA Instruments AR-G2 rheometer using a 40 mm diameter aluminum 2° cone. CAGE 1:2 was placed on the bottom plate, and the sample was equilibrated at 25 °C for 2 minutes prior to the experiment being conducted. A steady-state flow method was employed, with 5 points per decade being recorded from 1 to 100 Pa.

7.1.6 Conductivity

Conductivity was measured using a Horiba Laqua conductivity meter (model DS-71) with a 3574-10C probe. A sample of ionic liquid was introduced into the clean and dried probe and a reading was taken after the output stabilized. Conductivity was also measured using dielectric relaxation spectroscopy in the Segalman lab at UCSB. Briefly, CAGE was pressed between 2 stainless steel electrodes and through-plate spectroscopy was performed on a Biologic VSP-300, applying a sinusoidal voltage in an appropriate frequency range. DC conductivity was determined from the real component of the complex conductivity.

7.1.4 Assessment of insulin stability in CAGE

Overnight study. Samples containing human insulin (100 U, 3.5 mg) suspended in CAGE (1 mL) were incubated at physiological temperature (37 °C), room temperature (25 °C) or under refrigeration (4 °C) for 17 h in 1.5 mL micro centrifuge tubes. Samples were then centrifuged for 10 minutes at 10,000g and the CAGE was removed via pipetting, taking care not to disturb the soft insulin pellet in the bottom of the tube. The pellet was washed with 1 mL PBS and centrifuged again. PBS-CAGE was pipetted from the tube and washing/centrifugation steps were repeated until the insulin did not form a pellet in the tube during centrifugation. Circular

dichroism spectrophotometry (Jasco J-1500, Easton, MD, USA) was performed to collect spectra in the far-UV region (190 – 250 nm) using rectangular quartz cells with a 1mm path length (Starna Cells, 1-Q-1, Atascadero, CA, USA) loaded with 400 ul of sample. Insulin in PBS was used to collect a control spectrum.

Long-term study. Samples were prepared as for the overnight study and stored at RT or 4 °C. After 1 month, and approximately each month thereafter for 4 months total, samples were centrifuged for 10 min at 10 000 X g, the CAGE was removed via pipette, and the soft insulin pellet was washed with 1 mL phosphate buffered saline (PBS) and centrifuged again. PBS-CAGE was removed, and the washing/centrifugation steps repeated until the insulin did not form a pellet during centrifugation. To collect spectra in the far-UV region (190-250 nm) indicting protein secondary structures, circular dichroism spectrophotometry (Jasco J-1500, Easton, MD) was performed with rectangular quartz cells (1mm path length, Starna Cells, 1-Q-q, Atascadero, CA) loaded with 400 uL of sample.

7.2 Bacterial studies

7.2.1 Cell culturing

For all bacterial studies, 10 μL of *E. coli* from frozen stock was added to 5 mL TSB and incubated overnight at 37 °C. The next day, 20 μL of the culture was added to 5 mL of fresh media and allowed to grow at 37 °C until the OD_{600} reached 0.5 (4×10^8 CFU/mL). This culture protocol was used to prepare cells for the experiments described below.

7.2.2 Measurement of minimum bactericidal concentration (MBC)

E. coli cells were grown as described above. 20 μL of cell suspension was added to each well. 180 μL of 1:2 serial dilutions of CAGE 1:4, 1:2, 1:1 or 2:1, from 0.8 – 50 mM were added to duplicate or triplicate wells. Geranic acid was serially diluted for a test concentration range of 31 - 1000 mM, and choline bicarbonate was tested at 32 - 2000 mM. 200 μL of TSB was used as a negative control, and 180 μL media plus 20 μL cell suspension was added used as a positive control. The final cell concentration per well was approximately 4×10^7. The plate was incubated for 2 h at 37 °C. 20 ul was removed from each well and added to 10 mL TSB, and 50 μL of this culture was plated. Plates were incubated overnight at 37 °C. The minimum bactericidal concentration for each CAGE variant or the reactants geranic acid or choline bicarbonate was determined as the lowest concentration plate with no growth.

7.2.3 Kinetic study

A 96-well plate was seeded with 20 μL *E. coli* culture prepared as above, at either 1×, 2× or 4× the cell concentration used in the MBC study (~4×10^8 CFU/mL). 180 μL of either TSB media (untreated control), 6.5 or 13 mM 1:2 CAGE, diluted with TSB media, was added into 4 replicate wells. Blanks for each concentration contained 200 μL of TSB or the 1:2 CAGE dilutions only. The plate was incubated for 14 h at 37 °C in a BioTek Synergy NEO HTS microplate reader (Winooski, VT, USA) and the OD_{600} was measured every 30 minutes.

7.2.4 Flow cytometry studies

Following 2 h incubation in 1:2 CAGE at sub-lethal (8 mM) or lethal (26 mM) concentrations determined from MBC testing, cells were centrifuged at 3,000g for 5 min, the media removed.

PBS was added, and the samples were vortexed to break up the cell pellet, effectively washing media from the cells. The cells were recentrifuged, the supernatant decanted, and then resuspended in PBS. Propidium iodide (PI) was added (1 μL) and samples were incubated for 15 minutes at room temperature in the dark. Samples were analyzed for dead cells (positive for PI) using a BD FACSAria I (Franklin Lakes, NJ, USA) flow cytometer red laser emitting at 488 nm with a 695/40 filter. The positive PI signal gate was set to exclude 99.5% of unstained (background) cells. Culture viability was confirmed by overnight incubation of 10 μL of unstained cells in TSB media.

7.2.5 SEM imaging

E. coli cells were grown as described above. Following a 2 h incubation in each of the 4 CAGE variants at 37 °C, cell suspensions were centrifuged for 3000g for 5 minutes and the media was removed. To fix the cells, 2.5 % (w/v) glutaraldehyde was added and the cells refrigerated (4 °C) for 2 h. Cells were centrifuged, the supernatant removed, and dehydrated for 5 minutes in 30 and 70% (v/v) ethanol followed by 15 minutes in absolute ethanol. Multiple dehydration and washing steps were eliminated to retain cells. A 5 μL volume was applied to an SEM stub and then air dried overnight followed by coating with gold/palladium alloy for 90 s before analysis in a FEI Nova 650 FEG SEM (Hilsboro, OR) connected with Oxford Inca X-ray EDX system detector (Abingdon, Oxfordshire, UK). The images were taken at 40,000 X magnification at a beam voltage of 5 kV.

7.2.6 FTIR characterization of cell membrane changes

Cells were incubated for 2 or 24 h in 13 mM 1:2 CAGE. The OD_{600} was measured, and aliquots containing equal numbers of cells were made. An aliquot from a control sample incubated in medium only was prepared in the same way. The samples were centrifuged, the supernatant decanted, and MilliQ water added to wash the cells of residual CAGE/medium. The cells were recentrifuged, decanted and 20 μL of MilliQ water was added and mixed. The cells were dried overnight in a vacuum desiccator at room temperature. A Bruker Vertex 70 FTIR-ATR Spectrometer (Billerica, MA, USA) was used to analyze the cells by resuspending the cells in 10 μL of MilliQ water, placing the volume on the FTIR-ATR crystal, and allowing the water to evaporate while recording the spectrum every 3-5 minutes with spectral resolution at 4 cm^{-1} and sampling at 100 scans per sample. The final spectrum was deemed to be that which had no change in absorbance over the entire range (600-4000 cm^{-1}) for 3 cycles of sampling. FTIR spectra were automatically atmosphere corrected and a manual baseline correction was performed. To ensure that the peak height increases were not due to variations in cell concentration due to collection, dilution or measurement error, a set of control aliquots containing 0.5×, 1× and 1.5× the cell count was prepared and analyzed on the FTIR. The 1.5× standard spectrum was below both of the treated sample spectra, and it is very unlikely that the preparation and measurement errors could have resulted in any greater OD than 1.5× the control. The 24 h treatment was repeated for a total of n = 4 to check the reproducibility of the FTIR spectra. Frequency data was collected to 6 significant figures. Spectra were atmosphere and baseline corrected using the OPUS FTIR software. Peak centers were identified using the peak-finding function in Origin (Origin 2017, Northampton, MA, USA).

7.2.7 Antibacterial resistance challenges

Repeated 2 h challenge. The MBC study described above was repeated six times, using a colony from the highest concentration plate with growth from the previous test. The colony was grown overnight from the plate and subcultured to an OD_{600} of 0.5. This culture was challenged and plated, and the highest concentration plate with growth carried into the next round of testing. A colony from the ultimate round was also grown in TSB and prepared for FTIR as above.

24 and 48 h challenge. E. coli cells were grown as described above. 0.5 mL of cells at 4×10^8 CFU/mL was added to 4.5 mL of CAGE 1:2 diluted with TSB at a sub-lethal concentration of 13 mM for a final cell concentration of 4×10^7. Following incubation at 37 °C for with 24 or 48 h the culture OD_{600} was recorded, and the cells plated. A colony from this plate was grown overnight and subcultured in similar fashion to the wild-type and the MBC for 1:2 CAGE was determined. A control wild-type sample was included in the MBC study to check for run-to-run variations.

7.3 *In vitro* studies

7.3.1 Determination of IC_{50} for Caco-2 cells in 1:2 CAGE (96-well plate)

Human epithelial colorectal cells (Caco-2, ATCC) were seeded into 96-well microtiter plates and grown for 21 days to a fully differentiated and confluent monolayer in DMEM supplemented with 10% FBS and 1% P/S at 37 °C, 5% CO_2. Cell media was changed every third day in the first week and every alternate day in the remaining weeks. CAGE was diluted with DMEM to concentrations ranging from 25 to 3.125 mM. Three different sets of CAGE

dilutions were made (3 dilution replicates). The media was aspirated from each well containing the confluent Caco2 cells and each dilution was dispensed (100 uL per well) into 6 wells (6 cell replicates). Control wells were filled with media only. The cells were incubated at 37 °C, 5% CO_2 for 12, 24 or 48 h. At each time point, the CAGE:media mixture was aspirated from the wells, and cell viability was assessed using an MTT assay. MTT powder was mixed with media to a concentration of 0.5 mg/mL, added (100 uL) to each well, and incubated at 37 °C, 5% CO_2 for 4 h. The MTT solution was removed and 100 uL DMSO was added to each well. The plates were wrapped in foil and shaken for 20 min, then absorbance was read at 570 nm using a microplate reader (M220 Infinite Pro, Tecan Group Ltd, Morrisville, NY). Absorbance readings were normalized using the non-treated cell viability values and used to determine IC_{50} values with a dose response curve in Prism (v7.0, GraphPad).

7.3.2 Caco-2 monolayer growth for transport studies

A 3-day rapid Caco-2 growth system was used for transport studies. Cells were placed in Corning® BSM supplemented with MITO serum+ extender and seeded at density of 400,000 cells/mL on Millicell® PCF inserts placed inside 24-well plates. 500 μL of cells containing medium was placed in apical side while 1000 μL of cell free BSM was put in the basolateral side as per manufacturer recommendation. After 24h of incubation at 37 °C, 5% CO_2, the medium was replaced with same volume of EDM supplemented with MITO serum+ extender for 48h. TEER was measured and inserts with values above 150 ohms.cm^2, indicating sufficient tight junction integrity between cells, were used in the transport studies.

7.3.3 Transport studies (LY, Coumarin-6, FITC-Insulin, FITC-4 kDa dextran)

Before the start of the experiment, the medium in the transwells was replaced with DMEM without phenol red, FBS or P/S in both the apical (200 µL) and basolateral side (600 µL) and the cells were incubated for 30 minutes. Then the inserts were transferred to a new plate and the medium in the apical side was aspirated off and replaced with 200 µL of 500 µg/mL of either Lucifer Yellow (LY), 4 kDa FITC-dextran, FITC-insulin, or 5 µg/mL of coumarin-6 in DMSO (10% final concentration), prepared with or without 10, 25 and 50 mM CAGE and solubilized in DMEM free of phenol red, FBS and P/S. 600 µL of DMEM was then added to the basolateral chamber, and, a 100 µL aliquot was immediately withdrawn from the basolateral side and replaced with an equal volume of fresh DMEM. During the study, the transwell plates were placed inside an incubator at 37 °C, 5% CO_2 on a shaker rotating at 100 rpm and taken out at hourly intervals for sampling. After the end of study at 5h, fluorescence in the aliquots with LY, FITC-dextran, FITC-insulin and coumarin-6 were measured using a plate reader (BioTek Synergy NEO HTS microplate reader (Winooski, VT, USA) at 485/530 (LY), 495/520 (FITC) and 468/568 (Coumarin-6) nm excitation/emission wavelengths. Calibration standards for each transport marker molecule were prepared and analyzed to plot basolateral chamber concentration *vs* time plots.

7.3.4 TEER studies

Before the start of the experiment, the medium in the transwells was replaced with DMEM without phenol red, FBS or P/S in both the apical (200 µL) and basolateral side (600 µL) and the cells were incubated at least 30 minutes. The TEER was taken, and then the inserts were transferred to a new plate and the medium in the apical side was aspirated off and replaced

with 200 µL of either 0, 10, 25 and 50 mM CAGE, 10 mM sodium caprate or 0.03 w/v% PPS. 600 µl of DMEM was added to the basolateral chamber. During the study, the transwell plates were placed inside an incubator at 37 °C, 5% CO_2 on a shaker rotating at 100 rpm and taken out at hourly intervals to measure TEER. After 5h, the media in both chambers was replaced with DMEM and the transwell plates were incubated for an additional 19h without the test articles. At 24h, the TEER was measured for a final time. TEER was plotted as % change from initial value *vs* time.

7.3.5 Determination of IC_{50} for Caco-2 cells in 1:2 CAGE (inserts)

Caco-2 monolayers grown on inserts via the 3-day protocol and used in the TEER study were also analyzed for cell viability via an MTT assay at 5 and 24h.

7.3.6 CAGE-Mucin interaction

To determine the effect of CAGE on mucus viscosity, simulated mucus (SM) was made by mixing 2% dried pig mucin with saline and vortexed until the mucin was completely dissolved. Varying amounts of CAGE (0, 1, 5, 10, 25 and 50%) were added to the SM and viscosity was measured in triplicate across a shear rate range of 1 – 100 1/s at 25 °C using an AR-G2 rheometer with a 40 mm diameter aluminum 2° cone geometry (TA Instruments, New Castle, DE, USA).

7.4 *Ex vivo* studies

7.4.1 Visualization of protein-CAGE penetration in porcine skin

Ex vivo penetration into skin samples was measured according to previously published methods using Franz diffusion cells (FDC) (Karande, et al., 2006). The acceptor chamber of

an FDC was filled with PBS. Thawed porcine skin was positioned with the SC facing up and stretched to remove folds. Air bubbles were removed from the acceptor chamber and skin integrity was confirmed by measuring conductivity. The donor chamber was filled with 300 μL of CAGE, 50:50 (v/v) PBS:Ethanol, 50:50 (v/v) PBS:DGME, or PBS (control) containing FITC-BSA (1 mg/mL), FITC-OVA (1 mg/mL), or FITC-insulin (1 mg/mL, 29 U/mL) and the FDC were incubated for 24 h at 37 °C on a stir plate to provide mixing of the acceptor fluid. Protein loading in the samples were 0.1% w/w. Skin samples were harvested after 24 h, frozen in optimum cutting temperature compound and sectioned into 20 μm thick slices using a cryostat (CM1850 Leica Biosystems, Buffalo Grove, IL, USA). The tissue sections were mounted on positively charged glass slides to assist in adherence (Fisher Scientific, Pittsburgh, PA, USA) and imaged on a confocal microscope (Olympus Fluoview 1000 Spectral Confocal, Tokyo, Japan).

7.4.2 Quantitation of BSA-CAGE penetration through porcine skin

The experimental setup was similar to the one described above. BSA (1 mg/mL) was added to CAGE or PBS (control) and then ^{3}H-labeled BSA was added to produce a final concentration of 0.5 μCi/mL. 300 μL of this solution was added to the donor chamber and incubated for 6, 12, 24 or 48 h at 37 °C with stirring. For tests using DGME or ethanol, a 50:50 solution with PBS was prepared as previously described. BSA and ^{3}H-labeled BSA was added at the same concentrations noted above, and porcine skin-containing diffusion cells were dosed and incubated for 24 h. At the end of incubation period, the donor solution was removed, and the skin was washed with PBS and dried. Tape stripping was used to separate the SC from the epidermis. A total of ten strips, each corresponding to a SC layer, removes most of the SC. The

epidermis was removed from the dermis using a surgical scalpel. Samples from the SC, epidermis, and dermis layers and from the acceptor solution were incubated in Solvable™ (Perkin-Elmer, Waltham, MA, USA) overnight, and the concentration of the ^{3}H-labeled BSA was measured using a scintillation counter (Tri-Carb® 2100 TR, PerkinElmer, Waltham, MA, USA).

To measure insulin-CAGE penetration, FITC-insulin (1 mg/mL) was added to CAGE and 300 µL was applied to porcine skin in a diffusion cell and incubated for 24 and 48 h. Separate skin layers were obtained as above and samples were incubated in 3 mL of 1:1 PBS:methanol overnight, centrifuged (5 min, 14000g) and the concentration of FITC-insulin was measured in triplicate using a plate reader and 96-well black plate with a clear bottom (M220 Infinite Pro, Tecan, Männedorf, Switzerland).

7.4.3 Effect of CAGE on the stratum corneum

The SC was separated from full thickness skin using published methods (Zakrewsky, et al., 2014). In brief, the epidermis was separated from the dermis of a skin sample soaked in a heated water bath. Trypsin was used to digest the epidermis from the SC, which was then washed in PBS and dried at room temperature for 72 h. A control FTIR spectrum (Nicolet iS10 FTIR with a Smart Diamond ATR accessory) was obtained for each 1.5 x 1.5 cm SC piece, after which it was incubated for 24 h with 2 mL of PBS, one of the 4 CAGE variants alone or CAGE 1:2 with BSA or insulin added (1 mg/mL). The SC samples were rinsed by soaking in 5 mL PBS 3 times and dried for 72h at room temperature. A second FTIR spectrum was taken for each SC sample and compared to the pre-treatment spectrum to evaluate the effect of CAGE alone or protein-CAGE on the SC structure. FTIR spectra of human insulin powder, BSA

crystals and PBS were also collected for reference. Spectra were baseline and atmosphere corrected using the OMIC software provided with the Nicolet spectrometer and OriginPro (OriginLab, Northampton, MA, USA) for peak finding and peak deconvolution.

7.5 *In vivo* studies

All animal experiments were performed in accordance with the University of California Santa Barbara animal care committee guidelines and to the Guide for the Care and Use of Animals of the Institute of Laboratory Animal Resources, National Research Council.

7.5.1 Transdermal application of CAGE-Insulin

Male Wistar rats (non-diabetic) were fasted for 8h with free access to water. For transdermal test treatments (CAGE-Insulin, PBS-Insulin, CAGE alone), hair from a 5 cm x 5 cm square of skin on the back was clipped, and about 1 mL of a test treatment was applied. The application site was covered with gauze affixed with tape. 1 U/kg of insulin in saline was injected subcutaneously to another group of rats. Blood glucose levels were measured at least every 2h for 12 h with a commercial glucose meter and results were plotted as % change in blood glucose compared to the initial level *vs* time.

7.5.2 Intrajejunal dosing of CAGE and insulin

Male Wistar rats (non-diabetic) were fasted for 12h overnight with free access to water. The rats were anesthetized, the abdominal hair clipped, and the surgery area was prepped using 70% ethanol and betadine. An incision was made in the abdomen to expose the intestine. The jejunum was located and injected with 100 μL of insulin-CAGE or 100 μL of controls. The intestinal section was replaced in the abdomen and the incision closed. The initial blood

glucose was taken with a commercial glucose monitor after the intestine was exposed and every 0.5 h till the end of study. The anesthetized animals were kept on temperature controlled warming pads before the surgery and also covered with a towel after surgery and euthanized after 5h. The intestinal section around the injection site was removed for histological examination. A separate group of rats were subcutaneously injected with 2 U/kg insulin in saline for comparison of efficacy. The results were plotted as % change in blood glucose levels with respect to initial reading *vs* time. Pharmacokinetics of insulin-CAGE were evaluated by collecting around 250 μL blood in BD Vacutainer® red top tubes at 0, 1, 2, 3 and 5 h from rats injected intrajejunally and subcutaneously. Blood samples were left undisturbed at RT for at least 30 minutes to clot, followed by centrifugation at 2,000 g for 10 minutes. The clear supernatant (serum) was then collected into clean tubes, stored in ice during the procedure and then at -20 °C until an insulin concentration determination was performed via a human ELISA kit. Pharmacokinetic parameters including area under the curve (AUC) and % bioavailability were calculated from a serum insulin concentration *vs* time plot.

7.5.3 Oral dosing of CAGE and insulin

For the oral efficacy study, elongated size 9 capsules were filled with 80 μL of either 10 U/kg insulin-CAGE, neat CAGE or left empty. the capsules were then enterically coated at least two times with 12.5% w/v Eudragit® L-100 dissolved in isopropanol. Male Wistar rats (non-diabetic) were fasted for 8h overnight with free access to water. Following the fasting, the test articles in capsules (insulin-CAGE, neat CAGE or empty) or in liquid form (10 U/kg insulin-saline) were orally administered followed by a 5 mg/kg metoclopramide injection to induce gastric emptying. Initial blood glucose was recorded, and fasting was continued with free

access to water throughout the 12 h study. A separate group of rats were subcutaneously injected with 2 U/kg insulin in saline for comparison of efficacy. Blood glucose was evaluated hourly throughout the 12 h study as described in section 7.4.2, and the results plotted as % change in blood glucose compared to the initial level *vs* time.

For the sequential dosing study, following the fasting, 80 μL neat CAGE in enterically coated capsules was orally administered followed by a 5 mg/kg metoclopramide injection to induce gastric emptying. Blood glucose was recorded. 10 U/kg insulin powder in an enterically coated capsule was administered to the rats approximately 30 minutes later. Fasting was continued with free access to water throughout the study. Blood glucose was evaluated hourly throughout the 12 h study, and the results plotted as % change in blood glucose compared to the initial level *vs* time.

7.5.4 7-day toxicology study

Neat CAGE capsules, 10 U/kg insulin-CAGE capsules or 10 U/kg insulin in coated capsules were administered daily for a period of 7 days to non-diabetic male Wistar rats. All animals were euthanized on day 8 and their small intestinal tissue sections collected for tissue histology, which was performed by R. Chen of UCSB.

Chapter 8

8. Introduction to epitope mapping and antibody profiling

Antibodies are a major component of the adaptive immune system, active in humoral-based immunity to identify and neutralize pathogens by exhibiting high binding affinity to epitope regions of antigenic proteins (Janeway, et al., 2001). Since antibodies are highly specific for their target antigenic molecules, patient antibody repertoires can be screened using peptides that mimic their epitopes. The physicochemical properties of these ligands can be used to develop markers for disease to aid in diagnostic assays (Mayeux, 2004). Further, determining and characterizing antibody-targeted molecules specific to a disease can aid in understanding pathogenesis and identifying candidates for therapeutic intervention, such as tolerizing therapies (Shakya & Nandakumar, 2018). This portion of my dissertation uses the identification of antibody-binding peptides in two ways; improving epitope mapping strategies and investigating the humoral antibody repertoire of a model disease, Type 1 diabetes (T1D).

Mapping the interactions between antibodies and antigens can be accomplished using a variety of techniques that, at their heart, all display either random or non-random peptide libraries such that antibodies can selectively bind to their ligand (Paull, et al., 2018). Peptide libraries are a powerful protein engineering technique in the investigation of protein-protein interactions. Information gleaned from interactions can provide information to help establish a link between a protein's genotype and phenotype. Binding events are a function of the individual binding affinity of the peptide and its target and the avidity resulting from multiple copies of the peptide on the display surface.

Common surface display formats include microarrays, bacteriophages, and bacterial and yeast cells (Ullman, et al., 2011). Specifically, antigen arrays are a subset of the human proteome and do not include short peptides and membrane proteins (Meirsch, et al., 2013). Of the cellular approaches, phage display is currently the most widely used display method; however, peptide libraries displayed on the surface of bacteriophage are biased towards particular sequences that allow for rapid phage propagation (Mennuni, et al., 1997) (Farilla, et al., 2002) (Maglio, et al., 2009) (Larman, et al., 2013). Yeast-based libraries enable the display of large sequences that can have conformational aspects, better mimicking structural and therefore structural epitopes (Gai & Wittrup, 2007). Acellular systems are being developed as well; high-throughput sequencing has improved or enabled routes such as ribosomal peptide display (Heyduk & Heyduk, 2014), mRNA display (Ullman, et al., 2011) and B cell gene sequencing (Robinson, 2014). Peptide bacterial display libraries provide a large, unbiased library of epitope mimics (Hall & Daugherty, 2009). Fast replication, ease of manipulation and the ability to construct libraries with billions of clones make bacterial cells attractive candidates for display approaches (Rice, et al., 2006).

Antibody-peptide interactions are governed by receptor-ligand binding, obey Michaelis-Menten kinetics, and can be modeled using an expression determining the affinity of a peptide (P) for a target antibody (A):

$$P + A \leftrightarrow_{k_d}^{k_a} P \cdot A \qquad \frac{d[P \cdot A]}{dt} = k_a[P][A] - k_d[P \cdot A] \qquad [1]$$

Where k_a and k_d represent association (on-rate) and dissociation (off-rate) constants, respectively. At equilibrium:

$$K_{eq} = \frac{k_a}{k_d} = \frac{[P \cdot A]}{[P][A]} \qquad and \qquad K_D = \frac{1}{K_{eq}} \qquad [2]$$

The total number of peptides on a cell is represented by

$$[P_{tot}] = [P_{unbound}] + [P \cdot A] \quad [3]$$

This project used large, random (10^9 – 10^{10} independent transformants) 15-mer or 12-mer peptide libraries (X_{15} or X_{12}) displayed on the N-terminus of a transmembrane protein scaffold of *E. coli* (eCPX) (Rice & Daugherty, 2008) (Pantazes, et al., 2016). *E. coli* strain MC1061 [FaraΔ 139 D(ara-leu)7696 GalE15 GalK16 Δ (lac)X74 rpsL (StrR) hsdR2 (rK−mK +) mcrA mcrB1] was used with surface display vector pB33eCPX. The X_{15} library was constructed with NNS codons while the X_{12} library was constructed using trinucleotide oligomers to eliminate stop codons and normalize amino acid usage frequencies. In both libraries, the peptides are displayed on a protein scaffold made earlier in the Daugherty lab. The outer membrane protein OmpX, which is highly expressed in *E. coli,* was circularly permuted (CPX) (Rice, et al., 2006) and subsequently engineered using directed evolution to yield an enhanced CPX (eCPX) variant that exhibited display characteristics similar to the non-permuted parent and allowed bi-terminal display (Rice & Daugherty, 2008). The tethering of the peptide to a terminus via a flexible linker is thought to more closely approximate solution behavior (Rice, et al., 2006).

Libraries are most often screened via magnetic or fluorescence-based sorting methods (**Figure 8.1**) and can be combined with iterative selection (panning) to achieve objectives such as molecular recognition, affinity maturation or binding specificity (Hall & Daugherty, 2009).

Magnetic-activated cell selection (MACS) is an inexpensive sorting method that utilizes magnetic beads functionalized with proteins that bind to portions of the constant regions of antibodies (Fisher Scientific, 2018). Cells with peptides that have antibodies attached to them are selectively sorted by incubating the cells with beads and then applying a magnet. The

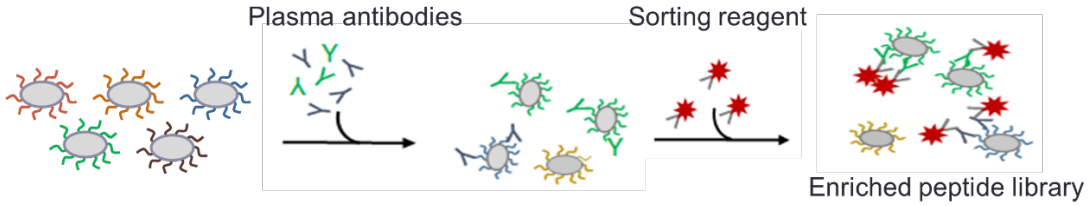


Figure 8.1. Simplified schematic for bacterial display library enrichment procedure. Using magnetic beads or a fluorophore that are attracted to antibodies provides a sorting medium to preferentially select bacterial cells displaying peptides that are bound to antibodies. The sorting protocol can be repeated to enrich the library with peptides having high binding affinities.

supernatant is removed, and the cells are washed several times to remove any unbound cells. The enriched library is grown overnight and can be used for subsequent enrichment steps to enrich the library in highly-binding peptides.

Flow cytometry has proven to be an invaluable tool in quantifying bacterial library enrichment after MACS. We can associate the fluorescent intensity data from Fluorescent Flow Cytometry (FFC) with equation [3] above using the maximum fluorescence value (F_{max}) to represent saturation binding of all peptides on a cell with antibodies, or the total peptide amount, P_{tot}. Using the same rationale, the mean cell fluorescence, MCF, can be related to the amount of P•A complex plus background fluorescence, F_{bkgd}:

$$F_{max} = a[P_{tot}] + F_{bkgd} \qquad MCF = a[P \cdot A] + F_{bkgd} \qquad [4]$$

where a is a proportionality constant. Combining these equations gives:

$$MCF - F_{bkgd} = \frac{(F_{max} - F_{bkgd})[A]}{K_{D,app} + [A]} \qquad [5]$$

The condition of a multivalent ligand (peptides on a cell surface) and a bivalent antibody can affect the avidity. (Krishnamurthy, et al., 2007). Since avidity will influence affinity, we note that K_D is an apparent value. F_{max} and $K_{D,app}$ can be determined by fitting data for several

different antibody concentrations. Avidity effects are accounted for by F_{max}. A low F_{max} and $K_{D,app}$ represents the best overall apparent affinity.

After a library has been enriched to the desired level, the plasmid containing the DNA sequences for the peptides is extracted from the *E. coli* cells. Using PCR, the portion of the plasmid encoding the peptide is amplified and sequenced via NGS. A library used to profile a serum sample typically has a few million sequences with several hundred thousand unique sequences. Using a computer algorithm developed in the Daugherty lab, redundant sequences are combined, resulting in a non-redundant sequence list that can be used to identify motifs via pattern discovery algorithms like MEME (Bailey & Elkan, 1994) or IMUNE (Pantazes, et al., 2016). These motifs can be used to identify epitopes on an antigen, if it is known, or to develop biomarkers for a cohort based on sensitivity (true positives) and specificity (true negatives) measures.

This study utilized bacterial display of a random library to conduct two studies. The first examined the use of random peptide libraries combined with next generation sequencing (NGS) to examine the ability of epitope prediction algorithms to utilize large NGS datasets. The existing algorithms were originally developed to use the much smaller sequence sets obtained from Sanger sequencing of highly screened libraries. The second study investigated the antibody repertoires of a group of newly diagnosed T1D patients in an effort to identify peptides with high specificity and sensitivity to the cohort.

Chapter 9

9. Development of a structural epitope mapping method for NGS

9.1 Motivation

Many antibodies bind structurally-defined epitopes within their antigens. The amino acid residues in these epitopes are discontinuous (i.e., not sequentially continuous) and rely on secondary and higher structures to create the binding surface. This discontinuity and conformational dependence significantly increases the difficulty of identifying discontinuous, as compared to continuous epitopes where sequence similarity can be used. Only a few studies have attempted to estimate how many epitopes might have a structural component; an early estimate by Barlow still widely cited suggested that less than 10% of epitope surfaces are composed of completely sequentially continuous residues (Barlow, et al., 1986). A more recent study of 47 proteins with discontinuous epitopes (i.e. comprised of several segments), found that more than 45% of the epitope segments were comprised of single residues, and the longest segments averaged 4 to 7 residues (Anderson, et al., 2006).

Experimental and computational methods have been developed to predict antibody interaction epitopes (Ahmed, et al., 2016), though definitive interface determination is generally reliant upon crystallography or NMR. Computational, or *in silico*, docking algorithms have proven useful when the structures of both antibody and antigen are known (Meng, et al., 2011) (Kuroda, et al., 2012) (Rapburger, et al., 2007) (Chakrabarti & Janin, 2002). Binding assays based upon ELISA and protein or peptide microarrays can identify potential epitope sections by assessing antibody binding to antigen fragments. Similarly, peptide display technologies have proven useful to identify peptide sequences, in random

(Daugherty, 2007) (Wentzel, et al., 2001) and antigen- or organism-derived libraries (Angelini, et al., 2015) that interact with an antibody of interest. Mimotopes, library derived peptides that mimic the antigen epitope, can help identify epitope residues for targeted studies such as mutagenesis, wherein reduced binding indicates the importance of a residue within the epitope (Hudson, et al., 2012) (Reimer, et al., 2005).

Epitope mapping via peptide display is dependent primarily upon library design, enrichment methods, determination of peptide sequences, and epitope prediction from sequence data. While libraries derived from protein sequences are common, large random peptide libraries (e.g., $>10^9$ members) can provide advantages in terms of their ability to yield peptides that mimic diverse structural epitopes. Typically, several rounds of selection or screening are performed to enrich binders to the antibody of interest. The enriched library typically consists of a few highly-represented sequences that can be identified via sequencing the encoding DNA. When mapping suspected structural epitopes, algorithms that seek to match these mimotopes with residue paths along the antigen's surface are typically employed. Examples include PepSurf, EpiSearch and Pep-3D-Search (Mayrose, et al., 2006) (Negi & Braun, 2009) (Huang, et al., 2008). Applied against a benchmark set of known epitopes, the algorithms typically report less than 50% average sensitivity (defined as the percentage of true positive residues in a predicted set) and precision (defined as the ratio of true to false interface residues in a predicted set) (Negi & Braun, 2009) (Huang, et al., 2008) (Sun, et al., 2011) (Chen, et al., 2012).

9.2 Study design

The wide availability of massively parallel or next generation sequencing (NGS) provides a potential means to improve mapping algorithm performance. Sanger sequencing provides high quality, low error reads of small DNA sequences, aspects which have traditionally been considered necessary for epitope mapping. Recently, NGS has been coupled with random peptide libraries in studies aimed at identifying immunogenic peptides (Heyduk & Heyduk, 2014) (Christiansen, et al., 2015). Other studies utilized NGS with antigenic fragment libraries to map epitopes (Domina, et al., 2014). Each study developed a unique computational method for manipulation of the NGS datasets.

Based on these studies, we hypothesized that large NGS datasets could provide a more complete set of mimotopes that would improve the ability of current computational mapping methods to identify epitope residues. To investigate this idea, a large random peptide library displayed on *E. coli* was enriched for antibody binding via magnetic-activated cell sorting (MACS), and antibody-binding sequences were determined using NGS to develop a method and assess the benefit of using large datasets in structural epitope mapping. The method identifies antibody-binding residues of a known antigen using currently available epitope mapping algorithms.

9.3 Antibody selection

Two monoclonal antibodies, trastuzumab (Herceptin®) and bevacizumab (Avastin®), were selected to benchmark the protocol because their structures complexed with antigen have been previously determined (trastuzumab/HER2 PDB ID 1n8z, Bevacizumab/VEGF-A PDB ID 1bj1). The interface for each antibody-antigen complex was determined using PyMOL v1.3

with the InterfaceResidues python script, which employs a cutoff value for the difference in the solvent-accessible areas of each protein to determine interface residues. Using a cutoff value of 0.75, PyMOL predicted the HER2 portion of the trastuzumab/HER2 interface contains 22 residues and VEGF-A portion contains 17 residues. In both antigens, the interface region contains non-contiguous sequences (**Figure 9.1**). These regions, considered the binding epitopes for this study, agree with reported interface regions for trastuzumab/HER2 (Cho, et al., 2003) and bevacizumab/VEGF-A (Muller, et al., 1998) (**Table 9.1**).

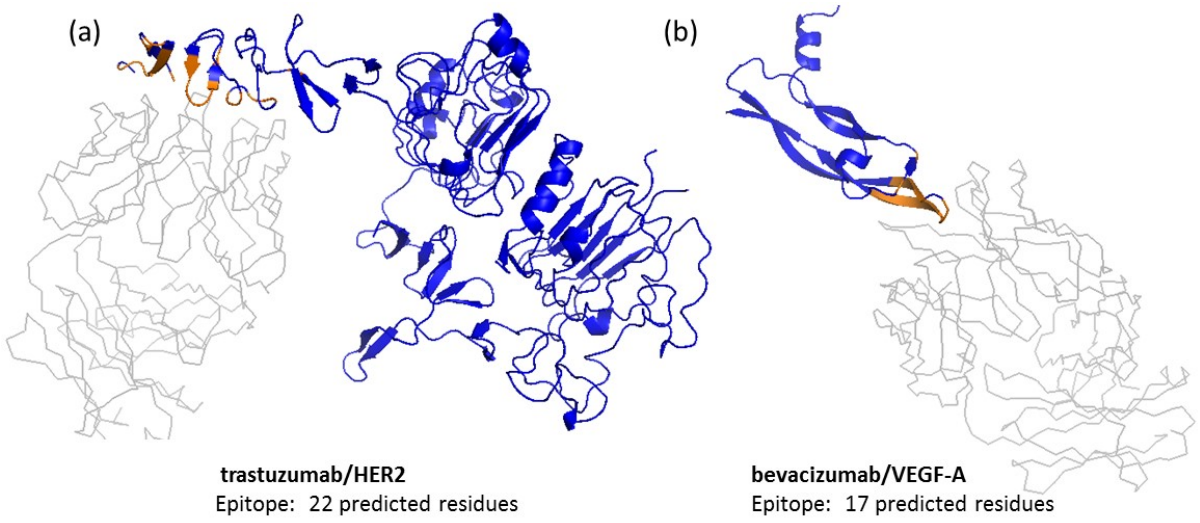


Figure 9.1. NGS-based epitope prediction (Nepitope) was utilized to predict the epitopes of therapeutic monoclonal antibodies (trastuzumab and bevacizumab). (a) trastuzumab binds to human epidermal growth factor receptor HER2 (PDB 1n8z), and (b) bevacizumab binds to vascular endothelial growth factor VEGF-A (PDB 1bj1). Antigen in blue, antibody in grey, and antigen interface residues predicted by PyMOL in brown. Interface residues compared well with reported epitope residues.

Table 9.1. Interface residues from PyMOL vs. published epitope residues

trastuzumab/HER2 (PDB 1n8z)		bevacizumab/VEGF-A (PDB 1bj1)	
PyMOL	Cho *et al.*, 2003	PyMOL	Muller *et al.*, 1998
540			17, 21
557-558	557-561	45, 48	45, 47, 48
560-561		79-84	79-84
569		86-94	86-94
570-573	570-573		
575, 579, 591			
593, 596, 598, 600-603	593-603		
604-605			

To confirm that the selected antibodies did not bind to *E. coli*, an induced cell culture expressing only the library scaffold was incubated with each antibody, diluted to 25 nM, as described in Methods, 12.1. As expected, neither antibody bound to the *E. coli* cells expressing only the eCPX scaffold.

9.4 Library enrichment

Magnetic selection enriches the library for antibody-binding peptides using magnetic beads functionalized with a blend of proteins A and G (Pierce), which bind to the constant region of an antibody. **Figure 9.2** summarizes the experimental protocol; additional detail is provided in **Figure 12.1** in the experimental methods.

Screening rounds were conducted until the library pool exhibited >50% enrichment via flow cytometry (**Figure 9.3**). The libraries from the final round of screening were plated for Sanger sequencing. Peptides binding to trastuzumab and bevacizumab were selected from a bacterial

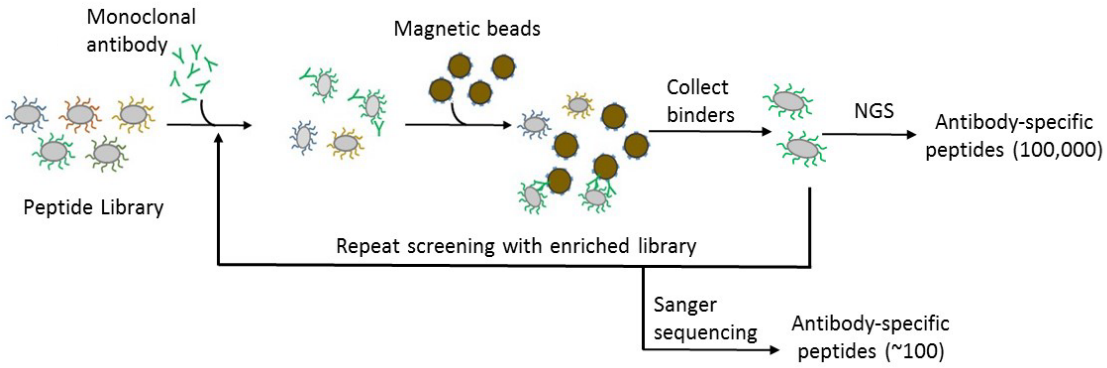


Figure 9.2. NGS-based methodology for mapping structural epitopes using bacterial display. Enrichment of random peptide bacterial display libraries for antibody-binding peptides using magnetic activated cell sorting followed by identification via NGS or Sanger sequencing.

display 12-mer peptide library ($8x10^9$) using 2-3 cycles of MACS, yielding enriched libraries containing 55% and 45% binding members, respectively.

Enriched libraries from the final MACS rounds resulted in 67 and 58 unique sequences obtained by Sanger sequencing for trastuzumab and bevacizumab, respectively. The NextSeq runs for each antibody-screened library from the final (n) and penultimate (n-1) MACS round resulted in 5-11 x 10^6 total reads representing 0.3 – 2 x 10^5 unique peptides.

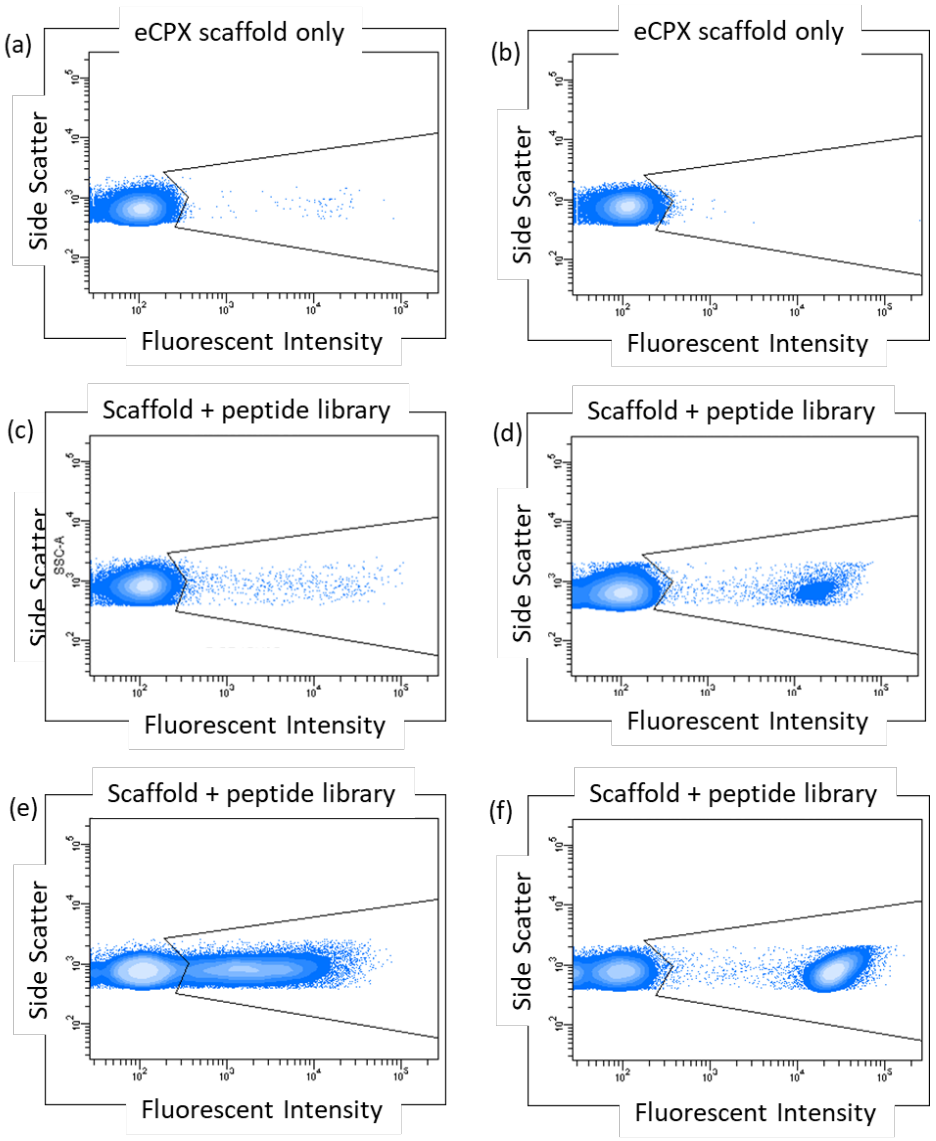


Figure 9.3. Flow cytometry is used to determine the number of MACS rounds required for >50% library enrichment to the target antibody. MACS-screened libraries are enriched in binders to the monoclonal antibody, compared to a cell population expressing only the scaffold. Top: example fluorescent signal from *E coli* expressing only the eCPX scaffold incubated with trastuzumab (a) or bevacizumab (b). The gate is set such that 99.5% of the cells are outside (negative) the gate. Middle: fluorescent signal from *E coli* expressing the peptide library after 1 MACS rounds for trastuzumab (c) and 2 MACS rounds for bevacizumab (d). Bottom: signal after 2 MACS rounds for trastuzumab (e) and 3 MACS rounds for bevacizumab (f).

9.5 Motif discovery

Enriched library pools from all rounds were processed for NGS analysis. Datasets from NGS were processed using IMUNE (Pantazes, et al., 2016) and used for motif discovery via MEME

(Bailey & Elkan, 1994). Sets of 5,000 sequences rank-ordered by observation from NGS datasets were evaluated using MEME to create motifs representing the sequence sets (**Figure 9.4**). MEME runtime scales cubically with the number of input sequences; 5,000 sequences results in a manageable runtime of several hours. MEME outputted consensus motifs, logo plots, list of sequences containing the motif, and several characteristic values including E-value which represents the likelihood of a motif being found in a random mix of the given amino acids in a dataset. Motifs with an expected value less than 0.05 were evaluated further. The

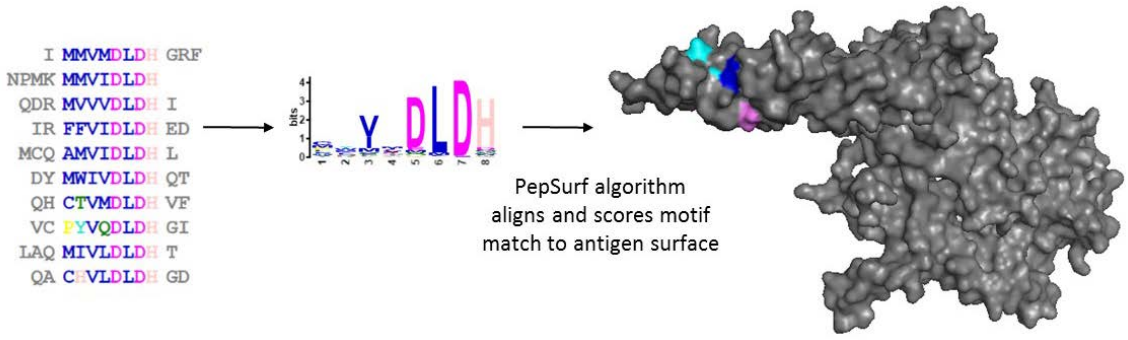
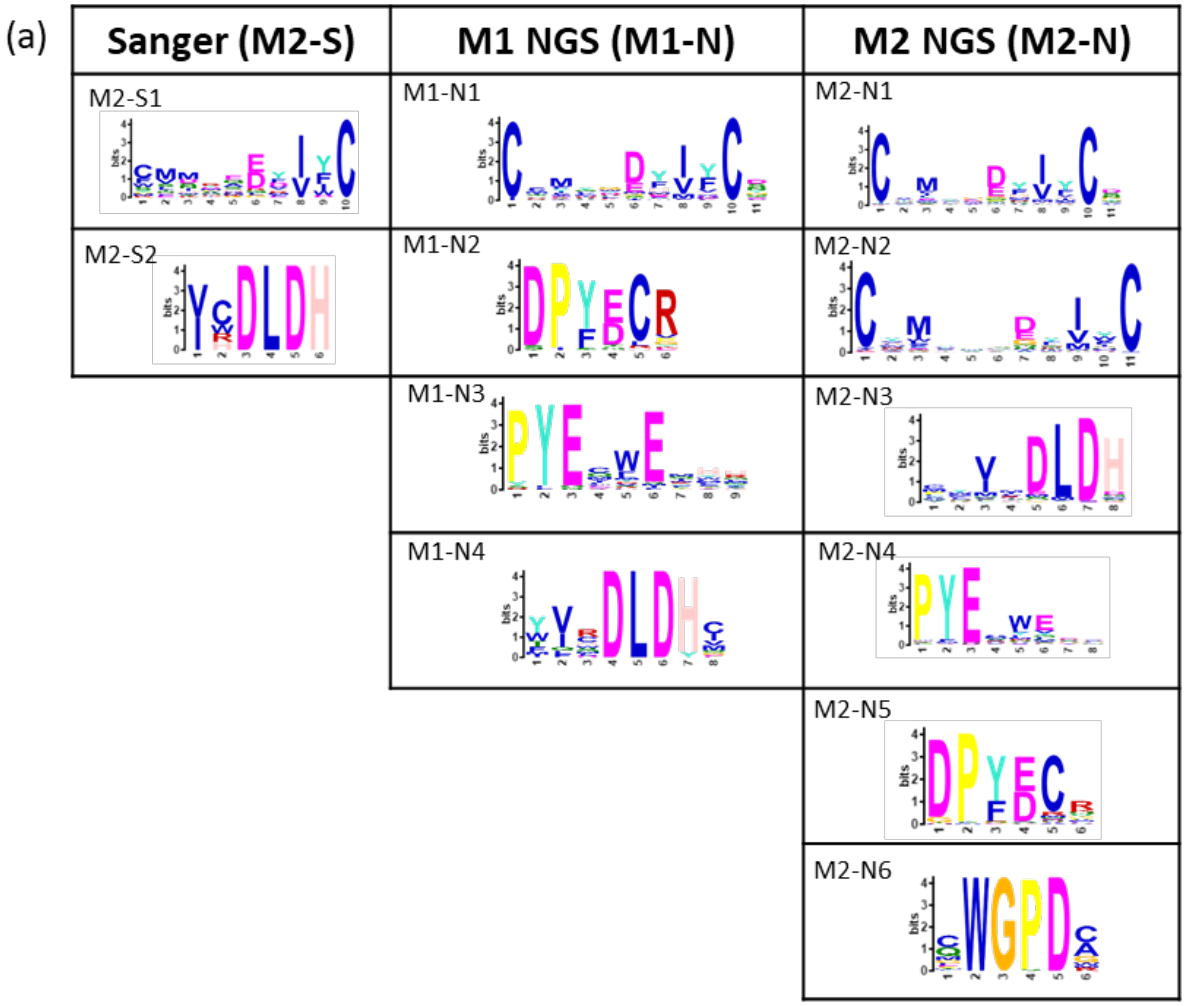


Figure 9.4. NGS-based methodology for mapping motifs to protein surfaces. NGS datasets are too large for currently available prediction algorithms, so to manage the large datasets from NGS, motif discovery via MEME was used to reduce a set of 5,000 of the most observed sequences to motifs (pattern groups) to allow input into PepSurf, which is built for small sets of mimotopes. The example shown is PepSurf mapping peptides from a MEME motif onto HER2.

consensus motif is a list of the most probable residues at each position based on a position-specific letter probability matrix (PSPM). In the consensus motif, MEME lists amino acids with values greater than or equal to 0.2 and records an X for positions where the amino acids have probabilities less than 0.2. Motif logo plots are constructed to show all residues in a position, scaling the size of the one-letter amino acid abbreviation with probability and listing amino acids top to bottom in a column (position in the motif) by their score for those with probabilities ≥0.2, and alphabetically for those with probabilities < 0.2. The PSPM was used to select amino acids for ambiguous (denoted as X) residue positions in between defined

positions by lowering the threshold to 0.15 or 0.10 and using residues with probabilities above this threshold. The sequences obtained from Sanger sequencing were similarly clustered with MEME to obtain representative motifs for comparison with NGS-derived motifs. Finally, a list of possible variants was generated for each motif for input to PepSurf.

MEME clustering of Sanger sequence sets revealed two significant (E-value <0.05) motifs. Motifs sets obtained from NGS data for both monoclonals included two motifs similar to the Sanger motifs along with additional unique motifs (**Figure 9.5**).



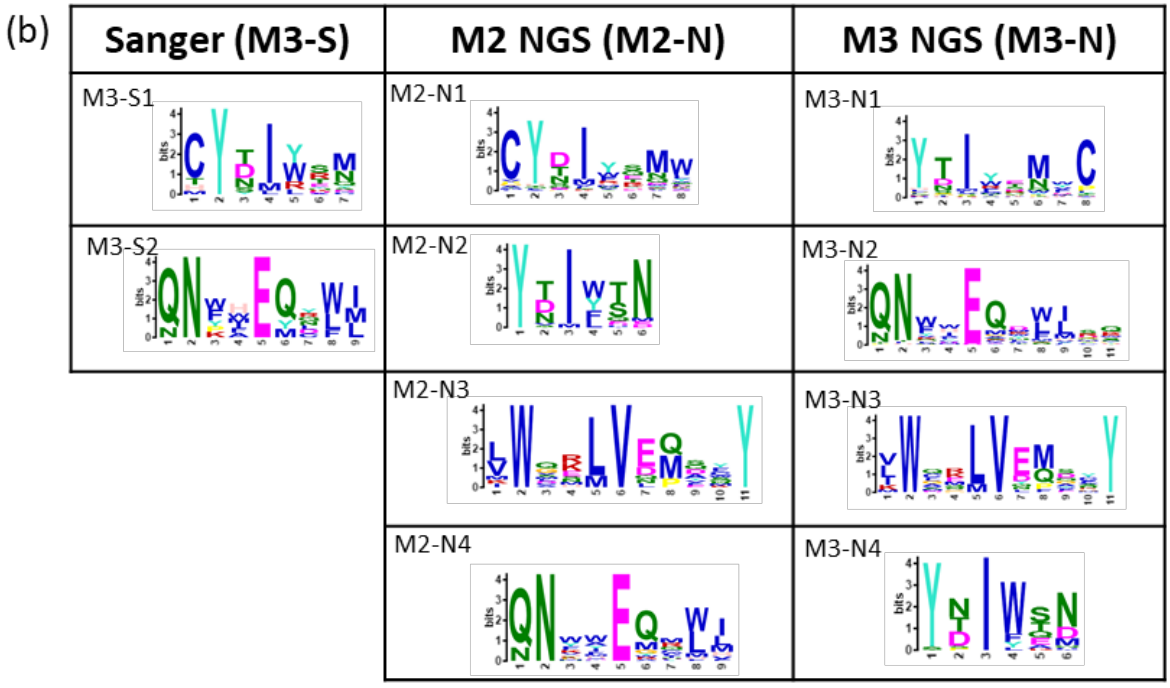


Figure 9.5. Motifs from trastuzumab and bevacizumab-screened libraries. Motifs from Sanger (S) sequencing of select colonies and from the 5,000 most-observed NGS (N) sequences for (a) trastuzumab and (b) bevacizumab. Bacterial colonies for Sanger sequencing were selected after the final round of MACS (M) for each monoclonal antibody. NGS sequencing was performed on libraries from the final (n) and (n-1) screening rounds. Residues for ambiguous positions were selected by setting a threshold value (p) lower than 0.2 (the default) in the MEME position-specific probability matrix.

9.6 Computational mapping of motifs

Four web-based computational methods to map mimotopes to conformational epitopes were evaluated; PepSurf, Mapitope algorithms (Mayrose, et al., 2006), Pep-3D-Search and MimoPro (Negi & Braun, 2009) (Huang, et al., 2008). PepSurf was selected for this study because it 1) allows users to describe the library type, which alters the similarity matrix accordingly and 2) accepts input sets of several hundred peptides. Both PepSurf and Mapitope met these criteria; PepSurf was selected because it provided improved values for the prediction metrics used in this study compared to Mapitope (Huang, et al., 2008) (Sun, et al., 2011).

PepSurf minimally requires an input set of peptides (or motif variants, as used here) from experimental screening, the PDB ID or file and chain identifier(s) for the antigen. The peptides can be weighted if desired. Additionally, the library type (NNK, NNS, etc.) can be selected to best represent the library used to generate the peptide list. We chose the library type "RANDOM_AA" since the random peptide library construction used synthetic trinucleotide codons to ensure equivalent usage of each amino acid. Briefly, PepSurf applies each inputted peptide to the solvent accessible surface of the antigen, determining the best path (Figure 2b). The user can reduce the value of the "best path probability threshold" to cut down the run time; we left it at the default of 0.95. The algorithm then creates (Huang, et al., 2008) and scores up to 3 clusters of antigen residues on the surface that best fit a grouping of peptides.

The highest-scoring cluster from PepSurf was compared to the interface residues determined from PyMOL modeling, employing commonly used performance indicators of coverage/sensitivity, precision and specificity (Negi & Braun, 2009) (Huang, et al., 2008) (Chen, et al., 2012):

$$Sensitivity\ (Coverage) = \frac{T_P}{Interface\ residues} = \frac{T_P}{T_P + F_N}$$

$$Precision = \frac{T_P}{T_P + F_P}$$

$$Specificity = \frac{T_N}{F_P + T_N}$$

Where true positives (T_P) are correctly predicted interface residues, true negatives (T_N) are correctly predicted non-interface residues, false positives (F_P) are residues incorrectly predicted by PepSurf to be in the interface, and false negatives (F_N) are interface residues not predicted by PepSurf.

The metric of precision is similar to a signal-to-noise ratio and indicates the likelihood of the prediction set containing more correct than incorrect information about possible binding regions; perfect prediction, equal to one, implies that no noise (i.e. false positives) exists in the predicted residue set. Various input types were evaluated for interface prediction in PepSurf including the motif variant lists, combinations of motif variant lists, sequence lists from MEME that were used to develop the motif, and entire Sanger sequence sets. The top NGS sequences, each with at least 50,000 observations were also used as whole sequences (unweighted or weighted with observations) and fragmented into 4- to 8-mer sets. Since the number of unique sequences after our sorting was higher than in reported peptide library studies (usually less than 100), we included sets of mimotopes from reported studies (**Table 9.2**) for trastuzumab/HER2 (Reimer, et al., 2004) and bevacizumab/VEGF-A (Li, et al., 2013) (Kobayashi & Shibui, 2011) to allow comparison.

Table 9.2. Published trastuzumab and bevacizumab mimotopes

trastuzumab-screened library (Frequency) from Reimer *et al.*, 2004	bevacizumab-screened library from Li *et al.*, 2013, Kobayashi 2011*
QMWAPQWGPD (51)	DHTLYTPYHTHP
KLYWADGELT (3)	NHFGKFLDALAG
KLYWADGEFT (18)	WLEMHWPAHS*
VDYHYEGTIT (1)	
VDYHYEGAIT (8)	

Systematically defining and applying rules for motif usage based on knowledge about the respective antigen and the strength of residues in the motif resulted in improved prediction. For example, combining and running similar motifs (e.g. trastuzumab NGS motifs 1 and 2) in PepSurf resulted in higher sensitivity and precision than running each motif individually. The most represented motifs (those with the highest number of sequences contributing to them) were M2-N1 for trastuzumab and M3-N1 for bevacizumab; these motifs resulted in the best

PepSurf prediction, suggesting that the higher frequency of a given motif in a sequence set is related to antibody binding strength. Omitting leading/trailing cysteines in the bevacizumab motifs resulted in improved coverage; while for trastuzumab motifs including the flanking cysteines (trastuzumab Sanger and NGS motifs 1) resulted in significantly better coverage. This could be due to the large number of cysteines in and near the interface regions in HER2 compared to VEGF-A. For one bevacizumab motif, M2-N3 and M3-N3, truncating a slightly ambiguous residue position at the front end of the motif and limiting ambiguous resides to no more than two resulted in better interface prediction than using the entire motif. We also tested combinations of motifs as inputs to PepSurf. The highest coverage and precision for each antibody-antigen pair resulted from mapping each motif through PepSurf separately, then combining the motif results into a panel. This was accomplished by creating a cumulative list of true and false positive residues from the clusters and using these values along with values for true and false negative residues to calculate the performance indicators sensitivity, specificity and precision in similar fashion to the individual motifs. Collectively, these observations suggest that independent PepSurf runs using motifs followed by combining the outputs results in the best coverage and precision.

To compare the effectiveness of NGS-enhanced mapping with conventional mapping following Sanger sequencing, sensitivity and precision (Negi & Braun, 2009) (Huang, et al., 2008) (Sun, et al., 2011) (Chen, et al., 2012) were calculated (**Figure 9.6, Table 9.3).** Sensitivity, or epitope coverage, was 18% or 82% for trastuzumab and 47% or 76% for bevacizumab using Sanger or NGS-based mapping on the enriched library pools, respectively.

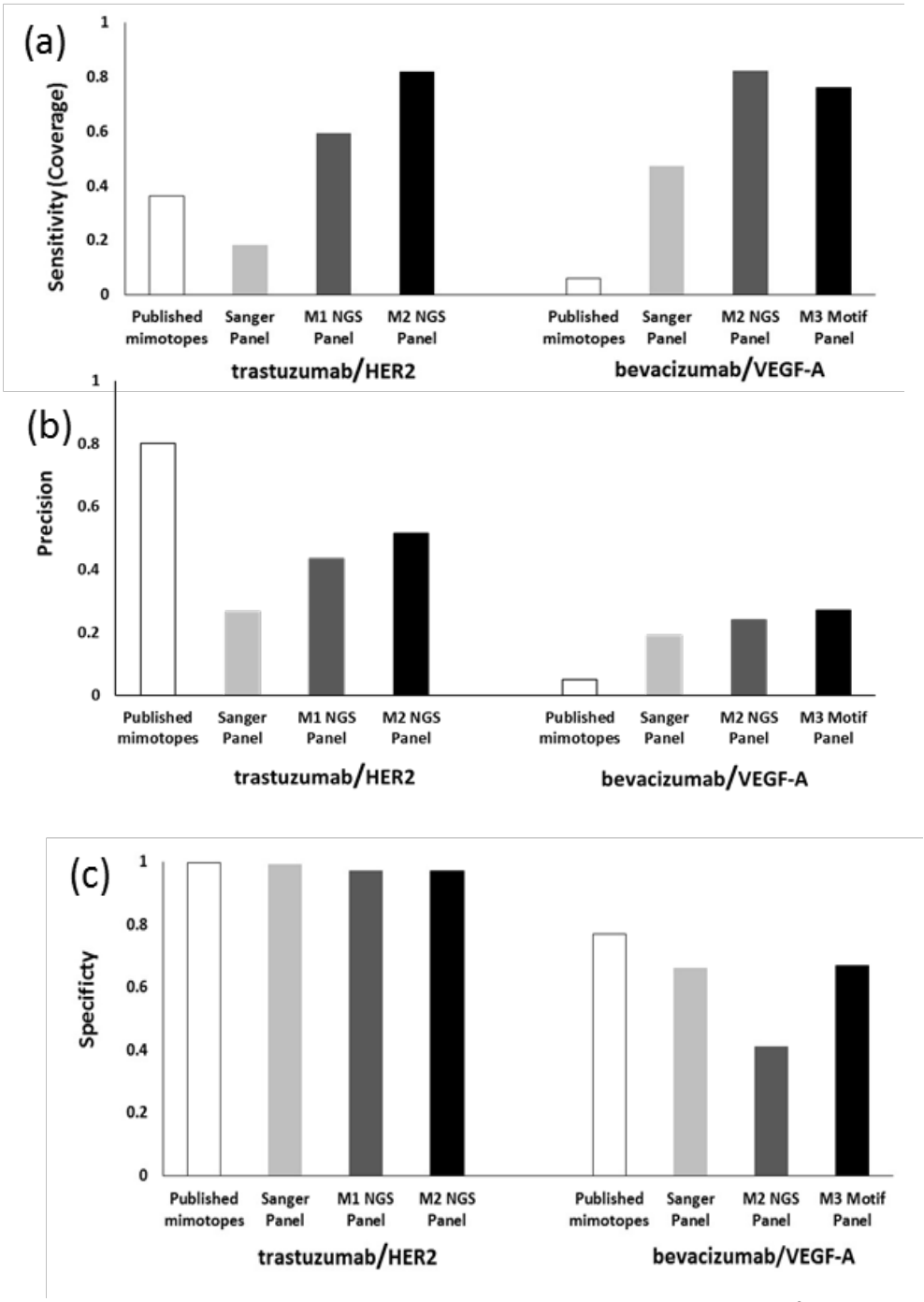


Figure 9.6. Epitope prediction results. Three commonly used indicators for measuring epitope prediction. (a) Sensitivity (coverage) is the measure of how completely the predicted cluster correlates to the known epitope. NGS motif panels provided the best prediction of coverage. (b) Prediction precision, the ratio of correct to total residues predicted, provides a "signal-to-noise" ratio. NGS showed improvement in precision compared to Sanger sequencing for both mAbs but less so for bevacizumab. For trastuzumab, published mimotopes (Reimer *et al*., 2004) provided the best precision. (c) Specificity is the measure of the method's ability to predict true negatives, minimizing false negatives. For trastuzumab, all methods provided better specificity than bevacizumab. PepSurf is a deterministic algorithm, so no deviation is reported for multiple runs on identical input data.

PepSurf is a deterministic algorithm; each time it runs with the same input, it yields the same result, so there is no deviation to report (i.e. no error bars).

Table 9.3. Performance metrics for motif sets from libraries from the final MACS round, sequenced via Sanger or NGS, and published mimotopes

	trastuzumab-screened libraries			bevacizumab-screened libraries		
	Sanger	NGS	Published mimotopes (Reimer, et al., 2004)	Sanger	NGS	Published mimotopes (Li, et al., 2013)
Sensitivity	18	82	36	47	76	6
Specificity	99	97	99	66	67	77
Precision	0.27	0.51	0.80	0.19	0.27	0.05

Published mimotopes from biopanning experiments for trastuzumab and bevacizumab resulted in sensitivity values of 36% and 6%. Specificity was >95% for both Sanger and NGS trastuzumab motif sets as well as the published mimotopes; lower specificity (<90%) was observed for bevacizumab motifs. The predicted residues were mapped to the respective antigen models (**Figure 9.7**). In general, all bevacizumab input sets (Sanger, NGS and reported mimotopes) resulted in more false positives than trastuzumab motifs. It is interesting to note that the specificity of bevacizumab is lower than trastuzumab; some possibilities are the antigen size difference (107 for bevacizumab vs. 607 for trastuzumab), homogeneity of the bevacizumab structure, or the fact that the epitope is a larger percentage of the whole molecule (16%) than for trastuzumab (4%). Others have attempted to correlate antigen size with prediction accuracy but were unable to definitively identify a pattern (Sun, et al., 2011). The highly-looped structure of the epitope region of trastuzumab compared to the flatter bevacizumab surface may affect prediction results. Precision for the trastuzumab and bevacizumab Sanger motifs were 0.27 and 0.19, while NGS motifs resulted in higher ratios of

0.51 and 0.27. Published mimotopes for trastuzumab resulted in a precision of 0.80 and for bevacizumab, a precision of 0.05.

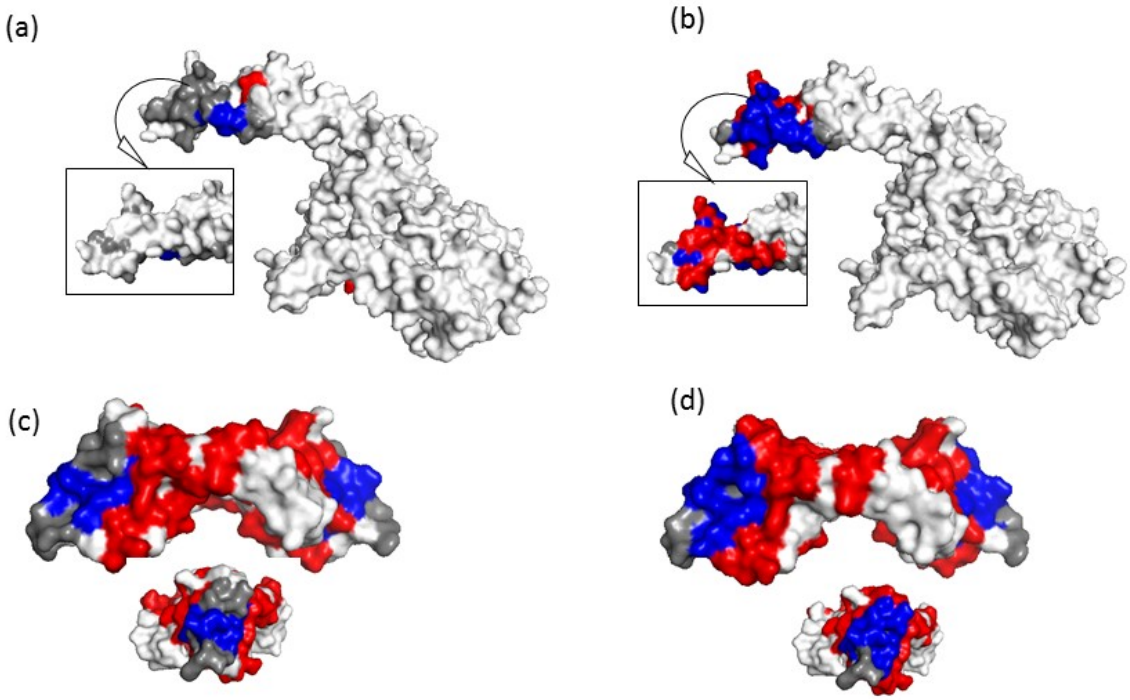


Figure 9.7. Structural epitope prediction using NGS and prior methods. Epitope prediction results of final round screening (a, c) Sanger and (b, d) NGS motif panels. For trastuzumab (HER2) (a) and (b) show 2 sides of the HER2 region of interest. Inset is the backside of the region. For bevacizumab (VEGF-A) (c) and (d) show the side and end view (inset) of VEGF-A shown. The compilation of NGS motifs into a panel provided superior sensitivity (coverage) but poor specificity (high false positives). Epitope residues colored grey were not predicted (false negatives), red residues were incorrectly identified as part of the epitope (false positives), and blue residues were correctly identified (true positives). White residues, the rest of the of antigen, are true negatives.

Motifs from the penultimate round (first round for trastuzumab, second round for bevacizumab) were similar to those from the final rounds. When run through PepSurf in a similar fashion to final round motifs, we found that the cumulative indicators (**Table 9.4**) were 59% sensitivity, 0.97 specificity and 0.43 precision for the trastuzumab-screened library.

Table 9.4. Performance metrics for motif sets from MACS rounds using NGS sequencing

	trastuzumab-screened libraries		bevacizumab-screened libraries	
	MACS 1	MACS 2	MACS 2	MACS 3
Sensitivity	0.59	0.82	0.82	0.76
Specificity	0.97	0.97	0.62	0.67
Precision	0.43	0.51	0.24	0.27

It is notable that the MACS1 trastuzumab-screened library resulted in fewer motifs than MACS2. For the bevacizumab-screened library, metrics were sensitivity of 82%, specificity of 0.62 and precision of 0.24. Performance indicators for individual motifs are listed below (**Table 9.5).**

Table 9.5. Performance indicators for PepSurf clusters based on MEME motifs

M1 trastuzumab-screened library

	Sensitivity (%)	Specificity (%)	Precision
N1	14	99	0.50
N2	32	100	1.00
N3	3	99	0.63
N4	0	98	0.00
Cumulative panel	59	97	0.43
Panel with constraint	9	100	1.00

M2 trastuzumab-screened library

	Sensitivity (%)	Specificity (%)	Precision
S1	0	98	0.00
S2	18	99	0.67
Cumulative panel	18	99	0.27
NGS 1/2	64	97	0.48
N3	27	99	0.55
N4	23	99	0.63
N5	32	100	1.00
N6	32	99	0.78
Cumulative panel	82	97	0.51
Panel with constraint	59	99	0.62

M2 bevacizumab-screened library

	Sensitivity (%)	Specificity (%)	Precision
N1	18	95	0.50
N2	29	79	0.23
N3	24	85	0.40
N4	53	78	0.32
Cumulative panel	82	62	0.24
Panel with constraint	24	89	0.33

M3 bevacizumab-screened library

	Sensitivity (%)	Specificity (%)	Precision
S1	47	82	0.35
S2	18	70	0.10
Cumulative panel	47	66	0.19
N1/4	65	79	0.37
N2	24	78	0.18
N3	29	90	0.42
Cumulative panel	76	67	0.27
Panel with constraint	35	90	0.46

To reduce the high number of false positives found on the VEGF-A antigen (**Figure 9.7c and d)**, the motif panel from the final bevacizumab library was subjected to an additional constraint. Namely, only residues that were predicted by at least two of the four motifs in the panel were considered (**Figure 9.8a**). This increased the precision and specificity but reduced the epitope coverage compared to using the entire NGS motif panel (**Figures 9.8b, 7c**). Notably, residues from each epitope region were still identified.

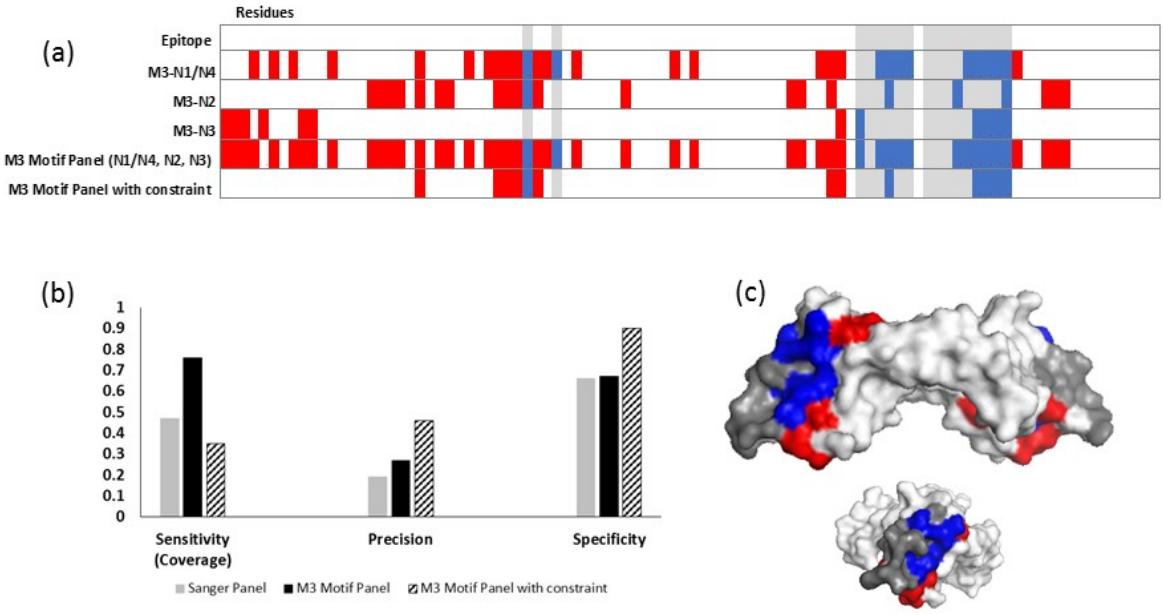


Figure 9.8. Increasing panel threshold improves specificity and precision but reduces epitope coverage for bevacizumab. (a) Heat map showing the predicted residues for each NGS motif (N1/4, N2, N3) from the final MACS round (M3), combined motif panel, and added constraint panel. (b) Adding a constraint to use only residues that were predicted by more than one motif in the panel provided improved precision and specificity but lowered epitope coverage. (c) Residues predicted with the constraint mapped onto VEGF-A (side and end views). Epitope residues colored grey were not predicted (false negatives), red residues were incorrectly identified as part of the epitope (false positives), and blue residues were correctly identified (true positives). White residues, the rest of the of antigen, are true negatives.

A similar result was found for HER2 (**Figure 9.9**). To assess whether motif matching to interface residues was significant, i.e. more than random chance, two bevacizumab-screened motifs were tested against the HER2 structure, and two trastuzumab-screened motifs were

tested against the VEGF-A structure. All non-related motifs resulted in zero sensitivity; that is no interface residues were correctly identified.

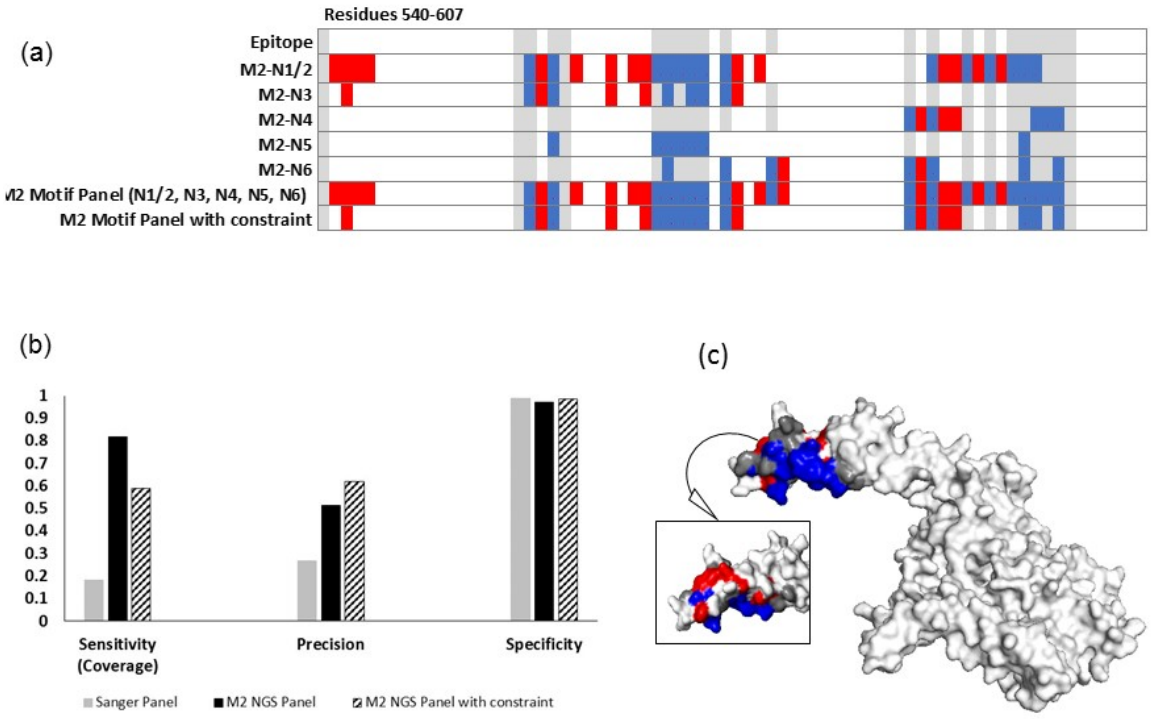


Figure 9.9. Increasing panel threshold improves specificity and precision but reduces epitope coverage for trastuzumab. (a) Heat map showing the predicted residues for each motif, combined motif panel, and added constraint panel (residues 1-540 are true negatives). (b) The compilation of NGS motifs into a panel including residues that appeared in more than one motif in the panel provided improved precision and slightly improved specificity, but lowered epitope coverage**.** (c) Residues predicted with the constraint mapped onto HER2 (2 sides). Epitope residues colored grey were not predicted (false negatives), red residues were incorrectly identified as part of the epitope (false positives), and blue residues were correctly identified (true positives). White residues, the rest of the of antigen, are true negatives.

9.7 Computational mapping of sequence sets

Inputting the entire Sanger sequence sets into PepSurf did not identify interface residues in the best cluster for either trastuzumab or bevacizumab. The most abundant NGS sequences had a sensitivity of 0.0 and 0.12 for the trastuzumab and bevacizumab-screened libraries, respectively. It is well-established that display library methods can result in the selection of

target unrelated peptides (TUPs), even when depletion against the reagents is performed, and TUPs can become enriched and dominate the library when several screening rounds are required. Use of NGS reduces the number of rounds required, helping to avoid TUP enrichment. Additionally, using NGS allows one to computationally remove undesired sequences, e.g. peptides that bind to constant IgG regions. Finally, using the sequence sets identified for each motif in MEME resulted in reduced predictive power; only two of the six trastuzumab and one of four bevacizumab NGS motifs had a sensitivity greater than 0 when the representative sequence set was used as input to PepSurf.

9.8 Summary

In the present study using two monoclonal antibodies directed towards structural epitopes, sequences from NGS datasets improved the quality and accuracy of structural epitope prediction when compared to prior approaches. While individual peptides from Sanger and NGS were unable to correctly predict interface residues, motifs discovered from the most observed set of 5,000 sequences in NGS datasets resulted in prediction clusters with improved sensitivity and precision when compared to Sanger sequences picked from library colonies screened with the same experimental protocol. Compared to mimotopes reported in literature, NGS data contained motifs that exhibited substantially improved sensitivity. Motifs from the penultimate MACS round also exhibited improved sensitivity and similar precision compared to the reported mimotopes, suggesting that in some instances, the use of NGS can reduce the number of required selection or screening rounds.

Methods for epitope mapping include X-ray crystallography, NMR spectroscopy, and a variety of antigen manipulation/antibody binding analysis techniques (Gershoni, et al., 2007).

The method described here provides a means to use NGS datasets with currently available prediction software to identify candidate epitope sites, reducing the number of target regions for fine mapping via costly and time-consuming mutagenesis methods such as alanine substitution (Kronqvist, et al., 2010). The protocol offers additional advantages. It uses web-based, well-documented motif generation (MEME) and epitope prediction algorithms. For structural epitopes, the use of a random peptide library in the method is superior to peptide scanning (overlapping peptides based on the antigen's linear sequence). Finally, MACS is a less expensive and easier to master screening technique than FACS.

Use of NGS with prediction methods can help reduce development time for diagnostics, therapeutics and vaccines that require detailed knowledge of antigenic epitopes (Gershoni, et al., 2007) (Rojas, et al., 2014). For the described protocol, the most time-consuming activity is library generation. Once made, a random peptide library can be used over and over for different antibodies, compared to a specific antigen-fragment library. The time to complete the workflow, including screening, sequencing, motif generation and mapping requires 10-15 working days.

Building informatics tools specifically suited to process NGS data from epitope mapping experiments is an important goal to enable improved identification. PepSurf is limited by the linear scaling of run time to the number and length of peptides (Mayrose, et al., 2006); a run of 1500 6-mer peptides took over a week to complete. New algorithms should be purposefully designed to run the large input sequence lists generated by NGS, ideally in a matter of hours. Motif generation may remain the best way to cluster sequence data prior to mapping to an antigen surface, but a motif generator built within the framework of an epitope prediction tool would streamline the workflow.

Chapter 10

10. Investigation of the antibody profile in Type 1 Diabetes

10.1 Motivation

Biomarkers are substances associated with a physiological sate of an organism, such as a diseased state or response to a treatment. To be useful, they must be detectable, measurable and an accurate indicator of a specific endpoint, such as the onset of clinical disease symptoms. Antibodies are common biomarkers for disease diagnosis; in autoimmune diseases such as rheumatoid arthritis and multiple sclerosis, autoantibodies produced in response to self-antigens are particularly useful for early detection, as they often appear before the onset of clinical symptoms (DeMarshall, et al., 2017) (DeMarshall, et al., 2015). Most recently, the role of autoantibodies as early biomarkers of cancer has become the focus of many studies (Xu, et al., 2014) (Xie, et al., 2011) (Patel, et al., 2011) (Belousov, et al., 2008) (Chen, et al., 2007). Perhaps the most interesting autoantibodies are natural autoantibodies (isotypes IgM, IgA and IgG), which are found in healthy humans with a diversity that is influenced by age, gender and the presence of disease (Nagele, et al., 2013). While their exact role is still unclear, it is thought that they help to clear cell debris from apoptosis to maintain tissue homeostasis and could ultimately provide potent, personal biomarkers to early cell or tissue damage (DeMarshall, et al., 2015).

In some autoimmune diseases, autoantibodies also correlate with the course and severity of the disease (Capeda & Reveille, 2004) (Hamaguchi, 2010) (Wielosz, et al., 2014). In Type 1 Diabetes (T1D), the number, type, concentration, and age at which autoantibodies appear varies among patients but correlates with risk for a genetically susceptible population

(Chiarelli, 2005) (Orban, et al., 2009) (Ziegler, et al., 2013) (Jin & She, 2012) (Steck, et al., 2011). T1D is an autoimmune disease mediated by the complex interaction of genetic and environmental factors (The American Diabetes Association, 2013) (Van Belle, et al., 2011). An estimated 5-10% of diagnosed diabetics have T1D, including an estimated 1-2 million Americans. T1D is primarily diagnosed in childhood (age 0-19) (Van Belle, et al., 2011) but can occur in adults. The incidence varies geographically; the highest reported rates are in Scandinavia and the lowest in Asia (The American Diabetes Association, 2013) (The Diamond Project Group, 2006) (Springer Science & Business Media, 1994), and may also exhibit a socioeconomic gradient (Springer Science & Business Media, 1994). The risk in the general US population is 1 in 300 (0.3%), climbing to 1 in 20 (5%) if a first-degree relative has T1D (The American Diabetes Association, 2013).

The exact etiology of T1D is unknown but results in the loss of approximately 85-90% of the insulin-producing pancreatic islet cells called beta cells (Lippincott, 2007) (Eisenbarth, 2010). T1D exhibits heterogeneous autoimmunity, partly due to genetics and demonstrated by different rates of islet cell decline (The American Diabetes Association, 2013), which can be categorized to some extent by the presence of circulating autoantibodies that bind to islet cell molecules and appear years before the onset of clinical symptoms. While their exact role is unknown, it's unlikely that autoantibodies are the cause of beta cell destruction. Studies showed that maternal transfer of islet autoantibodies to the fetus was insufficient to cause T1D (Koczwara, et al., 2004) (Ziegler, et al., 2013), and not all autoantibody-positive subjects become diabetic (Jin & She, 2012). Additionally, autoantibodies can be transient or low in concentration, are not detected in 100% of the population, and decline over time after diagnosis (Roep & Peakman, 2012) (Tridgell, et al., 2011) (Wenzlau, et al., 2010).

T1D-related autoantibodies identified to date are IgG isotypes, the most prevalent class in the human antibody repertoire. A test for islet cell autoantibodies (ICA), which represents reactivity to several different autoantigens, is a labor-intensive immunofluorescence assay that uses healthy post-mortem pancreatic tissue tested against T1D plasma. A positive result indicates the presence of at least one autoantibody to islet cells (Orban, et al., 2009) (Scherbaum, et al., 1989). Currently, 4 diabetes-associated specific autoantibodies have been validated (**Table 10.1**). A diagnostic sensitivity of 98% is achieved when all 4 autoantibodies are assayed individually (Wenzlau, et al., 2007); when combined in a single, multi-autoantigen assay, the panel of 4 exhibits 92% sensitivity and 99% specificity (Tiberti, et al., 2011).

Table 10.1. Known T1D autoantibodies

Autoantibody	Diagnostic Sensitivity (%)†	Autoantibody titer††	Appearance
Insulin (IAA)	50 (Schlosser, et al., 2010)	Not reported*	Usually 1st found in infants (The American Diabetes Association, 2013) and before exogenous insulin
Glutamic Acid Decarboxylase (GAD65A)	83 (Törn, et al., 2008)	0.2 – 2,000 nM (Mayo Medical Laboratories, 2015)	Usually the second autoantibody to develop in infants (The American Diabetes Association, 2013)
Islet Antigen-2 (IA-2A) and IA-2Abeta	72 (Törn, et al., 2008) 50 (Schlosser, et al., 2011)	Not reported*	Varying
Zinc transporter 8 (Zn-T8A)	70 (Lampasona, et al., 2011)	A 1:10 plasma dilution was used in assay development (Wenzlau, et al., 2007) *	Usually later than the others (Wenzlau, et al., 2007)

†Median values of sensitivity from several labs, adjusted for 95% specificity.
††All are listed as >0.02 nM for a positive result via serum testing (Burtis, 2012).
*Use of relative units (RU), baselined using negative controls, is common for assays.

Positive ICA assays can be followed by negative assays for the four specific autoantibodies results, suggesting the presence of as-yet unidentified autoantibodies. Autoantigens for the known autoantibodies have been characterized (**Table 10.2**); all are

present in islet cells. Other molecules have been suggested but not validated as autoantigenic (Arvan, et al., 2012) (LeRoith, 2004).

Table 10.2. Autoantigens to known T1D autoantibodies

Autoantibody	Autoantigen	Location(s)	Epitopes
IAA	Insulin, native and exogenous (The American Diabetes Association, 2013), proinsulin, preproinsulin (Arvan, et al., 2012)	Beta cell secretory granules, blood, liver (Arvan, et al., 2012).	Unclear. B9-B23 likely (Arvan, et al., 2012).
GAD65A	Enzyme GAD65 involved in synthesis of GABA, an inhibitory neurotransmitter believed involved in paracrine signaling in glucagon secretion regulation (Wendt, et al., 2004).	GAD65 is stored in synaptic-like vesicles in islet beta cells. Functional consequence is unclear (Arvan, et al., 2012) (Sorenson, et al., 1991).	GAD65-NH2 (1-100), -mid (235-442), -COOH (436-585). B and T cell antigenic determinants on -COOH terminus. -NH2 could be result of epitope spreading (Arvan, et al., 2012).
IA-2A and IA-2Abeta	Fragment of intracellular portion of tyrosine phosphatase-like protein encoded by PTPRN gene. IA-2Abeta binds to a second tryptic fragment; seen together so just IA-2A is assayed (Arvan, et al., 2012).	Membrane of secretory granules. Function unknown, possibly in regulation of secretory granule content (Arvan, et al., 2012).	Diverse. Within the cytoplasmic PTP-like and juxtamembrane (605-682) domains of IA-2. Epitope spreading seen early (Ziegler, et al., 2013).
Zn-T8A	Transmembrane transporter for zinc, which complexes with insulin to form storage crystals. Zinc may be paracrine or autocrine modulator of glucagon and insulin secretion. (Arvan, et al., 2012) (Chimienti, et al., 2006).	Membrane of secretory granules with both C and N terminus believed to be on cytoplasmic side (Chimienti, et al., 2006).	B-cell repertoire restricted to a few ZnT8 epitopes and is truly self-reactive as opposed to arising as a bystander response to a foreign antigen (Eisenbarth, 2010). Largely –COOH terminal epitopes (vs. –NH2) (Lampasona, et al., 2010).

Despite the lack of a proven role in disease genesis, autoantibodies are the most studied and well-known aspect of T1D, and several studies have shown that when combined with genetic testing their presence can predict those at risk of developing T1D and the time to onset of clinical symptoms (Ziegler, et al., 2013). For example, one study found the risk of

developing T1D in a cohort of genetically at-risk children tested periodically from birth to 20 years of age was 0.5% for those with no autoantibodies, rose to about 3% when one autoantibody was present and was 16% for 2 autoantibodies, 40% for 3 and 50% for 4 (**Figure 10.1**) (Orban, et al., 2009). Another study found that for IAA and IA-2 but not GAD65A autoantibodies, titer was predictive of risk (Ziegler, et al., 2013). Some researchers believe the multiple autoantibodies associated with higher risk of T1D suggest a pathological role, perhaps promoting inflammation as beta cell autoantigens are exposed (Ziegler, et al., 2013) (Arvan, et al., 2012).

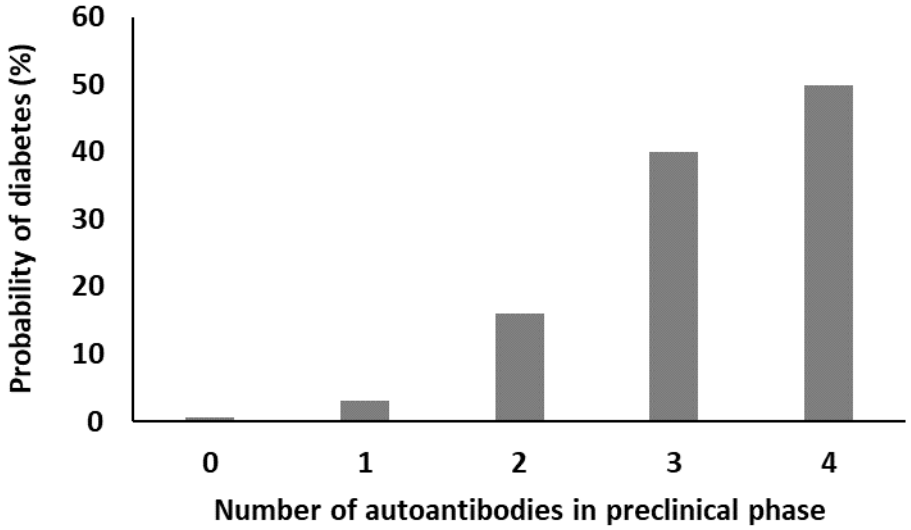


Figure 10.1 Autoantibody type and number correlate to risk of T1D for genetically at-risk populations. Some T1D patients are negative for the known autoantibodies, while others that are autoantibody-positive do not develop the disease. Data summarized from Ziegler (Ziegler, et al., 2013).

Although the percentage varies depending on the cohort (Wenzlau, et al., 2007) (Ziegler, et al., 2013) (Wang, et al., 2007), all the studies contained some T1D patients that remained negative for the 4 known autoantibodies suggesting that the T1D antibody profile is incomplete. Thus, the discovery of additional antibodies present in T1D patients could improve diagnostic sensitivity of at-risk populations and allow earlier invention.

10.2 Study design

The demonstrated predictive value of autoantibodies in T1D and the occurrence of negative patients suggested that antibody profiling of T1D blood samples could provide additional useful biomarkers. Plasma samples used in screening were provided by Dr. Outi Varaala from the National Institute for Health and Welfare in Helsinki, Finland. Autoantibody positivity and gene typing information was also provided. Volumes of approximately 100 μL of blood plasma from 49 newly diagnosed T1D patients and 49 age and gender matched non-T1D controls were aliquoted to ~20 μL and stored at -20 °C. Antibodies in T1D patients may appear at low titer, as evidenced by reported concentrations of 2 to 2,000 nM for GAD65A (Mayo Medical Laboratories, 2015) and 1:10 serum concentrations used in assay development for ZnT8A (Wenzlau, et al., 2007). For comparison, total IgG concentration in children's blood is on the order of 50,000 nM (Heuer, et al., 2008). We used 1:100 plasma dilutions (500 nM) in screening studies, based on past experience in the Daugherty lab on other diseases.

This study used a random 12-mer peptide library expressed on E. coli as discussed in Chapter 8. Random peptide libraries are well suited for antibody profiling, as they represent an unbiased pool of potentially reactive peptides; coupled with next generation sequencing (NGS), a robust set of sequence data can be produced. Due to time constraints, 42 of the 98 available samples were analyzed in this study, 21 T1D and 21 matched non-T1D controls.

Two rounds of MACS were used to enrich the naïve X_{15} or X_{12} library to antibodies in the T1D and non-T1D plasma samples. Enrichment was confirmed via flow cytometry. The enriched libraries were sequenced using NGS and a non-redundant sequence list was created. This list represents the unique sequences that bind to antibodies in each patient or control

plasma sample. Motifs were built from patterns found in the peptide sequences using IMUNE (Identifying Motifs Using Next generation sequencing Experiments), a C++ program developed by Dr. Robert Pantazes, a post-doctoral researcher in our lab (Pantazes, et al., 2016). The IMUNE program performs motif discovery via pattern development, looking through sequence data to find regular expressions that are statistically significantly represented in one group (e.g. T1D patients) but not in another (e.g. the non-T1D control group). IMUNE can screen the millions of sequences that result from NGS, overcoming a limitation of other motif discovery methods which are currently limited to thousands (Pantazes, et al., 2016). An algorithm was developed to further screen the large motif list from IMUNE to find those motifs that were most sensitive to the T1D sample set. Two control sets were used to determine specificity. The matched non-T1D samples comprised one control set, while a database of all sequences compiled for a variety of diseases in our lab served as a second control set in an effort to investigate the specificity of the motifs with a larger non-matched population.

10.3 Autoantibody analysis of the T1D cohort

Autoantibody positivity data for the 49 DM samples were provided by Jarno Honkanen of the Diabetes and Obesity Research Program at Children's Hospital in Helsinki. The samples were tested for GAD65A, IA-2A, IAA, ZnT8A and ICA autoantibodies (**Figure 10.2**). Four patients (8%) tested negative for all autoantibodies including the broad ICA test and 4 patients were ICA negative but positive for one of the 4 specific autoantibodies. Either result could be due to transient or low-level autoantibodies or actual autoantibody negativity; patients with clinical symptoms but negative for autoantibodies are typically retested (The American Diabetes Association, 2013), but it is not known if this cohort was retested. In an effort to increase the

potential to discovery new biomarkers, the group of assumed autoantibody negative patients was purposefully included in the sample set.

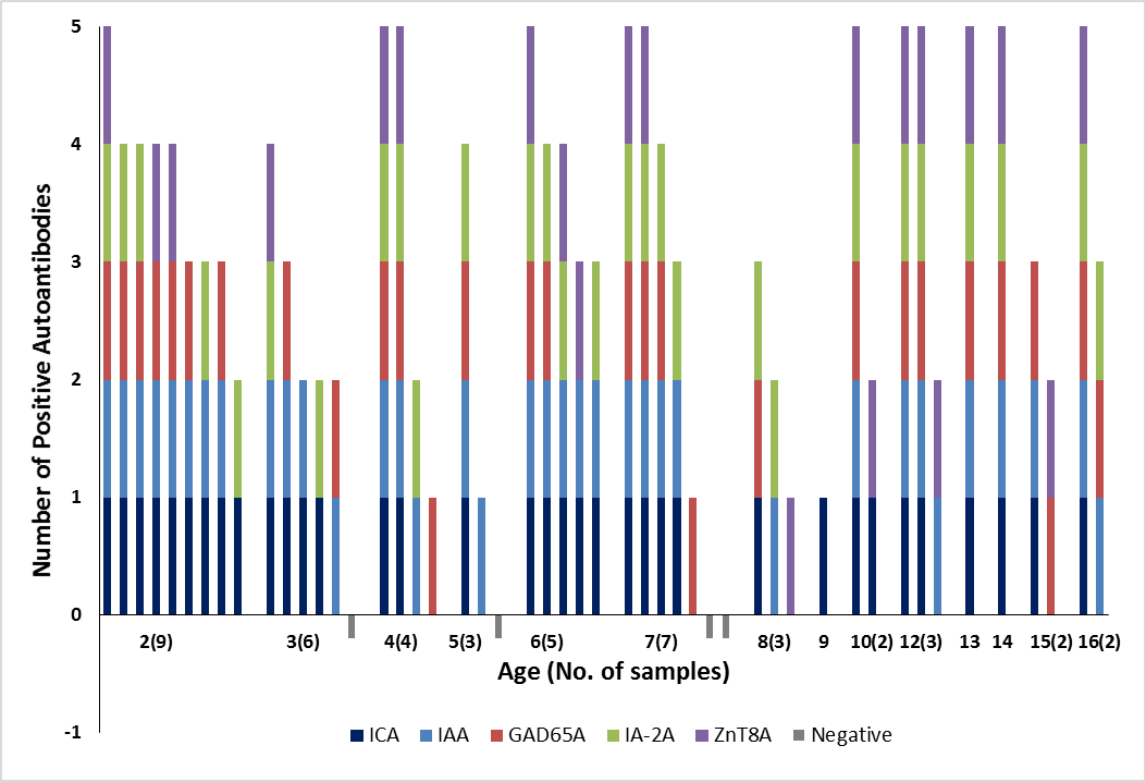


Figure 10.2. Autoantibody positivity for 49 T1D samples in the cohort. 8% of the samples were negative for all 5 autoantibodies, 2% were only ICA positive. Overall, 71% were positive for ICA, 73% for IAA, 59% for GAD65A, 57% for IA-2A and 43% for ZnT8A. Column order from the bottom is ICA/IAA/GAD65A/IA-2A/ZnT8A. Completely negative samples are shown as negative columns.

For the 21 samples used in the study, the mean age was 7.3 years (compared to 6.6 years for the entire cohort), and the gender split was 9/8 females/males (vs 17/32 for the entire

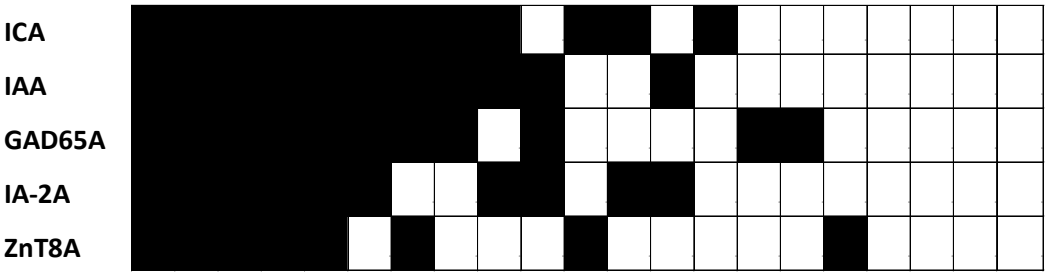


Figure 10.3. Autoantibody positivity for 21 T1D samples in the study. 20% of the samples were negative (white) for all 5 autoantibodies, 5% were only ICA positive. Overall, 57% were positive for ICA, 52% for IAA, 52% for GAD65A, 48% for IA-2A and 38% for ZnT8A.

cohort). The autoantibody positivity (black boxes denote positivity) of the samples used in this study is shown in **Figure 10.3**. 20% of the sample set is negative for all 5 autoantibodies compared to 8% of the entire cohort due to the decision to include all 4 autoantibody-negative samples in the sample set.

10.4 Library enrichment

To isolate peptides specific for each sample, all 42 plasma samples were individually used to screen either a naive X_{15} or X_{12} library via MACS. A dilution of 1:100 plasma was used in each round. All 42 samples in the discovery set exhibited greater than 50% enrichment, determined by flow cytometry, after two MACS rounds.

10.5 Sequencing

Plasmid DNA from the final populations from MACS screening for each of the 42 samples was extracted using a miniprep kit, amplified and pooled before being sequenced with either the Genewiz Ion Torrent system or the UCSB BNL's Illumina NextSeq 500 system. An average of 2.7 million total sequences and between 2 - 13 x 10^5 unique sequences per sample from a starting naive library population of 8 x 10^9. IMUNE was used to extract the sequences from the NGS output file, combine redundant sequences and those sequences with 3 or fewer amino acid differences (about 25% of the sequences were combined with others). Many of the sequences were only observed once. Both purposeful and experimental sources of bias can contribute to sequence bias such as amplification bias in PCR (the latter). NGS, with error rates >1%, can result in the disappearance or conversion of sequences (Matochko & Derda, 2013) (Quail, et al., 2012).

10.6 Motif discovery

The IMUNE algorithm (Pantazes, et al., 2016) was used for motif discovery of T1D-specific motifs. Briefly, IMUNE looks for patterns with user-defined lengths and definition (e.g. known amino acid residues in a position) that are in a group of sequences. IMUNE then builds motifs from the patterns by combining similar ones (**Figure 10.4**). IMUNE can selectively find

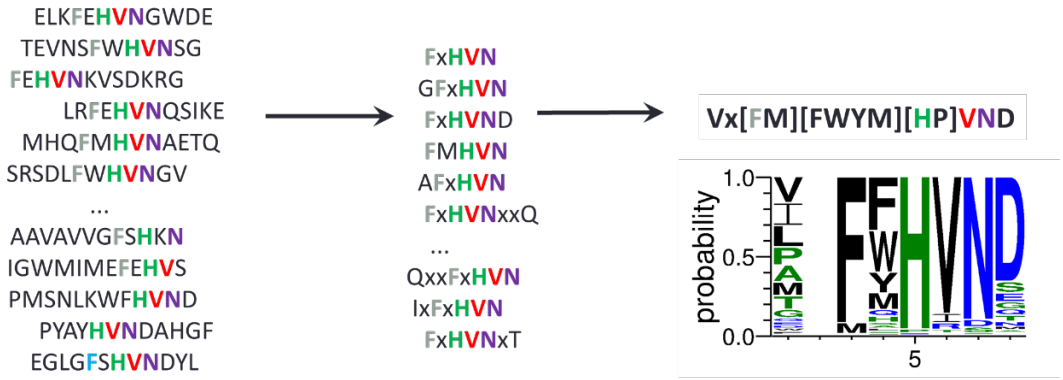


Figure 10.4. IMUNE motif discovery scheme. From patterns appearing in sequence sets, IMUNE builds motifs. Those motifs are then computationally screened for specificity and sensitivity to the disease group.

patterns, i.e. ones that are present in a disease group and not in a control group, based on user-determined sensitivity and specificity constraints. For each motif, IMUNE determines the number of times a motif is observed in a sample sequence set and compares it to the expected number of observations (via probability given the amino acid makeup of the library) to create an enrichment value:

$$\text{Enrichment (E)} = \frac{\text{Actual Observations}}{\text{Expected Observations}}$$

Using the 21 T1D compared against the 21 non-T1D sample sequence sets, 350 motifs were identified by IMUNE with the constraints of 20% sensitivity (patterns must be present in

at least 20% of the 21 T1D samples) and 100% specificity (patterns cannot be in any of the 21 non-T1D samples). The low sensitivity threshold was selected based on the knowledge that known autoantibodies to T1D exhibit low sensitivity across T1D cohorts (Jin & She, 2012). 99-100% specificity is a commonly used threshold for biomarkers.

10.7 Sensitivity and specificity of motifs to T1D

The number of found motifs (350) was too large to manually curate, so to identify motifs with the highest specificity to the T1D cohort a simple enrichment-based screening algorithm was devised (**Figure 10.5**). Briefly, motifs were deemed enriched in the T1D group if the mean enrichment of the motif in the T1D group was at least 10X greater than the mean enrichment in the non-T1D group. That set of motifs was then screened for enrichment greater than 5X in each sample (T1D and non-T1D) compared to the mean enrichment of the non-T1D control group. Motifs that met this enrichment threshold in a sample were denoted with a black cell on a heat map and those that did not were denoted with as white on the map (**Figure 10.6**). The motifs (rows) were then sorted for specificity (high to low) followed by sensitivity. Thus, the upper portion of the heat map represents motifs that are 100% specific and most highly sensitive to T1D samples (columns).

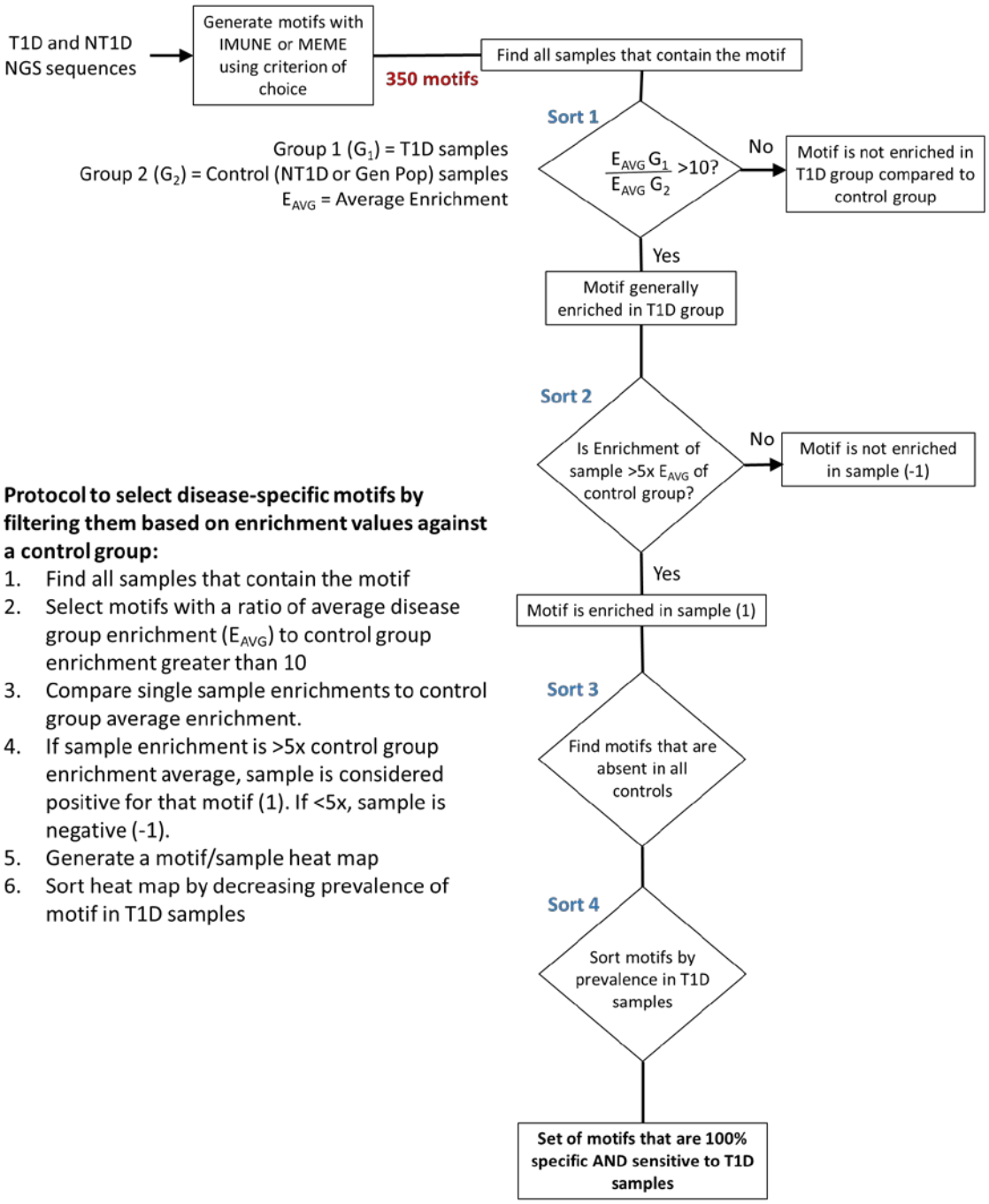


Figure 10.5. Decision tree for selecting sensitive and specific T1D motifs. The 350 motifs from IMUNE were screened by enrichment value. Motifs with mean enrichment values at least 10X in the T1D group compared to a control group were considered enriched in the T1D group. These motifs were then screened for high enrichment in the T1D samples compared to the mean enrichment in the control group. The selection algorithm provides binary data for each motif/sample pair that can be visualized using heat maps to determine motifs with high specificity and sensitivity to T1D compared to a control group.

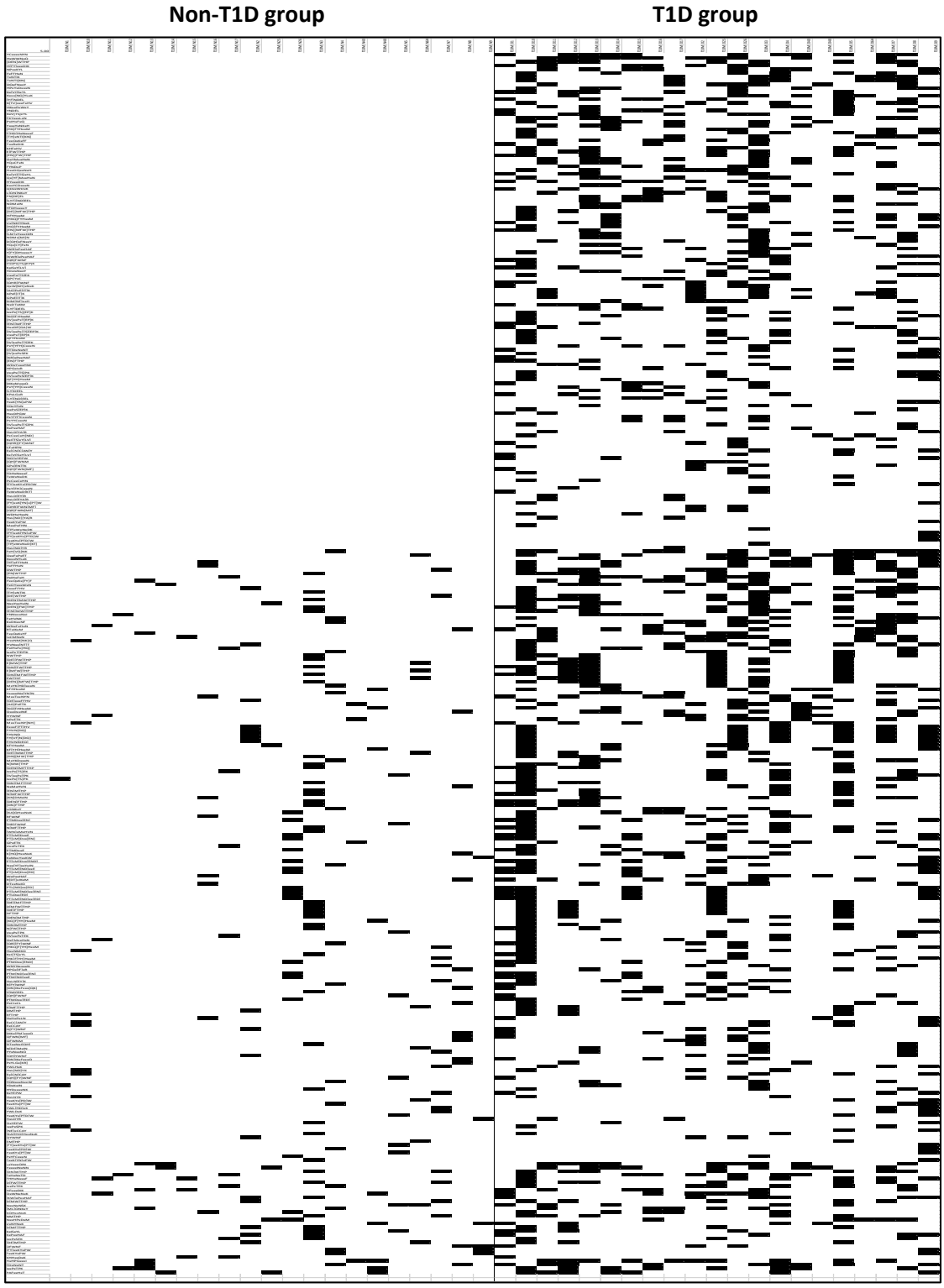


Figure 10.6. Heat map of motifs with greater than 10x enrichment compared to the average enrichment of the non-T1D control group, sorted by decreasing specificity and then decreasing sensitivity. The region of most interest is the upper section which represents motifs that are 100% specific and most highly sensitive to T1D samples (black boxes). Motifs are listed in rows, samples in columns.

Using this method, 133 motifs were identified that were 100% specific and between 5 and 52% sensitive (i.e. found in between 1 and 11 samples) to the T1D samples (**Figure 10.7**). Comparing the best motif sensitivities to that of the known autoantibodies in the sample set (**Figure 10.3**), we see similar values for IAA and GAD65A (52%) and ICA (general islet cell autoantibody, 57%)

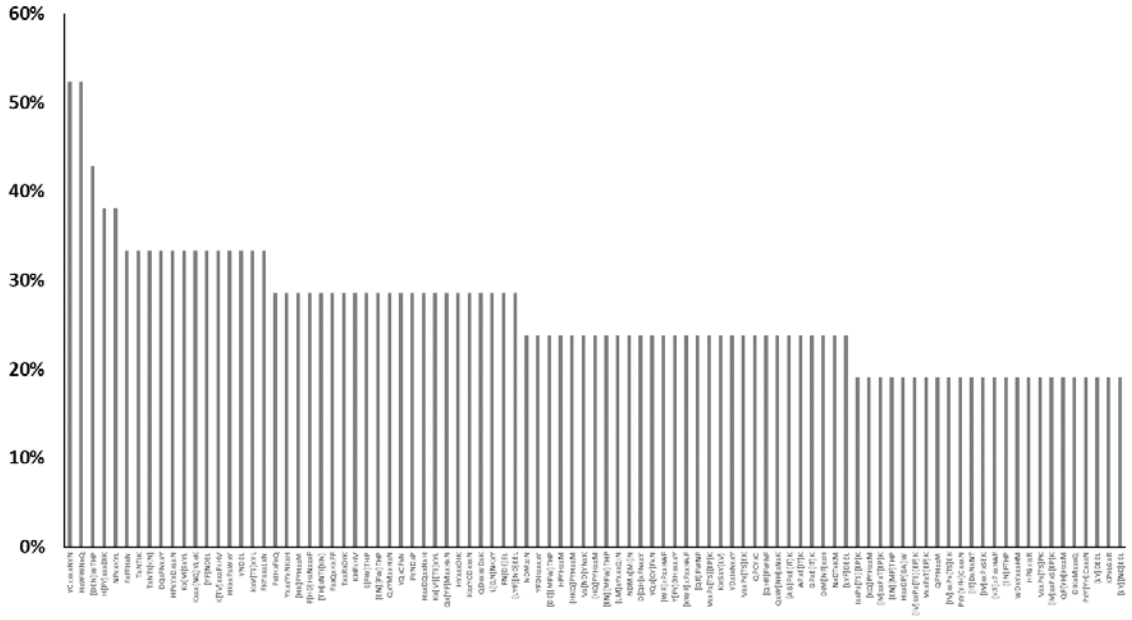


Figure 10.7. Motifs with greater than 19% sensitivity to T1D samples with 100% specificity against a non-T1D control group. The best motifs were >50% sensitive to T1D samples, similar to the sample set sensitivity to the known T1D autoantibodies.

Of the two most sensitive motifs, YCxxxxNYN, while long (9 residues) had a significant number of ambiguous residues (denoted as 'x'), while the other HxWWNxQ, was slightly shorter and more defined. Many of the motifs were similar, encouraging grouping:

TxNTIK TxNTI[KN] [TH]xNTI[KN]
Kx[VI]SxYL KxV[TS]xYL Kx[VI][TS]xYL KxISxY[LV]
[DEN]WTHP E[FW]THP

There was no apparent correlation between the occurrence of these found motifs and the autoantibody positivity; i.e. motifs were not significantly more or less present in samples that were positive for one or more of the known autoantibodies. The two top motifs were found in 3 of the 4 autoantibody negative samples, but also in 8 other samples. The fourth negative sample had the lowest overall number of motifs, all with < 18% sensitivity.

10.8 Specificity of T1D motifs to a larger control group

To determine how sensitive the T1D motifs were when a larger control group was used, the IMUNE motifs were screened against a larger general population proxy control group created from a database of 235 other disease or control samples (**Figure 10.8**). Due to known comorbidity with T1D, celiac samples were excluded. It should be noted that this control group cannot be labelled as non-T1D since the samples in the database, with the exception of the T1D and non-T1D ones, were not screened for T1D, and it's possible that some patients in the database have T1D. That said, the mean age of the database population is skewed to adult

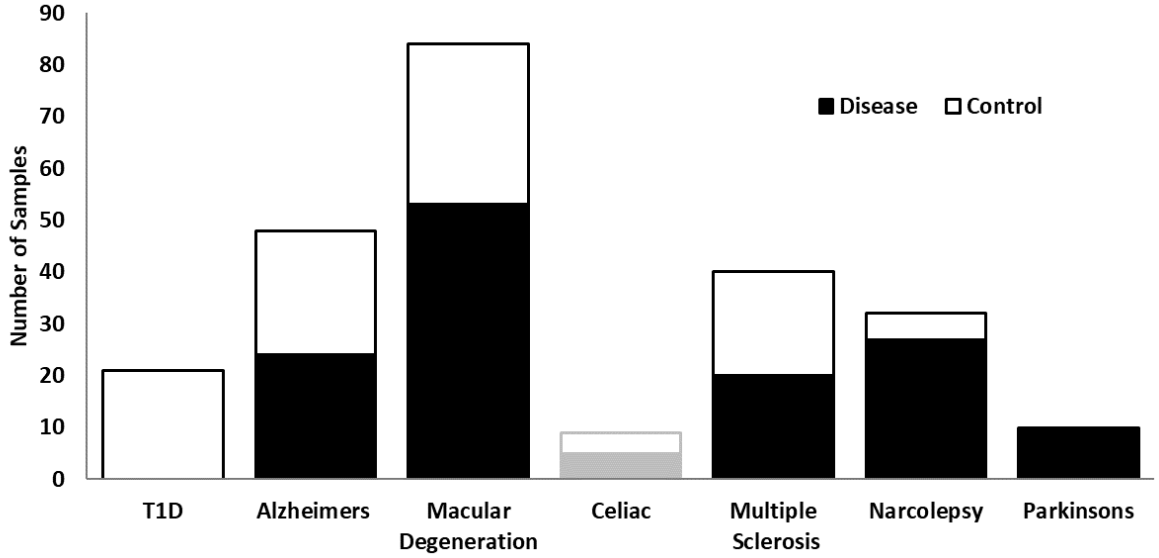


Figure 10.8. Sample groups in the general population proxy control group. The celiac disease and control groups were excluded due to known co-morbidity with T1D. The T1D control samples were included.

patients due to the type of diseases represented (e.g. multiple sclerosis, macular degeneration, Alzheimer's) and therefore seems a reasonable proxy for a population that is not newly T1D diagnosed (and therefore with low antibody titers, which decrease with disease duration). The same protocol was used (**Figure 10.5**) in finding motifs sensitive to the non-T1D control group in the previous section, resulting in a heat map representing all the motifs with enrichment >5x compared to the control group (**Figure 10.9**).

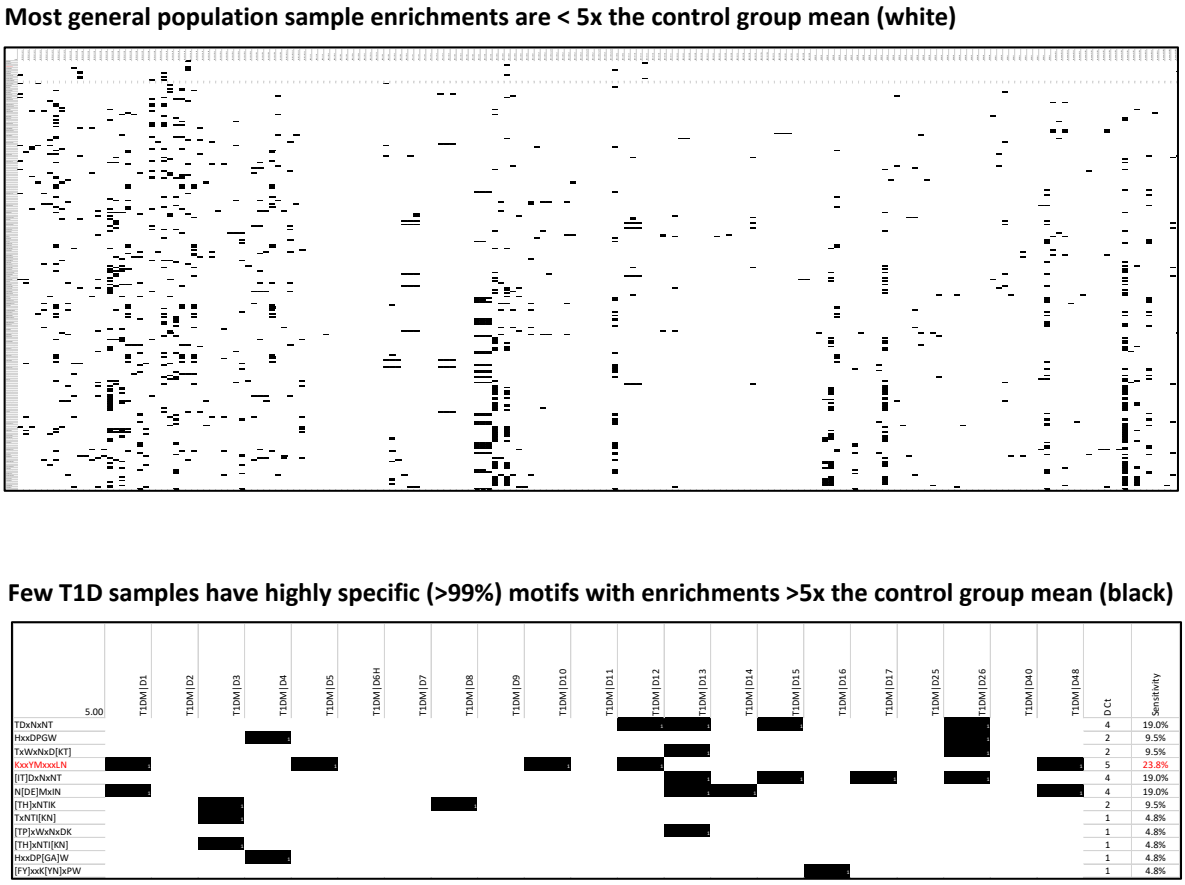


Figure 10.9. Heat map of motifs with greater than 10x enrichment to the T1D group than the general population proxy control group, sorted by decreasing specificity and then decreasing sensitivity. The best specificity was 99% (above the white line in the General Population heat map) and sensitivity within the 99% specific group ranged from 5 to 23.8% (bottom heat map).

Comparing the T1D motifs against this larger general population, 7 motifs were ≥ 99% specific and up to 20% sensitive (**Figure 10.10**). There were 3 general motif patterns, and three motifs, [IT]xNxNT, TxWxNxD[KT] and [TP]xWxNxDK retained their sensitivity to T1D samples against the larger control group enrichment mean, although the specificity dropped slightly. Two of these motifs were related; the one with the less ambiguous first position, TxWxNxD[KT] being more sensitive. There do not appear to be any published studies comparing T1D autoantibody specificity against a general population group to contrast these results with, but it is encouraging that some motifs retained their specificity and sensitivity against a much larger (10x) control group, suggesting that biomarkers could be found with higher sensitivity that might be used for general population screening.

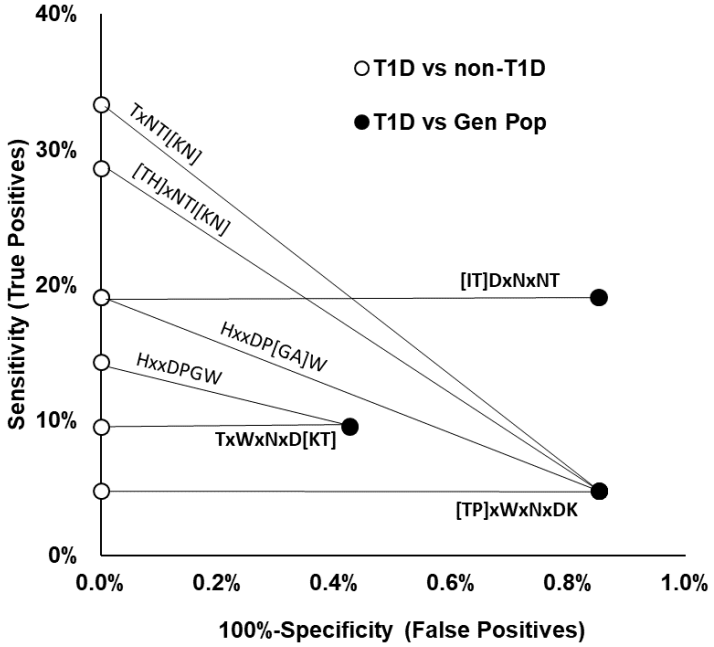


Figure 10.10. The best motifs against the general population proxy control group were up to 33% sensitive and 100% specific to T1D samples. The best motif ranged from 99.2 to 100% specific and 5 to 19% sensitive to T1D against the general population control group, and 100% specific and between 5 to 33% sensitive to T1D against the non-T1D control group. Two motifs did not change sensitivity between the two control groups but did drop slightly in specificity.

None of the found motifs were similar to the published epitopes for the known autoantibodies. To look for the possibility that the motifs were other epitopic regions of the autoantibodies, PepSurf (Mayrose, et al., 2006) was used to map the best motif ([IT]DxNxNT) onto the two autoantigens with crystal structures in The RCSB Protein Data Bank (insulin and GAD65). PepSurf found one cluster with residues in the region 89-442, which overlaps with region of a structural epitope (244-585) reported by Myers using phage display (Myers, et al., 2000). To investigate whether the top motif could be associated with new autoantigenic molecules, [IT]DxNxNT was searched against the protein sequence database UniProtKB via the ScanProsite search engine hosted by ExPASy (Sigrist, et al., 2012), or BLAST to search the NCBI protein databases (Altschul, et al., 1990), including homo sapiens, bacteria and virus taxonomies.

10.9 Summary

Bacterial display peptide libraries were used to screen plasma samples from 21 T1D patients and 21 matched non-T1D samples. Libraries were sequenced and computationally sorted for motifs with high specificity to T1D compared to non-T1D matched controls. The method found several motifs that were 100% specific and up to 52% sensitive to T1D vs. non-T1D samples; this sensitivity compared well with sensitivities of the known autoantibodies in this sample set. Comparing these motifs against a larger general population proxy control group found 7 motifs that were ≥99% specific but no more than 20% sensitive; three of the motifs retained their sensitivity to T1D sample even when compared against this 10x larger control set, suggesting that these motifs could be further developed to increase sensitivity to T1D samples. There were no similarities between the most sensitive found motifs and either the known autoantibodies

or other antigens, based on a protein database search, nor were there correlations between age, gender or HLA-associated risk rankings.

Chapter 11

11. Future Directions

Combined, the two studies presented in this section showcase the power of combining bacterial peptide display libraries with next generation sequencing for antibody-related tasks. For each area, potential future work is outlined below.

11.1 Epitope mapping

While the current epitope mapping algorithms were successfully used with NGS data by creating motifs to satisfy the input restrictions, building informatics tools specifically suited to process NGS data from epitope mapping experiments should be the focus of future work. PepSurf is limited by the linear scaling of run time to the number and length of peptides (Mayrose, et al., 2006); a run of 1500 6-mer peptides took over a week to complete. New algorithms should be purposefully designed to run the large input sequence lists generated by NGS, ideally in a matter of hours. Motif generation may remain the best way to cluster sequence data prior to mapping to an antigen surface, but a motif generator built within the framework of an epitope prediction tool would streamline the workflow.

11.2 Antibody repertoire analysis

Antibody profiling has been proven to provide insight into a variety of disease pathologies and provide biomarkers for diagnosis and therapeutic endpoints. The challenge to use it for some autoimmune diseases, such as T1D lies in the complex nature of the disease. Researchers are only now beginning to understand the true heterogeneity of diabetes, suggesting additional classification beyond Type 1 and 2, such as maturity-onset diabetes of the young (MODY),

latent autoimmune diabetes of adults (LADA), and idiopathic diabetes (Karalliedde & Gnudi, 2014). This recognition of different pathologies of diabetes suggests a profiling method going forward; that of grouping samples by some measure (e.g. autoantibody positivity, HLA risk level) to conduct antibody profiling.

Inherent to the success of any repertoire screening method is the successful identification of peptides that mimic the known autoantigenic epitope regions. Specifically, the development of strategies designed to increase the sensitivity for T1D antibody-reactive peptides while maintaining or increasing the specificity. While the results shown here demonstrate that the current protocol can identify disease-specific peptides with some cross-reactivity to new samples, being able to consistently identify reactive peptides that are highly similar to known autoantibody ligands will greatly improve the protocol reliability and will provide validity to subsequent results. These screening strategies can then be used to characterize antibody profiles in subgroups and, if possible, the entire T1D group of samples currently available to us. DM (or Control) antibody reactive peptides identified from screening can be valuable in several ways: 1) to help further characterize the heterogeneity of T1D pathology by defining additional markers for subgroups, 2) to provide peptide sequences useful in identifying potential antigenic molecules, and 3) to identify potential peptide sequences for T cell screening.

Chapter 12

12. Experimental Methods for peptide display and sequencing

Materials

Trastuzumab and bevacizumab used in the epitope mapping study were generously provided by Dr. Daniel Greenwald of the Cancer Center of Santa Barbara. Plasma samples were provided by Dr. Outi Varaala from the National Institute for Health and Welfare in Helsinki, Finland. Tween-20, NaCl, arabinose, glucose, tryptone, yeast extract, chloramphenicol, glycerol and other stock chemicals were purchased from Sigma Aldrich. Protein A/G beads were purchased from Pierce. α-IgG-PE was purchased from Jackson ImmunoResearch. The plasmid miniprep kits were purchased from Qiagen, Nextera XT indexing primers were purchased from Illumina. Agencourt AMPure XP beads were purchased from Beckman Coulter.

12.1 Antibody concentrations used in the studies

Epitope mapping study. Both antibodies were used at a concentration of 25 nM and added to the cells for a final volume of 500 μL. Antibodies had similar kD values in the nM range, and 25 nM was chosen because it provided good library enrichment after the first round of MACS.

T1D antibody repertoire analysis study. Plasma was used at a 1:100 dilution.

12.2 Depletion of *E. coli* binding antibodies

12.2.1 Epitope mapping study

To confirm that the selected antibodies did not bind to *E. coli*, cultures were grown overnight at 37 °C with shaking (250 rpm) in LB (10 g tryptone, 5 g yeast extract, 10 g/L NaCl) supplemented with 34 μg/mL chloramphenicol (CM). Cells were inoculated at 1:50 into 5 mL LB/CM and grown to an OD_{600} between 0.4 – 0.6 at 37 °C with 250 rpm shaking, and then induced with 0.04% wt/vol L(+)-arabinose and allowed to grow for another hour. A 5 μL aliquot of cells was centrifuged at 3,000 relative centrifugal force (rcf) at 4°C for 5 minutes and decanted. Cells were incubated with antibody diluted to 25 nM in 40 μL PBST (PBS + 0.05% Tween 20) at 4°C for 45 minutes on an orbital shaker (20 rpm). Following antibody incubation, the cells were washed to remove unbound antibodies by centrifuging, decanting and resuspending the cells in 40 μL cold PBST 3 times. To the final resuspension, α-IgG-PE (anti-human goat IgG conjugated to phycoerythrin) was added at 1:100, and the cells were incubated at the same conditions. Following a single wash step, the cells were resuspended in 500 μL cold PBS and analyzed for fluorescence using flow cytometry (BD Biosciences FACSARIA I with a FACSARIA II flow cell). There was no appreciable binding, therefore no depletion step was required.

12.2.2 T1D antibody repertoire analysis study

To remove *E. coli* binding antibodies from plasma samples prior to library screening, an induced culture of cells expressing the library scaffold alone was incubated with diluted (1:100) plasma. eCPX cultures grown overnight at 37 °C with vigorous shaking (250 rpm) in LB/CM and 0.2% glucose were collected by centrifugation, inoculated in fresh LB /CM, grown to an OD_{600} between 0.4 – 0.6 at 37 °C with 250 rpm shaking, and then induced with 0.04% wt/vol L(+)-arabinose and allowed to grow for another hour. After induction, cells were

centrifuged at 3,000 rcf for 5 min, washed once with cold PBST (PBS + 0.05% Tween 20), and resuspended in 1 mL PBST containing plasma diluted 1:25. Samples were incubated overnight at 4 °C with gentle mixing on an orbital shaker (20 rpm). Antibodies that bound to *E. coli* or the eCPX scaffold were removed by centrifugation of the incubated culture at 3,000 rcf for 5 minutes twice, recovering the plasma supernatant after each centrifugation. Depleted plasma was stored at 4 °C for up to 2 weeks during use.

12.3 Magnetic-Activated Cell Selection (MACS)

Prior to use, the beads were decanted via magnetic separation and washed two times in a 3x volume of PBST. A library frozen stock (1.3 mL) in 15% glycerol containing 10^{11} cells was added to 500 mL LB/CM and allowed to grow at 37 °C until OD_{600} reached 0.4 - 0.6. The culture was induced with 0.04% arabinose and allowed to grow another hour. The OD_{600} was measured to calculate the required volume for aliquots to contain 6-7 times the starting library diversity (usually about 40 mL). The aliquots were centrifuged for 15 minutes at 3,000 rcf and 4°C. To clear any cells that bind to protein A/G, the supernatant was removed, and cells were resuspended in 1000 μL PBST containing 50 μL washed beads at a ratio of 1:100 bead-to-cells. The ratio is based on the size and enrichment of the library; for the first round of MACS, the naïve library will contain many non-binding peptides, so a high ratio of beads to cells can be used. For subsequent rounds, there will be more binders, so a lower ratio (1:1) is used. The cells were incubated in 2 mL microcentrifuge tubes for 45 min at 4 °C on an orbital shaker at 20 rpm. Following incubation, the tubes placed in a magnetic rack on ice for 5 minutes and the supernatant was collected in a fresh tube, and again magnetized. The supernatant collected

from this second separation contained cells that do not bind to protein A/G. These cells were centrifuged for 5 minutes at 3,000 rcf and 4 °C and decanted.

Following incubation at the same conditions noted above, the cells were centrifuged, decanted and washed with 500 μL cold PBST. This was repeated 2 more times to remove any unbound antibodies. The final wash liquor was removed from centrifuged cells and magnetic beads were added at 1:10 for a 500 μL final volume. Following another incubation, the cells were placed on a magnetic rack on ice for 5 minutes and the supernatant was removed and discarded. The cells were washed four times with 500 μL PBST to remove any cells not attached to beads. The final cell library was incubated overnight in 20 mL LB/CM with 0.4% glucose added. A 10 μL sample of the library was removed, diluted and plated on LB/CM agar to estimate the diversity of the enriched library.

Subsequent MACS rounds were performed to increase library enrichment to antibody-binding peptides. Cells from the overnight growth were added at 1:50 to 15 mL LB/CM and grown to an OD_{600} of 0.4 - 0.6 at 37 °C. Following a one-hour induction with arabinose, an aliquot of cells containing >10x the library diversity (calculated from the overnight plates) was centrifuged and resuspended in 250 μL PBST containing 5 μL beads for at least a 1:1 bead-to-cell ratio to clear any protein A/G binding peptides. Magnetic separation on ice (5 min) followed by washing with cold PBST was performed two time to recover unbound cells, which were incubated with each antibody diluted to 25 nM in PBST for a final volume of 250 μL. Following incubation, the cells were centrifuged, and washed three times with 250 μL PBST. The final wash liquor was removed from centrifuged cells and 5 μL magnetic beads in PBST were added for a 250 μL final volume. Following incubation, cells were placed on a magnetic

rack on ice for 5 minutes and the supernatant was removed and discarded. The cells were magnetized and washed four times with 250 μL cold PBST to remove any cells not bound to beads. The final cell library was sampled for plating and incubated overnight in the same manner as above. The detailed protocol is shown in **Figure 12.1.**

12.4 Flow cytometry

The enriched library pools from MACS were analyzed for antibody binding via flow cytometry using a BD FACSARIA I/II to determine library enrichment. After overnight growth, the cells were subcultured at 1:50 into 15 mL LB/CM and grown to an OD_{600} of 0.4 - 0.6 at 37 °C. Following a one-hour induction with arabinose, a volume of cells >20x the estimated diversity was centrifuged and resuspended in 40 μL of 25 nM antibody. The cells were incubated for 45 minutes and washed as described in the MACS section above using 40 μL PBST. Then, cells were resuspended in α-IgG-PE diluted 1:100 in PBST. Following a 45 minute, 4 °C incubation, the cells were washed with cold PBST to remove unbound PE and resuspended in 500 μL cold PBS for flow cytometry analysis. A sample containing *E. coli* cells expressing the eCPX scaffold (but without a peptide) was used to determine the background signal by drawing a gate that excluded 99.5% of these cells in a plot of fluorescence vs. side scatter. This same gate was then applied to the antibody-screened library and cells within the gate (exhibiting a fluorescence greater than background) were considered enriched in antibody binders. The detailed protocol is shown in **Figure 12.1.**

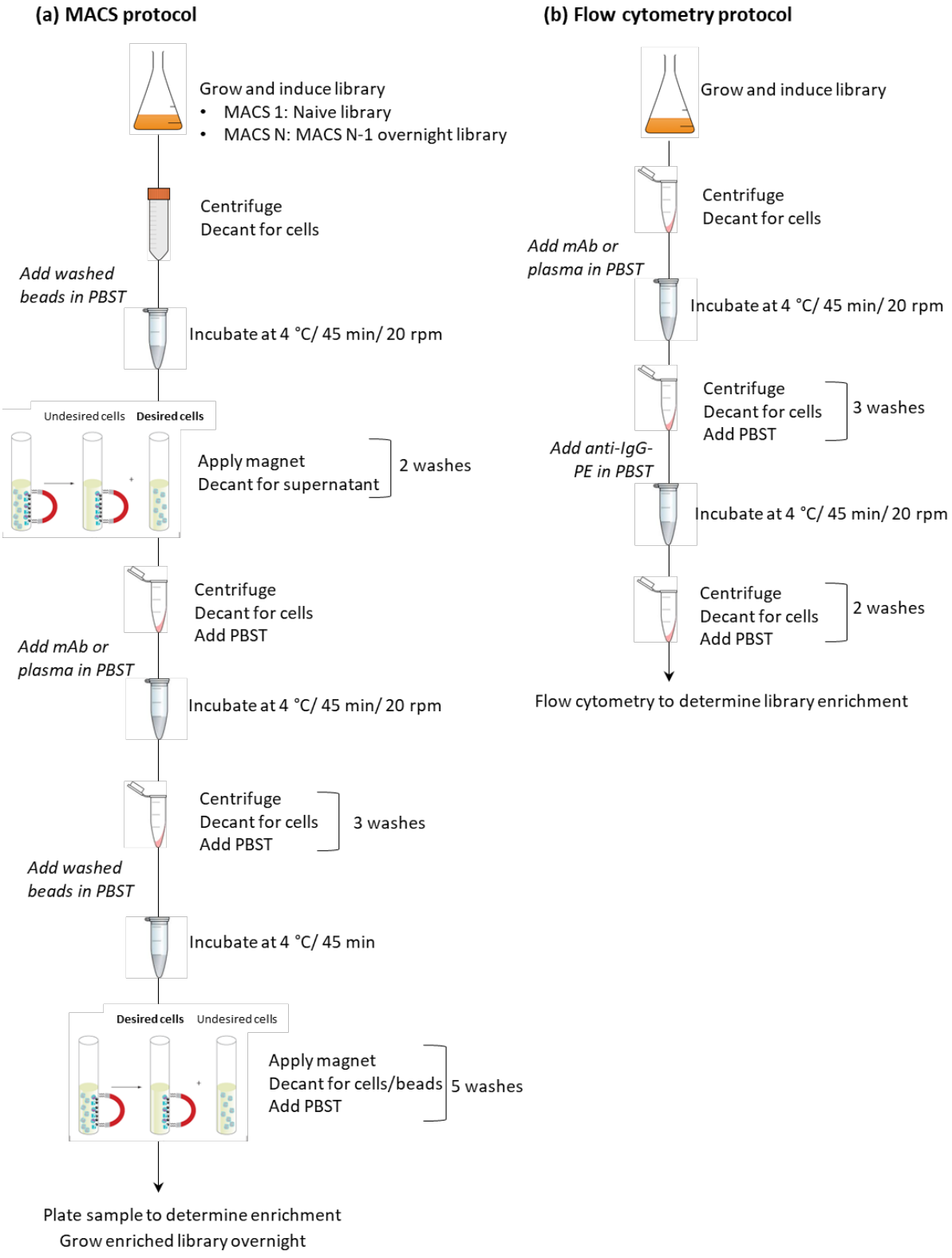


Figure 12.1. Detailed protocols for (a) MACS and (b) flow cytometry. One wash is defined as centrifugation, decanting and resuspension of cells or beads in PBST.

12.5 Sequencing

12.5.1 Sanger (dideoxy) sequencing

76 colonies were picked from LB/CM plates grown overnight from the final MACS round for each antibody. Individual colonies were transferred to a labeled, fresh LB/CM plate and incubated overnight at 37 °C, and peptide sequences were determined using Sanger DNA sequencing followed by translation using Geneious Pro v5.3.6.

12.5.2 Next Generation Sequencing

Libraries from MACS were prepared for NGS as previously described (Pantazes, et al., 2016). Briefly, plasmids were extracted from cells grown overnight using a plasmid miniprep kit. The random peptide region was amplified via two PCR steps. In the first PCR (a gradient starting at 72 °C annealing temperature and decreasing 0.5 °C per cycle for 14 cycles followed by 15 rounds at 65 °C annealing temperature), primers with adapters specific to the sequencing platform (Illumina) and annealing regions flanking the region of the eCPX scaffold. To identify each library, a second PCR (8 rounds, 70 °C annealing temperature) adds a unique nucleotide barcode to each amplicon using Illumina Nextera XT indexing primers. Clean-up using Agencourt AMPure XP beads (Beckman Coulter) followed each of the PCR steps. The purity of each amplicon was confirmed via gel electrophoresis and concentration was quantified using a fluorophore-based DNA high-sensitivity reagent (Qubit). Finally, amplicon libraries were normalized to one concentration and pooled for sequencing on an Illumina NextSeq 500. The pooled amplicon library purity was analyzed via a Bioanalyzer 2100 (Agilent) and the concentration measured using Qubit. A 75-cycle high-output flow cell with a single end read

was used with dual indexing. After sequencing, the samples were de-multiplexed using provided sample identities linked to the Nextera XT indices.

12.5.3 Generation of a non-redundant sequence list

Using the IMUNE processor developed in the Daugherty lab and described elsewhere (Pantazes, et al., 2016) the generated DNA sequences were translated into a set of unique peptide sequences for each antibody. First, the processor searches for the constant upstream and downstream annealing regions of the DNA and translates the intervening sequence to generate a peptide sequence. Then, identically matching sequences are combined. Finally, peptide sequences with 3 or fewer differences, which are considered mutations from sequencing errors, are combined. The value of 3 is based on a statistical analysis, described in Pantazes *et al.*, that shows the probability of finding 2 or more sequences with 10 or more identical positions is extremely low; therefore, sequences with 9 or more identical positions (3 or fewer mutations) are combined. Mutations outside the 12-mer peptide window are not counted.

Bibliography

Abdul-Sada, A. K. et al., 1995. *Process and catalysts for the alkylation of aromatic hydrocarbons.* World, Patent No. 95 21806.

Abdul-Sada, A. K., Seddon, K. R. & Stewart, N. J., 1995. *Ionic liquids of ternary melts.* World, Patent No. 95 21872.

Adawiyah, N., Moniruzzaman, M., Hawatulaila, S. & Goto, M., 2016. Ionic liquids as a potential tool for drug delviery systems. *Med Chem Commun,* Volume 7, pp. 1881-1897.

Agatemor, C., Ibsen, K. N., Tanner, E. E. L. & MItragotri, S., 2017. Ionic liquids for addressing unmet needs in healthcare. *Bioengineering & Translational Medicine,* pp. 1-19.

Ahmed, T. A., Eweida, A. E. & Sheweita, S. A., 2016. B-cell epitope mapping for the design of vaccines and effective diagnostics. *Trials in Vaccinology,* Volume 5, pp. 71-83.

Altschul, S. F. et al., 1990. Basic local alignment search tool. *J Mol Biol,* Volume 215, pp. 403-410.

Amde, M., Liu, J. F. & Pang, L., 2015. Environmental application, fate, effects and concerns of ionic liquids: a review. *Environ Sci Technol,* Volume 49, pp. 12611-12627.

Anderson, P. H., Nielsen, M. & Lund, O., 2006. Prediction of residues in discontinuous B-cell epitopes using protein 3D structures. *Protein Science,* Volume 15, p. 2558.

Angelini, A. et al., 2015. Protein engineering and selection using yeast surface display. In: B. Lui, ed. *Methods in Molecular Biology.* s.l.:s.n., p. 1319.

Anonymous, 1979. Geranic acid. *Food and Cosmetics Toxicology,* 17(Supplement), pp. 785-786.

Araujo, F. et al., 2017. Functionalized materials for multistage platforms in the oral delivery of biopharmaceuticals. *Progress in Materials Science,* Volume 89, pp. 306-344.

Arbit, E. & Kidron, M., 2017. Oral insulin delivery in a physiologic context: Review. *J Diabetes Sci Technol,* 11(4), pp. 825-832.

Arning, J. et al., 2008. Qualitative and quantitative structure activity relationships for the inhibitory effects of cationic head groups, functionalized side chains and anions of ionic liquids on acetylcholinesterase. *Green Chem,* Volume 10, pp. 47-58.

Arvan, P., Pietropaolo, M., Ostrov, D. & Rhodes, C. J., 2012. Islet autoantigens: structure, function, localization, and regulation. *Cold Spring Harbor perspectives in medicine,* 2(8), p. a007658.

Ashworth, C. R., Matthews, R. P., Welton, T. & Hunt, P. A., 2016. Doubly ionic hydrogen bond interactions within the choline chloride-urea deep eutectic solvent. *Phys Chem Chem Phys,* Volume 18, pp. 18145-18160.

Azevedo, J. R., Letourneau, J. J., Espitalier, F. & Ré, M. I., 2014. Solubility of a new cardioactive prototype drug in ionic liquids. *J. Chem Eng Data,* Volume 9, pp. 1766-1773.

Bailey, T. L. & Elkan, C., 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc of the Second Intl Conf on Intelligent Systems for Mol Bio,* pp. 28-36.

Banerjee, A. et al., 2018. Ionic liquids for oral insulin delivery. *Submitted to Proc Nat Acad Sci.*

Barlow, D. J., Edwards, M. S. & Thornton, J. M., 1986. Continuous and discontinuous protein antigenic determinants. *Nature,* 322(21), p. 747.

Barry, B. W., 2004. Breaching the skin's barrier to drugs. *Nat Biotechnol,* Volume 22, p. 165.

Bartels, J., Sonoja, G. E., Evans, C. M. & Segalman, R. A., 2017. Decouping mechanical and conductive dynamics of polymeric ionic liquids via a trivalent anion additive. *Macromolucules,* 50(22), pp. 8979-8987.

Belousov, P. V. et al., 2008. Cancer-associated antigens and antigen arrays in serological diagnostics of malignant tumors. *Biochemistry (Mosc),* 73(5), pp. 562-572.

Benson, H. A., 2005. Transdermal drug delivery: penetration enhancement techniques. *Curr Drug Deliv,* Volume 2, p. 23.

Bhattacharya, G. et al., 2017. X-ray reflectivity study of the interaction of an imidazolium-based ionic liquid with a soft supported lipid membrane. *Langmuir,* Volume 33, pp. 1295-1304.

Bhattacharya, G. et al., 2018. Thermodynamics of interaction of ionic liquids with lipid monolayer. *Biophysical Rev.*

Bingham, R. J. & Ballone, P., 2012. Computational study of room-temperature ionic liquids interacting with a POPC phospholipid bilayer. *J Phys Chem B,* 116(36), pp. 11205-11216.

Bolli, G. B., 2002. Clinical strageties for controlled peaks and valleys: type 1 diabetes. *Intl J Clin Practice (Supplement),* Volume 1129, pp. 65-74.

Bos, J. D. & Meinardi, M. M., 2000. The 500 Dalton rule for the skin penetration of chemical compounds and drugs. *Exp Dermatol,* Volume 9, p. 165.

Bouwstra, J. A. & Ponec, M., 2006. The skin barrier in healthy and diseased state. *Biochim Biophys Acta,* Volume 1758, p. 2080.

Brange, L. & Langkjoer, J., 1993. Insulin structure and stability. *Pharm Biotechnol,* Volume 5, pp. 315-350.

Brayden, D. J., Gleeson, J. & Walsh, E. G., 2014. A head-to-head multi-parametric high content analysis of a series of medium chain fatty acid intestinal permeation enhancers in Caco-2 cells. *Eur J Pharm Biopharm,* Volume 88, pp. 830-839.

Brown, J., Ross, T., McMeekin, T. & Nichols, P., 1997. Acid habituation of Escherichia coli and the potential role of cyclopropane fatty acids in low pH tolerance. *Int J Food Microbiol,* Volume 37, pp. 163-173.

Brown, M. B., Martin, G. P., Jones, S. A. & Akomeah, F. K., 2006. Dermal and transdermal drug delivery systems: current and future prospects. *Drug Deliv,* Volume 13, p. 175.

Bruno, B. J., Miller, G. D. & Lim, C. S., 2013. Basics and recent advances in peptide and protein drug delivery. *Ther Deliv,* 4(11), pp. 1443-1467.

Burtis, C., 2012. In: C. Burtis, ed. *Tietz Textbook of clinical Chemstiry and Molecular Diagnostics.* s.l.:Elsevier Health Sciences.

Cameron, D., Casal, H. & Mantsch, H., 1980. Characterization of the pretransition in 1,2-dipalmitoyl-sn-glycero-3-phosphocholine by Fourier transform infrared spectroscopy. *Biochemistry* , Volume 19, pp. 3665-3672.

Capeda, E. J. & Reveille, J. D., 2004. Autoantibodies in systemic sclerosis and fibrosing syndromes: Clinical indications and relevance. *Cur Opin Rheumatolgy,* 16(6), pp. 723-732.

Chakrabarti, P. & Janin, J., 2002. Dissecting protein-protein recognition sites. *Proteins: Structure, function and Genetics,* Volume 47, pp. 334-343.

Chen, W., Guo, W. W., Huang, Y. & Ma, Z., 2012. PepMapper: A collaborative web tool for mapping epitopes from affinity-selected peptides. *PLOS One,* 7(5).

Chen, Y. & Ganzle, M., 2015. Influence of cyclopropane fatty acids on heat, high pressure, acid and oxidative resistance in Escherichia coli. *Int J Food Microbiol* , Volume 222, pp. 16-22.

Chen, Y. et al., 2007. Autoantibodies to Ca2+ binding protein Calnuc is a potential marker in colon cancer detection. *Intl J Oncology,* 30(5), pp. 1137-1144.

Chiarelli, F., ed., 2005. Diabetes in Childhood and Adolescence. Karger Publishers.

Chimienti, F. et al., 2006. In vivo expression and functional characterization of the zinc transporter ZnT8 in glucose-induced insulin secretion. *J Cell Sci,* 119(20), pp. 4199-4206.

Cho, H. S. et al., 2003. Structure of the extracellular region of HER2 alone and in complex with the Herceptin Fab. *Letters to Nature,* Volume 421, pp. 756-759.

Choi, S., Ware Jr., W., Lauterbach, S. & W.M., P., 1991. Infrared spectroscopic studies on the phosphatidylserine bilayer interacting with calcium ions: Effect of cholesterol. *Biochemistry,* Volume 30, pp. 8563-8568.

Christiansen, A. et al., 2015. High-throughput sequencing enhanced phage display enables the identification of patient-specific epitope motifs in serum. *Sci Rep,* Volume 5, p. 12913.

Corte, L. et al., 2015. FTIR metabolomic fingerprint reveals different modes of action exerted by structural variants of N-alkyltropinium bromide surfactants on Escherichia coli and Listeria innocua cells. *PLOS One,* 10(1), p. e115275.

Cortijo, M., Alonso, A., Gomez-Fernandez, J. & Chapman, D., 1982. 44. Cortijo, M.; Alonso, A.; Gomez-Fernandez, J.C.; ChapmanIntrinsic protein-lipid interactions. Infrared spectroscopic studies of gramicidin A, bacteriorhodopsin and Ca2+-ATPase in biomembranes and reconstituted systems. *J Mol Biol* , Volume 157, pp. 597-618.

Cowie, J. M. G., 1991. *Polymers: Chemistry and Physics of Modern Materials.* 2nd ed. s.l.:CRC Press.

Cross, J., 1994. In: J. Cross & E. J. Singer, eds. *Introduction, Cationic Surfactants, Analytical and Biological Evaluation Surfactant Science Series.* New York: Marcel Dekker.

Curt, J. R. & Pringle, R., 1969. Viscosity of gastric mucus in duodenal ulceration. *Gut,* 10(11), pp. 931-934.

Daugherty, P. S., 2007. Protein engineering with bacterial display. *ScienceDirect,* Volume 17, pp. 474-480.

Davis, K. L., Hollister, L. E., Vento, A. L. & Simonton, S., 1978. Choline chloride in methylphenidate- and apomorphine-induced stereotypy. *Life Sciences,* 22(24), pp. 2171-2177.

DeMarshall, C. et al., 2017. Autoantibodies as diagnostic biomarkers for the detection and sutbyping of multiple sclerosis. *J Neuroimmunology,* Volume 309, pp. 51-57.

DeMarshall, C. et al., 2015. Utility of autoantibodies as biomarkers for diagnosis and staging of neurodegenerative diseases. *Intl Rev Neurobiology* Volume 122, pp. 1-51.

Desbois., A. & Smith, V., 2010. Antibacterial free fatty acids: activities, mechanisms of action and biotechnological potential. *Appl Microbiol Biotechnol,* Volume 85, pp. 1629-1642.

Docherty, K. & Kulpa Jr, C., 2005. Toxicity and antimicrobial activity of imidazolium and pyridinium ionic liquids. *Green Chem,* Volume 7, pp. 185-189.

Domina, M. et al., 2014. Rapid profiling of the antigen regions recognized by serum antibodies using massively parallel sequencing of antigen specific libraries. *PLOS One,* 9(12).

Egorova, K. S., Gordeev, E. G. & Ananikov, V. P., 2017. Biological activity of ionic liquids and their application in pharmaceutics and medicine. *Chem Rev,* Volume 117, pp. 7132-7189.

Eisenbarth, G., 2010. *Immunoendocrinology: Scientific and Clinical Aspects.* s.l.:Springer Science & Business Media.

Emel'yanenko, V. N., Boeck, G., Verevkin, S. P. & Ludwig, R., 2014. Volatile times for the very first ionic liquid: Understanding the vapor pressures and enthalpies of vaporization of ethylammonium nitrate. *Chem Eur J,* Volume 20, pp. 11640-11645.

EngineeringToolbox, 2018. *Absolute viscosity liquids.* [Online]
Available at: engineeringtoolbox.com/absolute-viscosity-liquids-d_1259
[Accessed 15 April 2018].

Farilla, L. et al., 2002. Application of phage display peptide library to autoimmune diabetes: identification of IA-2/ICA512bdc dominant autoantigenic epitopes.. *European journal of immunology,* 32(5), pp. 1420-1427.

Fields, M., Hutson, G. V., Seddon, K. R. & Gordon, C. M., 1998. *Ionic liquids as solvents.* World, Patent No. 98 06106.

Fisher Scientific, 2018. *pierce-protein-ag-magnetic-beads.* [Online]
Available at: https://www.fishersci.com/shop/products/pierce-protein-ag-magnetic-beads/p-4531522
[Accessed 5 May 2018].

Fonte, P., Araujo, F., Reis, S. & Sarmento, B., 2013. Oral insulin delivery: How far are we? *J Diabetes Sci Technol,* 7(2), pp. 520-531.

Fonte, P. et al., 2016. Polymer-based nanoparticles for oral insulin delivery: Revisited approaches. *Biotechnol Adv,* 34(1), p. 64.

Friedman, M., Henika, P., Levin, C. & Mandrell, R., 2004. Antibacterial activities of plant essential oils and their components against Escherichia coli O157:H7 and Salmonella enterica in apple juice. *J Agric Food Chem,* Volume 52, pp. 6042-6048.

Fujita, K., MacFarlane, D. R. & Forsyth, M., 2005. Protein solubilizing and stabilising ionic liquids. *Chem Commun,* pp. 4804-4806.

Gai, S. A. & Wittrup, K. D., 2007. Yeast surface display for protein engineering and characterization. *Curr Opin Structural Biology,* Volume 17, pp. 467-473.

Gershoni, J. M. et al., 2007. Epitope mapping: the first step in developing epitope-based vaccines. *BioDrugs,* 21(3), pp. 145-56.

Graenacher, C., 1931. *Cellulose Solution.* US, Patent No. 1943176A.

Gupta, V., Hwang, B. H., Doshi, N. & Mitragotri, S., 2013. A permeation enhancer for increasing transport of therapeutic macromolecules across the intestine. *J Contr Rel,* 172(2), pp. 541-549.

Hadebe, S. I., Ngubane, P. S., Serumula, M. R. & Msabayane, C. T., 2014. Transdermal delivery of insulin by amidated pectin hydrogel matrix patch in streptozotocin-induced diabetic rats: effects on some selected metabolic parameters. *PLOS One.*

Hall, S. S. & Daugherty, P. S., 2009. Quantitative Specificity-Based Display Library Screening Identifies Determinants of Antibody-Epitope Binding Specificity. *Protein Science* , Volume 18, pp. 1926-1934.

Hamaguchi, Y., 2010. Anutoantobody profiles in systemic sclerosis: Predictive value for clinical evaluation and prognosis. *J Dermatology,* 37(1), pp. 42-53.

Hartmann, D. & Silva Pereira, C., 2013. A molecular analysis of the toxicity of alkyltributylphosphonium chlorides in Aspergillus nidulans. *New J Chem,* Volume 37, pp. 1569-1577.

Heuer, L. et al., 2008. Reduced levels of immunoglobulin in children with autism correlates with behavioral symptoms. *Autism Research,* 1(5), pp. 275-283.

Heyduk, E. & Heyduk, T., 2014. Ribosome display enhanced by next generation sequencing: A tool to identify antibody-specific peptide ligands. *Analytical Biochemistry,* Volume 464, pp. 73-82.

Holland, D. N. R. a. P. M. ed., 1991. In: *Cationic Surfactants: Physical Chemistry.* New York: Marcel Dekker.

Hough-Troutman, W. L. et al., 2009. Ionic liquids with dual biological function: sweet and anti-microbial; hydrophobic quaternary ammonium-based salts. *New J Chem,* Volume 33, pp. 26-33.

Hough, W. L. et al., 2007. The third evolution of ionic liquids: active pharmaceutical ingredients. *New J Chem,* Volume 31, pp. 1429-1436.

Hsu, P., Jefferies, D. & Khalid, S., 2016. Molecular dynamics simulations predict the pathways via which pristine fullerenes penetrate bacterial membranes. *J Phys Chem B,* Volume 120, p. 11170− 11179.

Huang, K. et al., 2004. How insulin binds: the beta-chain alpha-helix contacts and the L1 beta-helix of the insulin receptor. *J. Mol. Biol.,* Volume 341, p. 529.

Huang, Y. X. et al., 2008. Pep-3D-Search: a method for B-cell epitope prediction based on mimotope analysis. *BMC Bioinformatics,* Volume 9, p. 538.

Hua, Q. X. et al., 2002. Mechanism of insulin chain combination. Assymetric roles of A-chain alpha-helices in disulfide pairing. *J Biol Chem,* Volume 277, p. 43443.

Hudson, E. P., Uhlen, M. & Rockberg, J., 2012. Multiplex epitope mapping using bacterial surface display reveals both linear and conformational epitopes. *Sci Rep,* 2(706), p. 1.

Janeway, C. A., Travers, P. & Walport, M., 2001. The humoral response. In: *Immunobiology: The Immune System in Health and Disease.* New York: Garland Science.

Javadzadeh, Y., Adibkia, K. & Hamishekar, H., 2015. In: N. Dragicevic & H. Maibach, eds. *Percutaneous Penetration Enhancers Chemical Methods in Penetration Enhancement.* s.l.:Springer, p. 195.

JDRF, 2018. *what-is-tid/facts.* [Online].

Jeong, S. et al., 2012. Elucidation of molecular interactions between lipid membranes and ionic liquids using model cell membranes. *Soft Mater,* Volume 8, pp. 5501-5506.

Jeong, S. et al., 2012. Elucidation of molecular interactions between lipid membranes and ionic liquids using model cell membranes. *Soft Mater,* Volume 8, pp. 5501-5506.

Jin, Y. & She, J., 2012. Novel Biomarkers in Type 1 Diabetes. *Rev Diabet Stud,* Volume 9, pp. 224-235.

Kalia, Y. N., Merino, V. & R. H. Guy, 1998. Transdermal drug delivery. Clinical aspects. *Dermatol. Clin,* Volume 16, p. 289.

Kalia, Y. N., Naik, A., Garrison, J. & Guy, R. H., 2004. Iontophoretic drug delivery. *Adv Drug Deliv Rev,* Volume 56, p. 619.

Karalliedde, J. & Gnudi, L., 2014. Diabetes mellitus, a complex and heterogeneous disease, and the role of insulin resistance as a determinant of diabetic kidney disease. *Nephrol Dial Transplant,* Volume 31, pp. 206-213.

Karande, P. et al., 2005. Design principles of chemical penetration enhancers for transdermal drug delivery. *Proc Nat Acad Sci,* Volume 102, p. 4688.

Karande, P., Jain, A. & Mitragotri, S., 2004. Discovery of transdermal penetration enhancers by high-throughput screening. *Nat Biotechnol,* Volume 22, p. 192.

Karande, P., Jain, A. & Mitragotri, S., 2006. Relationships between skin's electricial impedance and permeability in the presence of chemical enhancers. *J Contr Rel,* Volume 110, pp. 307-313.

Karande, P. & Mitragotri, S., 2009. Enhancement of transdermal drug delivery via synergistic action of chemicals. *Biochim Biophys Acta,* 1788(11), pp. 2362-73.

Karsdal, M. A., Byrjalsen, I., Riis, B. J. & Christiansen, C., 2008. Optimizing bioavailability of oral administration of small peptides through pharmacokinetic and pharmacodynamic

parameters: the effect of water and timing of meal intake on oral delivery of Salmon Calcitonin. *BMC Clin Pharmacol,* Volume 8, p. 5.

Kaushik, D. & Michniak-Kohn, B., 2010. Percutaneous penetration modifiers and formulation effects: thermal and spectral analyses. *AAPS PharmSciTech,* Volume 11, p. 1068.

Kharroubi, A. & Darwish, H., 2015. Diabetes mellitus: The epidemic of the century. *World J Diabetes,* 6(6), pp. 850-867.

Kim, Y. C., Park, J. H. & Prausnitz, M. R., 2012. Microneedles for drug and vaccine delivery. *Adv Drug Deliv Rev,* Volume 64, p. 1547.

Klähn, M. & Zacharias, M., 2013. Transformations in plasma membranes of cancerous cells and resulting consequences for cation insertion studied with molecular dynamics. *Phys Chem Chem Phys ,* Volume 15, pp. 14427-14441.

Klinguer, C. et al., 2001. *Vaccine,* Volume 19, p. 4236.

Kobayashi, T. & Shibui, T., 2011. Isolation of mimotopes to the anti-human VEGF antibody Bevaxizumab by mRNA display using random peptide libraries and the vaccination of a rabbit. *Adv Biosci & Biotech,* Volume 2, pp. 97-102.

Koczwara, K., Bonifacio, E. & Ziegler, A. G., 2004. Transmission of maternal islet antibodies and risk of autoimmune diabetes in offspring of mothers with type 1 diabetes. *Diabetes,* 53(1), pp. 1-4.

Kohanski, M., DePristo, M. & Collins, J., 2010. Sub-lethal antibiotic treatment leads to multidrug resistance via radical-induced mutagenesis. *Mol Cell,* 37(3), pp. 311-320.

Konsoula, R. & Barile, F. A., 2005. Correlation of in vitro cytotoxicity with paracellular permeability in caco-2 cells. *Toxicol in Vitro,* 19(5), pp. 675-684.

Kragl, U., Eckstein, M. & Kaftzik, N., 2002. Enzyme catalysis in ionic liquids. *Current Opinion in Biotechnology,* Volume 13, pp. 565-571.

Krishnamurthy, V. M. et al., 2007. Dependence of Effective Molarity on Linker Length for an Intramolecular Protein-Ligand System. *Journal American Chemical Society,* Volume 129, pp. 1312-1320.

Kronqvist, N. et al., 2010. Staphylococcal surface display in combinatorial protein engineering and epitope mapping of antibodies. *Recent Pat Biotechnol,* Volume 4, pp. 171-182.

Krug, S. M., 2013. Sodium caprate as an enhancer of macromolecule permeation across tricellular tight junctions of intestinal cells. *Biomaterials,* 34(1), pp. 275-282.

Kumar, A., 2013. Prevention of insulin self-aggregation by a protic ionic liquid. *Royal Society of Chemistry Advances,* Volume 3, pp. 362-367.

Kumar, V., Parmar, V. S. & Malhotra, S. V., 2007. Enhanced solubility and selective benzoylation of nucleosides in novel ionic liquid. *Tetrahedron Lett,* Volume 48, pp. 809-812.

Kuma, S. et al., 2015. Peptides as skin penetration enhancers: mechanism of action. *J Control Rel,* Volume 199, p. 168.

Kupper, T. S. & Fuhlbrigge, R. C., 2004. Immune serveillance in the skin: mechanisms and clinical consequences. *Nat Rev Immunology,* Volume 4, pp. 211-222.

Kuroda, D., Shirai, H., Jacobson, M. P. & Nakamura, H., 2012. Computer-aided antibody design. *Protein Eng Des & Selection,* 25(10), pp. 507-521.

Kwiatek, M. A. et al., 2009. Effects of clonidine and sumatriptan on postprandial gastric volume response, antral contraction waves and emptying: an MRI study. *Neurogastroenterol Motil,* 21(9), p. 928.

Lai, S. K., Wang, Y. Y., Wirtz, D. & Hanes, J., 2009. Micro- and macrorheology of mucus. *Adv Drug Deliv Rev,* 61(2), pp. 86-100.

Lampasona, V. et al., 2010. Zinc transporter 8 antibodies complement GAD and IA-2 antibodies in the identification and characterization of adult-onset autoimmune diabetes non-insulin requiring autoimmune diabetes (NIRAD) 4. *Diabetes Care,* 33(1), pp. 104-108.

Lampasona, V. et al., 2011. Diabetes antibody standardization program: first proficiency evaluation of assays for autoantibodies to zinc transporter 8. *Clinical Chemistry,* 57(12), pp. 1693-1702.

Larman, H. B. et al., 2013. PhIP-Seq characterization of autoantibodies from patients with multiple sclerosis, type 1 diabetes and rheumatoid arthritis. *Journal of autoimmunity,* Volume 43, pp. 1-9.

Lee, D. et al., 2009. Permeability of model stratum corneum lipid membrane measured using quarz crystal microbalance. *Langmuir,* Volume 25, p. 5752.

Lee, W. R. et al., 2002. The effect of laser treatment on skin to enhance and control transdermal delivery of 5-fluorouracil.*J Pharm Sci,* Volume 91, p. 1613.

LeRoith, D., 2004. *Diabetes mellitus: a fundamental and clinical text.* D. LeRoith; S. I. Taylor; J. M. Olefsky ed. s.l.: Lippincott Williams & Wilkins..

Ling, M. H. & Chen, M. C., 2013. Dissolving polymer microneedle patches for rapid and efficient transdermal delivery of insulin to diabetic rats. *Acta Biomater,* 9(11), pp. 8952-8961.

Lin, M. et al., 2004. Discrimination of intact and injured Listeria monocytogenes by Fourier transform infrared spectroscopy and principal component analysis. *J Agric Food Chem,* Volume 52, pp. 5769-5772.

Lippincott, 2007. *Diabetes Mellitus: A Guide to Patient Care.* Lippincott, Williams & Wilkins ed. s.l.:Wolters Kluwer Health.

Liu, F. & Huang, L., 2002. Development of non-viral vectors for systemic gene delivery. *J. Controlled Release,* Volume 78, p. 259.

Li, W. et al., 2013. Mimotope vaccination for epitope-specific induction of anti-VEGF antibodies. *BMC Biotechnology,* Volume 13, p. 77.

Lopes, L. B., Garcia, M. T. J. & Bentley, M. V., 2015. Chemical penetration enhancers. *Ther Deliv.*

Maglio, M. et al., 2009. Majority of children with type 1 diabetes produce and deposit anti-tissue transglutaminase antibodies in the small intestine. *Diabetes,* 58(7), pp. 1578-1584.

Ma, H. et al., 2017. Modeling diversity in structures of bacterial outer membrane lipids. *J Chem Theory Comput,* Volume 13, p. 811−824.

Ma, H., Irudayanathan, F., Jiang, W. & Nangia, S., 2015. Simulating Gram-negative bacterial outer membrane: A coarse grain model. *J Phys Chem B,* Volume 119, p. 14668−14682.

Ma, H., Khan, A. & Nangia, S., 2017. Dynamics of OmpF trimer formation in the bacterial outer membrane of Escherichia coli. *Langmuir.*

Mantsch, H. & McElhaney, R., 1991. Phospholipid phase transitions in model and biological membranes as studied by infrared spectroscopy. *Chem Phys Lipids,* Volume 57, pp. 213-226.

Matochko, W. L. & Derda, R., 2013. Error analysis of deep sequencing of phage libraries: peptides censored in sequencing. *Computational and Mathematical Methods in Medicine,* p. 491612.

Mayeux, R., 2004. Biomarkers: potential uses and limitations. *NeuroRx,* Volume 1, pp. 182-188.

Mayo Medical Laboratories, 2015. *Reference Laboratory Services for Health Care Organizations Test Catalog. Evaluation for DM Type 1.* [Online]
Available at: http://www.mayomedicallaboratories.com/index.html
[Accessed 15 May 2015].

Mayrose, I. et al., 2006. Epitope mapping using combinatorial phage-display libraries: a graph-based algorithm. *Nuclei Acids Res,* 35(1), pp. 69-78.

Meirsch, S. et al., 2013. Serological autoantibody profiling of type 1 diabetes by protein arrays. *Journal of Proteomics,* Volume 94, pp. 486-496.

Meng, X., Zhang, H., Mezei, M. & Cui, M., 2011. Molecular docking: A powerful approach for structure-based drug discovery. *Cur Comput Aided Drug Des,* 7(2), pp. 146-157.

Mennuni, C. et al., 1997. Identification of a novel type 1 diabetes-specific epitope by screening phage libraries with sera from pre-diabetic patients. *Journal of Molecular Biology,* 268(3), pp. 599-606.

Micaelo, N. M. & Soares, C. M., 2008. Protein structure and dynamics in ionic liquids. insights from molecular dynamics simulation studies. *J Phys Chem B,* 112(9), pp. 2566-2572.

Mikszta, J. A. et al., 2002. Improved genetic immunization via micromechanical disruption of skin-barrier function and targeted epidermal delivery. *Nat Med ,* Volume 8, p. 415.

Mitragotri, S., Blankschtein, D. & Langer, R., 1995. Ultrasound-mediated transdermal protein delivery. *Science,* Volume 269, p. 850.

Mizuuchi, H., Jaitely, V., Murdan, S. & Florence, A. T., 2008. Room temperature ionic liquids and their mixtures: potential pharmaceutical solvents. *Eur J Pharm Sci,* Volume 33, pp. 326-331.

Moghadam, S. H. et al., 2013. Effect of chemical permeation enhancers on stratum corneum barrier lipid organizational structure and interferon alpha permiability. *Mol Pharm,* Volume 10, p. 2248.

Moniruzzaman, M., Tahara, Y., Kamiya, N. & Goto, M., 2010. Ionic liquid-assisted transdermal delivery of sparingly soluble drugs. *Chem Commun,* Volume 47, pp. 1452-1454.

Monti, D. et al., 2017. Ionic liquids as potential enhancers for transdermal drug delivery. *Int J Pharm,* Volume 516, pp. 45-51.

Morrison, H. G., Sun, C. C. & Neervannan, S., 2009. Characterization of thermal behavior of deep eutectic solvents and their potential as drug solubilization vehicles. *International J Pharmaceutics,* Volume 378, pp. 136-139.

Muheem, A. et al., 2016. A review on the strategies for oral delviery of proteins and peptides and their clinical perspectives. *Saudi Pharm J,* Volume 24, pp. 413-428.

Muller, Y. A. et al., 1998. VEGF and the Fab fragment of a humanized neutralizing antibody: crystal structure of the complex at 2.4 A resolution and mutational analysis of the interface. *Structure,* 6(9), pp. 1153-1167.

Mura, P., Faucci, M. T., Bramanti, G. & Corti, P., 2000. Evaluation of transcutol as a clonazepam transdermal permeation enhancer from hydrophilic gel formulations. *Eur J Pharm Sci,* Volume 9, p. 365.

Myers, M. A. et al., 2000. Conformational epitopes on the diabetes autoantigen GAD65 identified by peptide phage display. *J Immunology,* 165(7), pp. 3830-3838.

Nagele, E. P. et al., 2013. Natural IgG autoantibodies are abundant and ubiquitous in human sera, and their number is influenced by age, gender, and disease. *PLOS One,* 8(4), p. e60726.

Negi, S. S. & Braun, W., 2009. Automated detection of conformational epitopes using phage display peptide sequences. *Bioinformatics & Biology Insights,* Volume 3, pp. 71-81.

Niu, M. et al., 2012. Hypoglycemic activity and oral bioavailability of insulin-loaded liposomes containing bile salts in rats: the effect of cholate type, particle size and adminstered dose. *Eur J Pharma & Biopharma,* Volume 81, pp. 265-272.

Olokoba, A., Obateru, O. & Olokoba, L., 2012. Type 2 diabetes mellitus: A review of current trends. *Oman Med J,* 27(4), pp. 269-273.

Orban, T. et al., 2009. Pancreatic islet autoantibodies as predictors of type 1 diabetes in the Diabetes Preventions Trial-Type 1. *Diabetes Care,* Volume 32, pp. 2269-2274.

Pantazes, R. J. et al., 2016. Identification of disease-specific motifs in the antibody specificity repertoire via next-generation sequencing.. *Sci Rep,* Volume 6, p. 30312.

Papadimitriou, K. et al., 2008. RNA arbitrarily primed PCR and Fourier Transform infrared spectroscopy reveal plasticity in the acid tolerance response of Streptococcus macedonicus. *J Appl Environ Microbiol,* 74(19), pp. 6068-6076.

Paschoal, V. H., Faria, L. F. O. & Ribeiro, M. C. C., 2017. Vibrational spectroscopy of ionic liquids. *Chem Rev,* 117(10), pp. 7053-7112.

Patel, K. et al., 2011. Enhancement of a multianalyte serum biomarker panel to identify lymph node metastases in non-small cell lung cancer with circulating autoantibody biomarkers. *Intl J Cancer,* 129(1), pp. 133-142.

Pathan, I. B. & Setty, C. M., 2009. Chemical penetration enhancers for transdermal drug delivery systems. *Tropical J Pharm Res,* 8(2), pp. 173-179.

Paudel, K. S. et al., 2010. Challenges and opportunities in dermal/transdermal delivery. *Ther Deliv,* Volume 1, p. 109.

Paull, M. L., Johnstion, T., Ibsen, K. N. & Daugherty, P. S., 2018. A general approach for identifying protein epitopes targeted by antibody repertoires using whole proteomes. *In preparation.*

Pendleton, J. & Gilmore, B., 2015. The antimicrobial potential of ionic liquids: A source of chemical diversity for infection and biofilm control. *Int J Antimicrob Agents,* 46(2), pp. 131-139.

Percival, S. L., McCarty, S. M. & Lipsky, B., 2014. Biofilms and wounds: An overview of the evidence. *Adv Wound Care,* 4(7), pp. 373-381.

Pernak, J. et al., 2007. Synthesis and properties of chiral imidazolium ionic liquids with a (1R,2S,5R)-(-)-menthoxymethyl substituent. *New J Chem,* Volume 31, pp. 879-892.

Pernak, J. & Skrzypczak, A., 1996. 3-alkylthiomethyl-1-ethylimidazolium chlorides. Correlations between critical micelle concentrations and minimum inhibitory concentrations. *Eur J Med Chem,* Volume 31, pp. 901-903.

Pernak, J., Sobaszkiewicz, K. & Mirska, I., 2003. Anti-microbial activities of ionic liquids.. *Green Chem,* Volume 5, pp. 52-56.

Pernak, J. et al., 2007. Choline-derivative-based ionic liquids. *Chem Eur J,* Volume 13, pp. 6817-6827.

Petkovic, M. et al., 2010. Novel biocompatible cholinium-based ionic liquids-toxicity and biodegradability. *Green Chem,* Volume 12, pp. 643-649.

Petrocci, A. N., 1983. Disinfection, Sterilization and Preservation. In: S. S. Block, ed. Philadelphia: Lea & Febiger.

Pettis, R. J. et al., 2011. Intradermal microneedle delivery of insulin lispro acheives faster insulin absorption and insulin action than subcutaneous injection. *Diabetes Technol Ther,* 13(4), pp. 435-442.

Peyrot, M., Rubin, R., Kruger, D. & Travis, L., 2010. Correlates of insulin injection omission. *Diabetes Care,* 33(2), pp. 240-145.

Plechkova, N. V. & Seddon, K. R., 2008. Applications of ionic liquids in the chemical industry. *Chem Soc Rev,* Volume 37, pp. 123-150.

Prausnitz, M. R., 2004. Microneedles for transdermal drug delivery. *Adv Drug Deliv Rev,* Volume 56, p. 581.

Prausnitz, M. R., Bose, V. G., Langer, R. & Weaver, J. C., 1993. Electroporation of mammalian skin: a mechanism to enhance transdermal drug delivery. *Proc Natl Acad Sci,* Volume 90, p. 10504.

Prausnitz, M. R. & Langer, R., 2008. Transdermal drug delivery. *Nat Biotechnol,* Volume 26, p. 1261.

Prausnitz, M. R., Mitragotri, S. & Langer, R., 2004. Current status and future potential of transdermal drug delivery. *Nat Rev Drug Discov,* Volume 3, p. 115.

Quail, M. A. et al., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, Pacific Biosciences and Illumina MiSeq sequencers. *BMC genomics,* 13(1), p. 341.

Rajkowska, K. et al., 2016. Quaternary ammonium biocides as antimicrobial agents prtecting historical wood and brick. *ACTA ABP Biochimica Polonica,* 63(1), pp. 153-159.

Rantwijk, F. v., Lau, R. M. & Sheldon, R. A., 2003. Biocatalytic transformations in ionic liquids. *TRENDS in Biotechnology,* 21(3), pp. 131-138.

Rapburger, R., Lukas, A. & Mayer, B., 2007. Identification of discontinuous antigenic determinants on proteins based on shape complementarities. *J Mol Recognit,* Volume 20, pp. 113-121.

Reimer, A. B. et al., 2004. Generation of peptide mimics of the epitope recognized by trastuzumab on the oncogenic protein Her-2/neu. *J of Immunol,* Volume 173, pp. 394-401.

Reimer, A. B. et al., 2005. Matching of trastuzumab (Herceptin®) epitope mimics onto the surface of Her-2/neu – a new method of epitope definition. *Molecular Immunology,* Volume 42, p. 1121.

Reis, O., Winter, R. & Zerda, T., 1996. The effect of high external pressure on DPPC-cholesterol multilamellar vesicles: a pressure-tuning Fourier transform infrared spectroscopy study. *Biochim Biophys Acta,* Volume 1279, pp. 5-16.

Rengstl, D., Fischer, V. & Kunz, W., 2014. Low-melting mixtures based on choline ionic liquids. *Phys Chem Chem Phys,* Volume 16, pp. 22815-22822.

Rice, J. J. & Daugherty, P. S., 2008. Directed Evolution of a Biterminal Bacterial Display Scaffold Enhances Display of Diverse Peptides. *Protein Engineering, Design & Selection,* Volume 21, pp. 435-442.

Rice, J. J. et al., 2006. Bacterial Display Using Circularly Permuted Outer Membrane Protein OmpX Yields High Affinity Peptide Ligands. *Protein Science,* Volume 15, pp. 825-836.

Richmond, J. M. ed., 1990. *Cationic Surfactants.* New York: Marcel Dekker.

Robinson, W. H., 2014. Sequencing the functional antibody repertoire - diagnostic and therapeutic discovery. *Nature Reviews Rheumatology,* Volume 11, pp. 171-182.

Roep, B. O. & Peakman, M., 2012. Antigen targets of type 1 diabetes autoimmunity. *Cold Spring Harbor perspectives in medicine,* 2(4), p. a007781.

Rojas, G., Tundidor, Y. & Infante, Y. C., 2014. High throughput functional epitope mapping: Revisiting phage display platform to scan target antigen surface. *mAbs,* 6(6), pp. 1368-1376.

Sadowski, T. et al., 2017. Large-scale human skin lipidomics by quantitative, high-throughput shotgun mass spectrometry. *Sci Rep,* Volume 7, p. 43761.

Sarmento, B. et al., 2007. Oral bioavailablility of insulin contained in polysaccharide nanoparticles. *Biomacromolecules,* Volume 8, pp. 3054-3060.

Scherbaum, W., Trischler, G. & Pfeiffer, E. F., 1989. Non-human primate pancreas as a substrate for the detection of islet-cell antibodies in human sera. Diabetes Research & Clinical Practice 7(1):1-5.. *Diabetes Research & Clinical Practice,* 7(1), pp. 1-5.

Schlosser, M. et al., 2011. Diabetes Antibody Standardization Program First evaluation of assays for autoantibodies to IA-2β. *Diabetes Care,* 34(11), pp. 2410-2412.

Schlosser, M. et al., 2010. Diabetes Antibody Standardization Program: evaluation of assays for insulin autoantibodies. *Diabetologia,* 53(12), pp. 2611-2620.

Seddon, K. R., 1999. *The International George Papatheodorou Symposium Proceedings.* Patras, Institute of Chemical Engineering and HIgh Temperature Chemical Processes, pp. 131-135.

Shaji, J. & Patole, V., 2008. Protein and peptide drug delivery: Oral approaches. *Indian J Pharm Sci,* 70(3), pp. 269-277.

Shakya, A. K. & Nandakumar, K. S., 2018. Antigen-specific tolerization and targeted delivery as therapeutic strategies for autoimmune diseases. *Trends in Biotechnology.*

Sheng, J. et al., 2016. Enhancing insulin oral absorption by using mucoadhesive nanoparticles loaded with LMWP-linked insulin conjugates. *J Contr Rel.*

Siew, A. et al., 2011. Enhanced oral absorption of hydrophobic and hydrophilic drugs using quaternary ammonium palmitoyl glycol chitosan nanoparticles. *Mol Pharmaceutics,* Volume 9, pp. 14-28.

Sigrist, C. J. A. et al., 2012. New and continuing developments at PROSITE. *Nucleic Acids Res.*

Simovic, S., Song, Y., Nann, T. & Desai, T. A., 2015. Intestinal absorption of fluorescently labeled nanoparticles. *Nanomedicine,* 11(5), pp. 1169-1178.

Siopa, F. et al., 2016. Choline-based ionic liquids: Improvement of antimicrobial activity. *ChemistrySelect,* Volume 1, pp. 5909-5916.

Smeden, J. v. et al., 2014. Combined LC/MS-platform for analysis of all major stratum corneum lipids, and the profiling of skin substitutes. *Biochimica et Biophysica Acta,* Volume 1841, pp. 70-79.

Soares, S., Costa, A., Fonte, P. & Sarmento, B., 2017. In: Y. Rosen, P. Gurman & N. Elman, eds. *Drug Delivery: An Integrated Clinical and Engineering Approach.* s.l.:Taylor and Francis.

Sommer, M., Munck, C., Toft-Kehler, R. & Andersson, D., 2017. Prediction of antibiotic resistance: time for a new preclinical paradigm?. *Nat Rev Microbiol,* Volume 15, pp. 689-696.

Sonaje, K. et al., 2010. Biodistribution; pharmacohynamics and pharmacokinetics of insulin analogues in a rat model: oral delviery using pH-responsive manoparticles vs. subcutaneous injection. *Biomaterals,* Volume 31, pp. 6849-6858.

Sorenson, R. L., Garry, D. G. & Brelje, T. C., 1991. Structural and functional considerations of GABA in islets of Langerhans: β-cells and nerves. *Diabetes,* 40(11), pp. 1365-1374.

Springer Science & Business Media, 1994. *Standardization of Epidemiologic Studies of Host Susceptibility.* J. Dorman ed. s.l.:s.n.

Stark, A. & Seddon, K., 2007. In: A. Seidel, ed. *Kirk-Othmer Encyclopaedia of Chemical Technology.* Hoboken: John Wiley & Sons, Inc., pp. 836-920.

Steck, A. K. et al., 2011. Age of islet autoantibody appearance and mean levels of insulin but not GAD or IA-2 autoantibodies predict age of diagnosis of type 1 diabetes. *Diabetes Care,* Volume 34, pp. 1397-1399.

Steingoetter, A. et al., 2006. Effects of posture on the physiology of gastric emptying: a magnetic resonance study. *Scand J Gastroenterol,* 41(10), pp. 1155-64.

Suchodolski, J., Feder-Kubis, J. & Krasowska, A., 2017. Antifungal activity of ionic liquids based on (-)-menthol: a mechanism study. *Microbiol Res,* Volume 197, pp. 56-64.

Sun, P. et al., 2011. Epitope prediction based on random peptide library screening: Benchmark dataset and prediction tools evaluation. *Molecules,* Volume 16, pp. 4971-4993.

Tamm, L. & Tatulian, S., 1997. Infrared spectroscopy of proteins and peptides in lipid bilayers. *Q Rev Biophys,* 30(4), pp. 365-429.

Tanko, L. B. et al., 2004. Safety and efficacy of a novel salmon calcitonin (sCT) technology-based oral formulation in healthy postmenopausal women: acute and 3-month effects on biomarkers of bone turnover.. *J Bone Miner REs,* 19(9), pp. 1531-8.

Tanner, E. E. L., Ibsen, K. N. & Mitragotri, S., 2018. Transdermal insulin delivery using choline-based ionic liquids (CAGE). *Submitted to J Cont Rel.*

The American Diabetes Association, 2013. *Juvenile Diabetes Research Foundation sourcebook for Type 1 diabetes.* A. Peters; L. Laffel ed. s.l.:s.n.

The Diamond Project Group, 2006. The Diamond Project Group (2006) Incidence and trends of childhood Type 1 diabetes worldwide 1990-1999. *Diabet Med,* Volume 23, pp. 857-866.

Thomas, B. J. & Finnin, B. C., 2004. The transdermal revolution. *Drug Discov Today,* Volume 9, p. 697.

Tiberti, C. et al., 2011. Detection of four diabetes specific autoantibodies in a single radioimmunoassay: an innovative high-throughput approach for autoimmune diabetes screening. *Clinical & Experimental Immunology,* 166(3), pp. 317-324.

Tischer, M., Pradel, G., Ohlsen, K. & Holzgrabe, U., 2012. Quaternary ammonium salts and their antimicrobial potential: Targets or nonspecific interactions?. *ChemMedChem,* Volume 7, pp. 22-31.

t'Kindt, R. et al., 2012. Profiling and characterings skin ceramides using reversed-phase liquid chromatorgraphy-quadrupole time-of-flight mass spectrometry. *Anal Chem,* Volume 84, pp. 403-411.

Törn, C. et al., 2008. Diabetes Antibody Standardization Program: evaluation of assays for autoantibodies to glutamic acid decarboxylase and islet antigen-2. *Diabetologia,* 51(5), pp. 846-852.

Torrecilla, J., García, J., Rojo, E. & Rodríguez, F., 2009. Estimation of toxicity of ionic liquids in leukemia Rat Cell Line and Acetylcholinesterase enzyme by principal component analysis neural networks and multiple linear regressions. *J Hazard Mater,* Volume 164, pp. 182-194.

Tridgell, D. M., Spiekerman, C., Wang, R. S. & Greenbaum, C. J., 2011. Interaction of onset and duration of diabetes on the percent of GAD and IA-2 antibody-positive subjects in the type 1 diabetes genetics consortium database. *Diabetes Care,* Volume 34, pp. 988-993.

Trommer, H. & Neubert, R. H., 2006. Overcoming the stratum corneum: the modulation of skin penetration. A review. *Skin Pharmacol Physiol,* Volume 19, p. 106.

Tsai, J. et al., 2004. Feasibility of rapid quantitation of stratum corneum lipid content by Fourier transform infrared spectrometry. *Spectroscopy,* Volume 18, p. 423.

Tsai, Y., Ma, S., Nishimura, S. & Ueda, I., 1989. Infrared spectra of phospholipid membranes: interfacial dehydration by volatile anesthetics and phase transition. *Biochim Biophys Acta,* Volume 1022, pp. 245-250.

U.S. Food and Drug Administration Report, 2016.

Ullman, C. G., Frigotto, L. & Cooley, R. N., 2011. In vitro methods for peptide display and their applications. *Briefings in Func Genomics,* 10(3), pp. 125-134.

Van Belle, T., Coppieters, K. & Von Herrath, M., 2011. Type 1 diabetes etiology; immunology and therapeutic strategies. *Physiological Rev,* Volume 91, pp. 79-118.

Van Oosten, B. & Harroun, T., 2016. A MARTINI extension for Pseudomonas aeruginosa PAO1 lipopolysaccharide. *J Mol Graph Model,* Volume 63, p. 125−133.

Walden, P., 1914. Molecular weights and electrical conductivity of several fused salts. *Bull Acad Imper Sci,* pp. 405-422.

Wang, J. et al., 2007. Prevalence of autoantibody-negative diabetes is not rare at all ages and increaees with older ager and obesity. *J Clin Endocrinology & Metabolism,* 92(1), pp. 88-92.

Wendt, A. et al., 2004. Glucose inhibition of glucagon secretion from rat alpha cells is mediated by GABA released from neighboring beta cells. *Diabetes,* Volume 53, pp. 1038-1045.

Wentzel, A., A, C., Adams, T. & Kolmar, H., 2001. Display of passenger proteins on the surface of Escherichia coli K-12 by the Enterohemorrhagic E. coli Intimin EaeA. *J Bacteriology,* Volume 283, p. 24.

Wenzlau, J. M. et al., 2007. The cation efflux transporter ZnT8 is a major autoantigen in human type 1 diabetes. *Proc Natl Acac Sci,* Volume 104, pp. 17050-17045.

Wenzlau, J. M. et al., 2010. Kinetics of the post-onset decline in zinc transporter 8 autoantibodies in type 1 diabetic human subjects. *J Clin Endocrinol Metab,* Volume 95, pp. 4712-4719.

Wielosz, E., Dryglewska, M. & Majdan, M., 2014. Serological profile of patients with systemic sclerosis. *Advances in Hygiene and Experimental Medicine,* Volume 68, pp. 987-991.

Wilkes, J. S. & Zaworotko, M. J., 1992. Air and water stable 1-ethyl-3-methylimidazolium based ionic liquids. *J Chem Soc, Chem Commun,* pp. 965-967.

Williams, A. C. & Barry, B. W., 2004. Penetration enhancers. *Adv Drug Deliv Rev,* Volume 56, p. 603.

Williams, H. D. et al., 2014. Ionic liquids provide unique opportunities for oral drug delivery: structure optimization and in vivo evidence of utility. *Chem Commun,* Volume 50, pp. 1688-1690.

Wong, C. Y., Martinez, J. & Dass, C. R., 2016. Oral delivery of insulin for treatment of diabetes: Status quo, challenges and opportunities. *J Pharm Pharmacol,* 68(9), pp. 1093-1108.

World Health Organization, 2016. Global report on diabetes.

Xie, C. et al., 2011. A novel multiplex assay combining autoantibodies plus PSA has potential implications for classification of prostate cancer from non-malignant cases. *J Translational Medicine,* 9(43).

Xu, Y. et al., 2014. Autoantibodies as potential biomarkers for the early detection of esophageal squamous cell carcinoma. *The American J Gastroenterology,* Volume 109, pp. 36-45.

Yano, T., Nakagawa, A., Tsuji, M. & Noda, K., 1986. *Life Sci,* Volume 39, p. 1043.

Yoo, B. et al., 2016. Molecular mechanisms of ionic liquid cytotoxicity probed by an integrated experimental and computational approach. *Sci Rep,* Volume 6, p. 19889.

Yu, Y. & Nie, Y., 2011. Toxicity and antimicrobial activities of ionic liquids with halogen anion. *J Environ Prot,* Volume 2, pp. 298-303.

Zakrewsky, M. et al., 2016. Choline and geranate deep eutectic solvent as a broad-spectrum antiseptic agent for preventative and therapeutic applications. *Adv Health Mat.*

Zakrewsky, M. et al., 2014. Ionic liquids as a class of materials for transdermal delivery and pathogen neutralization. *Proc Natl Acad Sci,* Volume 111, pp. 13313-13318.

Zeisel, S. H. & Costa, K. A. d., 2009. Choline: An essential nutrient for public health. *Nutr Rev,* 67(11), pp. 615-623.

Zhao, B. et al., 2015. Biocompatible deep eutectic solvents based on choline chloride: Characterization and application to the extraction of rutin from Sophora japonica. *ACS Sustain Chem Eng,* Volume 3, pp. 2746-2755.

Zhao, G. et al., 2013. Biofilms and inflammation in chronic wounds. *Adv Wound Care,* 2(7), pp. 389-399.

Zheng, W. et al., 2011. Effect of cation symmetry on the morphology and physicochemical propertires of imidazolium ionic liquids. *J Phys Chem B,* Volume 115, pp. 6572-6584.

Ziegler, A. G. et al., 2013. Seroconversion to multiple islet autoantibodies and risk of progression to diabetes in children. *JAMA,* Volume 23, pp. 2473-2479.

Zoumpopoulou, G. et al., 2010. Detection of changes in the cellular composition of Salmonella enterica serovar typhimurium in the presence of antimicrobial compound(s) of Lactobacillus strains using Fourier transpform spectroscopy. *Intl J Food Microbiol,* Volume 144, pp. 202-207.