The Dissertation Committee for Jason Robert Cantor Certifies that this is the approved version of the following dissertation:

Novel Strategies Towards Engineering Therapeutic Enzymes with Reduced Immunogenicity for Cancer Therapy

Committee:

George Georgiou, Supervisor

Brent L. Iverson

Jennifer A. Maynard

Charles B. Mullins

Christian P. Whitman

Novel Strategies Towards Engineering Therapeutic Enzymes with Reduced Immunogenicity for Cancer Therapy

by

Jason Robert Cantor, B.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2010

Dedication

To my loving parents,

Sheldon and Marlene Cantor.

Anything that I have accomplished to this point or hope to achieve in future endeavors would not have been possible without your endless encouragement and support, and I am forever grateful.

Acknowledgements

I feel remarkably fortunate to have had the opportunities that have brought me to this stage in my life and for the people that have helped shape my journey along the way.

I want to first recognize those individuals who have been especially instrumental in guiding my development as a scientist. I would like to thank Dr. Kelvin Lee for introducing me to the world of research by providing me with an opportunity to work in his lab as an undergrad at Cornell. I also want to thank Dr. Matt DeLisa for many useful discussions during my senior year that ultimately helped shape my decision for graduate school destination. Matt enthusiastically recommended I work in the lab from which he had just completed his own postdoctoral studies, and so I ultimately left behind the 10-month winters of Ithaca, NY for the 10-month summers of Austin, TX.

I will be forever grateful for the training I have received both in science and otherwise (like how to write an email in less than 2000 words) that I have received over the past 6+ years from my adviser, Dr. George Georgiou. He has provided a resourceful work environment, an opportunity to develop and flex my scientific imagination, and has shown great compassion and patience for my growth as a scientist and as a person. I also want to thank Dr. Brent Iverson for his perpetual enthusiasm in imparting both scientific knowledge and encouragement throughout my graduate career. I also want to express thanks to Dr. Thomas Forsthuber at UTSA for his willingness to collaborate with us and for our many useful conversations throughout the past year. I will still get you that Starbucks card before I leave the Texas state lines.

I would also like to acknowledge all the BIGG lab members I have had the pleasure of working beside and building friendships with during my time here, especially: Karl Griswold, Navin Varadarajan, Jamie Link, Yariv Mazor, Yorgos Skretas, Everett

Stone, Mark Pogson, Tae Hyeon Yoo, Karthik Veeravalli, Jack Borrok, and Sai Reddy. You have all offered invaluable scientific mentoring, encouragement during the bleak times, treasured friendships, and an environment where we were not afraid to tell jokes, karaoke to the radio, or grab a beer. And further to Karl: I extend my gratitude for your initial advice and training when I was first learning the ropes, even while you yourself were busy writing your dissertation and interviewing. I honestly couldn't have had a better-matched mentor.

I was also fortunate enough to make a number of close friends both in the UT department of chemical engineering and beyond, who have helped me retain my wits during the times I was able to get outside of MBB. Drew and Melissa Sexton, John McLees, Ryan Fitzpatrick, Paul Anderson, Brian Talarico, Adam Johnson, Greg Beatty, Saul and Melody Lee, Danielle Smith, Terry Farmer, and Jason and Niki Davis are among those I will certainly keep in touch with once I leave UT.

There is not a chance that I would be at this stage in my life without the love and encouragement of my family. I will never forget the sacrifices my parents made solely because they wanted the best for my brother and I, and they constantly went above and beyond the call of duty to guide us along the way. There's not enough space to write just how appreciative I truly am.

The final thank you is reserved for the love of my life, Kendra. You have put up with long hours spent in the lab, countless grumblings about experiments gone south, and living in the limbo of not knowing when I would finish or where we'll end up. You have seen me at both my best and worst for the past two and a half years, yet you have stood by me. I can't tell you how much that means or how excited I am to continue to build our life together.

Novel Strategies Towards Engineering Therapeutic Enzymes with Reduced Immunogenicity for Cancer Therapy

Publication No._____

Jason Robert Cantor, PhD

The University of Texas at Austin, 2010

Supervisor: George Georgiou

Heterologous enzymes have been investigated for a variety of therapeutic applications, including the treatment of a number of cancers that are sensitive to the systemic depletion of specific amino acids. One such example is acute lymphoblastic leukemia (ALL) for which enzyme-mediated L-Asparagine (L-Asn) depletion by the *Escherichia coli* L-Asparaginase II (EcAII) has been proven critical for treatment. However, the repeated or prolonged therapeutic administration of such enzymes is restricted by their immunogenicity, which frequently results in the generation of anti-enzyme antibodies that may in turn mediate a variety of adverse hypersensitivity reactions and neutralization of the enzymes themselves. Thus, while the therapeutic efficacy of asparaginase is well established, a significant number of patients still develop adverse immune responses to the enzyme. Here, we have developed and explored novel strategies towards engineering an asparaginase with reduced immunogenicity for ALL therapy.

First, we identified and investigated human enzymes that putatively shared functional similarity to asparaginase with the long-term aim of engineering such enzymes to acquire biochemical and pharmacological properties requisite for eventual therapeutic application. In one study, we described the bacterial expression and characterization of the human asparaginase-like protein 1 (hASRGL1). We presented evidence that hASRGL1 exhibited an activity profile consistent with enzymes previously designated as β-aspartyl peptidases, which had only been previously identified in plants and bacteria. Similar to non-mammalian β-aspartyl peptidases, hASRGL1 was revealed to be an N-terminal nucleophile (Ntn) hydrolase whereby Thr168 serves as the essential Ntn for both intramolecular processing and catalysis. In a second study, we described the optimized bacterial expression and biochemical characterization of the human N-terminal asparagine amidohydrolase 1 (hNTAN1). We demonstrated that hNTAN1 catalysis is dependent upon direct involvement of a thiol group, and subsequently identified Cys75 as an essential residue that may act as the catalytic nucleophile. Further, we presented the first description of hNTAN1 kinetics, secondary structure composition, and thermal stability.

Second, we devised and validated a novel therapeutic deimmunization approach by combinatorial T-cell epitope removal using neutral drift. We showed that combinatorial saturation mutagenesis coupled with a robust neutral drift screen enabled the isolation of engineered EcAII variants that contained multiple amino acid substitutions yet exhibited catalytic efficiencies nearly indistinguishable to that of the parent enzyme. Three regions of EcAII were computationally identified as putative T-cell epitopes and then subjected to saturation mutagenesis at 4 positions (per region) believed to be critical for MHC-II binding. The resulting libraries were then sequentially subjected

to a neutral drift FACS screen in order to isolate EcAII mutants that retained wild-type function. Pools of neutral drift variants were then computationally evaluated for MHC-II binding and those that displayed scores indicative of compromised binding were purified and biochemically characterized. Finally, T-cell activation assays and antibody titers in HLA-transgenic mice were used to evaluate T-cell epitope removal and immunogenicity, respectively. Ultimately, we revealed that mice immunized with an EcAII neutral variant containing 8 amino acid substitutions – 3 of which were non-phylogenetically conserved – within computationally predicted T-cell epitopes, displayed a significant 10-fold reduction in serum anti-EcAII IgG titer relative to mice similarly immunized with the parent enzyme.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction

1.1 THE PROMISE AND CHALLENGES OF PROTEIN THERAPEUTICS

The establishment and progression of recombinant DNA technologies over the past several decades have allowed for the development of a variety of protein therapeutics, which have already imparted a tremendous impact on both modern medicine and the pharmaceutical industry as a whole. As of 2008, the US Food and Drug Administration (FDA) had approved more than 130 different proteins for clinical administration; over 95 of which were produced via recombinant technologies, while many more are currently in development [1]. The emergence and continued increase in recombinant protein therapeutics is due in large part to the several advantages they offer relative to small-molecule drugs. Primarily, proteins can achieve an assortment of complex functions unmatched by simple small-molecules, and further, can do so with exquisite specificity in many cases, thus reducing their potential to interfere with normal biological processes. As such, proteins may sometimes provide the only viable treatment option for a diverse range of pathological states including various cancers, diabetes, and microbial infections. Moreover, recombinant DNA technologies afford the ability to modify a target protein towards novel or enhanced function, altered substrate specificity, increased expression yield, improved pharmacokinetic parameters, or any number of other biochemical and pharmacological properties for which protein variants may be interrogated.

Recently, Leader et al. [1] categorized this rapidly growing class of pharmaceuticals into four majors groups: (I) protein therapeutics with enzymatic or regulatory activity (e.g. enzymes, hormones, and cytokines); (II) protein therapeutics with special targeting activity (e.g. monoclonal antibodies and soluble receptors); (III) protein vaccines; and (IV) protein diagnostics. In total, biopharmaceuticals generated over $80B in 2008 sales and that total was expected to approach nearly $100B in 2010 [2]. Furthermore, it was recently estimated that by 2014, approximately half of the top 100 selling ethical drugs would be protein therapeutics.

Despite the therapeutic value realized for many proteins, potential therapeutics of this type have failed far more often than they have succeeded to date due to a number of challenges faced during their development and use with respect to both safety and clinical efficacy. To address these issues and better exploit the immense potential that protein therapies offer, several factors need to continue to be taken into consideration.

One concern is the choice of recombinant protein expression host, as the system must allow for purification and storage of the protein in a therapeutically active and stable form and in sufficient quantities. In some cases, protein stability and physiological function are dependent upon specific post-translational modifications such as glycosylation, phosphorylation, and proteolytic cleavage. Thus, while cost and time considerations ostensibly favor prokaryotic recombinant expression systems such as *Escherichia coli*, these hosts lack the machinery to appropriately modify a number of important protein therapeutics, such as full-length monoclonal antibodies. Conversely, mammalian cell lines that are capable of carrying out specific human-like post-translational modifications have several drawbacks including low protein yields, long fermentation times, and heterogeneous products. Recently, Hamilton et al. [3]

demonstrated a significant breakthrough by reporting the humanization of the glycosylation pathway in the yeast *Pichia pastoris* in order to secrete a human glyocoprotein with the appropriately desired N-glycosylation pattern.

Pharmacokinetic/pharmacodynamic (PK/PD) aspects such as solubility, stability, bioavailability, route of administration, and *in vivo* half-life may all significantly influence the successful application of a protein therapy. Proteins are large molecules with diverse and typically complex physicochemical properties that may dramatically influence their intracellular distribution or entry into systemic circulation. Further, they may be susceptible to degradation by proteolytic enzymes distributed throughout the body, or may be eliminated via an array of clearance mechanisms (e.g. renal filtration) [4]. For example, due to the presence of gastrointestinal proteases, most protein therapeutics cannot be orally administered. Instead, intravenous and subcutaneous are the two most frequently used routes of administration. In some cases, however, the biological efficacy and clinical safety of a protein therapeutic may be dependent upon selecting an optimal route of administration [5]. Of the strategies used to increase the *in vivo* half-life of protein therapeutics, chemical attachment of a polyethylene glycol (PEG) moiety (PEGylation) is one of the most common [6]. Besides increasing half-life, PEGylation sometimes reduces immunogenicity and increases the stability of the therapeutic payload by sterically shielding the conjugated protein from proteolysis. However, PEGylation may compromise protein function [7], may result in the accumulation of PEG hydrates that can interfere with normal glomerular filtration, and may elicit PEG-specific antibodies [8, 9] that adversely affect treatment. Another strategy that has been broadly explored for improving protein stability is microencapsulation using erythrocyte carriers.

While this approach has also shown promise, it too is encumbered by a number of disadvantages that have been summarized elsewhere [10].

More recently, a novel approach for addressing (and even tuning) protein half-life was demonstrated by Schellenberger et al. [11]. These authors used a genetic fusion of a long unstructured polypeptide to increase the hydrodynamic radius and thus improve the half-life of green fluorescent protein (GFP) in rats from between 3-fold to 12-fold (relative to unmodified GFP) depending upon the length of the fusion. Alternatively, others have focused on enhancing protein stability through strategies aimed at enhancing solubility via targeted mutagenesis of key amino acids implicated in aggregation [12, 13].

Repeated or prolonged administration of protein therapeutics, especially those of non-human origin, often leads to the induction of undesirable host anti-drug antibodies that in turn may mediate a variety of adverse effects including: hypersensitivity reactions, inactivation and clearance of the protein itself, and in severe cases anaphylactic shock [14]. Over the next several sections, the immensely critical challenge of overcoming protein therapeutic immunogenicity will be discussed in great detail.

1.2 THE IMMUNOGENICITY OF BIOPHARMACEUTICALS

Nearly all protein therapeutics – whether derived from human or non-human sources – elicit some level of host antibody response upon administration. Consequently, all exogenous proteins have the potential for immunogenicity [15]. The immune system has evolved to not react with self-antigens (autoimmunity) through a process referred to as tolerance [16], whereby lymphocytes that recognize self-antigens are depleted before they develop into fully immunocompetent cells.

The principle of tolerance to self-antigens by definition implicates that human-derived proteins should not elicit an immune response when administered to a patient. Nonetheless, immune responses to autologous proteins have been frequently observed. For example, factor VIII [17] and interferon (IFN)-β [18] can elicit antibody responses in patients lacking the ability to produce the respective protein (and therefore have not been subjected to tolerance) or as a result of the formation of neo-epitopes due to protein aggregates produced during formulation. Other human therapeutic proteins that have been observed to elicit immune responses include recombinant IFN-β [19], erythropoietin (EPO) [20], and fully human or 'humanized' monoclonal antibodies [21]. The immunogenicity of self-proteins must result from a breach of the immunological tolerance mechanisms, though the molecular details of this process likely vary from protein to protein and are not well understood.

Proteins of non-human origin generally elicit much higher immune responses relative to human protein drugs. Non-human proteins, including many enzymes and engineered variants of various other human proteins, commonly possess functional and/or pharmacological properties for which a human-derived counterpart is either inferior or unavailable altogether. However, the immunogenicity of heterologous proteins has limited their development and overall potential for use in a variety of therapeutic settings, especially when repeated administration is required. Therefore, the need for safe and effective non-human protein therapeutics is critical to the continued growth of the biopharmaceutical industry.

The need to address the immunogenicity of both native and engineered protein therapeutics has been a driving force towards the maturation of more sophisticated approaches to predict, detect, and ultimately limit immunogenicity while maintaining

requisite therapeutic functions. To better understand how these strategies have progressed, it may be useful to first provide an overview of the sequence of events that occur in the immune recognition of protein therapeutics.

1.2.1 Immune recognition

Heterologous proteins predictably exhibit considerable immunogenicity (i.e. recognized immunologically as non-self) more often than not. To elicit an antibody response, a protein therapeutic must typically interact with several types of immune cells, including antigen presenting cells (APCs), B-cells, and T-cells [22]. Each cell type recognizes different features of the exogenous protein, and therefore preventing or modifying these recognition events represents a means to significantly affect immunogenicity.

The sequence of events that lead to B-cell activation and the subsequent production of antibodies can be divided into either T-cell independent (Ti) or T-cell dependent (Td) scenarios. Briefly, Ti activation of B-cells occurs when structural features of an exogenous molecule, such as polymeric repeats, induce the direct activation of a subset of B-cells [23]. Antibodies produced by direct B-cell activation are mostly low affinity, high avidity IgM and limited low affinity IgG isotypes, as affinity maturation and class-switching require T-cell help [24]. However, while Ti activation of B-cells has been implicated for some protein therapeutics [25], the focus here will be on Td activation because the latter is much more significant for protein therapeutics and is associated with more potent antibody responses, antibody isotype switching, and the development of B-cell memory [24].

In order to induce a Td antibody response to a protein therapeutic, several events must be coordinated, usually in specialized regions of secondary lymphoid organs (e.g. spleen and lymph nodes) [26]. The first step in this process is antigen uptake by specialized APCs (e.g. dendritic cells and macrophages) via pinocytosis, receptor-mediated endocytosis, or phagocytosis. The efficiency of uptake varies significantly with the route of administration, the aggregation state of the protein, and its receptor-binding specificity. Intravenous injection will typically result in the antigen directly entering secondary lymphoid tissue via afferent lymphatics. Subcutaneous injection will likely activate Langerhans cells in the skin [27]. In both cases, the internalized antigen is processed intracellularly by a variety of proteases (e.g. cathepsins and endopeptidases) into short linear peptides that bind to major histocompatability complex (MHC)-II molecules and are then presented in complex on the surface of the APC. (Briefly, we note here that the MHC is referred to as the human leukocyte antigen (HLA) complex in humans.) These peptide fragments (13-25 amino acids) bound to MHC-II are recognized by CD4+ T-cells through interactions with the T-cell receptor (TCR), and thus these peptide fragments are referred to as T-cell epitopes. In addition to the TCR interaction with this MHC-II epitope complex (referred to here as T-cell signal 1, T1), T-cell activation further requires co-stimulation through engagement of CD28 via CD80 (referred to here as T-cell signal 2, T2), which are expressed in large amounts on the surface of APCs. Once fully activated, these T-cells proliferate and produce an array of cytokines. Specific encounter of the same antigen by naïve B-cells via interaction with IgM and IgD receptors on the B-cell surface results in similar processing and presentation in the context of MHC-II at the B-cell surface (referred to here as B-cell signal 1, B1). Analogous engagement of CD28/CD80 (referred to here as B-cell signal 2, B2) results in

the secretion of cytokines by the T-cell and engagement of CD40 with CD40 ligands (referred to here as B-cell signal 3, B3). These additional co-stimulatory actions allow the B-cell to mature into either an IgG-secreting plasma cell or a memory B-cell (referred to here as B-cell signal 4, B4). The affinity-matured, secreted antibodies may then bind to conformational regions of the whole antigen, which are thus referred to as B-cell epitopes. The sequence of events for Td activation of B-cells is summarized in **Figure 1.1**.

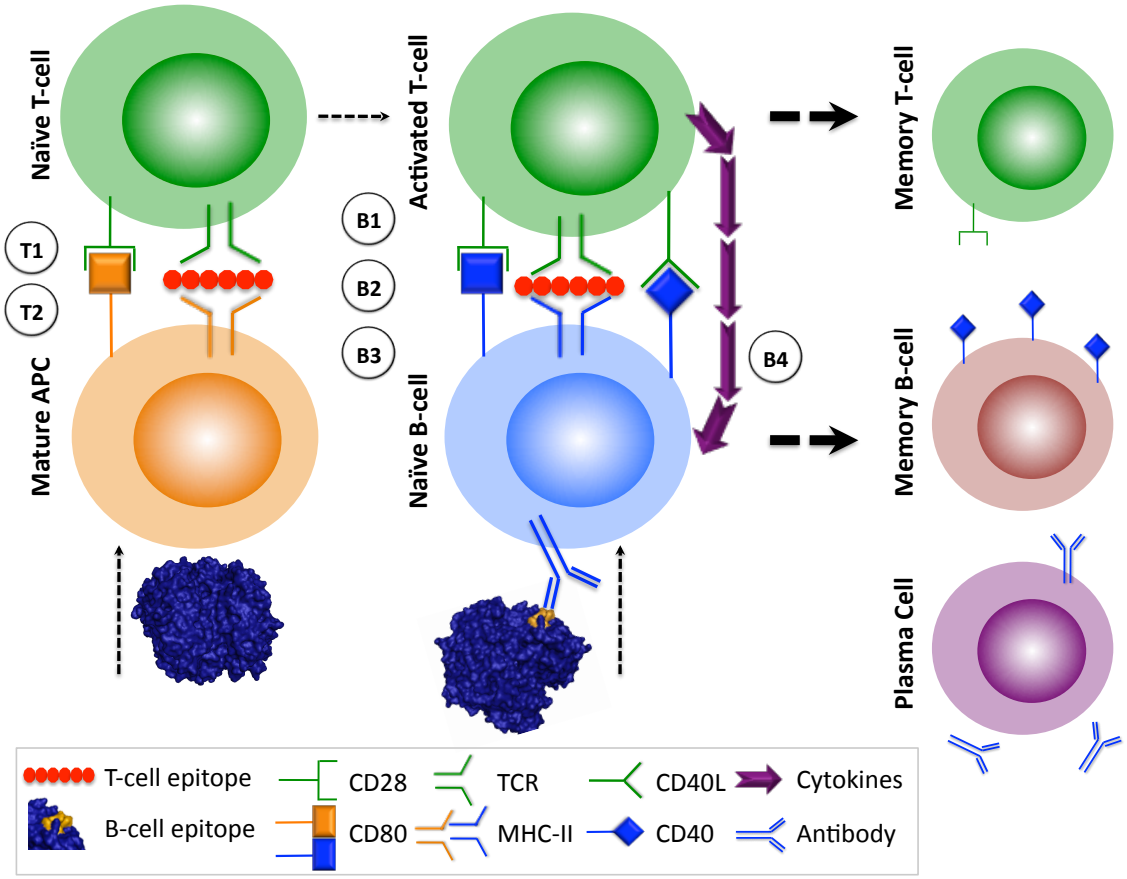


Figure 1.1 Activation of CD4+ T-cells and Td antibody response. The first step is non-specific antigen uptake by a professional antigen-presenting cell (APC) such as a dendritic cell. The mature APC processes the antigen into peptides, which are then presented to naïve CD4+ T-cells (T-helper; T_H) in complex with MHC-II on the surface of the APC (T1). To fully activate the T-cell, this recognition even must be accompanied by the additional co-stimulatory interaction between CD80 and CD28 (T2). An interaction between B-cell surface-bound IgM (and/or IgD) receptors and antigen is required to initiate activation of the naïve B-cell (B1). The B-cell then internalizes the antigen-Ig complex (specific uptake) and presents processed peptides in the context of MHC-II (analogous to other APCs) to T-cells. An additional co-stimulatory interaction between CD80/CD28 (B2) results in the secretion of cytokines from the T-cells and subsequently, the increased expression and engagement of CD40/CD40L (B3). Ultimately, the sum of these synergistic signals allow the B-cell to mature and become either a dedicated, antibody-producing plasma cell or memory B-cell (B4).

A number of other factors may also influence the immunogenicity of protein therapeutics, including: dose, frequency and route of administration, formulation (e.g. the propensity to aggregate under specific storage conditions), and contaminants that may arise from downstream manufacturing and processing. Furthermore, immunogenicity can be affected by a number of patient-specific characteristics, which may include: immunocompetence, state of tolerance to a specific antigen, allelic variations (e.g. MHC-II haplotype), and genetic deficiencies. Each of these extrinsic factors have been reviewed in detail elsewhere [14] and – given the scope of the work described in later chapters – will not be discussed further here.

1.2.2 Prediction and assessment of immunogenicity

Several approaches can be used to assess the potential immunogenicity of a protein therapeutic and thereby attempt to minimize, or ideally eliminate, the risk of neutralizing antibodies and the adverse effects that they in turn mediate. Here, three types of prediction and assessment approaches will be briefly described: *in silico*, *in vitro,* and *in vivo*.

In silico methods for immunogenic epitope prediction

As described in the previous section, TCRs on CD4+ T-cells recognize antigenic peptides presented in complex with MHC-II molecules on the surface of APCs. The general architecture of the MHC-II peptide-binding grooves is maintained both within and between different isotypes and species [28]. Consequently, the peptide backbone conformation of T-cell epitopes is largely sequence-independent and shares essentially no relationship to its conformation in the context of the whole protein. The MHC-II binding groove contains four well-defined pockets that accommodate the side chains of the P1, P4, P6, and P9 residues within a core 9mer region of the T-cell epitope and these key residues largely determine binding affinity and specificity [28] (**Figure 1.2**). Accordingly, a number of MHC binding motif-based computational methods have been developed and evaluated [29] over the past several years to rapidly screen protein sequences for the presence of potential MHC-II binding peptide fragments, in large part owing to the establishment of numerous MHC-II-epitope databases.

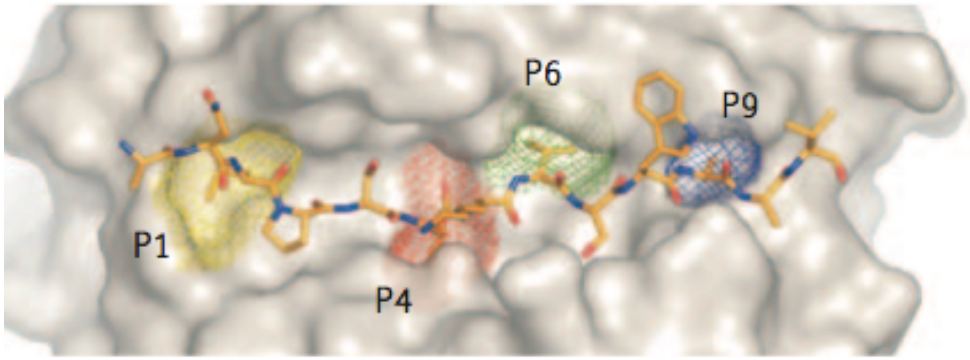


Figure 1.2 Depiction of MHC-II peptide binding groove. The solvent-accessible surface of HLA-DQ6.2 is shown with a bound hypocretin peptide, for which the key anchor positions (P1, P4, P6, P9) are highlighted. This figure is reproduced from Jones, E.Y., et al., *MHC class II proteins and disease: a structural perspective*. Nat Rev Immunol, 2006. **6**(4): p. 271-82 with permission from the Nature Publishing Group.

Each of these methods can screen large numbers of sequences for putative epitopes. However, the use of *in silico* methods typically results in overprediction, i.e. a number of non-binding peptides are predicted to bind (false positives). Conversely, a number of true binders may in fact be predicted as non-binders (false negatives). Such uncertainties in T-cell epitope prediction are likely attributable to the complex relationship between antigen processing, MHC-II binding, and TCR recognition and signaling [30]. Despite these shortcomings, *in silico* prediction models have been successfully applied both as described in Chapter 4 and elsewhere [31, 32]. Nonetheless, computational methodologies should continue to become more sophisticated as the collection of additional high quality training data becomes available.

B-cell epitope mapping *in silico* is not nearly as developed as the T-cell epitope prediction algorithms described above, and further studies will be necessary to determine whether computational prediction of epitopes recognized by the antibody repertoire is

plausible at all [33]. The complexity in predicting B-cell epitopes is in large part due to their predominantly conformational, and thus discontinuous nature. Consequently, such algorithms must rely on hydrophilicity or chain flexibility scales for example, which to date have performed inadequately.

In vitro methods for immunogenic epitope prediction

Many of the limitations of *in silico* methods described above can be overcome by directly measuring the binding of putative T-cell epitope peptides to MHC-II or by detecting antigen-specific activation of T-cells using a variety of in vitro culture systems.

HLA binding assays using soluble MHC-II molecules, for example, may be used to evaluate the affinity of a given peptide sequence for HLA alleles through quantitative competition with a fluorescently labeled MHC ligand [34]. However, this technique is somewhat low-throughput, is limited by the number of alleles that can be tested, and requires the purification of MHC-II molecules via a complex procedure that requires culturing large volumes of homozygous B lymphoblastoid cell lines [35].

In another approach, T-cells may be isolated from human blood (peripheral blood mononuclear cells, PBMCs) and subsequently stimulated with either whole protein, or using an overlapping library of peptides spanning the length of the same protein, in order to evaluate T-cell response and locate T-cell epitopes. This methodology has been successfully applied for a number of protein therapeutics including factor VIII [36], EPO [37], and IFN-α [38].

Protein-specific T-cell responses may be monitored via a variety of methods. One commonly used approach is monitoring the dilution of the fluorescent dye

carboxyfluorescein diacetate succinimidyl ester (CFSE) across successive cell generations [39] using fluorescence activated cell sorting (FACS). More commonly, T-cell proliferation is monitored by quantifying the release of specific cytokines (e.g. IL-2 or IFN-γ) as measured by standardized ELISPOT [40]. The primary advantage of the latter method is that additional information can be gathered about the nature of the T-cell response with respect to localization of key immunogenic epitopes. However, these types of studies require large sample populations in order to ensure that a diverse range of MHC-II haplotypes (representative of human populations) are being investigated. Briefly, the loci constituting the MHC are highly polymorphic, meaning that many alternative forms of the gene, or alleles, exist at each locus among the population. The genes of the MHC loci lie close together and therefore most individuals inherit the alleles encoded by these closely linked loci as two sets, one from each parent. Each set of alleles is referred to as a haplotype. Each heterozygous individual will thus have two MHC haplotypes, one in each chromosome (one of paternal origin and one of maternal origin).

In general, strictly *in vitro* approaches for T-cell epitope prediction are faced with the limitation that not every peptide capable of binding to MHC-II will be a T-cell epitope [37]. Similar to computational prediction methods, *in vitro* identification techniques are again limited by the multifaceted relationship between antigen processing, MHC-II binding, and TCR recognition and signaling. Therefore, direct peptide-MHC-II binding (or peptide challenge of isolated T-cells) cannot account for the possibility that certain test peptides may not be recognized by the TCR following formation of the MHC-II-peptide complex, or whether such peptides are generated at all during antigenic processing. Further, even overlapping panels of peptides may not overcome the potential for artificial truncation of certain T-cell epitopes.

In addition to the identification of T-cell epitopes, efforts to develop approaches to accurately identify antigenic B-cell epitopes are also in progress. To date, the ability to identify B-cell epitopes through *in vitro* approaches has been challenging due in large part to the same inherent limitations described earlier in the context of *in silico* B-cell epitope prediction. Recently however, Nagata and Pastan [41] described and validated a novel B-cell epitope identification approach called immune complex capture ELISA (ICC-ELISA) which interrogates antigen-antibody complexes in a conformation-dependent manner. In this manner, the authors were able to identify putative B-cell epitopes of *Pseudomonas* exotoxin A 38 (PE38) by measuring the antibody binding affinity differences between the wild-type PE38 and variants of the protein in which mutations were introduced into highly exposed amino acids in the protein crystal structure.

In vivo methods for immunogenic epitope prediction

In general, animal models are not suitable for the evaluation of protein therapeutics because their MHC-II molecules have significantly different peptide binding preferences. Consequently, peptide presentation observed in these animal models may be significantly different from that which would be expected for humans [42]. For this reason, mice – the most common experimental animal model for these purposes – exhibit different immune responses to antigenic challenges compared to humans [43]. To address this problem, researchers have constructed a number of transgenic mouse strains that express common human HLA molecules. T-cell responses in these mice correlate well with T-cell responses observed in infected or vaccinated humans [44-47]. Protein-specific

T-cell responses in such mice can be identified by isolating lymphocytes and evaluating their proliferation to antigen stimulation *in vitro* using standard methods [48, 49].

1.2.3 Approaches to reduce immunogenicity

A number of methods have been developed for reducing protein therapeutic immunogenicity. The optimization of extrinsic factors such as those described in section 1.2.1, e.g. more rigorous evaluation of protein handling during downstream storage and manufacturing [50-52], can have a significant effect on protein immunogenicity. Additionally, strategies for engineering the protein sequence to remove epitopes likely to elicit an adverse immune response have been developed and will be discussed here in some detail.

The most intuitively obvious means to overcome the immunogenicity of heterologous proteins in particular is to replace them with proteins from human sources or at least chimeras containing primarily human sequences. Such proteins would be expected to have already undergone tolerance induction (at least partially so in the case of protein chimeras) and thus elicit far less drastic, if any, adverse antibody responses. Perhaps the best example of this approach is the history of insulin in the treatment of diabetes mellitus type I (DM-1) and type II (DM-2). In 1922, insulin was first purified from bovine for life-saving daily injection for DM-1 patients [53], but problems including the immunological reaction triggered by the non-human protein limited its widespread use. However, with the advent of recombinant DNA technology, human insulin could be produced recombinantly in *E. coli* in large quantities. Recombinant human insulin was less immunogenic as might be expected, and in fact, human insulin

was the first commercial recombinant protein therapeutic, approved by the US FDA in 1982. [54].

For therapeutic antibodies, for which both structure and function are well understood, immunogenicity reduction has been focused primarily on the strategy of humanization whereby key functional residues from a heterologous antibody (e.g. murine) are grafted onto a human antibody framework [55] in order to simultaneously maintain antigen binding while eliminating a significant fraction of the non-human amino acid sequence. The efficacy of this approach is manifested by the plethora of both FDA-approved chimeric and humanized antibodies [56]. While these approaches have been largely successful, 'humanized' antibodies may possess residual immunogenicity because they still contain segments of non-human antibodies. Thus, methods for the generation of fully human antibodies have been developed, for example using mice encoding the human antibody repertoire [57], or via the cloning of antibodies from vaccinated patients [58].

In some cases, in fact, efforts have been made to engineer human proteins with functionality similar to that of a heterologous therapeutic. Recently, Stone et al. [59] demonstrated that the catalytic efficiency of the human enzyme L-Arginase I could be improved 10-fold simply by replacing the two Mn^{2+} ions normally present in the enzyme with Co^{2+}. This modified human arginase was shown to have cytotoxicity identical to that of the bacterial arginine deiminase, an enzyme in phase II clinical trials for a number of cancers, which however has been shown to potently elicit the generation of neutralizing antibodies. Nevertheless, many heterologous proteins do not have an obvious human-derived counterpart that possesses – or can be engineered to attain – functional and/or pharmacological properties necessary for successful therapeutic administration.

Another approach for reducing immunogenicity is the masking of potentially immunogenic epitopes via covalent conjugation with polyethylene glycol (PEG) [60]. As mentioned in section 1.1, PEG also acts to increase the serum half-life of its protein payload. PEGylation has been successfully utilized for a number of approved therapeutic proteins, including adenosine deaminase, IFN-α, and asparaginase, among others [1]. However, as described in section 1.1, PEGylation is still associated with several drawbacks including potential compromise of protein function [7] or the induction of adverse PEG-specific antibodies [8, 9].

Alternatively, protein immunogenicity can be ameliorated by mutating sequences likely to be recognized as either B-cell epitopes or T-cell epitopes (see section 1.2.1). However, as highlighted in section 1.2.2, the identification, and thus elimination, of B-cell epitopes is exceedingly difficult due to the lack of accurate and robust methods to identify these immunogenic regions, in large part because of their conformational and often discontinuous nature. The identification of B-cell epitopes is further complicated by our incomplete knowledge of naïve antibody repertoires and how these might vary across different human populations. Further, because the human antibody response is typically polyclonal, identification of antigenic epitopes has an added level of complexity given the need to distinguish (and account for) the unique specificities of any number of individual antibodies within the polyclonal pool.

In contrast, the disruption of T-cell epitopes by site-directed mutagenesis, aimed at compromising MHC-II binding in order reduce immunogenic responses, has been attempted for a number of therapeutic protein candidates including staphylokinase [61], IFN-α2b [62], IFN-β [63], EPO [37], factor VIII [36], and β-lactamase [64]. Further, T-cell epitope removal has shown promise in reducing the immunogenicity of humanized

and chimeric antibodies [65-67]. For nearly all of these examples, T-cell epitopes were initially identified using T-cells pools isolated from either human blood or from non-transgenic mice [63]. Typically, alanine scanning was used to guide the incorporation of amino acid substitutions at key anchor MHC-II binding positions to reduce binding affinity. In more sophisticated approaches, structural models and/or sequence homology to protein homologues were used to aid more plausible substitutions aimed at not disrupting protein folding and function.

It is worth noting here that the classification of this strategy as T-cell epitope removal (and similarly, the classification of methods described in 1.2.2 as being for T-cell epitope identification) is not completely accurate despite the generally accepted designation. By definition, the residues that are predicted or identified as being critical for MHC-II binding in fact comprise what is referred to as the MHC 'agretope', which is distinct from the set of residues that comprise the TCR binding determinant – the T-cell epitope [22] (**Figure 1.2**). To date, the capability to identify (or predict) the actual T-cell epitopes of protein therapeutic candidates has been lacking despite its potential utility. However, for clarity, the general approach described here will continue to be referred to as T-cell epitope mutagenesis, so as not to be interpreted discretely relative to the literature to date.

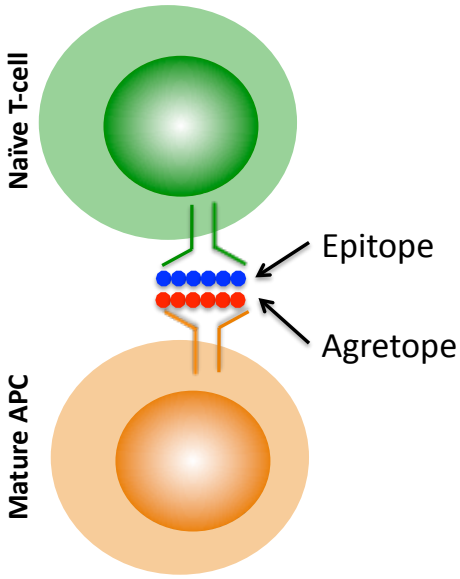


Figure 1.2 Distinction between T-cell epitope and MHC-II agretope. The mature APC processes antigen into peptides, which are then presented to naïve CD4+ T-cells in complex with MHC-II on the surface of the APC. The epitope (blue circles) is defined as the peptide region recognized by the T-cell receptor (TCR); the agretope (red circles) is defined as the peptide region that binds to the MHC-II molecules.

In general, the two primary factors that must be considered in attempting to disrupt a protein therapeutic epitope are: (i) the possibility that novel epitopes are inadvertently created, and (ii) that modifications made must not disrupt the biological function of the protein. The former can be evaluated using HLA-transgenic mice, while the latter significantly complicates the engineering of proteins for which extensive reengineering may be necessary to compromise MHC-II binding across multiple T-cell epitopes.

1.3 L-ASPARAGINASE

L-Asparaginase (L-Asparagine amidohydrolase EC 3.5.1.1) has been used principally in the treatment of acute lymphoblastic leukemia (ALL) for over 40 years [68-75]. Its therapeutic mechanism has been attributed to the exploitation of a specific metabolic discrepancy between normal human cells and certain leukemic cells [76]. Whereas normal cells are capable of producing the nonessential amino acid L-Asn via the enzyme L-Asparagine synthetase (AS), leukemic cells lack or express very low levels of AS and therefore require uptake of L-Asn from serum for cellular proliferation [77, 78]. L-Asparaginase catalyzes the hydrolysis of L-Asn to L-aspartic acid (L-Asp) and ammonia, resulting in the systemic depletion of 'tumor-essential' L-Asn from the serum [79-82], which in turn induces apoptosis of ALL lymphoblasts [83-85].

The chemotherapeutic potential of the enzyme initially stemmed from key observations and discoveries by Kidd [86] and Broome [68] in the 1950s and 1960s, which revealed that antilymphoma activity detected in guinea pig sera was due to asparaginase. Just a few years later, Campbell, Mashburn, and Wriston demonstrated that asparaginase isolated from *E. coli* exhibited antineoplastic activity similar to that found in the guinea pig sera [87, 88]. This key finding provided a practical source for the production of large quantities of the enzyme for preclinical and clinical investigations, ultimately resulting in the FDA-approval of the native *E. coli* asparaginase (ELSPAR) in 1978 as a component of a multi-agent chemotherapeutic regimen for the treatment of patients with ALL.

Since that time, asparaginase has been isolated and characterized from various microorganisms, including many gram-negative bacteria, mycobacteria, yeasts, and even plants and the plasma of certain vertebrates [89, 90]. However, not all isolated enzymes have been shown to have useful antitumor activity. In fact, asparaginases are generally classified as either bacterial-type or plant-type on the basis of sequence homology, structural features, and other biochemical properties [91].

Plant-type asparaginases share a high degree of amino acid sequence similarity (60-70%) with the aspartylglucosaminidase (AGA) (EC 3.5.1.26) family [91], with both enzyme groups belonging to the N-terminal nucleophile (Ntn) hydrolase protein superfamily [92-97]. Plant-type asparaginases from plant [96], cyanobacteria [97], bacteria [97], and human [98] (see Chapter 2) display no hydrolytic activity toward N^4-(β-*N*-acetyl-D-glucosaminyl)-L-asparagine (GlcNAc-L-Asn), despite their high degree of sequence similarity to AGAs.

Bacterial-type asparaginases can be further subdivided into types I and II, as defined by characteristics such as cellular localization, K_M values, and oligomeric form. Descriptions of the distinct isozymes expressed in *E. coli* can be used to illustrate the distinguishing features between the two subtypes. Type I *E. coli* asparaginase (EcAI) is cytoplasmic, is expressed constitutively, functions as a homodimer [99], and has been characterized to have a relatively high L-Asn K_M value ($K_M \sim 10^{-3}$ M), whereas type II *E. coli* asparaginase (EcAII) is periplasmic, is expressed primarily under anaerobic conditions, functions as a heterodimer, and has a much lower K_M value with L-Asn ($K_M \sim 10^{-5}$ M). However, Campbell and Mashburn found that only the high affinity type II enzyme has antitumor activity [88]. Recently, Hejazi et al. showed that *E. coli* also harbors a gene that encodes a plant-type asparaginase (EcAIII) [97] though its exact

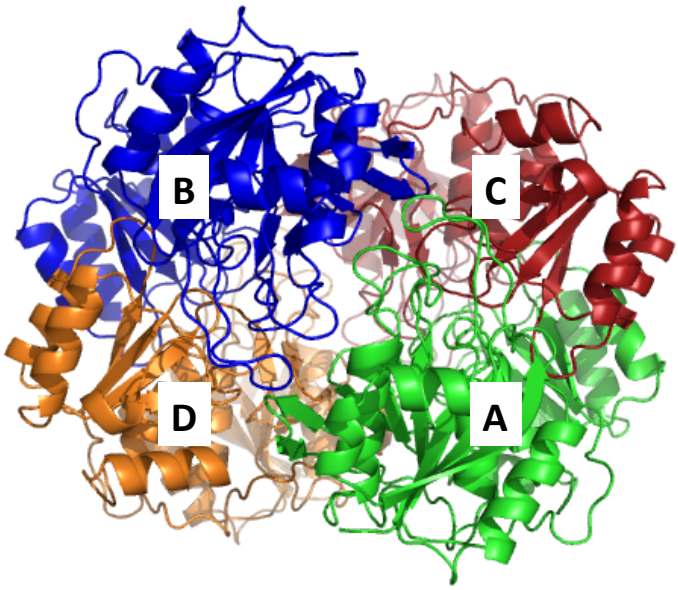
physiological role is unknown. Additional type II asparaginases with antitumor activity have been isolated from a number of bacterial sources (e.g. *Wolinella succinogens* [100] and *Serratia marcescens* [101]), however; clinically available asparaginase is derived from only two sources: *E. coli* and *Erwinia caratovora*.

1.3.1 Biochemistry of chemotherapeutic asparaginase

L-Asparaginase II from derived from *E. coli* (EcAII) is the FDA-approved clinical benchmark for the treatment of ALL and is available in both native and PEGylated formulations [102].

EcAII (PDB code 3ECA) is a tetramer with molecular weight ~140kDa, comprised of four identical 326 residue subunits and four non-cooperative active sites (**Figure 1.3A**). According to the nomenclature introduced for EcAII [103], the four subunits are labeled A, B, C, and D. There is a single disulfide bond in each subunit, connecting Cys77 to Cys105. Two active sites are formed at the dimerization interface between the A and C subunits, while the remaining two are formed at the interface between the B and D subunits. While the enzyme should be enzymatically competent as a dimer, the formation of the tetramer has been demonstrated as essential for catalytic activity. Sanches et al. [104] have suggested that the role of the tetramer formation is to guarantee proper folding of the enzyme, as it keeps the hydrophobic portions of the protein buried inside its core, away from solvent contact, thus providing stabilization to the structure.

(A)



(B)

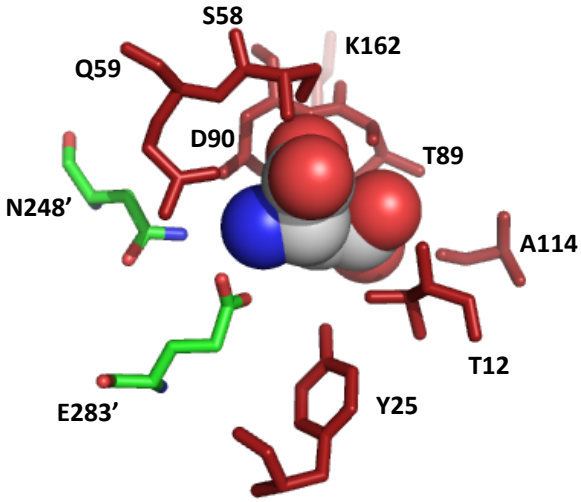


Figure 1.3 (A) Crystal structure of *E. coli* L-asparaginase II (PDB 3ECA) showing the nomenclature of the 4 subunits. (B) One of 4 non-cooperative EcAII active sites. In this case, one of the active sites formed at the A/C interface is shown. Residues colored red and numbered without prime designation are located in the C subunit, while residues colored green and numbered with prime designation are located in the A subunit. Bound L-Asp product is shown as spheres. Images were created using PyMol [105]

The location and identification of the EcAII active site residues was elucidated in studies that employed the use of binding ligands such as L-Asp [103, 104]. The binding pocket involves residues of both subunits of a given intimate dimer (e.g. subunits A and C). Residues Thr12, Tyr25, Ser58, Gln59, Thr89, Asp90, Ala114, and Lys162 from one subunit, and Asn248 and Glu283 from the other, form the binding pocket (**Figure 1.3B**).

The currently accepted mechanism for the hydrolysis of L-Asn into L-Asp and ammonia by EcAII is via a covalently bound intermediate involving a β-aspartyl intermediate, thus implicating a double displacement (or 'ping pong') mechanism [106] (**Figure 1.4**). In the first step, a nucleophilic group of the enzyme attacks the C-γ of L-Asn leading to a tetrahedral intermediate which is subsequently broken down to form an acyl-enzyme intermediate with EcAII covalently bound to the substrate and results in the elimination of ammonia. In the subsequent step, the intermediate is attacked by a second nucleophile (typically water), leading to the hydrolysis of the intermediate, yielding L-Asp and free enzyme. Given the identification of residues located at the active site, it has been proposed that the hydrolysis carried out by EcAII is mechanistically similar to that of serine proteases, whose activity is dependent upon a specific 'catalytic triad'. It has been shown, however, that none of the three histidines in EcAII are required for catalysis [107], therefore eliminating the possibility of a classical Ser-His-Asp catalytic triad. Studies have revealed that mutation of either Thr12 [108] or Thr89 [109] diminishes enzymatic activity to < 0.01% relative to the parent enzyme. A synopsis of the efforts and various hypotheses made over the past two decades in attempting to establish the key mechanistic residues involved in the hydrolysis carried out by EcAII, as well as a current postulation for the presence of two catalytic triads (with either Thr12 or Thr89 acting as

the attacking nucleophile) involved uniquely in the initial acylation and deacylation steps, respectively, has been described extensively by Sanches et al. [106].

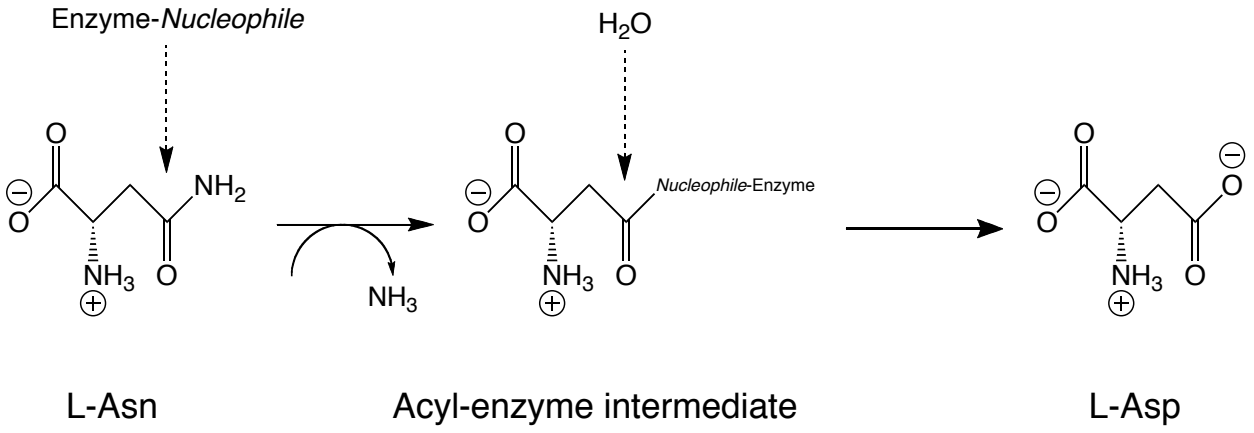


Figure 1.4. Schematic of the reaction catalyzed by L-asparaginases.

EcAII also possess a promiscuous glutaminase activity, in that it is able to catalyze the analogous hydrolysis of L-Gln to L-Glu and ammonia. However, the maximal rate of L-Gln hydrolysis is ~3% of that for L-Asn (K_M (L-Gln) = 3mM) [90]. While L-Gln depletion may potentially enhance the therapeutic activity of EcAII [110], it has also been suggested to contribute to additional clinical toxicities associated with the enzyme [111]. Thus, the optimal level (and role) of L-Gln hydrolysis by EcAII remains unclear.

1.3.2 Clinical profile and immunogenicity of L-Asparaginase

ALL is a heterogeneous group of disorders that result from the clonal proliferation and expansion of malignant lymphoid cells in bone marrow, blood, and other organs [112]. Nearly 4000 cases of ALL are diagnosed annually in the US, approximately two-thirds of which are in children and adolescents, making ALL the most common cancer in these age groups [75]. Pui et al. [75] reported that in the 1990s, the overall cure rate for childhood ALL was approximately 80% in developed countries, thanks in part to the continued optimization in the usage of various antileukemic agents and elucidation of various prognostic factors for risk-directed therapy. Unfortunately, treatment of adult ALL has been far less successful, though overall cure rates have improved to nearly 40%. The poor outcome in adult ALL has been variously attributed to increased drug resistance, poorer tolerance of treatment, reluctance to accept certain temporary toxicities, and less effective treatment regimens.

The role of L-Asparaginase as a potent antileukemic agent for the treatment of ALL has been well established. Though initially evaluated for its clinical efficacy as a single agent [113], asparaginase has since become a standard component of combination chemotherapy treatment of ALL, typically administered along with some combination of a glucocorticoid, vincristine, and methotrexate. Today, asparaginase is clinically available in three preparations. The native and PEGylated formulations of EcAII are both FDA-approved for first-line treatment of ALL, while the native preparation from *Er. caratovora* (ErA) is typically used in the USA only for patients allergic to the *E. coli* enzyme as it is non-cross-reactive with anti-EcAII [90]. However, besides being

immunogenic itself, Duval et al. [114] demonstrated that ErA is also clinically inferior to EcAII with respect to both event-free survival and overall survival rates at 6 years.

The toxicities of asparaginase include those related to the inhibition of protein synthesis, such as: (i) liver dysfunction resulting from decreased levels of serum albumin, fibrinogen, and lipoprotein levels, and (ii) coagulation abnormalities caused by low levels of certain clotting factors [115]. Additional toxicities that may result from asparaginase treatment include: (i) acute pancreatitis, (ii) nonspecific gastrointestinal toxicity, and (iii) neurotoxicity [90].

However, the most prominent limitation associated with asparaginase as a protein therapeutic is its immunogenicity, as the enzyme has been shown to elicit adverse antibody responses in a large fraction of patients. The development of antibodies against EcAII (both formulations) can reduce its activity, leading to failure of L-Asn depletion upon re-administration of the drug. The reported frequency of anti-EcAII against the native enzyme is as high as 60-70% in children and 79% in adults, while the development of anti-EcAII against the PEGylated enzyme has been reported in up to 12% of EcAII-naïve children and in 4-15% of EcAII-naïve adults [116]. Armstrong et al. [9] recently revealed that treatment with the PEGylated enzyme could elicit neutralizing anti-PEG antibodies as well. The development of adverse antibody responses has also been shown to mediate a wide variety of clinical hypersensitivity reactions that range from mild local allergic reactions to full blown anaphylaxis [115]. Moreover, Asselin et al. [117] showed in a study nearly 20 years ago that the serum half-lives for EcAII preparations in patients who had a previous hypersensitivity reaction to EcAII were significantly decreased relative to EcAII-naïve patients: (i) 1.82 days versus 5.73 days for PEG-EcAII, and (ii) undetectable (enzyme cleared too quickly) versus 1.28 days for native EcAII. Ultimately,

though PEGylation has proven to reduce immunogenicity and prolong serum persistence relative to native EcAII, the induction of antibodies that contribute to both neutralization and clinical hypersensitivity reactions are still elicited towards PEG-EcAII in a significant fraction of patients.

Asparaginase displaying reduced immunogenicity could prove critical in the safer and more effective treatment of both newly diagnosed and relapsed ALL patients. Interest in this area is evidenced by the continued efforts towards the isolation and characterization of alternative bacterial asparaginases [118], as well as ongoing investigations into alternative strategies to protect asparaginase from immune surveillance [119]. However, attempts to investigate human sequences that putatively share functional similarity to asparaginase have been lacking. Such sequences may ultimately represent useful enzyme engineering targets for replacement of the bacterial, immunogenic benchmark therapeutic, given that they will have likely already undergone tolerance induction and consequently have a much lower propensity to elicit adverse antibody responses. In Chapters 2 and 3, we detail the recombinant expression and subsequent biochemical characterization of two such human enzymes for which experimental evidence of their putative functions had been previously lacking.

1.4 ENZYME ENGINEERING BY NEUTRAL DRIFT

Enzyme engineering describes the process of altering the protein structure of an existing enzyme, typically by modifying the amino acid sequence, in order to generate enzymes with improved or novel properties. Over the past decade, it has become a powerful tool for tailoring various enzyme properties in large part due to the enormous impact imparted by advancements made in bioinformatics, protein chemistry, molecular biology, and recombinant DNA technologies. Examples of the successful application of enzyme engineering to date have been extensively reviewed elsewhere [120-122]. Kazlauskas and Bornscheur [123] recently summarized and highlighted some of the main considerations and strategies applied in enzyme engineering today. One objective of the work detailed in this dissertation (Chapter 4) was to reduce immunogenicity of the benchmark therapeutic EcAII by T-cell epitope removal through adaptation of a novel directed evolution strategy referred to as neutral drift.

The primary focus of molecular and experimental evolutionary studies to date has been adaptive evolution, or specifically, the acquisition of new protein functions [124]. However, it has become clear that only a fraction of mutations that accumulate in most naturally evolving proteins are driven by selection for an entirely new function [125]. Instead, most natural genetic diversity is the outcome of neutral, non-adaptive drifts, i.e., the gradual accumulation of sequence changes that occur under selection to maintain the function and structure of the parent protein [124]. Here, neutrality refers to the notion that numerous genotypes (sequence redundancy) can accommodate the same phenotype, resulting in what is referred to as neutral networks. Each sequence in a neutral network

has accumulated mutations under the constraint that they do not affect the parent protein's existing function, and thus appear as neutral under these non-adaptive conditions. However, these mutations could still facilitate future adaptations under different constraints by altering latent promiscuous functions.

Neutral drifts of two different enzymes, paraoxonase 1 (PON1) [126] and cytochrome P450 [125], were performed by applying iterative rounds of mutagenesis and selecting for each enzyme's native (or primary) activity. These studies demonstrated how the potential for adaptation, and thus new functions, dramatically increases when the neutral network of a protein expands. Almost half of the 311 neutral PON1 variants characterized exhibited significant changes in promiscuous activities, specificities, or inhibition, and further, a number of such variants with just one to two mutations could mediate a new phenotype. Similar variations in promiscuous activities were observed in the P450 drift as well.

Bershstein et al. [124] later performed a neutral drift of TEM-1 β-lactamase to evaluate how the fitness (activity, stability, etc.) of a protein changes as mutations gradually accumulate. Sequence analysis of the drifting libraries indicated that mutations that act as 'global suppressors' were enriched during the drift. In other words, such mutations suppressed the effect of a broad range of destabilizing mutations, and thereby dramatically enhanced TEM-1 evolvability, i.e. afforded a higher number of mutations to accumulate without deleterious consequences, and increased the probability of a new function emerging [127]. The experimental proof that neutral drifts can prompt the emergence of mutational robustness, predicted previously by van Nimwegen et al. [128], yields the interesting notion that the incorporation of stabilizing residues via application of a neutral drift may allow for the generation of protein variants that comprise highly

evolvable starting points in other directed evolution studies. Thus, rather than starting from the random mutagenesis of a parent gene and selecting immediately for an adaptive function, neutral drift can first be performed in order to create a library with increased stability with large, non-deleterious variation, which possesses a considerably increased potential for evolving new functions [127].

For our purposes, as described in greater detail in Chapter 4, neutral drift of EcAII was performed not to evolve new or promiscuous functions, but instead exploited in a novel approach to establish an EcAII neutral network via iterative screening for variants with equivalent asparaginase activity, in order to isolate neutral mutants with reduced propensity to bind MHC-II.

Chapter 2: The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with β-aspartyl peptidase activity

2.1 INTRODUCTION

L-Asparaginases (L-asparagine amidohydrolases, EC 3.5.1.1) catalyze the hydrolysis of L-Asn to L-Asp and ammonia. Enzymes with L-Asparaginase activity are classified into either the bacterial-type or plant-type subfamilies based on sequence homology, structural features, and other biochemical properties. Considerable interest has been devoted to bacterial-type asparaginases for over 40 years, in part because of their antineoplastic properties [68, 100, 101, 129]. In contrast, plant-type asparaginases have not been studied as extensively and less is known about their structural and kinetic properties. Plant-type asparaginases share a high degree of amino acid sequence similarity (60-70%) with the aspartylglucosaminidase (AGA) family (EC 3.5.1.26) [91], with both enzyme groups belonging to the N-terminal nucleophile (Ntn) hydrolase protein superfamily [92-95, 97, 130]. Enzymes in this superfamily [131] are translated as inactive precursors which undergo an autocatalytic intramolecular activation step that exposes a nucleophilic residue (Thr, Ser, or Cys) at the N-terminus of the newly generated β subunit. The N-terminal nucleophile also acts as the catalytic residue during the activation step. The core folding pattern shared by Ntn-hydrolases consists of a conserved αββα structure consisting of two antiparallel β-sheets between flanking α-helical layers [132, 133].

Plant-type asparaginases from plant (*Lupinus luteus* [130] and *Arabidopsis thaliana [97]*), cyanobacteria [97] (*Synechocystis sp. PCC 6803* and *Anabaena sp. PCC*

7120), and bacteria [97, 130] (*E. coli*) display no hydrolytic activity towards N^4-(*β-N*-acetyl-D-glucosaminyl)-L-asparagine (GlcNAc-L-Asn), despite their high sequence similarity to AGAs. In addition to the hydrolysis of L-Asn, they also display significant and often higher activity towards β-aspartyl peptides, which has led to the suggestion that these enzymes be more accurately classified as β-aspartyl peptidases (EC 3.4.19.5) [97, 130, 134]. Formation of isoaspartyl peptide bonds is one of the most common sources of non-enzymatic protein damage under physiological conditions, as it introduces a kink in the protein backbone that can disrupt normal folding, leading to altered susceptibility to proteolysis or loss of function [135]. The β-aspartyl peptidases speculatively function to degrade these detrimental isoaspartyl peptides in the cell, as they would otherwise go unnoticed by α-peptide bond specific peptidases. So far, 13 enzymes have been verified or putatively designated as β-aspartyl peptidases, but none have been identified from mammals.

A putative L-asparaginase alternatively designated asparaginase-like protein 1 (ASRGL1), Glial asparaginase (GLIAP), or CRASH [136-139], was previously cloned from rat and human cDNA libraries [136, 138, 139]. Owing to its high sequence homology to a variety of asparaginases and AGAs, ASRGL1 was classified as an asparaginase though direct experimental evidence of its ability catalytically hydrolyze L-Asn has been lacking. Herein, we describe the bacterial expression and characterization of the human ASRGL1 (hASRGL1) and demonstrate that this enzyme is an Ntn hydrolase that displays an activity profile consistent with other previously studied β-aspartyl peptidases, thus revealing hASRGL1 as the first mammalian enzyme of the β-aspartyl peptidase family.

2.2 MATERIALS AND METHODS

Materials

Oligonucleotides were purchased from Integrated DNA Technologies (Coralville, IA). Restriction enzymes, *Vent* DNA polymerase, T4 DNA ligase, and dNTPs were from New England Biolabs (Ipswich, MA). *o*-phthalaldehyde (OPA) reagent was from Agilent Technologies (Santa Clara, CA). Difco 2xYT growth medium was from Becton Dickinson. (Franklin Lakes, NJ) The β-aspartyl peptides, β-L-Asp-L-Phe, β-L-Asp-L-Ala, β-L-Asp-L-Leu, and β-L-Asp-L-Lys, as well as L-aspartic acid β-(7-amido-4-methylcoumarin), L-Asp β-methyl ester, and GlcNAc-L-Asn were purchased from Bachem (Torrance, CA). The β-aspartyl peptide β-L-Asp-L-Phe methyl ester was from Sigma-Aldrich (St. Louis, MO). All other reagents were from Sigma-Aldrich (St. Louis, MO) unless otherwise noted.

Molecular Biology Methods

A gene encoding the human asparaginase-like protein 1 (hASRGL1) (927 bp) with an N-terminal His_6 affinity tag was assembled synthetically using a set of 36 codon-optimized overlapping oligonucleotides designed by the program DNAWorks [140]. *NcoI* and *EcoRI* restriction sites were incorporated into the outermost 5' and 3' oligonucleotides respectively. The gene assembly PCR reaction consisted of the oligonucleotide mix, ThermoPol buffer, dNTPs, and *Vent* DNA polymerase and was carried out at an initial 95ºC for 2 min, followed by 30 cycles at 95ºC for 1 min, 60ºC for 1 min, and 72ºC for 2 min, followed by a 72ºC polishing step for 10 min. After a subsequent amplification reaction with the outermost oligonucleotides, the resulting DNA product was gel purified

(Qiagen), digested with *NcoI/EcoRI* and ligated into pET28-a (Novagen) The gene insert in the resulting plasmid, pASRGL1, was sequenced and then ultimately transformed into *E. coli* BL21 (DE3) for subsequent expression. In addition, a hASRGL1-Thr168Ala variant was constructed by overlap-extension PCR using the pASRGL1 plasmid as template and the primer pairs shown in **Table 2.1**. The point-mutant gene was cloned analogously to hASRGL1, resulting in the plasmid pASRGL1-T168A.

Primer	Sequence
T7 promoter	5'-TAATACGACTCACTATAGGG-3'
ASRGL1-T168A-R	5'-CAGGGCGACAGCGCCGAC*TGC*CCCAAGGTTCTTCTGGCA - 3'
ASRGL1-T168A-F	5'- TGCCAGAAGAACCTTGGG*GCA*GTCGGCGCTGTCGCCCTG - 3'
T7 terminator	5'-GCTAGTTATTGCTCAGCGG-3'

Table 2.1 Primer pairs used for the construction of hASRGL1-T168A via overlap-extension PCR.

Expression and purification

E. coli BL21 (DE3) cells containing pASRGL1 plasmid was cultured overnight at 37ºC in 2xYT medium supplemented with 30 μg/mL kanamycin and used to inoculate fresh medium (1:100 dilution). When the absorbance at 600nm (A_{600}) reached 0.5-0.7, the cells were transferred to a 25ºC and allowed to equilibrate for 20 min, at which point the culture was supplemented with IPTG to a final concentration of 1mM to induce protein expression. After 16 hr of incubation at 25ºC, the cells were harvested by centrifugation at 10,000x*g* for 10 min. The cell pellet was resuspended in binding buffer (50mM Tris-HCl, 100mM NaCl, 10mM imidazole, pH 8) and placed on ice. The cells were then lysed by three passes through a French pressure cell and subsequently pelleted at 40,000x*g* for

45 min. The resulting supernatant (soluble fraction) was decanted, diluted 1:1 in binding buffer and mixed with 1mL of pre-equilibrated nickel-nitrilotriacetic acid (Ni^{2+}-NTA) resin. After incubation for 90 min at 4ºC with gentle rotation, the solution was applied to a 5mL polypropylene column (Pierce). The resin was washed with 25 bed volumes binding buffer and 25 bed volumes wash buffer (50mM Tris-HCl, 100mM NaCl, 25mM imidazole, pH 8) before the resin was incubated with 4mL elution buffer (50mM Tris-HCl, 100mM NaCl, 250mM imidazole, pH 8) for 10 min and collected drop-wise. All purification steps were performed at 4ºC. The elution fraction was applied to an Amicon Ultra 10K MWCO filter, buffer exchanged against activity buffer (50mM HEPES, 100mM NaCl, pH 7.4), mixed with glycerol (final 10% v/v), and finally snap-frozen with liquid nitrogen and stored at -80ºC. An identical protocol was followed for expression of hASRGL1-T168A. Protein concentrations were determined using a calculated extinction coefficient of 22,190 $M^{-1}cm^{-1}$ [141]. The expression of hASRGL1 and hASRGL1-T168A were confirmed by Western Blotting as described previously [142] using mouse monoclonal anti-polyhistidine peroxidase.

In vitro Processing

To evaluate the autocatalytic processing of hASRGL1, aliquots of both the purified wild-type and Thr168Ala enzyme variants were incubated at 37ºC [143-146] and at various times, aliquots were withdrawn and analyzed by SDS/PAGE on a 4-20% precast Tris-Glycine gel (NuSep Ltd) run under reducing conditions and stained with GelCode Blue (Thermo Scientific). To quantitatively assess the effect of 37ºC incubation on intramolecular processing, gel band intensities were measured using a densitometry imaging program (Quantity 1, Bio-Rad).

Activity Assays

Catalytic activity was qualitatively determined using the fluorometric substrate L-aspartic-acid-β-7-amido-4-methylcoumarin (AspAMC). Briefly, aliquots of soluble crude cell lysate fractions were normalized to equal pre-lysis A_{600} and then diluted ten-fold to 100μL in 50mM HEPES, 100mM NaCl, pH 7.4. 1μL of 10mM AspAMC in DMSO was added and mixed by pipetting in a 96 well plate (Nunc). The increase in fluorescence was monitored using a 360/40nm excitation filter and 460/40 emission filter (Synergy HT Fluorescent Platereader, BioTek) for 10 min at 25ºC.

The kinetics of hASRGL1 hydrolysis were determined with freshly purified enzyme that was first incubated at 37ºC for 48 hr and then stored at 4ºC until needed. The formation of L-Asp was determined following *o*-phthalaldehyde (OPA)-derivatization and HPLC analysis essentially as described by Agilent Technologies [147]. Reactions of hASRGL1 (2-4 μM total enzyme) with substrate (concentrations from 0-5x K_M) were carried out at 37ºC in 50mM HEPES, 100mM NaCl, pH 7.4, to a total volume of 100μL, and were subsequently quenched with 5μL 12% (w/v) trichloroacetic acid. An aliquot of the quenched reaction was then mixed with a molar excess (relative to substrate) of OPA reagent and brought to a final volume of 100μL with borate buffer. The resulting solutions were analyzed by HPLC using an Agilent ZORBAX Eclipse AAA Column (C18 reverse phase, 5μm, 4.6 x 150 mm). All reactions were done at least in triplicate and the observed rates were fit to the Michaelis-Menten equation using the program Kaleidagraph (Synergy).

2.3 RESULTS

Construction and expression of a synthetic hASRGL1 gene

The human ASRGL1 gene (GenBank accession no: BC093070) contains a number of rare *E. coli* codons such as the L-arginine codons AGA and AGG, whose presence has been shown to be detrimental to the expression of several recombinant proteins [148]. To circumvent expression problems due to rare codons, a synthetic hASRGL1 fused to a 5' sequence encoding a His_6 affinity tag was constructed by PCR gene assembly using codon-optimized oligonucleotides. The optimized hASRGL1 gene was expressed in *E. coli* BL21 (DE3) and protein synthesis was confirmed by Western blot analysis with an anti-His tag antibody (**Figure 2.1**). A protein band of the expected M.W. for the full-length protein (~33 kDa) was detected in both the soluble and insoluble fractions. In addition, a lower M.W. band (~18 kDa) was observed in the soluble fraction of the wild-type enzyme only. Activity assays using the fluorometric substrate AspAMC indicated the presence of L-Asn hydrolytic activity in cells expressing hASRGL1, but not in cells expressing hASRGL1-T168A (data not shown).

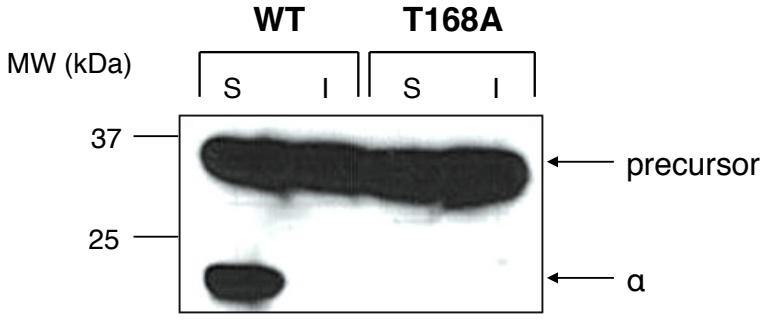


Figure 2.1 Western blot analysis of BL21 (DE3) cells expressing human ASRGL1 (hASRGL1) or hASRGL1-T168A. Samples corresponding to an equal number of cells were loaded in each lane. S denotes the soluble whole cell lysate fraction; I denotes the insoluble whole cell lysate fraction.

In Vitro Processing

hASRGL1 purification by immobilized metal ion affinity chromatography (IMAC) yielded 30mg/L of protein with a purity > 90% as determined by SDS/PAGE. Incubation of purified wild-type enzyme at 37ºC over time resulted in a gradual decrease of intensity in the band observed at ~33 kDa concomitant with an increase of intensity in the bands observed at ~18 kDa and ~15 kDa suggesting either specific proteolysis or intramolecular processing (**Figure 2.2**). Amino acid sequence alignment with other characterized β-aspartyl peptidases (**Figure 2.3**) identified Thr168 as the putative hASRGL1 N-terminal nucleophile requisite for the intramolecular processing characteristic of Ntn hydrolases. Consistent with this hypothesis, the hASRGL1-T168A variant was not processed to the lower M.W. bands as seen in both crude cell lysate and purified enzyme samples (**Figures 2.1, 2.2**). Approximately equivalent intensities of precursor and processed subunit gel bands were observed following incubation of hASRGL1 at 37ºC for 48 hr, however; longer incubation times did not further enhance processing. Similarly, the rate of AspAMC hydrolysis by hASRGL1 increased over time following incubation of the enzyme at 37ºC up to 48 hr, but the rate did not increase appreciably at later times (**Figure 2.4**).

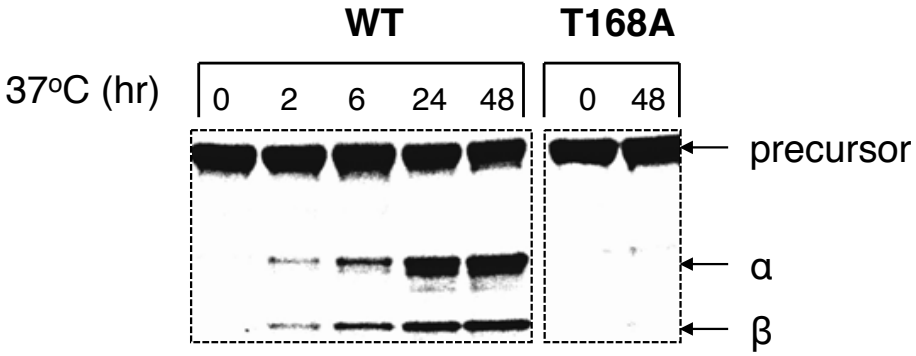


Figure 2.2 SDS-PAGE of human ASRGL1 (hASRGL1) and its Thr168Ala point mutant (hASRGL1-T168A) following in vitro incubation at 37 _C over time. Samples corresponding to an equiva- lent mass of total enzyme were loaded in each lane.

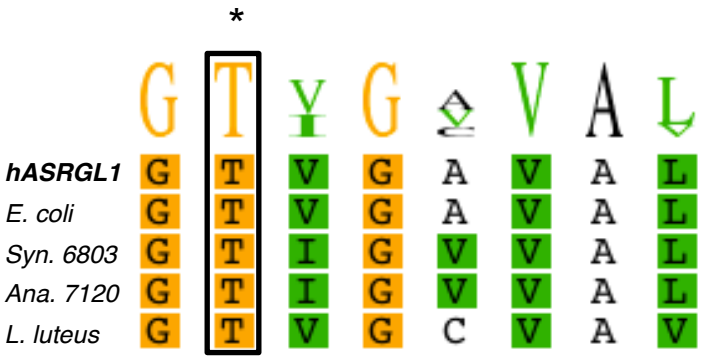


Figure 2.3 Sequence alignment of human asparaginase-like protein 1 (hASRGL1) with plant-type asparaginases. The asterisk indicates the conserved autoproteolytic cleavage site and corresponds to residue Thr168 for the hASRGL1 sequence. Residue conservation across the alignment is denoted by the degree of shading. The sequences of the plant-type asparaginases correspond to the following UniProtKB numbers: *E. coli* (P37595), *L. luteus* (Q9ZSD6), *Ana.7120* (Q8YQB1), and *Syn.6803* (P74383).

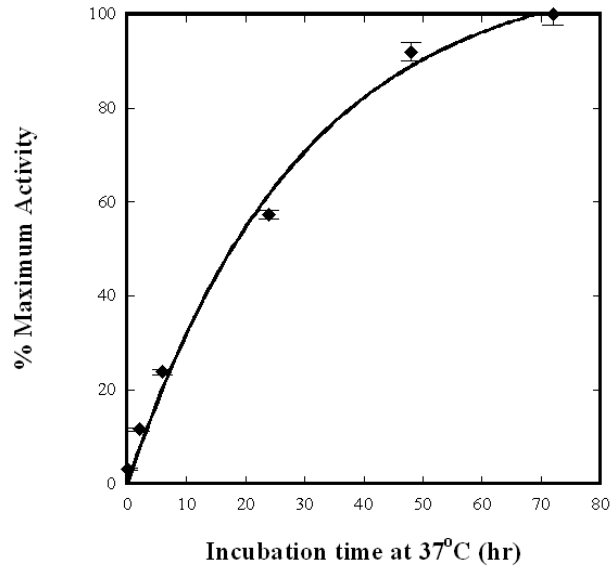


Figure 2.4 Progression of intramolecular processing of human ASRGL1 (hASRGL1) following in vitro incubation at 37C as determined by relative AspAMC hydrolysis rate over time. At various time points, equivalent aliquots of enzyme were withdrawn and analyzed using the fluorometric AspAMC activity assay as described in Experimental Procedures with the final time point serving as the maximum rate observed. The intramolecular proces- sing reaction as evaluated using this approach exhibited a $t_{1/2}$ of 20 $\pm$ 3 h to the maximum level of observed in vitro processing.

The mass spectrum of the protein incubated at 37ºC for 48 hr contained three prominent peaks of molecular masses 33,023 ± 11, 14547 ± 5, and 18495 ± 6 Da corresponding to the predicted M.W.s of the precursor enzyme minus the N-terminal methionine residue (33,023 Da) and the α (expected mass 18,499 Da) and β (expected mass 14,551Da) subunits likely to be generated following intramolecular processing of the precursor at the putative Thr168 cleavage site.

Kinetic Analysis

We determined the kinetic parameters of hydrolysis against a variety of potential hASRGL1 substrates including L-Asn; the β-aspartyl dipeptides: β-L-Asp-L-Phe, β-L-Asp-L-Phe-methyl ester, β-L-Asp-L-Ala, β-L-Asp-L-Leu, and β-L-Asp-L-Lys; L-Asp-β-methyl ester, and GlcNAc-L-Asn (**Table 2.2**). The calculation of k_{cat} was based on the observation that 50% of the total enzyme used in a given reaction was in the processed active form based on gel densitometry analysis.

Substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M ($mM^{-1}s^{-1}$)
L-Asp β-methyl ester	7.7 ± 0.3	0.4 ± 0.004	21 ± 1.0
β-L-Asp-L-Phe	3.0 ± 0.1	0.4 ± 0.04	8 ± 1.3
β-L-Asp-L-Ala	6.0 ± 0.2	1.0 ± 0.1	6 ± 0.8
β-L-Asp-L-Leu	5.0 ± 0.2	1.2 ± 0.1	4 ± 0.5
L-Asn	6.9 ± 0.2	3.4 ± 0.3	2 ± 0.3
β-L-Asp-L-Phe methyl ester	3.5 ± 0.1	1.6 ± 0.2	2 ± 0.3
β-L-Asp-L-Lys	5.3 ± 0.4	5.2 ± 1.2	1 ± 0.3
GlcNAc-L-Asn	-	n.d.	-
L-Gln	-	n.d.	-

Table 2.2 Summary of kinetic parameters for hASRGL1-catalyzed hydrolysis of various substrates. Reactions were all conducted at 37C and pH 7.4. *n.d.*, Not within the detection limit.

The activity profile of hASRGL1 was observed to be very similar to that of enzymes designated as isoaspartyl aminopeptidases with secondary L-asparaginase activity [97, 130] – classified as the β-aspartyl peptidase family (EC 3.4.19.5). The set of characterized enzymes in this family display a millimolar K_M for L-Asn, but a lower K_M

for a variety of isoaspartyl dipeptides. Moreover, these enzymes are unable to hydrolyze GlcNAc-L-Asn and L-Gln, consistent with the results seen for hASRGL1. In contrast, bacterial type II L-asparaginases such as the well-studied enzymes from *E. coli* and *Er. chrysanthemi*, possess a micromolar K_M value for L-Asn and relatively little to no activity towards β-aspartyl peptides [97, 149]. In addition, hASRGL1 was able to hydrolyze L-Asp β-methyl ester with a specificity constant greater than those for all other substrates tested.

2.4 DISCUSSION

The biochemical and structural features of plant-type asparaginases have only begun to be elucidated over the last decade. Many of these enzymes have been shown, or strongly suggested, to belong to the N-terminal nucleophile (Ntn) hydrolase superfamily [97, 130]. ASRGL1 has been cloned from both rat and human cDNA libraries [136, 138, 139], and on account of its high sequence homology to a variety of asparaginases and AGAs, it was classified as an L-asparaginase, though direct experimental evidence of its ability to hydrolyze L-Asn had not been demonstrated. Previously, ASRGL1 was shown to be concentrated in the cytosol and abundantly expressed in the brain, testis, and liver [138, 139].

Western blot, mass spectrometry and mutagenesis analyses revealed that, as its sequence homology to plant-type asparaginases implicated, hASRGL1 belongs to the Ntn hydrolase family for which Thr168 serves as the critical residue for intramolecular processing and catalytic activity. Kinetic analysis revealed that the enzyme displays both L-asparaginase and β-aspartyl peptidase activities, however, fails to hydrolyze either L-

Gln or GlcNAc-L-Asn, consistent with other characterized plant-type asparaginases [97, 130]. The k_{cat}/K_M values for β-aspartyl peptides containing a hydrophobic amino acid were to 2-4-fold higher relative to L-Asn, whereas substitution of a basic amino acid into the β-aspartyl position resulted in a 2-fold reduction in k_{cat}/K_M, again relative to L-Asn. We also found that an L-Asp β-methyl ester is hydrolyzed by this enzyme with a k_{cat}/K_M that is an order of magnitude higher than that for L-Asn, possibly because methanol is a superior leaving group relative to ammonia (L-Asn). Further, we found that the K_M value for β-L-Asp-L-Phe methyl ester (aspartame) was 4-fold higher relative to that for β-L-Asp-L-Phe, suggesting the enzyme prefers β-aspartyl dipeptides relative to longer β-aspartyl peptides, which may be related to the fact that such dipeptides are released following digestion of peptides containing a β-aspartyl linkage by certain carboxypeptidases [150]. Given that the k_{cat}/K_M values for all the β-aspartyl dipeptides assayed were within 8-fold of each other, it appears reasonable to suggest that hASRGL1 is a general β-aspartyl dipeptidase, though it is possible that the optimal substrate for this enzyme has not yet been identified.

The human lysosomal AGA is the only other known mammalian enzyme for which β-aspartyl peptidase activity had been previously observed [151]. However; the reported k_{cat}/K_M values of AGA for a variety of such substrates are much lower relative to those for hASRGL1.

In contrast, AGA displays activity towards GlcNAc-L-Asn unlike hASRGL1 and other β-aspartyl peptidases. This difference in catalytic specificity is likely a result of the divergence within the active site residues located near the L-Asn side chain moiety [97, 139] A model of the hASRGL1 structure based on *E. coli* isoaspartyl aminopeptidase/L-asparaginase was created from Phyre [152] and aligned to human AGA (PDB: 1APZ;

with bound L-Asp) using PyMol [105]. (**Figure 2.5**). Thus, while a near exact superimposition is observed in residues that bind the common L-aspartyl terminus of the substrate, differences are apparent in residues near the product side chain. In this region, whereas hASRGL1 predominantly contains Gly residues, human AGA contains Trp34 and Phe301, which are crucial for the binding of a sugar moiety.

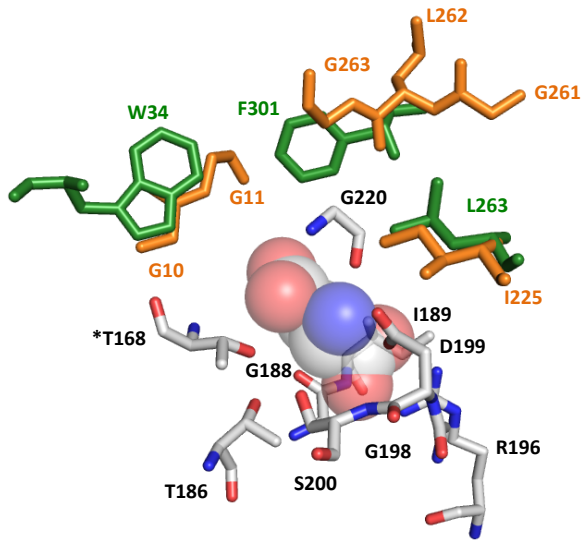


Figure 2.5 Predicted human ASRGL1 (hASRGL1) active site overlaid against the human aspartylglucosaminidase (AGA) active site. A model of hASRGL1 structure based on *E. coli* isoaspartyl amino- peptidase/L-asparaginase (EcAIII) was obtained from Phyre [152] (Job code 45ae21e95bf1f71f, SCOP code c2zakA, E-value of 7.5 x 10^{-41}, identity of 35%, estimated precision of 100%) and aligned with human AGA (Protein Data Bank entry 1APZ, with bound L-aspartate) using PyMol (34). Bound L-Asp is shown as spheres. Amino acids conserved between the two structures are colored by CPK and are numbered by hASRGL1 sequence (the asterisk indicates the nucleophilic Thr). Amino acids shaded green are specific to human AGA, and those shaded orange are specific to hASRGL1.

The physiological function of β-aspartyl peptidases from plant, cyanobacteria, and bacteria has been proposed to be related to the hydrolysis of isoaspartyl peptides [97, 130]. Formation of isoaspartyl peptide bonds is one of the most common sources of non-enzymatic protein damage under physiological conditions, as it introduces a kink in the protein backbone that can disrupt normal folding, leading to altered susceptibility to proteolysis, loss of function, or the potential to elicit autoimmunity [135]. Interestingly, the enzyme Protein L-Isoaspartyl Methyltransferase (PIMT, EC 2.1.1.77) was found to have a high degree of specificity for L-isoaspartyl residues [153, 154]. PIMT catalyzes the S-adenosyl-L-methionine (AdoMet)-dependent methylation of the α-carboxyl of an L-isoaspartyl site. Enzymatic methylation is followed by spontaneous ester hydrolysis ultimately resulting in a mixture of isoaspartyl (~70%) and aspartyl (~30%) linkages. This isoaspartyl product can then re-enter the methylation/demethylation cycle so that eventually there is a predominant shift towards the aspartyl linkage product [135]. The action of PIMT has fostered the idea that the enzyme serves an important intracellular repair function to keep isoaspartyl levels low. However, the methyltransferase activity is highly influenced by both the local sequence and solution environment around the isoaspartyl modification [155, 156] and thus in some cases, a PIMT-catalyzed isoaspartyl repair cannot take place. Presumably, in these situations, the isoaspartyl-containing proteins must be degraded, though the proteolytic quality control machinery is comprised of α-peptide bond proteases. Therefore, elimination of the remaining β-aspartyl peptides may be carried out by specialized isoaspartyl peptidases – the previously postulated role for plant-type asparaginases [130].

Though PIMT is present in all mammalian tissues examined to date, levels of the enzyme are notably higher in the brain and testis similar to the distribution of hASRGL1

[138, 139, 157]. While PIMT has been proposed to be important for the repair of damaged proteins within mature spermatozoa in the testes, its activity in the brain has proven to be crucial through the investigation of PIMT-deficient mice, which demonstrated the link between isoaspartate accumulation and neurological abnormalities [158-160]. Mammalian β-aspartyl peptidases may play a synergistic role with PIMT in these particular tissues, which perhaps possess a critical need for the repair and/or degradation of isoaspartyl-damaged proteins. Furthermore, evidence for the existence of mammalian β-aspartyl peptidases may corroborate a proposed mechanism through which such enzymes could account for the steady state level of isoaspartyl-damaged proteins observed in PIMT-deficient mice despite the continuous and spontaneous generation of such proteins [161].

Chapter 3: Expression, characterization, and identification of cysteines critical to the activity of human protein N-terminal asparagine amidohydrolase (hNTAN1)

3.1 INTRODUCTION

Targeted proteolytic degradation is essential for the regulation of a variety of cellular processes by achieving the correct balance between protein folding and the degradation of misfolded or damaged proteins [162]. In eukaryotes, the ubiquitin (Ub)-proteasome system (UPS) is the primary pathway for the temporally-controlled elimination of targeted proteins and the N-end rule pathway is a subset of this system that relates the *in vivo* half-life of a protein to the identity of its N-terminal residue [163, 164]. Degradation signals (degrons) that can be targeted by the N-end rule pathway are of two discrete subsets: N-terminal degrons (N-degrons) and internal degrons [163, 165]. The pathway has been observed in all organisms examined from prokaryotes to eukaryotes, though the former exists in a Ub-independent manner [166]. In eukaryotes, an N-degron consists of three determinants: a destabilizing N-terminal residue (**Figure 3.1**), one (or more) internal L-Lys residues (the site of poly-Ub chain formation), and a conformationally flexible region in the vicinity of this internal L-Lys [167, 168]. An N-degron can be generated from a precursor, called a pre-N-degron, through specific protease-mediated post-translational modifications [169].

The N-end rule has a hierarchic structure (**Figure 3.1**). In eukaryotes, N-terminal L-Asn and L-Gln function as tertiary destabilizing residues through their deamidation by N-terminal amidohydrolases (Nt-amidases) into the secondary destabilizing N-terminal residues L-Asp and L-Glu, respectively [170-172]. The secondary destabilizing activity of

N-terminal L-Asp and L-Glu requires their conjugation to L-Arg by the *ATE1*-encoded Arg-tRNA protein transferase (R-transferase) [173-175]. In mammals and other eukaryotes that produce nitric oxide (NO), N-terminal L-Cys can also function as a tertiary destabilizing residue through its oxidation to Cys-sulfinate or Cys-sulfonate in an O_2- or NO-dependent manner, whereby the oxidized L-Cys can be arginylated by ATE1 [169, 174]. Together with other primary destabilizing residues, the N-terminal L-Arg allows for protein substrate recognition by specific E3 Ub ligases and subsequent degradation [176].

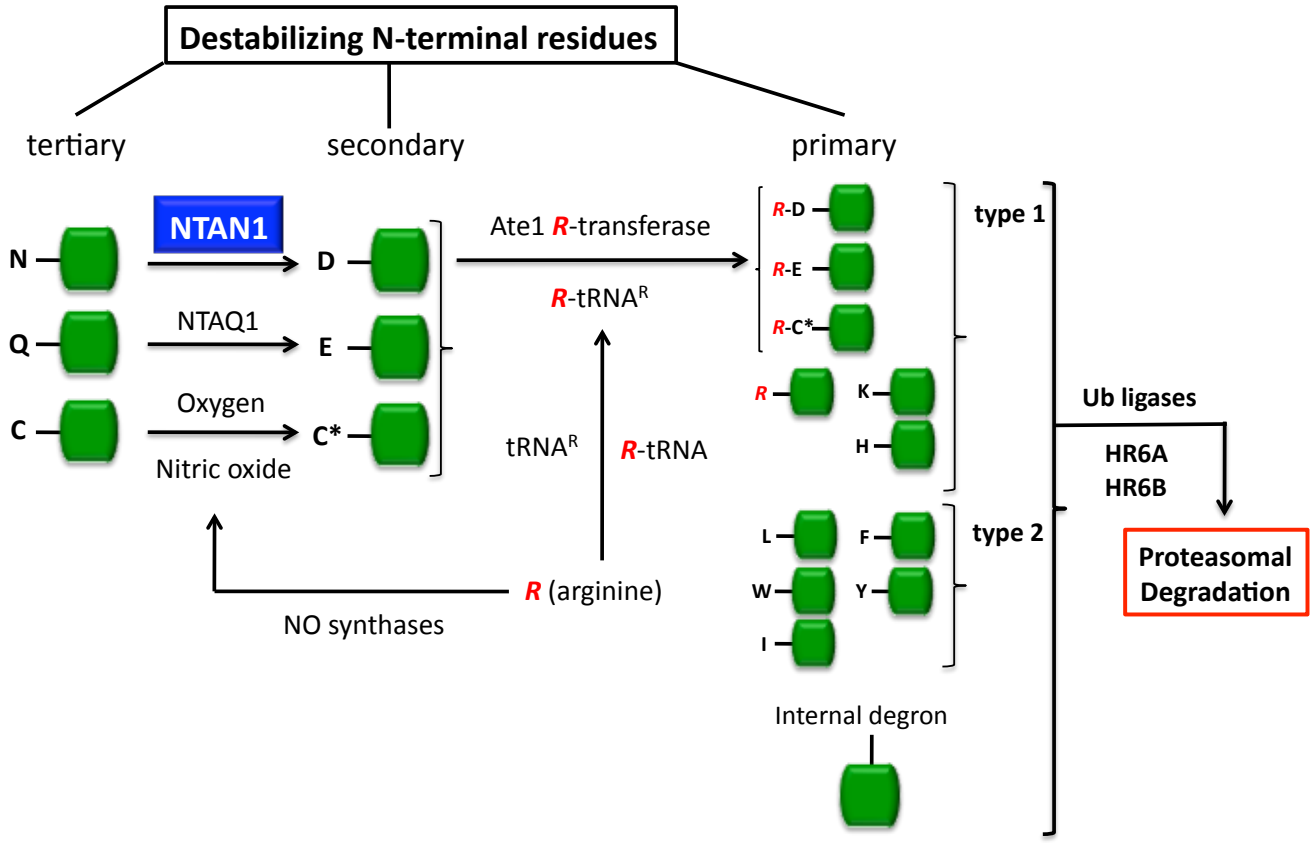


Figure 3.1 The Mammalian N-end rule pathway. N-terminal residues are indicated by single-letter abbreviations. Green boxes denote the remainder of the protein substrate. C* denotes oxidized N-terminal L-Cys (either Cys-sulfinate or Cys-sulfonate). The Cys oxidation requires nitric oxide (NO) and oxygen or its derivatives. "Tertiary", "secondary", and "primary" denote mechanistically distinct subsets of destabilizing N-terminal residues (See "Introduction"). The enlarged and blue-colored "NTAN1" denotes the N-terminal asparagine amidohydrolase characterized in this chapter.

A number of genetic studies with knock-out mice have implicated involvement of the N-end rule pathway in cardiac development and signaling, angiogenesis [173, 177], meiosis [178], DNA repair [179], neurogenesis [177], pancreatic functions [180], learning and memory [171, 181], female development [178], muscle atrophy [182], and olfaction [183]. Other functions of the pathway include: (i) selective degradation of misfolded proteins [184], (ii) sensing of nitric oxide, oxygen, and heme [173-175, 185], (iii) regulation of short peptide import [186, 187], (iv) fidelity of chromosome segregation [188], (v) regulation of apoptosis [189], and (vi) regulation of leaf senescence in plants [190].

In contrast to *Saccharomyces cerevisiae* which possesses a single enzyme (Nta1) that can deamidate either L-Asn or L-Gln at the N-terminus of a protein substrate [170], vertebrates possess one distinct enzyme for the specific deamidation of N-terminal L-Asn (NTAN1) and one for the specific deamidation of N-terminal L-Gln (NTAQ1) [172]. Previously, NTAN1-deficient mice were reported to be impaired in spontaneous activity, spatial memory, and a socially conditioned exploratory behavior [171]. However, the underlying molecular mechanisms explaining these behavioral profiles have not yet been determined.

NTAN1 has been previously purified from porcine liver through an 8-step purification [191] and the *Ntan1* gene has been cloned from a mouse cDNA library and then evaluated for activity complementation in *S. cerevisiae* Δ*nta1* [192]. Herein, we

describe the optimization of bacterial expression and subsequent biochemical characterization of the human NTAN1 (hNTAN1), which to the best of our knowledge represents the first example of bacterial recombinant expression of a mammalian N-end rule enzyme. The sequence of hNTAN1 is highly conserved among vertebrates, but is dissimilar to other amidases, including both Nta1 and NTAQ1, and shows no significant sequence similarity to any entries in the protein data bank. Similar to Nta1 and NTAQ1, we demonstrate that hNTAN1 catalysis is dependent upon involvement of a thiol group and we identify an essential Cys residue that is implicated to act as a nucleophile in catalysis. Further, we describe the first analysis of hNTAN1 kinetics, secondary structure composition, thermal stability, and reveal that hNTAN1 activity appears to be dependent upon exposure of a strictly conserved Pro2 following methionine aminopeptidase (MAP) cleavage of the initiation L-Met.

3.2 MATERIALS AND METHODS

Materials

Synthetic oligonucleotides were purchased from Integrated DNA Technologies (Coralvillle, IA). Restriction enodnucleases, *Vent* DNA polymerase, T4 DNA ligase, and dNTPs were from New England Biolabs (Ipswich, MA). Nickel-nitrilotriacetic acid (Ni^{2+}-NTA) resin was from QIAGEN (Valencia, CA). Strep-Tactin Superflow high capacity suspension and D-Desthiobiotin were obtained from IBA GmbH (St. Louis, MO). Anti-FLAG M2 magnetic beads, FLAG peptide, Val^5-angiotensin II (AII), and Asn^1Val^5-angiotensin II (N^1-AII) were purchased from Sigma-Aldrich (St. Louis, MO). Gln^1Val^5-angiotensin II (Q^1-AII) and Ac-Asn^1Val^5-angiontensin-II (Ac-N^1-AII) were

purchased from Abgent Inc. (San Diego, CA). Difco 2xYT growth medium was from Becton Dickinson (Franklin Lakes, NJ). All other reagents were obtained from Sigma-Aldrich (St. Louis, MO) unless otherwise noted.

Construction of hNTAN1 Expression Plasmids

A gene encoding the human protein N-terminal asparagine amidohydrolase (hNTAN1) (933bp) with a fused N-terminal His_6 affinity tag was synthetically assembled using a set of codon-optimized overlapping oligonucleotides designed by the program DNAWorks as described previously [98]. *NcoI* and *EcoRI* restriction sites were incorporated into the outermost sense and antisense oligonucleotides respectively. Additional sense primers (**Table 3.1**) were designed to: (i) introduce the hNTAN1 initiation codon within an *NcoI* site thereby eliminating the N-terminal His_6 affinity tag (NTAN-N-NoHis) and (ii) modify the ribosome binding site (RBS) and incorporate an upstream *XbaI* site in order to improve the translation initiation rate of hNTAN1 in *E. coli* (NTAN-RBS2). Additional antisense primers were designed to append a C-terminal His_6 affinity tag (NTAN-C-H), a C-terminal strepII affinity tag (NTAN-C-SII), and a C-terminal FLAG-strepII tandem affinity tag (NTAN-C-FSII), all followed by a *NotI* site. Lastly, a sense primer (NTAN-MP-RBS2) was designed to eliminate the GGA codon immediately following the initiation ATG (within an *NcoI* site) that was incorporated to maintain a correct open reading frame in the cassettes described above. Individual PCR products for each cassette were digested accordingly and ligated into pET28-a (Novagen). The RBS and the N- and C-terminal amino acid sequences of each hNTAN1 construct created are listed in **Table 3.2**. Plasmids were sequenced and ultimately transformed into *E. coli* BL21 (DE3) for expression analysis. The cassettes for six additional single hNTAN1 point mutant

variants, Cys75Ala, Cys75Ser, Cys75Thr, Cys89Ala, Cys93Ala, and Cys160Ala, were constructed by overlap-extension PCR using pNTAN2FLAGSII plasmid as template and the primer pairs shown in **Table 3.1**.

Primer Name	Nucleotide Sequence (5' → 3')
NTAN-N-NoHis	CATGCATGCCATGGGACCTCTTTTAGTGGAAGGCCGTCGTGTGCGTCTGCCGCGTAGCGC
NTAN-RBS2	CCCCTCTAGAAATAATTTTGTTTAACTTTAAGAACAATAATAAGGAGATAA GAAATGGGACCTCTTTTAGTGGAAGGCCGTCGTGTG
NTAN-C-H	TTTTCCTTTTGCGGCCGCTTATTAATGATGATGGTGGTGATGGCTGGAG CCTCCGCTACCCGGGCTGCTAATTTTTTCCCACAGACC
NTAN-C-SII	TTTTCCTTTTGCGGCCGCTTATTATTTTTCAAACTGCGGATGGCTCCAC GCGCTGCTACCCGGGCTGCTAATTTTTTCCCACAGACC
NTAN-C-FSII	TTTTCCTTTTGCGGCCGCTTATTATTTTTCAAACTGCGGATGGCTCCACGCGCTCTTGTC GTCATCGTCTTTGTAATCGCTACCGCCACCGCTACCCGGGCTGCTAATTTTTTCCCACAGACC
NTAN-MP-RBS2	CCCCTCTAGAAATAATTTTGTTTAACTTTAAGAACAATAATAAGGAGATAAGAAA TGCCTCTTTTAGTGGAAGGCCGTCGTGTG
NTAN-75AFor	AGCATCCTGGGCAGCGATGATGCCACCACCGCGCATATTGTGGTCTTACGGCACACCGGC
NTAN-75ARev	GCCGGTGTGCCGTAAGACCACAATATGCGCGGTGGTGGCATCATCGCTGCCCAGGATGCT
NTAN-75SFor	AGCATCCTGGGCAGCGATGATGCCACCACCAGCCATATTGTGGTCTTACGGCACACCGGC
NTAN-75SRev	GCCGGTGTGCCGTAAGACCACAATATGGCTGGTGGTGGCATCATCGCTGCCCAGGATGCT
NTAN-75TFor	AGCATCCTGGGCAGCGATGATGCCACCACCACCCATATTGTGGTCTTACGGCACACCGGC
NTAN-75TRev	GCCGGTGTGCCGTAAGACCACAATATGGGTGGTGGTGGCATCATCGCTGCCCAGGATGCT
NTAN-89AFor	GTCTTACGGCACACCGGCAACGGCGCGACCGCGTTAACCCATTGCGATGGGACCGATACC
NTAN-89ARev	GGTATCGGTCCCATCGCAATGGGTTAACGCGGTCGCGCCGTTGCCGGTGTGCCGTAAGAC
NTAN-C93AFor	GGCAACGGCGCGACCTGCTTAACCCATGCGGATGGGACCGATACCAAAGCGGAAGTCCCG
NTAN-C93ARev	CGGGACTTCCGCTTTGGTATCGGTCCCATCCGCATGGGTTAAGCAGGTCGCGCCGTTGCC
NTAN-C160AFor	GAAGACGATATTCACCTGGTGACCCTGGCGGTGACCGAACTGAACGACCGCGAAGAAAAC
NTAN-C160Arev	GTTTTCTTCGCGGTCGTTCAGTTCGGTCACCGCCAGGGTCACCAGGTGAATATCGTCTTC

Table 3.1 Summary of primers used for construction hNTAN1 plasmids.

Plasmid	RBS	Amino Acid Sequence	Yield (mg/L)	hNTAN1 Purity
pHisNTAN [a]	AAGGAGA	MGGSSHHHHHHSSGPLLV…SGPS	2	30 %
pNTANHis	AAGGAGA	MGPLLV…SGPSGGSSHHHHHH	2	30 %
pR1NTANSII	AAGGAGA	MGPLLV…SPGSSAWSHPQFEK	1 [b]	30 %
pR2NTANSII	AACAATAATAAGGAGAT AAGAA	MGPLLV…SPGSSAWSHPQFEK	1	70 %
pNTANFLAGSII	AACAATAATAAGGAGAT AAGAA	MGPLLV… SPGSGGGSDYKDDDDKSAWSHPQ FEK	0.2 [c]	> 98%
pNTAN2FLAGSII	AACAATAATAAGGAGAT AAGAA	MPLLV… SPGSGGGSDYKDDDDKSAWSHPQ FEK	0.2	> 98 %

Table 3.2 hNTAN1 constructs and their relevant expression characteristics in *E. coli* BL21 (DE3). The sequences of the sense and antisense primers used for the modified termini and RBS are found in **Table 3.1.** The resultant amino acid sequences are shown where residues following the ellipsis are indicative of the C-terminus of the sequence. The purity of hNTAN1 was estimated by SDS-PAGE, where band intensity was estimated by densitometry imaging (Quantity 1, BioRad). [a] Trace activity could be detected for His_6-hNTAN1 through the addition of final purified product (~10μg total protein) to 1mM N1-AII for 90 min at 37ºC in 50mM Tris-HCl, 150mM NaCl, pH 7.5. Reactions were quenched with 12% TCA (w/v) and then evaluated by HPLC for the presence of a peak corresponding to the expected product AII. For all other constructs, the presence of hNTAN1 activity could be detected through the addition of final purified product (10-100ng total protein) to 100μM N1-AII for 1 min at 37ºC in the same buffer and evaluated similarly. [b] Binding capacity of strep-tactin superflow high capacity resin is reported by manufacturer as 3-15mg per 2mL 50% resin suspension. .[c] Binding capacity of anti-FLAG M2 magnetic beads is reported by manufacturer as ~60μg/100μL packed gel volume.

Expression and purification of recombinant hNTAN1

E. coli BL21(DE3) cells were cultured overnight at 37ºC in 2xYT medium supplemented with 30μg/mL kanamycin and used to inoculate 500mL fresh medium (1:100 dilution). When the absorbance at 600nm (A_{600}) reached 0.5-0.7, the cells were transferred to 25ºC and allowed to equilibrate for 20 min, at which point the culture was supplemented with IPTG to a final concentration of 1mM to induce protein expression. After 16 hr

incubation at 25ºC, the cells were harvested by centrifugation at 10,000*xg* for 10 min. Subsequent purification steps were dependent upon the affinity tag used for a particular protein preparation and are described as follows.

(i) His_6 affinity tag purification

The cell pellet was resuspended in binding buffer (50mM Tris-HCl, 150mM NaCl, 10mM imidazole, pH 8) and placed on ice. The cells were then lysed by three passes through a French pressure cell and subsequently pelleted at 40,000x*g* for 45 min. The resulting supernatant (soluble fraction) was decanted, diluted 1:1 in binding buffer and mixed with 1mL of pre-equilibrated Ni^{2+}-NTA resin. Following incubation for 90 min at 4ºC with gentle rotation, the solution was applied to a 5mL polypropylene column (Pierce). The resin was washed with 25 bed volumes binding buffer and 25 bed volumes wash buffer (binding buffer with 25mM imidazole) before the resin was incubated with 4mL elution buffer (binding buffer with 250mM imidazole) 10 min and collected drop-wise. All purification steps were performed at 4ºC. The elution fraction was applied to an Amicon Ultra 10K MWCO filter, buffer exchanged against storage buffer (50mM Tris-HCl, 150mM NaCl, 1mM EDTA, 0.5mM dTT pH 7.5) and stored at 4ºC.

(ii) Strep-II affinity tag purification

The cell pellet was resuspended in 15mL buffer W (100mM Tris-HCl, 150mM NaCl, 1mM EDTA, pH 8) and placed on ice. The cells were then lysed by three passes through a French pressure cell and subsequently pelleted at 40,000x*g* for 45 min. The resulting supernatant (soluble fraction) was decanted and applied to a 5mL polypropylene column (Pierce) that had been prepared as described elsewhere [193] with 300μL strep-tactin superflow high capacity resin. The column flow through was collected and re-applied to the column for 3 total passes. The resin was then washed with 30 bed volumes buffer W and 10 bed volumes buffer W2 (buffer W containing 100μM D-Desthiobiotin) before the resin was incubated with 10 bed volumes buffer E (buffer W containing 2.5mM D-Desthiobiotin) and collected drop-wise. All purification steps were performed at 4ºC. The elution fraction was applied to an Amicon Ultra 10K MWCO filter, buffer exchanged against storage buffer and stored at 4ºC.

(iii) FLAG-Strep-II tandem affinity tag purification

Steps were followed as for (ii) through the collection of eluant with buffer E, at which point the elution fraction was applied to an Amicon Ultra 10K MWCO filter, concentrated to a final volume of 1mL in buffer W, and then incubated with 200μL anti-FLAG M2 magnetic beads (prepared according to the manufacturer's instructions) for 1 hr at room temperature with gentle rotation. The magnetic beads were then collected using a magnetic separator to remove the supernatant, and washed with 30 bed volumes buffer W before the purified protein was collected by competitive elution with 5 bed volumes of buffer W containing 100μg/mL FLAG peptide by gentle rotation for 10 min at room temperature. The final eluant was applied to an Amicon Ultra 10K MWCO filter,

concentrated to a final volume of 100μL in storage buffer and stored at 4ºC. The expression and purity of hNTAN1 was evaluated by SDS-PAGE. Protein concentrations were determined using a calculated extinction coefficient of 32,430 $M^{-1}cm^{-1}$ [141].

Activity assay

For kinetic analyses, reactions of hNTAN1 (40-50nM purified enzyme) with substrate (concentrations from 0-5 x K_M) were carried out at 37ºC in activity buffer (50mM Tris-HCl, 150mM NaCl, pH 7.5) to a total volume of 100μL, and were subsequently quenched with 5μL 12% (w/v) trichloroacetic acid (TCA). For reactions with the substrate Asn^1Val^5-angiotensin II (N^1-AII), an aliquot of the quenched reaction was brought to a final volume of 100μL with activity buffer and then analyzed by HPLC on a Phenomenex C18 reverse-phase column using the following gradient: 5% ACN, 95% H_2O for 2 min, increasing to 21% ACN, 79% H_2O over 5 min, then increasing to 24.7% ACN, 75.3% H_2O over 8 min, and finally increasing to 95% ACN, 5% H_2O over 10 min before returning to 5% ACN, 95% H_2O over 3 min. Both solvents contained 0.1% trifluoroacetic acid (TFA). The product concentration was determined using the integration area at 280nm and a standard curve generated under the same conditions with Val^5-angiotensin II (AII). For the substrates L-Asn and L-Gln, quenched reactions were evaluated by *o*-phthalaldehyde (OPA)-derivitization and HPLC analysis as described previously [98]. All reactions were done at least in triplicate and the observed rates were fit to the Michaelis-Menten equation using the program Kaleidagraph (Synergy).

Effect of pH on catalysis

The effect of pH on hNTAN1 activity was determined using 50mM Bis-Tris, pH 6.0-6.5, 50mM Tris-HCl, pH 7.0-8.5, and 50mM CAPSO, pH 10. All buffers contained 150mM NaCl. Enzyme was incubated at 25ºC for 1.5 min in a buffer of interest and reactions were then initiated by the addition of preincubated enzyme (1nM) to 25μM N^1-AII at 37ºC in the same buffer. Reactions were carried out for 1.5 min and then quenched and analyzed by HPLC as described earlier.

Effect of metals, EDTA, D-desthiobiotin and iodoacetamide on hNTAN1

The effects of various divalent salts, EDTA, D-desthiobiotin and the alkylating agent iodoacetamide on hNTAN1 activity were determined in activity buffer. Enzyme (20nM) was incubated at 4ºC for 3 hr with either iodoacetamide (0-50mM), EDTA (1mM), D-desthiobiotin (3mM), or a divalent salt (1mM) and reactions were then initiated by the addition of the preincubated enzyme (1nM) to 100μM N^1-AII in 50mM Tris-HCl, 150mM NaCl, pH 7.5 at 37ºC. Reactions were carried out for 4 min and then quenched and analyzed by HPLC as described earlier.

Circular Dichroism (CD) Spectra and T_M measurements

CD spectra were measured at 200-260nm using a Jasco J-815 CD spectrometer equipped with a temperature-controlled cell holder set to room temperature. Results were generated using equimolar amounts (15μM) of each enzyme in 50mM Tris-HCl, 150mM NaCl, 1mM EDTA, 0.5mM dTT, pH 7.5 of purified hNTAN1-FLAG-strepII variants and the CD spectra obtained were representative of 3 accumulations at each wavelength at a scanning speed of 100nm/min. For determination of thermal stability, CD signal (in mdeg) was monitored at 200-260nm as temperature was increased from 20ºC to 80ºC at a

rate of 1ºC/min. Points were collected at 5ºC intervals and reflect three accumulations at each temperature. The data were then fit to sigmoidal curves using Kaliedagraph and the melting temperature (T_M) was designated as the temperature corresponding to the inflection point of the fitted curve.

3.3 RESULTS

Construction of a synthetic hNTAN1 gene

The human NTAN1 gene (GenBank accession no: NM_173474) contains a number of rare *E. coli* codons, including 13 such L-Arg codons (AGA, AGG, and CGA), whose presence has been shown to be detrimental to the expression of several recombinant proteins [194]. To avoid expression problems owing to rare codons, PCR gene assembly using codon-optimized oligonucleotides was employed to construct a synthetic hNTAN1 fused to a 5' sequence encoding a His_6 affinity tag. The optimized hNTAN1 gene was expressed in *E. coli* BL21 (DE3) and protein synthesis was confirmed by Western blot analysis with an anti-His tag antibody as described previously [142]. A protein band of the expected M.W. for the full-length protein (~35kDa) was detected in both the soluble and insoluble fractions (**Figure 3.2**). However, purified His_6-hNTAN1 was found to be minimally active towards the substrate N^1-AII (**Table 3.2**).

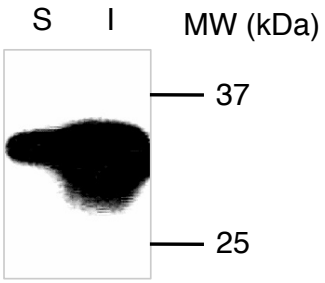


Figure 3.2 Western blot analysis of purified hNTAN1 expressed with N-terminal His_6 affinity tag. Samples corresponding to an equal number of cells were loaded in each lane. Western blotting was done using mouse monoclonal anti-polyhistidine peroxidase. S, soluble whole cell lysate fraction; I, insoluble whole cell lysate fraction

Optimization of expression plasmid

The plasmid pNTANHis was constructed to determine whether His_6-hNTAN1 inactivity resulted from the deleterious encoding of additional amino acids upstream of the native N-terminus of the enzyme. Despite both low expression yield and purity (**Figure 3.3**), recombinant hNTAN1-His_6 displayed much greater activity towards the substrate N^1-AII. Further characterization of hNTAN1-His_6 showed that a number of divalent metals, including Ni^{2+} and Co^{2+}, at least partially inhibited enzymatic activity. Thus, to afford a means of protein purification that was not dependent on immobilized metal ion affinity chromatography (IMAC), we replaced the His_6 affinity tag with a C-terminal strep-II affinity tag (pR1NTANSII) whereby EDTA was present throughout purification and gentle elution of recombinant hNTAN1-strepII was accomplished using low concentrations of D-Desthiobiotin at physiological pH. Although hNTAN1-strepII retained the previously observed activity, its expression yield and purity remained comparable to those observed for the IMAC-purified versions of the enzyme (**Table 3.2**).

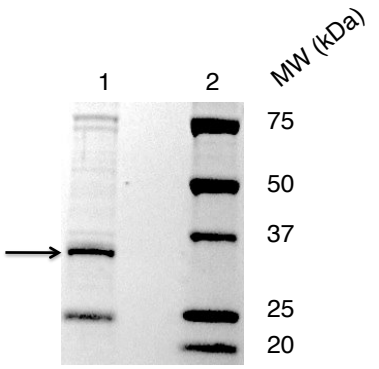


Figure 3.3 SDS-PAGE analysis of purified hNTAN1 expressed with C-terminal His_6 affinity tag. 1: Final product following recombinant expression and purification from construct pNTANHis. Arrow denotes NTAN1 band which was equivalent to ~30% of total protein in the lane as determined by band densitometry. 2: M.W. marker.

To improve expression yield, we utilized an RBS calculator [195] to engineer an optimized RBS designed to improve the translation initiation rate of the synthetic hNTAN1 coding sequence in *E. coli*. Upon expression of hNTAN1-strepII using the engineered RBS (pR2NTANSII) shown in **Table 3.2**, we saw a 2-3-fold increase in hNTAN1 yield and concomitantly improved hNTAN1 purity equivalent to 70% of total purified protein (**Figure 3.4**), representing over a 2-fold increase in purity relative to the previous constructs. To further improve purity, we constructed a C-terminal FLAG-strepII tandem affinity tag (pNTANFLAGSII) for which gentle elution of FLAG-bound enzyme was achieved by competitive FLAG peptide binding at physiological pH. The tandem affinity tag increased purity of recombinant hNTAN1 to near homogeneity, as determined by SDS-PAGE (**Figure 3.5**), without significantly sacrificing the yield obtained from the single tag purification strategies.

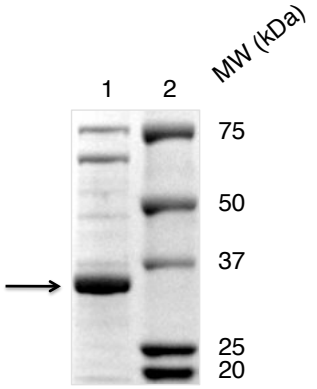


Figure 3.4 SDS-PAGE analysis of purified hNTAN1 expressed with engineered RBS and C-terminal strep-II affinity tag. 1: Final product following recombinant expression and purification from construct pR2NTANSII. Arrow denotes NTAN1 band which was equivalent to ~70% of total protein in the lane as determined by band densitometry. 2: M.W. marker.

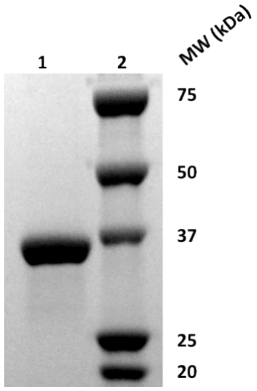
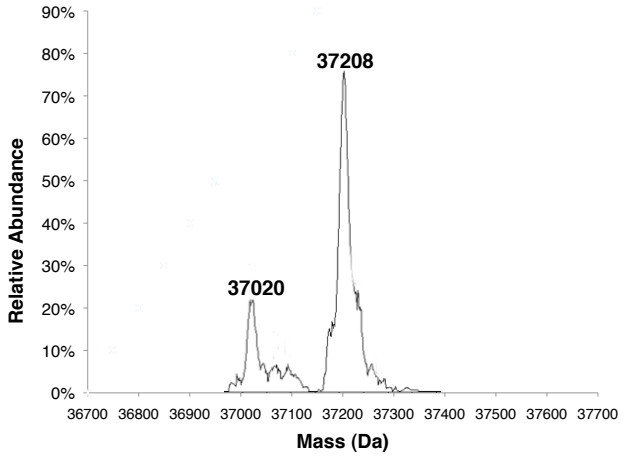


Figure 3.5 SDS-PAGE analysis of purified hNTAN1 expressed with engineered RBS and tandem C-terminal FLAG-strepII affinity tag. 1: Final product following recombinant expression and purification from construct pNTANFLAGSII. Equivalent homogeneity was observed from the modified construct pNTAN2FLAGSII. 2: M.W. marker.

ESI-MS analysis of hNTAN1

The mass spectrum of purified hNTAN1-FLAG-strepII contained two prominent peaks of molecular masses 37208 $\pm$ 12 and 37020 $\pm$ 12 (**Figure 3.6A**), corresponding to the predicted molecular weights of the encoded enzyme (37,214 Da) and the enzyme minus the N-terminal Met and Gly residues (37,026 Da) respectively. Though *E. coli* methionine aminopeptidase (MAP) is highly efficient when Gly occupies P1' of a protein substrate, its catalytic efficiency is severely compromised when Pro occupies P2' of that same substrate [196]. Therefore, incorporation of the non-native encoded Gly2 during assembly of the synthetic hNTAN1 gene appeared to hinder MAP processing, resulting in a ~3:1 ratio of unprocessed to processed enzyme despite overnight protein induction. As described in the next section, failure to expose the native NTAN1 Pro2 upon MAP processing may compromise hNTAN1 activity, further implicating the sensitivity of the hNTAN1 N-terminus initially seen in analysis of His_6-hNTAN1. To obtain pure MAP-processed hNTAN1, a modified construct was created (pNTAN2FLAGSII) in which the artificial Gly2 was eliminated. It is worth noting that *E. coli* MAP is highly efficient against substrates in which Pro occupies P1' and Leu occupies P2' [196] given that the native NTAN1 contains Pro2 and Leu3. The mass spectrum of the modified enzyme ultimately contained the 1 expected prominent peak of molecular mass 37022 $\pm$ 12 (**Figure 3.6B**). Therefore, the construct pNTAN2FLAGSII was used for all extensive biochemical characterization of hNTAN1 described.

(A)



(B)

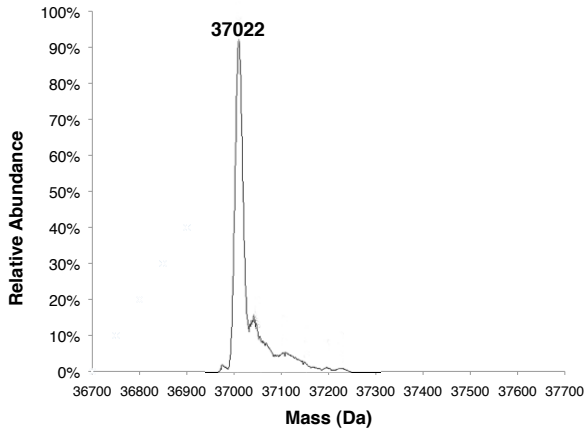


Figure 3.6 Deconvoluted ESI-MS spectra of hNTAN1. (A) Wild-type hNTAN1 with C-terminal FLAG-strepII tandem affinity tag and containing non-native Gly2. (B) Wild-type hNTAN1 with C-terminal FLAG-strepII tandem affinity tag without non-native Gly2. The peak at mass 37208 corresponds to newly synthesized full-length protein following removal of the formyl group from the initiation Met. The peaks at masses 37020 (and 37022) correspond to the MAP-processed protein with Pro2 of the native NTAN1 sequence at the N-terminus.

Characterization of hNTAN1 activity

We obtained the hNTAN1 kinetic parameters of hydrolysis against the substrate N^1-AII using purified hNTAN1-FLAG-strepII preparations from both pNTANFLAGSII and pNTAN2FLAGSII (**Table 3.3**). Although the K_M was essentially equivalent between the two enzymes, the hNTAN1 lacking the artificial Gly2 exhibited a nearly 3-fold higher k_{cat} When taken with the mass spectra of each preparation described in the previous section, it appeared that the action of MAP to expose the strictly conserved NTAN1 Pro2 (**Figure 3.7**) at the N-terminus may be important for activity. Moreover, hNTAN1 was unable to hydrolyze Q^1-AII, Ac-N^1-AII, L-Asn, and L-Gln, consistent with results previously described for NTAN1 purified from porcine liver [191].

Enzyme	K_M (μM)	k_{cat} (s^{-1})	k_{cat}/K_M ($M^{-1}s^{-1}$)
hNTAN1 (+Gly2)	31 ± 4.4	2.3 ± 0.07	7.4 ± 1.1 x 10^4
hNTAN1 (-Gly 2)	35 ± 3	6.6 ± 0.13	1.9 ± 0.17 x 10^5

Table 3.3 Kinetic parameters of hNTAN1 against the substrate N^1-AII. Kinetic parameters were derived from nonlinear fit of the Michaelis-Menten equation to initial rate measurements. Reactions were carried out at 37ºC in 50mM Tris-HCl, 150mM NaCl, pH 7.5. The parenthetical in the enzyme column depicts whether the enzyme contained a non-native encoded Gly2.

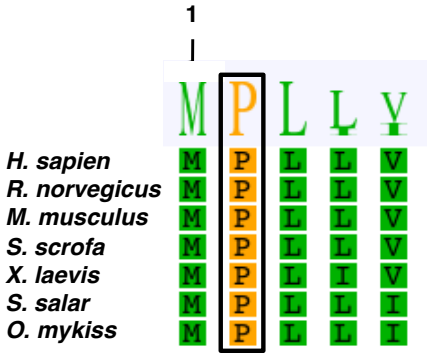


Figure 3.7 Sequence alignment of NTAN1 from various vertebrates. A multiple sequence alignment identified a strictly conserved Pro residue at position 2 among NTAN1 homologues, which is shown as boxed. Deduced amino acid sequences of the 7 NTAN1 homologues were aligned with the Geneious software. Residue numbering corresponds to the *H. sapien* NTAN1. The abbreviations are for the following organisms and corresponding UniProtKB numbers: *Homo sapien* – Q96AB6, *Rattus norvegicus* – Q5BK79, *Mus musculus* – Q64311, *Sus scrofa* – Q28955 *Xenopus laevis* – Q5BJ36, *Salmo salar* – B5X7Q7, *Oncorhynchus mykiss* – C1BHX6.

Effect of pH on hNTAN1 activity

The effect of pH on hNTAN1 activity against the substrate N^1-AII at a single concentration (approximately equivalent to the calculated K_M at pH 7.5) is shown in **Figure 3.8**. The pH-dependence of porcine NTAN1 activity performed under similar buffer conditions previously had revealed a general lack of effect of pH on catalytic activity over the range 6-8 using the same N^1-AII substrate [191]. Here, we observe a similar broad activity optimum over pH 6.0-7.5, followed by a sharp decline in activity at alkaline pH. This approximation of the pH dependence on hNTAN1 activity suggests that the ionization state of a single group within the enzyme is crucial for catalysis, resulting in a pH-dependence curve with a single descending limb p*K*a of 8.2 ± 0.06.

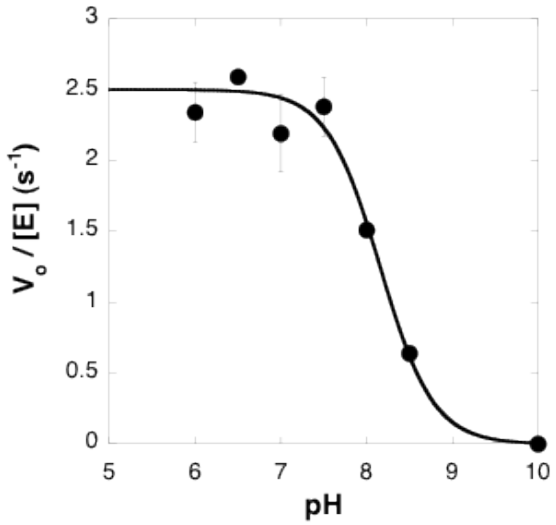


Figure 3.7 Effect of pH on hNTAN1 activity. Enzyme (20nM) was incubated at 25ºC for 1.5 min in the buffer of interest and reactions were initiated by the addition of the preincubated enzyme to 25μM N^1-AII at 37ºC in the identical buffer. Reactions were carried out for 1.5 min and then quenched and analyzed by HPLC. Activity was evaluated as the initial velocity of the reaction (as calculated using the area of the AII product peak) normalized by the enzyme concentration for each pH condition. All reactions were done in triplicate. Buffers used were 50mM Bis-Tris, pH 6.0-6.5, 50mM Tris, pH 7.0-8.5, and 50mM CAPSO, pH 10. All buffers contained 150mM NaCl.

Effect of metals, EDTA, D-desthiobiotin and iodoacetamide on hNTAN1activity

As summarized in **Table 3.4**, Zn^{2+} and Cu^{2+} were able to completely inhibit hNTAN1 activity, while Ni^{2+}, Co^{2+}, and Mn^{2+}, present at a final concentration of 50μM, only partially inhibited the enzyme. Conversely, Mg^{2+}, Ca^{2+}, EDTA, and D-desthiobiotin did not impart any noticeable effects on activity. Further, hNTAN1 activity gradually decreased to the point of complete abolishment upon incubation with increasing concentrations of the alkylating agent iodoacetamide (**Figure 3.8**).

Effector	% Activity Remaining
$MgCl_2$	100
$CaCl_2$	100
$CoCl_2$	58 ± 2.8
$MnSO_4$	40 ±14
$NiSO_4$	39 ± 4.2
$CuCl_2$	0 [a]
$ZnSO_4$	0
EDTA	100
D-desthiobiotin	100

Table 3.4 Effects of divalent metals, EDTA, and D-desthiobiotin on hNTAN1 activity. Enzyme (20nM) was incubated at 4ºC for 3 hr with either EDTA (1mM), D-desthiobiotin (3mM), or a divalent salt (1mM) and reactions were then initiated by the addition of the preincubated enzyme (1nM) to 100μM N^1-AII at 37ºC in 50mM Tris-HCl, 150mM NaCl, pH 7.5. Reactions were carried out for 4 min and then quenched and analyzed by HPLC. Activity reported is relative to the AII peak area generated from enzyme preincubated in buffer only. [a]Area of the AII peak was less than or equal to that observed in an HPLC trace of N^1-AII only at the equivalent reaction concentration.

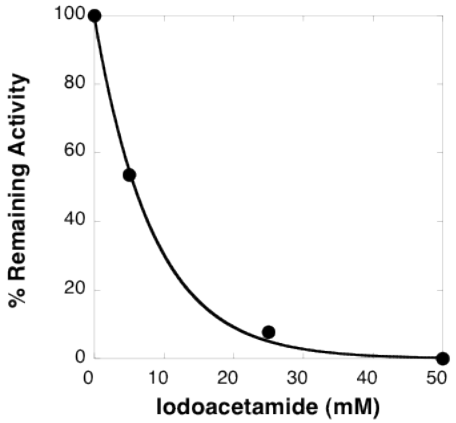


Figure 3.8 Effect of iodoacetamide on hNTAN1 activity. Enzyme (20nM) was incubated at 4ºC for 3 hr with varying concentrations (0-50mM) of iodoacetamide and reactions were then initiated by the addition of the preincubated enzyme to 100μM N^1-AII at 37ºC in 50mM Tris-HCl, 150mM NaCl, pH 7.5. Reactions were carried out for 4 min and then quenched and analyzed by HPLC. Activity reported is relative to the AII peak area generated from enzyme preincubated in buffer only. Data was fit to follow an exponential decay.

Identification of critical Cys residues for hNTAN1 activity

The elimination of hNTAN1 activity upon incubation with high concentrations of iodoacetamide led to the presumption that the enzyme utilizes a nucleophilic Cys, which to date had not been suggested or identified in any mammalian NTAN1. Homology modeling with six additional vertebrate NTAN1 sequences revealed 4 strictly conserved Cys residues, which we then targeted for individual Cys to Ala hNTAN1 point mutations (**Figure 3.9**). While the variants Cys89Ala and Cys160Ala displayed activity equivalent to wild-type, the Cys93Ala variant exhibited ~40% loss in activity relative to wild-type, and the Cys75Ala variant was completely inactive (**Table 3.5**). This result implicates Cys75 as an essential residue, and further, as a potential catalytic nucleophile, which would be consistent with results demonstrated previously for both Nta1 [192] and NTAQ1 [172]. Moreover, enzymatic activity could not be reconstituted in either Cys75Ser or Cys75Thr variants, suggesting strict residue inflexibility at the position with respect to catalysis.

Figure 3.9 Sequence alignment of NTAN1 from various vertebrates. A multiple sequence alignment identified 4 conserved Cys residues among NTAN1 homologues, which are shown as boxed. Residue numbers shown correspond to the human sequence. Deduced amino acid sequences of the 7 NTAN1 homologues were aligned with the Geneious software. Residue numbering corresponds to the *H. sapien* NTAN1. The abbreviations are for the following organisms and corresponding UniProtKB numbers: *Homo sapien* – Q96AB6, *Rattus norvegicus* – Q5BK79, *Mus musculus* – Q64311, *Sus scrofa* – Q28955 *Xenopus laevis* – Q5BJ36, *Salmo salar* – B5X7Q7, *Oncorhynchus mykiss* – C1BHX6.

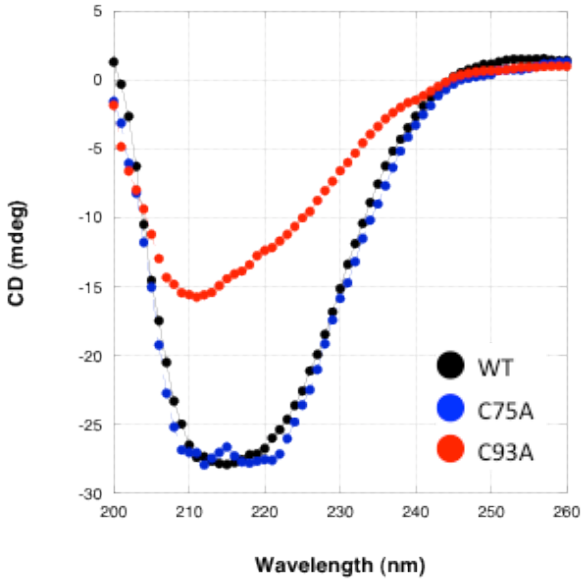
Enzyme	% Relative Activity (v. WT)
WT	100
C89A	100
C161A	100
C93A	62 ± 7
C75A	0 [a]
C75S	0
C75T	0

Table 3.5 Relative activity of hNTAN1 variants. Reactions of enzyme (40-50nM) with 100μM N^1-AII were carried out at 37°C for 1 min in 50mM Tris-HCl, 150mM NaCl, pH 7.5. Reactions were quenched and analyzed by HPLC. Activity reported is relative to the AII peak area generated by the wild-type (WT) enzyme. [a]Area of AII peak was less than or equal to that observed in an HPLC trace of N^1-AII only at the equivalent reaction concentration.

Circular Dichroism (CD) of select hNTAN1 variants

CD spectra were used to detect whether amino acid point mutations that had either compromised (Cys93Ala) or abolished (Cys75Ala) hNTAN1 activity caused changes in either protein secondary structure and/or thermal stability relative to the wild-type enzyme. Secondary structure evaluation was done using the Jasco software. Analysis of the wild-type hNTAN1 CD spectra revealed a secondary structure composition that was 28.5% helical, 22.7% sheet, 19.1% turn, and 29.8% random. This result is the first describing any exploration into the higher order structure of eukaryotic NTAN1. The mutation Cys75Ala resulted in no obvious change in CD spectra (**Figure 3.10A**), further supporting that the inactivity of this variant was strictly due to the removal of the potentially nucleophilic thiol group and not to changes in folding. In contrast, the Cys93Ala variant resulted in a large decrease in helical structure (from 28.5% to 16.2%) (**Figure 3.10B)** denoting that this conserved Cys is likely involved in proper folding, and also providing a rationale for the reduced activity of this variant.

(A)



(B)

Enzyme	% Helix	% Sheet	% Turn	% Random
WT	28.5	22.7	19.1	29.8
C75A	26.8	21.9	23.0	28.3
C93A	16.2	32.8	19.5	31.5

Figure 3.10 Circular dichroism analysis of wild-type, Cys75Ala, and Cys93Ala hNTAN1. CD spectra were collected at 25ºC using a Jasco-J815 CD spectrometer, and the spectra shown reflect 3 accumulations each. (A) Spectra of wild-type hNTAN1 (black dots), hNTAN1-C75A (blue dots), and hNTAN1-C93A (red dots). mdeg, millidegrees. (B) Summary of secondary structure analysis data for each variant as determined by the Jasco software.

The thermal stability of wild-type hNTAN1 and the point mutants Cys75Ala and Cys93Ala were determined by measuring the temperature dependence of their CD spectra in the far-UV region over a range of 20-80ºC (**Table 3.6**). All variants precipitated with increasing temperature indicative of irreversible thermal denaturation. Sigmoidal curves

fit to the CD signal at 210nm over this temperature range (**Figure 3.11**) indicated that both the wild-type and Cys75Ala variants exhibited a T_M of ~60ºC. However, in line with its distinguishing CD spectra and likely improper folding, corresponding analysis of the Cys93Ala data revealed a nearly 10ºC decrease in T_M (~52ºC) relative to the other two variants. Taken with the evaluation of its activity relative to wild-type, the data suggests that the Cys93Ala mutation likely imparts a destabilizing effect on protein folding.

Enzyme	T_M (ºC)
WT	60 ± 1
C75A	61 ± 1
C93A	52 ± 2

Table 3.6 T_M values of wild-type, Cys75Ala, and Cys93Ala hNTAN1. Equimolar amounts (15μM) of each enzyme in 50mM Tris-HCl, 150mM NaCl, 1mM EDTA, 0.5mM dTT, pH 7.5 were evaluated for thermal unfolding by observing CD signal at 200-260nm from 20 to 80ºC. The melting temperature (T_M) was determined as the inflection point of a sigmoidal curve fit to the data at 210nm (**Figure 3.11**).

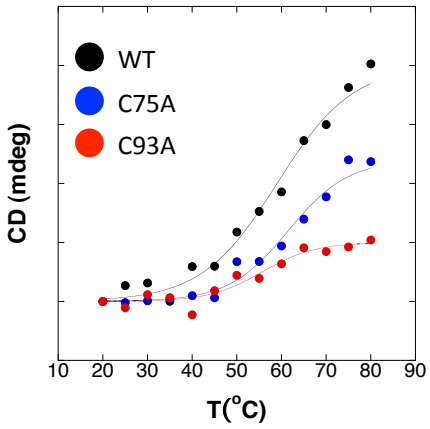


Figure 3.11 Thermal stabilities of wild-type, Cys75Ala, and Cys93Ala hNTAN1. Melting curves were determined using CD and measuring the signal (in mdeg) at 210nm in 5ºC intervals. Data were then fit to a sigmoidal curve and T_M was determined as the inflection point of each curve.

3.4 DISCUSSION

A number of biochemical and genetic studies over the past twenty years have begun to elucidate the implicated involvement of the N-end rule pathway in a number of physiological processes and have revealed many of the critical components involved in this subset of UPS-dependent proteolytic quality control. Proteolysis by the N-end rule pathway is dependent upon distinct types of degradation signals (degrons), including a set called N-degrons whose primary determinant is a destabilizing N-terminal residue. In vertebrates, N-terminal L-Asnand L-Gln can function as tertiary destabilizing residues of the N-end rule pathway through enzymatic deamidation by either NTAN1 (L-Asn-specific) or NTAQ1 (L-Gln-specific) to yield the secondary destabilizing N-terminal L-Asp and L-Glu, respectively.

NTAN1 has been purified from porcine liver [191] and was revealed to be a monomeric enzyme found predominantly in cytosolic fractions. In addition, the *Ntan1* gene has been cloned from a mouse cDNA library was then evaluated for complementation in *S. cerevisiase* Δ*nta1* [192]. Both studies demonstrated that NTAN1 was an L-Asn-specific Nt-amidase, though its sequence is dissimilar to those of other amidases, including the recently described NTAQ1 [172] and the *S. cerevisiase* Nta1 [170]. However, direct experimental evidence for the molecular mechanism of NTAN1 activity and its structural characteristics had not yet been demonstrated.

Though Western Blot and SDS-PAGE suggested soluble bacterial expression of human NTAN1 (hNTAN1) with an N-terminal affinity tag, the enzyme itself was minimally active. However, L-Asn-specific Nt-amidase activity was observed upon

expression of hNTAN1 with a C-terminal affinity tag, appearing to indicate that an unmodified hNTAN1 N-terminus is essential for catalysis. This point was further illustrated upon observing that post-translational exposure of hNTAN1 Pro2 upon MAP removal of the initiation Met was critical to activity. Optimization of the expression RBS and the nature of the C-terminal affinity tag for downstream purification ultimately allowed for the purification of ~0.2mg/L homogeneous hNTAN1 from *E. coli*.

Kinetic analysis revealed that the enzyme displays activity for the substrate N^1-AII, however, fails to deamidate L-Asn, L-Gln, Q^1-AII, or Ac-N^1-AII, corroborating the substrate selectivity described for NTAN1 purified from porcine liver [191]. However, relative to the kinetic parameters described for porcine NTAN1, we observed an approximately 10-fold lower K_M and 100-fold lower k_{cat} for the N^1-AII peptide, perhaps resulting from discrepancies in the quantitation of enzyme concentration or unknown enzyme cofactors not present in *E. coli*. Further, we observed a broad activity pH optimum in the range 6.0-7.5, with a sharp decline in activity upon shifting to more alkaline pH – indicative of a requirement for the correct ionization state of a single group within hNTAN1 for catalysis. Interestingly, while the pH optimum reported for porcine NTAN1 was similar to the one reported here, the broad relationship of pH v. activity was slightly different outside of this range and further; the authors of that study suggested a lack of direct thiol group involvement in activity based on the sum of their observed pH optimum and the p*K*a values of residues in that range. However, we demonstrated that the alkylating agent iodoacetamide was able to completely inhibit hNTAN1 activity according to an exponential decay across increasing concentrations of iodoacetamide. This finding suggested the existence of an hNTAN1 nucleophilic Cys, which was strongly implicated to be Cys75 upon activity analysis of a panel of Cys to Ala hNTAN1

point mutants which were constructed based on strict Cys conservation at 4 positions across seven vertebrate NTAN1 sequences. This finding was further supported by the observation that the Cys75Ala point mutant variant displayed no obvious change in its CD spectra or thermal stability relative to wild-type hNTAN1, suggesting that abolishment of activity in the point mutant was strictly due to the removal of a thiol group directly involved in activity and not to modifications in protein folding. Therefore, similar to both Nta1 and NTAQ1, NTAN1 appears to possess a critical nucleophilic Cys. The identification of this residue represents a key step towards understanding the molecular mechanism by which NTAN1 functions.

Though the activity and implicated involvement of NTAN1 and other components of the N-end rule pathway have been described for a variety of cellular processes, very little is known about: (i) the identity of actual physiological substrates targeted by the pathway, (ii) the exact molecular mechanisms or interactive processes underlying the observed N-end rule-targeted phenotypes in various genetic studies (including those of *Ntan1*$^{-/-}$ mice [171]), and (iii) the catalytic mechanism and nature of the active site structure of enzymes such as NTAN1 and NTAQ1 which share little sequence similarity to other amidases.

While it is conceivable to scan the proteome for amino-terminal sequences of the type Met-Asn (or Met-Gln), the prediction of physiological substrates targeted by NTAN1 (or NTAQ1) is inherently complicated due to the potentially vast contribution of endoproteolysis in generating N-terminal degrons of these types. This point was undoubtedly illustrated by the discovery of NTAN1-mediated degradation of *Drosophila* inhibitor of apoptosis 1 (DIAP1) in the regulation of apoptosis [189]. DIAP1 is cleaved by a caspase to generate a C-terminal fragment with N-terminal L-Asn, which is then

targeted through the NTAN1-ATE1 pathway in ultimately generating a primary destabilizing N-terminal Arg for subsequent substrate protein degradation. To date, DIAP1 remains the only known substrate of N-terminal Asn deamidation.

Recombinant expression and purification of N-end rule pathway components such as hNTAN1 should afford discovery of their individual molecular mechanisms through both rapid isolation and characterization, and more broadly, allow for the *in vitro* reconstruction of the pathway for the elucidation of potential physiological substrates. Ultimately, this sort of modeling may provide a means towards the biochemical and genetic dissection of the phenotypic results observed in *Ntan1*$^{-/-}$ and other N-end rule-targeted knockout mice. Further, recombinant expression of the human homologues of these components could present an opportunity to specifically investigate mechanisms and features of human pathological states associated with protein misfolding and aggregation, some of which may be associated with the N-end rule pathway.

Lastly, expression of N-end rule components either individually or in circuitous fashion in *E. coli* may afford the creation of novel *in vivo* screening and selection methods towards the identification of physiological substrates, with little genetic manipulation necessary given the near orthogonal hierarchial structure of the prokaryotic N-end rule pathway relative to eukaryotes [163].

Chapter 4: Therapeutic Enzyme Deimmunization by Combinatorial T-cell Epitope Removal Using Neutral Drift

4.1 INTRODUCTION

A variety of genetic and acquired human diseases can be treated by the systemic administration of enzymes catalyzing the depletion of metabolites that contribute to pathological states. Recombinant human enzymes are used extensively as replacement therapy for lysosomal storage disorders such as Gaucher's, Fabry's, and Pompe disease[1]. However, there are many diseases for which a human enzyme displaying the requisite catalytic and pharmacological properties for clinical use is unavailable. Therefore, heterologous enzymes, primarily of bacterial origin, have been evaluated for the treatment of a variety of disorders including phenylketonuria (PKU) [197], gout [198], and a number of cancers that are sensitive to enzyme-mediated, systemic depletion of amino acids. Examples of the latter include a large fraction of hepatocellular carcinomas and metastatic melanomas that become apoptotic under conditions where the non-essential amino acid L-Arg in serum is depleted[199], central nervous system cancers that respond to L-Met deprivation[200], and acute lymphoblastic leukemia (ALL) for which enzyme- mediated L-Asn depletion is a critical step in the current clinical treatment approach [76, 77, 79, 201].

A major impediment in the therapeutic application of heterologous enzymes is their immunogenicity, which results in the generation of anti-enzyme antibodies that in turn mediate a variety of adverse effects including hypersensitivity reactions, anaphylactic shock, and the inactivation and clearance of the enzyme itself[202]. Masking of immunogenic epitopes via covalent modification with polyethylene glycol

(PEG) can reduce protein immunogenicity [203], but eventually, antigen-specific and even PEG-specific, antibodies that contribute to therapeutic neutralization may be elicited towards PEGylated proteins [204, 205].

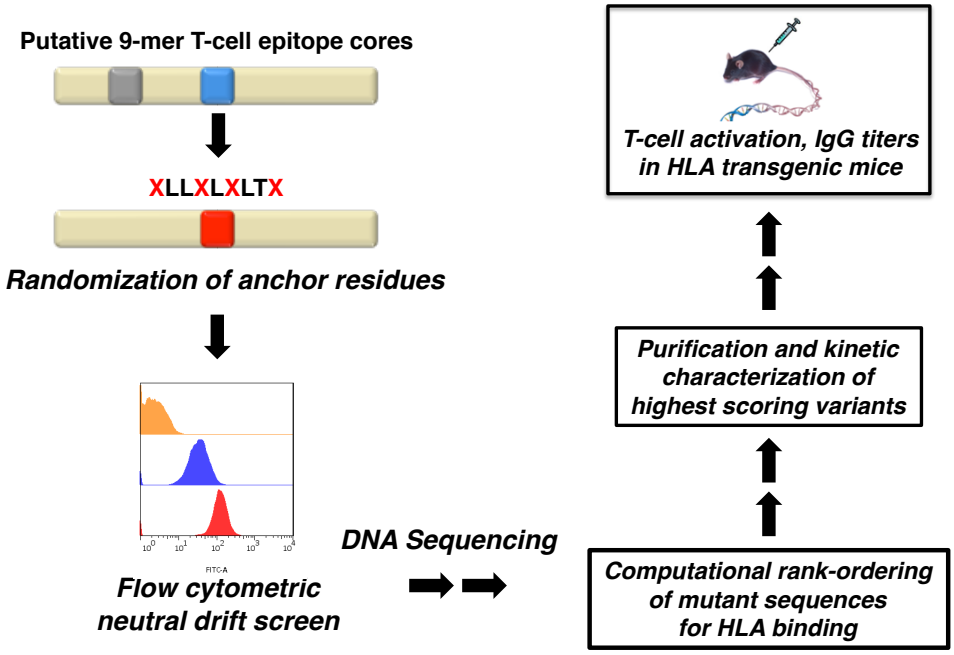
Protein immunogenicity can be ameliorated by mutating sequences likely to be recognized by the naïve antibody repertoire (B-cell epitopes) or sequences that are bound by the major histocompatability complex (MHC)-II and thus can elicit T cell-dependent (Td) immune responses. However, the identification and removal of B-cell epitopes is exceedingly difficult given their conformational nature, and is further complicated by our incomplete knowledge of the naïve antibody repertoire, and how they vary across different human populations. In contrast, there is extensive evidence from animal models, *in vitro* experiments, and early stage clinical studies, that the disruption of T-cell epitopes can reduce antibody responses in some therapeutic proteins[63, 64]. T-cell receptors on CD4+ T-cells recognize antigenic peptides (typically 13-25mers) presented in complex with MHC-II molecules on the surface of antigen presenting cells (APCs). The MHC-II binding groove contains four well-defined pockets that accommodate the side chains of the P1, P4, P6, and P9 residues within a core 9mer region of the T-cell epitope and these key residues largely determine binding affinity and specificity[206]. Although the MHC-II locus is highly polymorphic, assays using APCs from volunteers representative of the major MHC-II haplotypes in human populations have been deployed successfully to identify T-cell epitopes that contribute to protein immunogenicity in a large fraction of patients. Alternatively, a plethora of *in silico* methods[207] have been developed over the past several years for the prediction of sequences that bind to various MHC-II alleles.

The removal of T-cell epitopes by mutagenesis has been used with some success in reducing the immunogenicity of humanized and chimeric antibodies[65-67]. However, whereas these proteins contain, at most, a few relatively short potentially immunogenic sequences, heterologous enzymes that have not undergone immunological tolerance induction typically contain multiple T-cell epitopes, the removal of which thus necessitates extensive alteration of the polypeptide sequence in a manner that does not affect protein function. Further, enzyme catalysis is dictated not only by the active site residues, but also on a network of amino acids distributed throughout the protein[208]. For this reason, the introduction of multiple amino acid substitutions that disrupt MHC-II binding but do not affect catalytic activity represents a significant challenge. This is particularly problematic when deimmunization requires the replacement of amino acids that are phylogenetically conserved and consequently, substitutions at these positions could impact protein stability or catalytic efficiency.

In this chapter we take advantage of the evolutionary biology concept of neutral drift [209, 210] for the combinatorial deimmunization of a therapeutic enzyme without loss of function. Neutral drift refers to the accumulation of mutations under selective conditions that do not ultimately impact protein function. For deimmunization, putative T-cell epitopes are first identified computationally (or experimentally), key residues important for MHC-II binding are subjected to combinatorial randomization, and the resulting libraries are subjected to a neutral drift screen to isolate variants that retain wild-type (WT) function. The pools of neutral drift variants are evaluated for MHC-II binding and those that display scores indicative of compromised binding are purified and characterized biochemically. Finally, T-cell activation assays and antibody titers in transgenic mice homozygous for disease associated HLA alleles are used to evaluate T-

cell epitope removal and immunogenicity, respectively (**Figure 4.1A**). Although in this work we employed computational prediction of T-cell epitopes, the neutral drift screening methodology can be coupled to the experimental detection of sequences likely to bind MHC-II, using either haplotyped human peripheral blood mononuclear cell (PBMC) pools or relevant HLA-transgenic animals.

(A)



(B)

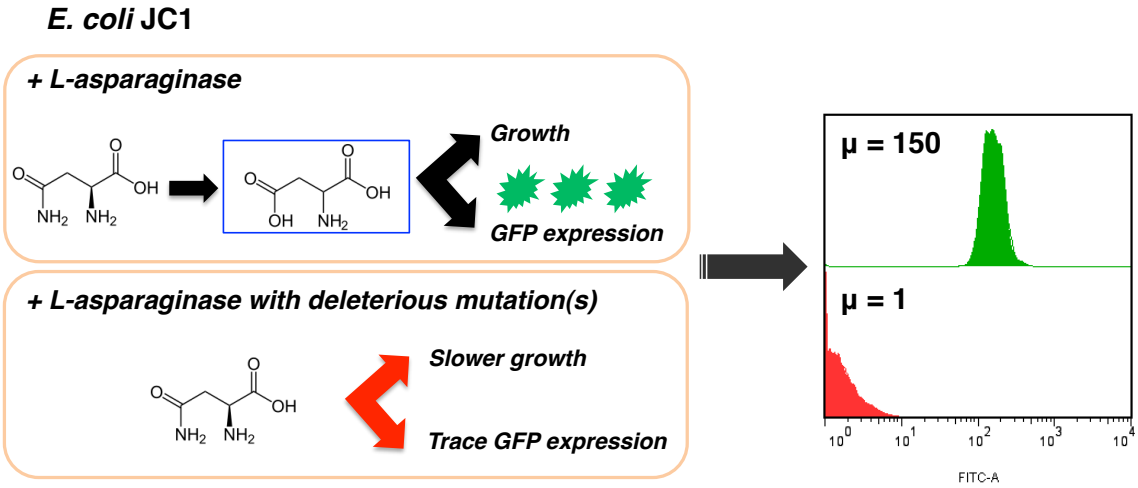


Figure 4.1 Deimmunization by combinatorial T-cell epitope removal using neutral drift. (A) Methodology for combinatorial T-cell epitope removal by neutral drift and subsequent evaluation of isolated variants using HLA-transgenic mice. (B) High-throughput neutral drift FACS screen for L-Asparaginase. μ: geometric mean fluorescence.

The implementation of the strategy described above is critically dependent upon a high-throughput technology for the rapid isolation of mutations that have minimal or no effect on function. Unfortunately, neutral drift screens for most enzymes – let alone those of therapeutic significance – have not been developed, necessitating the use of surrogate screening methods that interrogate proteins for stability and expression rather than catalytic function [210]. Manual assays, e.g. using a 96-well microtiter plate format, do not afford sufficient throughput for most purposes, while genetic selections based on complementation of auxotrophic strains to growth on selective media lacks a necessary degree of quantitation. Thus, in many instances, the expression of clones displaying significant differences in catalytic activity does not result in noticeable differences in colony formation [211]. Further, the extreme adaptability of biological systems can lead

to growth via mechanisms that bypass the action of the expressed heterologous protein [212-214].

We have developed a simple and robust neutral drift screen readily applicable to a variety of important therapeutic enzymes that catalyze the depletion of amino acids or other metabolites important for disease states. **Figure 4.1B** presents a schematic of the neutral drift screen as applied to the chemotherapetic enzyme L-Asparaginase II (EcAII, EC 3.5.1.1). EcAII has been a cornerstone component of chemotherapeutic protocols for the treatment of ALL for over 40 years [69-74, 215, 216]. In ALL, lymphoblasts lack or express low levels of L-asparagine synthetase (AS) and therefore require the uptake of L-Asn from serum for cell proliferation [76]. EcAII catalyzes the hydrolysis of L-Asn to L-Asp and ammonia with k_{cat}/K_M = 3.3 x 10^6 $M^{-1}s^{-1}$ (as calculated in this study) resulting in the systemic depletion of serum L-Asn [78, 79, 81, 82], which in turn induces apoptosis of ALL lymphoblasts [83-85]. However, antibody responses to EcAII have been reported in up to 60% of patients [116]. The primary strategy for managing the disease in patients with adverse immune responses to EcAII is treatment with the *Er. caratovora* L-Asparaginase II, which although is non-cross-reactive with anti-EcAII antibodies[217], is also highly immunogenic and clinically inferior to EcAII with respect to both event-free survival and overall survival rates at 6 years [218].

4.2 MATERIALS AND METHODS

Strains and Plasmids

Chromosomal gene deletions were performed using the λ-red recombinase system [219]. The asparaginase genes *ansA*, *ansB, iaaA*, the aspartate aminotransferase gene *aspC,* and the tyrosine aminotransferase gene *tyrB* were deleted from the chromosome of *E. coli* MC1061 (*F⁻ Δ(ara-leu)7697 [araD139]$_{B/r}$ Δ(codB-lacI)3 galK16 galE15 λ⁻ e14⁻ mcrA0 relA1 rpsL150(strR) spoT1 mcrB1 hsdR2(r⁻m⁺)*) resulting in *E. coli* JC1. Primers used for the deletion of each gene are found in **Table 4.1**. Briefly, each primer pair was used to amplify a PCR fragment containing the kanamycin resistance cassette of pKD13. Subsequently, the linear PCR product was used to replace the entire ORF of the targeted gene on the MC1061 chromosome. Colonies containing the correct gene deletions were transformed with the FLP recombinase plasmid pCP20 to remove the kanamycin resistance marker, and the pCP20 was then cured from the resulting strain as described previously[219]. The genes *ansA, ansB,* and *iaaA* were also deleted from the *E. coli* strain BL21 (DE3) (*F⁻ ompT gal dcm lon hsdS$_B$(r$_B$⁻ m$_B$⁻) λ(DE3 [lacI lacUV5-T7 gene 1 ind1 sam7 nin5*]) chromosome, resulting in *E. coli* JC2, used to express the EcAII variants. Where necessary, gene deletions were transferred to recipient strains via P1 transduction.

Primer Name	Nucleotide Sequence (5' → 3')
ansBFor	GTGCAGCACATATGTTACCCAATATCACCA
ansBRev	GGCGGGATCCTTAGTACTGATTGAAGA
T12AFor	GTGCAGCACATATGTTACCCAATATCACCATTTTAGCAACCGGCGGGGCCA
T7term	GCTAGTTATTGCTCAGCGG
tetFor	CTAGCTAGTCTAGAGCGGAGCCTATGGAAAAACGC
tetRev	CATGCCATGGCACTTTTCTCTATCACTGATAGGG
KanFor	GGGCCGGCAGGCGCTCCATTGCCCAGTCGGCTAAGGGATTTTGGTCATGAAC
KanRev	GCGTAGCGACCGAGTGAGCTAGCTTATTAGAAAAACTCATCGAGCATC
5' ansA KO	CCTCACGTATATACTTTTGCTCTTTCGATATCATTCATATCAATATCATGATTCCGGGGATCCGTCGACC
3' ansA KO	ACAGGGCGCGAGGGGGCATTACAGTCTCCTTAATCATCCGGCGTCAGTTCTGTAGGCTGGAGCTGCTTCG
5' ansB KO	CAGAGCTAAGGGATAATGCGTAGCGTTCACGTAACTGGAGGAATGAAATGATTCCGGGGATCCGTCGACC
3' ansB KO	AGCCCCGGCACGATACCGGGGCGAGGCGATTAGTACTGATTGAAGATCTGTGTAGGCTGGAGCTGCTTCG
5' iaaA KO	TGATATTTATAGCAAAAGTGGCGAACCACCCTTAATGGACGAATACTATGATTCCGGGGATCCGTCGACC
3' iaaA KO	CCGCCAGCACATTACCGGCATCAAGTTCATCACTGTGTGGCAACGGTGTCTGTAGGCTGGAGCTGCTTCG
5' aspC KO	TACCCTGATAGCGGACTTCCCTTCTGTAACCATAATGGAACCTCGTCATGATTCCGGGGATCCGTCGACC
3' aspC KO	TTTTCAGCGGGCTTCATTGTTTTTAATGCTTACAGCACTGCCACAATCGCTGTAGGCTGGAGCTGCTTCG
5' tyrB KO	GTTTATTGTGTTTTAACCACCTGCCCGTAAACCTGGAGAACCATCGCGTGATTCCGGGGATCCGTCGACC
3' tyrB KO	GCTGGGTAGCTCCAGCCTGCTTTCCTGCATTACATCACCGCAGCAAACGCTGTAGGCTGGAGCTGCTTCG
5' MSSA-NNS	GGTCGGCGCANNSCGTCCGNNSACGNNSATGAGCNNSGACGGTCCATTCAACCTG
3' MSSA-NNS	CAGGTTGAATGGACCGTCSNNGCTCATSNNCGTSNNCGGACGSNNTGCGCCGACC
5' INAS-NNS	CTGCCGAAAGTCGGCNNSGTTTATNNSTACNNSAACGCANNSGATCTTCCGGCTAAAGCACTG
3' INAS-NNS	CAGTGCTTTAGCCGGAAGATCSNNTGCGTTSNNGTASNNATAAACSNNGCCGACTTTCGGCAG
5' VQAQ-NNS	CCGCAAAAAGCGCGCNNSCTGCTGNNSCTGNNSCTGACGNNSACCAAAGATCCGCAGCAG
3' VQAQ-NNS	CTGCTGCGGATCTTTGGTSNNCGTCAGSNNCAGSNNCAGCAGSNNGCGCGCTTTTTGCGG
pCTKFor	CGATCAAACCACCTCCCCAGGTGGTTTTTTCGTTTACAGGGC
5' G57A	GGTAGTGAATATCGCGTCCCAGGACATGAACG
5' G57V	GGTAGTGAATATCGTGTCCCAGGACATGAACG
5' G57L	GGTAGTGAATATCCTGTCCCAGGACATGAACG
3' G57mp	GGTGACGTCACGGCCATCAAGCACGGTGTCATTCATC

Table 4.1 List of primers used for EcAII deimmunization.

Plasmids	Relevant Characteristics	Reference or Source
pET-28a	Kan^R , T7 promoter	Novagen
pHisEcAII	Kan^R , T7 promoter, encodes His6x-EcAII	This study
pET-26b	Kan^R , T7 promoter	Novagen
pPelbHisEcAII	Kan^R , T7 promoter, encodes sspelb-His6x-EcAII	This study
pPelbHisT12A	Kan^R , T7 promoter, encodes sspelb-His6x-EcAII(T12A)	This study
pCTK	Kan^R , *tet* promoter, CloDF ori	This study
pASK75	Amp^R, *tet* promoter	220
pCDF-1b	Sp^R, T7 promoter, CloDF ori	Novagen
pCTK-EcAII	encodes sspelb-His6x-EcAII in pCTK	This study
pCTK-T12A	encodes sspelb-His6x-EcAII(T12A) in pCTK	This study
pCTK-G57A	encodes sspelb-His6x-EcAII(G57A) in pCTK	This study
pCTK-G57V	encodes sspelb-His6x-EcAII(G57V) in pCTK	This study
pCTK-G57L	encodes sspelb-His6x-EcAII(G57L) in pCTK	This study
p28pelHisEcAII	encodes sspelb-His6x-EcAII in pET-28a	This study
p28pelHis1.1.C4	encodes sspelb-His6x-1.1.C4 in pET-28a	This study
p28pelHis2.2.G10	encodes sspelb-His6x-2.2.G10 in pET-28a	This study
p28pelHis3.1.E2	encodes sspelb-His6x-3.1.E2 in pET-28a	This study
pQE80L-GFP(11.3.3)	Amp^R, lac promoter, encodes GFP (11.3.3)	221

Table 4.2 List of plasmids used for EcAII deimmunization. pASK75 [220] pQE80L-GFP (11.3.3) [221]

All primers and plasmids used in this study are described in **Tables 4.1** and **4.2,** respectively. PCR reactions were carried out using *Vent* DNA polymerase (New England Biolabs) and oligonucleotides were synthesized by Integrated DNA Technologies. The *ansB* gene (mature sequence only) was amplified from the genomic DNA of *E. coli* K12

using the primers ansBFor/ansBRrev, digested with *NdeI-BamHI*, and cloned into pET-28a to generate plasmid pHisEcAII. Subsequently, plasmid pPelBHisEcAII was generated through subcloning the *NcoI-BamHI* digested fragment from pHisEcAII into pET-26b. In addition, a plasmid coding EcAII-T12A was generated by PCR using the pPelBHisEcAII plasmid as template and the primer pair T12AFor/T7 term, resulting in plasmid pPelBHisT12A.

To construct vector pCTK, firstly, the *tet* promoter region from vector pASK75 was amplified using the primers tetFor/tetRev, digested with *Xba*I-*Nco*I, and ligated into pCDF-1b. Next, the kanamycin resistance cassette from vector pET-28a was amplified using the primers KanFor/KanRev, digested with *BglI-BmtI*, and then cloned into pCDF-1b as well, ultimately generating the final pCTK vector.

Plasmids pCTK-EcAII and pCTK-T12A were generated by subcloning the *NcoI-NotI* digested fragments of pPelBHisEcAII and pPelBHisT12A respectively, into vector pCTK. Plasmids pCTK-G57A, pCTK-G57V, and pCTK-G57L were constructed using a 2-step protocol based on the QuickChange methodology. In the first step, megaprimers for each G57 mutant were amplified by PCR from the plasmid pCTK-EcAII using the primer pair 5'-G57X (X = A, V, L)/3'-G57mp. In the second step, the megaprimer was used in place of outside primers in a PCR reaction again using plasmid pCTK-EcAII as the template with the following cycling parameters: 95ºC-2 min, 16 cycles of 95ºC-30 sec, 55ºC-1 min, 72ºC-10 min, and a final polishing step at 72ºC-15 min. Each product was then digested with *DpnI* for 1 hr at 37ºC to eliminate the initial template plasmid.

In silico identification and characterization of EcAII T-cell epitopes

The primary sequence corresponding to the mature region of EcAII was screened for putative T-cell epitopes using the IEDB consensus prediction method. The 326 amino acid sequence was parsed into overlapping 15-mer peptide fragments and within each fragment, a 9-mer core region was identified and scored for predicted binding by a consensus percentile rank (CPR) in which a lower score (in arbitrary units) was indicative of a higher predicted binding affinity. Because the consensus method scoring is based on the outputs of individual MHC-II binding prediction methods, multiple 9mer cores were identified in some 15mer fragments. In these instances, the 9mer core selected was the one predicted by TEPITOPE (Sturniolo) [222], which served as the basis for ProPred [223] – the most accurate algorithm for epitope core identification among those evaluated by the developers of the IEDB consensus method [207]. Binding was further evaluated for 7 additional HLA-DR alleles which when taken with DRB1*0401 cover nearly 95% of human populations worldwide [224]. Three 9mer core regions that were scored with a consensus percentile rank (CPR) falling within the lowest 10% of the parsed peptide fragments as determined for binding to DRB1*0401 (CPR < 2) and that further received equivalently low scores for at least one other DRB1 allele were selected for neutral drift combinatorial mutagenesis.

Library construction

Oligonucleotides encoding degenerate NNS (N is A, T, G, C; S is G, C) codons at the sites corresponding to residues in positions P1, P4, P6, and P9 of each of the three 9mer core sequences chosen for mutagenesis were used for library construction and can be found in **Table 4.1**. For the first library, PCR with *Vent* DNA polymerase and pCTK-

EcAII as template was carried out to generate two fragments from the primer pairs pCTKFor/3' MSSA-NNS and 5' MSSA-NNS/T7term respectively. The DNA fragments obtained from these PCRs were electrophoresed and purified using a QIAGEN gel purification kit. Equimolar quantities of the two fragments were then mixed and subjected to overlap-extension PCR using the primers pCTKFor/T7term. The resulting 1.5-kb PCR product was digested with *NcoI-NotI* and ligated into pCTK-T12A digested with the same enzymes. The ligation mixture was then transformed into electrocompetent *E. coli* JC1 encoding plasmid pQE80L-GFP (11.3.3) [221], yielding ~ 10^7 individual transformants. The clones were pooled and stored in 15% glycerol at -80ºC in aliquots. The second and third libraries were constructed analogously, using the internal primers 5' INAS-NNS/3' INAS-NNS or 5' VQAQ-NNS/3' VQAQ-NNS in place of 5' MSSA-NNS/3' MSSA-NNS respectively.

FACS screening

M9 medium supplemented with 0.4% glucose, 3.5μg/mL thiamine, 1mM $MgSO_4$, 0.1mM $CaCl_2$, 160μg/mL of the amino acids L-Asp and L-Tyr, 80μg/mL of the 18 remaining amino acids, 30μg/mL kanamycin, and 200μg/mL ampicillin was inoculated with a frozen aliquot of *E. coli* JC1 transformed with pQE80L-GFP (11.3.3) [225] and either a library or a single mutant. Cultures were grown at 37°C to an A_{600} = 0.9-1.1, harvested by centrifugation (6000*xg*, 4ºC, 6 min), and washed twice with cold 0.9% NaCl. The cell pellets were resuspended in supplemented M9 medium containing 19 amino acids (no L-Asp, L-Tyr at 160μg/mL, remaining amino acids at 80μg/mL). GFP expression was induced following the media shift by addition of isopropyl-1-thio-β-D-galactopyranoside (IPTG) to a final concentration of 1mM. After 1 hr induction at 37°C, the cells were

harvested by centrifugation (6000*xg*, 4°C, 6 min), washed twice with PBS, and resuspended in PBS to a final A_{600} ~ 0.05-0.1 for flow cytometric analysis and cell sorting.

Flow cytometric analyses were performed with a FACSAria (BD Biosciences) using a 488nm solid-state laser for excitation and a 530/30 band pass filter for detection. The throughput rate of cells was adjusted to 4,000-5,000 events per second and ~10^7 cells were sorted each round in single cell mode except for the initial sort of each library which was done in purity mode. A gate in the fluorescence channel was set to recover the 4-5% most highly fluorescent cells, while additional gates were set based on both the forward- and side- scatter channels to exclude sorting non-single cell events. The sorted cells were collected in 0.5mL of 2xYT medium and then plated onto 2xYT medium supplemented with 30μg/mL kanamycin and 200μg/mL ampicillin. Following overnight growth at 30ºC, the clones were pooled and stored in 15% glycerol at -80°C in aliquots.

Colorimetric asparaginase activity assay

A colorimetric asparaginase assay using L-Aspartic Acid β-hydroxomate (AHA) was used to isolate active asparaginase clones from the final FACS-sorted population of each library. Following the final round of sorting, the polyclonal gene cassette of the collected population was amplified using primers pCTKFor/T7term, digested with *NcoI-NotI*, and subcloned into pET-28a digested with the same restriction enzymes. The ligation mixture was transformed into electrocompetent *E. coli* JC2 and single colonies were used to inoculate 90μL 2xYT supplemented with 30μg/mL kanamycin over two 96-well plates. Following 2 hr incubation with shaking at 350rpm at 37ºC, protein expression was induced by adding an additional 90μL 2xYT (30μg/mL kanamycin, 1mM IPTG) to each

well. After 3 hr induction at 37ºC, 120μL from each well was transferred to a fresh 96-well plate and cells were harvested by centrifugation (3500*xg*, 4ºC, 10 min). The cell pellets were then resuspended in 120μL B-PER Bacterial Protein Extraction Reagent (Thermo Scientific), incubated at 25ºC for 20 min with shaking at 350rpm, and subsequently pelleted (3500*xg*, 4ºC, 15 min). The resulting supernatants were transferred to a Ni^{2+}-NTA HisSorb Plate (QIAGEN) and stored at 4ºC overnight. After decanting the supernatant and rinsing twice with wash buffer (50mM Tris-HCl, 100mM NaCl, 25mM imidazole, pH 8), 50μL of 10mM AHA in activity buffer (50mM Tris-HCl, 100mM NaCl, pH 7.4) was added to each well. Following incubation with substrate at 25ºC for 20 min, 50μL color reagent (2% 8-hydroxyquinoline in ethanol/1M Na_2CO_3, 1:3 by volume) was added to each well and the plate was covered, heated in a 100ºC oven for 90 sec, and allowed to cool at 4ºC for 15 min. Activity within each well was quantitated by measuring the absorbance at 705nm (Synergy HT Fluorescent Platereader, BioTek).

Expression and purification of EcAII variants

E. coli JC2 harboring pET-28a encoding either WT EcAII or an isolated mutant EcAII (p28pelHisEcAII, p28pelHis1.1.C4, p28pelHis2.2.G10, and p28pelHis3.1.E2) were cultured overnight at 37ºC in 2xYT medium supplemented with 30μg/mL kanamycin and used to inoculate 250mL fresh medium (1:100). When the A_{600} reached 0.5-0.7, the cells were transferred to 25ºC and allowed to equilibrate for 20 min, at which point the culture was supplemented with IPTG to a final concentration of 1mM to induce protein expression. After 16 hr of incubation at 25ºC, the cells were harvested by centrifugation at 10,000*xg* for 10 min. The cell pellet was resuspended in binding buffer (50mM Tris-HCl, 100mM NaCl, 10mM imidazole, pH 8), lysed by three passes through a French

pressure cell, and subsequently pelleted at 40,000*xg* for 45 min. The resulting supernatant (soluble fraction) was decanted, diluted 1:1 in binding buffer and mixed with 0.5mL of pre-equilibrated nickel-nitrilotriacetic acid (Ni^{2+}-NTA) resin. After incubation for 90 min at 4ºC with gentle rotation, the solution was applied to a 5mL polypropylene column. The resin was then washed with 25 bed volumes binding buffer and 25 bed volumes wash buffer (50mM Tris-HCl, 100mM NaCl, 25mM imidazole, pH8) before the resin was incubated with 4mL elution buffer (50mM Tris-HCl, 100mM NaCl, 250mM imidazole, pH 8) for 10 min and collected drop-wise. The eluted fractions were concentrated using an Amicon Ultra 10K MWCO filter and purified into PBS by gel filtration on a Superdex 200 column (Amersham Pharmacia).

Determination of kinetic parameters

The kinetics of L-Asn hydrolysis were determined with freshly purified enzyme as described previously [98]. Briefly, reactions of each asparaginase variant (10-20nM enzyme) with L-Asn (0 to 5 x K_M) were carried out at 37ºC in 50mM Tris-HCl, 100mM NaCl (pH 7.4) in a total volume of 100μL, and were subsequently quenched with 5μL of 12% (w/v) trichloroacetic acid. An aliquot of the quenched reaction mixture was then mixed with a molar excess (relative to substrate) of *o*-phtalaldehyde (OPA) reagent and brought to a final volume of 100μL with borate buffer. The resulting solutions were analyzed by HPLC using an Agilent ZORBAX Eclipse AAA Column (C18 reverse phase, 5μm, 4.6mm x 150mm). All reactions were conducted at least in triplicate and the observed rates were fit to the Michaelis-Menten equation using Kaliedagraph (Synergy).

Asparaginase serum stability

Approximately 250μg of WT EcAII or variant 3.1.E2 was mixed with 1mL pooled human serum (Innovative Research) and incubated at 37ºC. At various time points, a 15μL aliquot from each sample was removed and used to set up triplicate reactions in which 5μL per aliquot was added to 50μL 10mM AHA in activity buffer in a microtiter plate. After allowing the reaction to proceed for 5min at room temperature, 50μL color reagent was added to each well, the plate was covered, heated in a 100ºC oven for 90 sec, and then allowed to cool at 4ºC for 15 min. Activity within each well was quantitated by measuring the absorbance at 705nm (Synergy HT Fluorescent Platereader, BioTek) and the average values for each enzyme-containing sample were then normalized by subtracting the average value measured for the control sample at the same time point.

Antigen preparation and EcAII peptide library

Endotoxin contamination of purified enzymes was reduced by a previously described phase separation technique using the detergent Triton X-114 [226]. Following 6-8 phase separation cycles, protein was buffer exchanged against sterile, commercially purchased 1x PBS (Gibco) using an Amicon Ultra 10K MWCO filter to remove any detergent that may have persisted within the solution. Enzyme preparations treated by this procedure retained normal activity. The endotoxin levels of the enzyme were determined by Limulus Amebocyte Lysate (LAL) assay and observed to be < 7 endotoxin units/100μg protein. A collection of 32 overlapping 20mer EcAII peptides (**Table 4.3**), staggered by 10 amino acids and spanning the entire primary sequence, were synthesized by GenScript. An additional set of peptides corresponding to sequences containing the engineered mutations in 3.1.E2 were also synthesized (Abgent).

Peptide Name	Peptide Sequence
WTp1-20	MLPNITILATGGTIAGGGDS
WTp11-30	GGTIAGGGDSATKSNYTVGK
WTp21-40	ATKSNYTVGKVGVENLVNAV
WTp31-50	VGVENLVNAVPQLKDIANVK
WTp41-60	PQLKDIANVKGEQVVNIGSQ
WTp51-70	GEQVVNIGSQDMNDNVWLTL
WTp61-80	DMNDNVWLTLAKKINTDCDK
WTp71-90	AKKINTDCDKTDGFVITHGT
WTp81-100	TDGFVITHGTDTMEETAYFL
WTp91-110	DTMEETAYFLDLTVKCDKPV
WTp101-120	DLTVKCDKPVVMVGAMRPST
WTp111-130	VMVGAMRPSTSMSADGPFNL
WTp121-140	SMSADGPFNLYNAVVTAADK
WTp131-150	YNAVVTAADKASANRGVLVV
WTp141-160	ASANRGVLVVMNDTVLDGRD
WTp151-170	MNDTVLDGRDVTKTNTTDVA
WTp161-180	VTKTNTTDVATFKSVNYGPL
WTp171-190	TFKSVNYGPLGYIHNGKIDY
WTp181-200	GYIHNGKIDYQRTPARKHTS
WTp191-210	QRTPARKHTSDTPFDVSKLN
WTp201-220	DTPFDVSKLNELPKVGIVYN
WTp211-230	ELPKVGIVYNYANASDLPAK
WTp221-240	YANASDLPAKALVDAGYDGI
WTp231-250	ALVDAGYDGIVSAGVGNGNL
WTp241-260	VSAGVGNGNLYKSVFDTLAT
WTp251-270	YKSVFDTLATAAKTGTAVVR
WTp261-280	AAKTGTAVVRSSRVPTGATT
WTp271-290	SSRVPTGATTQDAEVDDAKY
WTp281-300	QDAEVDDAKYGFVASGTLNP
WTp291-310	GFVASGTLNPQKARVLLQLA
WTp301-320	QKARVLLQLALTQTKDPQQI
WTp311-326	LTQTKDPQQIQQIFNQY

Table 4.3 Overlapping synthetic peptides of WT EcAII for T-cell activation assays.

Transgenic mice

HLA-DR4 (DRB1*0401) transgenic mice were generated as described previously [227] and bred under specific pathogen-free conditions at the University of Texas at San Antonio. Transgenic mice were injected at 6-10 weeks of age with the antigen in Complete Freund's Adjuvant (CFA). WT EcAII and variant 3.1.E2 were purified and

treated for endotoxin reduction (See Supplementary Methods). CFA was prepared by mixing Mycobacterium tuberculosis H37RA (Difco Laboratories, Detroit, MI) at 5mg/mL into IFA. Antigens were mixed with the adjuvant to yield a 2mg protein/mL emulsion, of which 50μL was injected subcutaneously as specified. Ten days later, popliteal and inguinal lymph nodes (LN) were removed and single cell suspensions were adjusted to 5 x 10^6 cells/mL in HL-1 media (BioWhittaker, Gaithersburg, MD.). Serum was obtained by terminal cardiac puncture. All animal care and experimental procedures were conducted according to guidelines of the Institutional Care and Use Committee (IACUC) at the University of Texas at San Antonio.

Cytokine measurements by ELISPOT and computer-assisted ELISPOT image analysis

Cytokine ELISPOT assays were performed as described previously [49]. Briefly, ELISPOT plates (Multiscreen IP, Millipore, Billerica, MA) were coated overnight with 2μg/mL IFN-γ-specific capture antibody (AN-18; eBioscience, San Diego, CA) diluted in PBS. The plates were blocked with 1% BSA in PBS for 1 hr at room temperature and then washed four times with PBS. LN cell suspensions were plated at 5 x 10^5 cells/well with either whole antigen or with EcAII overlapping peptides and incubated at 37ºC for 24 hr. Note that cells were plated with the whole antigen or overlapping peptides corresponding to the EcAII variant used to immunize the mouse from which they were isolated. Subsequently, the cells were removed by washing three times with PBS and four times with PBS/Tween, and IFN-γ-specific biotinylated detection Ab (R4-6A2; 0.5μg/mL, eBioscience) was added and incubated overnight. The plate-bound secondary antibody was then incubated with streptavidin-alkaline phosphatase (Dako, Carpinteria,

CA), and cytokine spots were visualized by 5-bromo-4-chloro-3-indolyl phosphate/NBT phosphatase substrate (KPL, Gaithersburg, MD). Image analysis of ELISPOT assays was performed on a Series 2 Immunospot analyzer and software (Cellular Technology, Cleveland, OH) as described previously. In brief, digitized images of individual wells of the ELISPOT plates were analyzed for cytokine spots based on the comparison of experimental (containing T-cells and APC with Ag or peptide) and control wells (T-cells and APC without Ag or peptide). Following the separation of spots that were touched or partially overdeveloped, nonspecific noise was gated out by applying spot size and circularity analysis as additional criteria. Spots that fell within the accepted criteria were highlighted and counted.

Detection of antigen-specific antibody titer by ELISA

Serum was obtained from mice immunized with either WT EcAII or 3.1.E2 by terminal cardiac puncture. Microtiter plates (eBioscience 44-2504-21) were coated overnight at 4ºC with 1μg of antigen (WT EcAII or 3.1.E2) in PBS and blocked for an additional 1 hr at room temperature with 1x assay diluent (eBioscience # 00-4202-56). Serial dilutions of sera were added to wells coated with the corresponding immunizing antigen and incubated for 2 hr at room temperature. The plates were washed and incubated with ImmunoPure goat anti-mouse IgG conjugated with horseradish peroxidase for 1 hr at room temperature. The plates were subsequently washed, incubated with tetramethylbenzidine (TMB) substrate for 15 min at room temperature before the reactions were stopped by addition of 2M H_2SO_4. The absorbance was read (450nm) using an ELISA microplate reader (μQuant; Biotek Instruments, Winooski, VT). End-

point titers were calculated by using an absorbance corresponding to a control well (PBS substituted for sera) as the cutoff value.

4.3 RESULTS

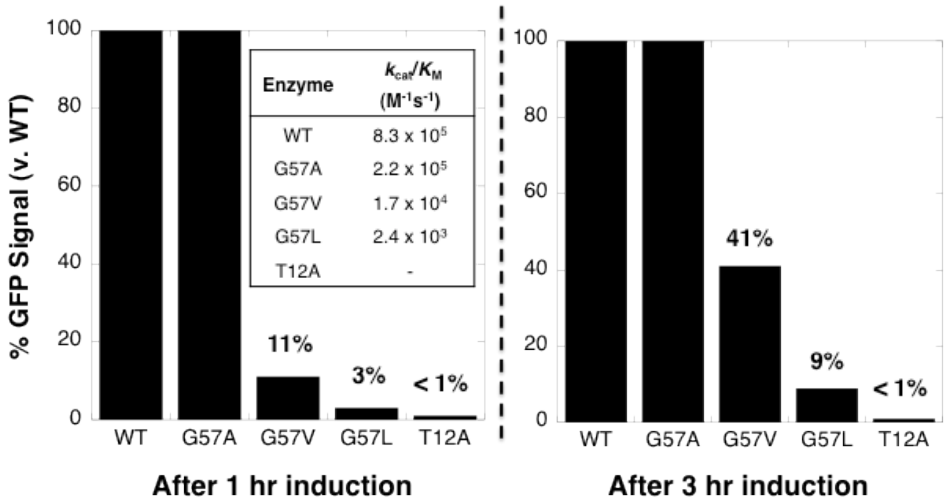
Development and validation of neutral drift screen

To develop a neutral drift screen for EcAII, we first constructed *E. coli* JC1 [MC1061 *ΔaspCΔtyrBΔansAΔansBΔiaaA*] in which the genes required for L-Asp biosynthesis (*aspC, tyrB*) and the three genes required for endogenous L-asparaginase enzymes (*ansA, ansB, iaaA*) were deleted. JC1 cells expressing a low level of recombinant EcAII formed normal size colonies when plated on minimal media plates with 19 amino acids (no L-Asp). In contrast, cells without plasmid or expressing the recombinant, inactive, EcAII-T12A point mutant formed pinpoint colonies, presumably because spontaneous hydrolysis of L-Asn provides a basal level of L-Asp for growth. The formation of pinpoint colonies by null mutants and the cross-feeding of L-Asp generated by low activity clones frustrated efforts to select neutral mutants on plates with selective media; in multiple attempts, similar size colonies formed by plating mutagenized enzyme were later found to encode enzymes with dramatically different L-Asn hydrolysis kinetics.

To enable an additional level of quantitation, cells were transformed with a plasmid expressing green fluorescent protein (GFP) under an isopropyl β-D-1-thiogalactopyranoside (IPTG) inducible promoter, grown to late exponential phase in media containing all 20 amino acids, washed and transferred for a short time period to media with 19 amino acids (no L-Asp) and IPTG to induce GFP synthesis. Following the addition of IPTG, GFP synthesis – and hence intracellular fluorescence – was

dependent on the availability of L-Asp, which in turn was proportional to EcAII enzymatic activity. Utilizing this assay, we found that intracellular GFP fluorescence correlated well with the activity of a panel of recombinantly expressed EcAII variants displaying up to two orders of magnitude differences in catalytic efficiency. For example, cells expressing EcAII-G57V exhibited a nearly 10-fold lower GFP fluorescence relative to cells expressing EcAII-G57A, which has approximately 15-fold higher catalytic efficiency [111] (**Figure 4.2A**). While the assay was able to easily eliminate low activity clones, EcAII-G57A (k_{cat}/K_M (L-aspartic acid β-hydroxomate (AHA)) = 2.4 x 10^6 $M^{-1}S^{-1}$) exhibited identical GFP fluorescence relative to WT EcAII (k_{cat}/K_M (AHA) = 8.2 x10^6 $M^{-1}S^{-1}$) indicating that the signal saturates for enzymes with k_{cat}/K_M within 3 to 4-fold of the WT enzyme catalytic efficiency. Nonetheless, enzymes with k_{cat}/K_M (L-Asn) > 10^6 M^{-1} s^{-1} would ideally be more than sufficient for therapeutic purposes. To validate the enrichment capabilities of the assay, three rounds of cell sorting produced a 6,000-fold enrichment of JC1 cells expressing WT EcAII from an initial mixture containing a 10,000-fold excess of *E.coli* JC1 expressing EcAII-T12A (**Figure 4.2B**).

(A)



(B)

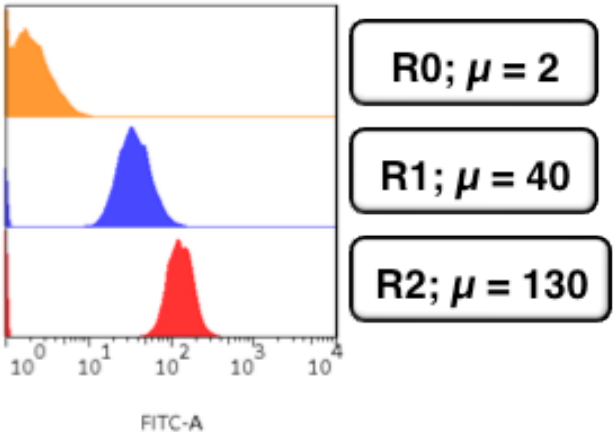


Figure 4.2 Validation of the neutral drift screen for cells expressing EcAII variants with high catalytic activity. (A) Relative GFP signal of a panel of *E.coli* JC1 cells expressing EcAII variants with different catalytic efficiencies for the hydrolysis of the L-Asn analog AHA [111] (B) Fluorescence histograms showing 3 round enrichment of JC1 cells expressing EcAII from a mixture containing a 1:10,000 excess of JC1 cells expressing EcAII-T12A. After 3 rounds of sorting, DNA sequencing revealed that 5 of 8 clones selected at random encoded EcAII. μ: geometric mean fluorescence.

Computational identification of putative EcAII T-cell epitopes

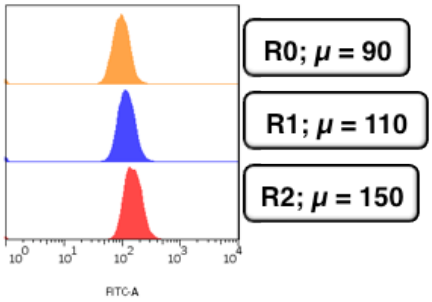
Putative EcAII T-cell epitopes were identified using the Immune Epitope Database (IEDB) consensus method [207]. The protein sequence was parsed into overlapping 15mer peptide fragments (staggered by one residue) and within each fragment, 9-mer core regions were scored for predicted binding first to HLA-DRB1*0401, which shows strong association with childhood ALL in males[228] , and then to an additional seven HLA-DR alleles that collectively cover nearly 95% of the human population [224]. Three 9-mer core regions scored with a consensus percentile rank (CPR) within the lowest 10% of the parsed peptide fragments for binding to DRB1*0401 (CPR < 2); the regions that further showed equivalently low scores for at least one other DRB1 allele were selected for T-cell epitope removal: M_{115}RPSTSMSA, I_{216}VYNYANAS, and V_{304}LLQLALTQ (designated M_{115}, I_{216}, and V_{304} where these three residues correspond to the respective P1 positions).

Neutral drift screening of EcAII libraries

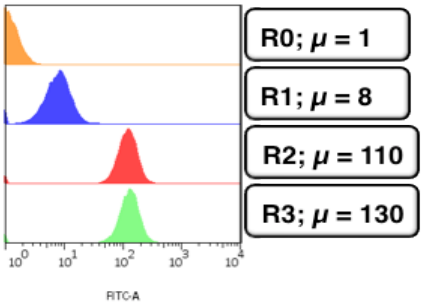
The P1, P4, P6, and P9 positions, which are most critical for the binding of peptides to the MHC-II binding groove [206], were subjected to saturation mutagenesis using the NNS (N = A, T, G, C; S = G, C) randomization scheme. Randomization and neutral drift screening were carried out sequentially, starting with M_{115} and continuing with I_{216} and finally V_{304} to: (i) ensure complete library coverage for each individual epitope; (ii) evaluate the relative plasticity of different regions of the protein to amino acid substitutions, and (iii) simplify the structural interpretation of any observed changes in the activity of isolated mutants.

10^7 transformants of the M_{115} library (predicted theoretical diversity ≈ 10.5^6) were subjected to 3 rounds of FACS screening until the mean cell fluorescence of the sorted population was comparable to that of cells expressing the WT enzyme (**Figure 4.3A**). In this instance, the high initial fluorescence of the library (μ = 90; **Figure 4.3A**) suggested that amino acid substitutions at the targeted sites were generally tolerated. Following the final round of sorting, 120 individual clones selected at random were assayed in microtiter well plates using the colorimetric asparaginase substrate AHA [229]. This secondary screen eliminated inactive clones that might have been inadvertently recovered during FACS enrichment. Active clones were found to display minor variations in AHA hydrolysis rate consistent with the notion that the FACS screen enriches variants with near WT activity. **Figure 4.4** shows the frequency of amino acid occupancy at M115, S118, S120, and A123. Interestingly, M115, which is absolutely conserved among the nearly 500 bacterial type II L-asparaginases in the database, could tolerate a variety of non-conservative substitutions. Analogous promiscuity was observed at both S120 and A123, which are also highly conserved phylogenetically. Evaluation of the isolated sequences using the IEDB consensus model revealed that the alteration of M_{115}RPSTSMSA to V_{115}RPPTRMSP results in over a 20-fold increase in CPR score for the DRB1*0401 allele as well as increases in the CPR scores for 5 other HLA-DR alleles (**Table 4.4**). The resulting enzyme variant, EcAII M115V/S118P/S120R/A123P (designated as clone 1.1.C4), having 4 amino acid substitutions, 3 of which were non-conservative, displayed catalytic properties (k_{cat} = 28 s^{-1}, K_M = 17μM, k_{cat}/K_M = 1.6 x 10^6 $M^{-1}s^{-1}$) nearly identical to those of the parental enzyme with only a 2-fold increase in K_M.

(A)



(B)



(C)

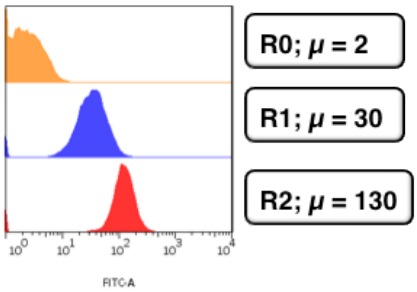


Figure 4.3 FACS histograms of 4 residue saturation libraries at the anchor positions in the predicted T cell epitope 9-mer peptides M_{115}, I_{216}, V_{304} by the neutral drift assay. Each library comprised of $>10^7$ transformants generated by randomizing the P1, P4, P6 and P9 positions of the respective T cell epitopes using the NNS scheme. (A) M_{115} library (B) I_{216} library (C) V_{304} library. μ: geometric mean fluorescence.

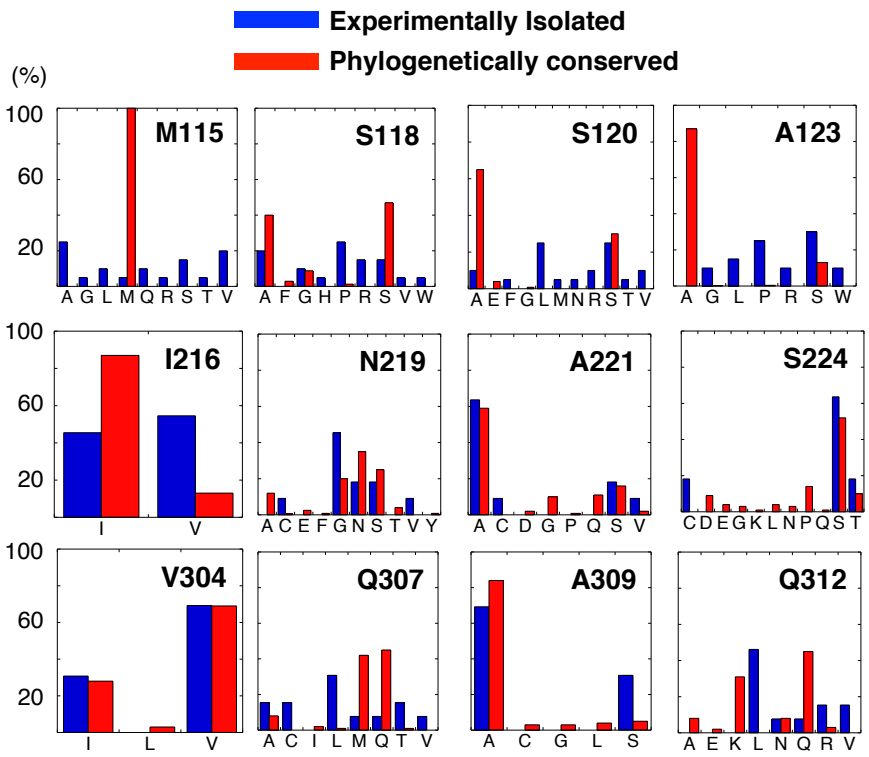


Figure 4.4 Residue plasticity of known bacterial type II asparaginases (IPR004550; n = 478) and of EcAII variants exhibiting WT activity isolated by neutral drift (n = 20 per library) at the 12 amino acids targeted for mutagenesis. Residue numbers correspond to positions in EcAII.

	Minimum CPR		
HLA Allele	MRPSTSMSA (VRPPTRMSP)	IVYNYANAS (VVYGYANAS)	VLLQLALTQ (VLLTLALTN)
DRB1*0101	18.48 (n/a)	5.6 (12.27)	6.9 (n/a)
DRB1*0301	2.29 (2.24)	1.07 (6.71)	6.71 (15.89)
DRB1*0401	1.40 (30.75)	0.27 (1.88)	0.79 (2.42)
DRB1*0701	18.53 (n/a)	0.56 (3.95)	n/a (18.91)
DRB1*0801	11 (9.9)	5.17 (9.9)	6.4 (9.9)
DRB1*1101	1.82 (6.94)	1.12 (7.58)	0.98 (0.94)
DRB1*1301	6.36 (6.41)	0.38 (2.93)	2.40 (13)
DRB1*1501	19.91 (27.58)	3.14 (11.83)	10.89 (9.32)
CPR ≤ 2	2	5	2

Table 4.4 Computational prediction of T-cell epitopes in WT EcAII and the 3.1.E2 mutant by IEDB consensus method. Minimum consensus percentile rank score (CPR) for the 3 targeted EcAII T-cell epitope core regions across 8 common HLA-DR alleles. The ALL-associated DRB1*0401 allele is shown in bold. Lower scores are indicative of higher predicted binding affinity. Sequences and scores in parentheses are for the 3.1.E2 mutant. n/a: not predicted to bind.

The 1.1.C4 variant was then used as a template to diversify the P1, P4, P6, and P9 positions in I_{216}VYNYANAS. The near background mean fluorescence of the initial library cell population (10^7 transformants) revealed that the overwhelming majority of amino acid substitutions at these residues are deleterious. Nonetheless, a population with near WT fluorescence was established after 4 rounds of FACS sorting (**Figure 4.3B**). In contrast to the high degree of plasticity observed in the M_{115} core region, mutagenesis of the MHC anchor positions in the I_{216} core yielded mostly conservative amino acid substitutions (**Figure 4.4**). One variant of 1.1.C4 containing two mutations, a conservative change at I216V and a non-conservative one a N219G, displayed a 7-fold

increase in CPR score of the modified I_{216} core region for DRB1*0401, and increases in the CPR scores for 7 other common MHC-II alleles (**Table 4.4**). Once again, these mutations did not affect the catalytic properties of the enzyme (k_{cat} = 21 s^{-1}, K_M = 19μM, k_{cat}/K_M = 1.1 x 10^6 $M^{-1}s^{-1}$). This variant, designated 2.2.G10, was then used as a template for mutagenesis of the V_{304}LLQLALTQ T-cell epitope core region (10^7 transformants). The final enzyme variant following three rounds of sorting (**Figure 4.3C**), designated 3.1.E2, further containing a non-conservative change at Q307T and a conservative Q312N substitution, showed a 3-fold increase in the CPR score for binding to DRB1*0401 and increased CPR scores for 4 other alleles (**Table 4.4**). 3.1.E2 contained a total of 8 amino acid substitutions (**Figure 4.5**), but retained a k_{cat} identical to the parent enzyme with just a 3-fold increase in K_M (k_{cat} = 24s^{-1}, K_M = 23μM, k_{cat}/K_M = 1.0 x 10^6 $M^{-1}s^{-1}$). 3.1.E2 further displayed slightly reduced (33%) specific activity towards L-Gln hydrolysis that may be of therapeutic benefit [111], was stable in serum for over 10 days (**Figure 4.6**); essentially identical to the WT EcAII [230], and could be expressed at a high yield (> 30mg/L shake flask culture) (**Figure 4.7**). Finally, we note that because recombinant EcAII folds and expresses well, no mutants with low activity but compensatory expression levels were isolated in screening each library.

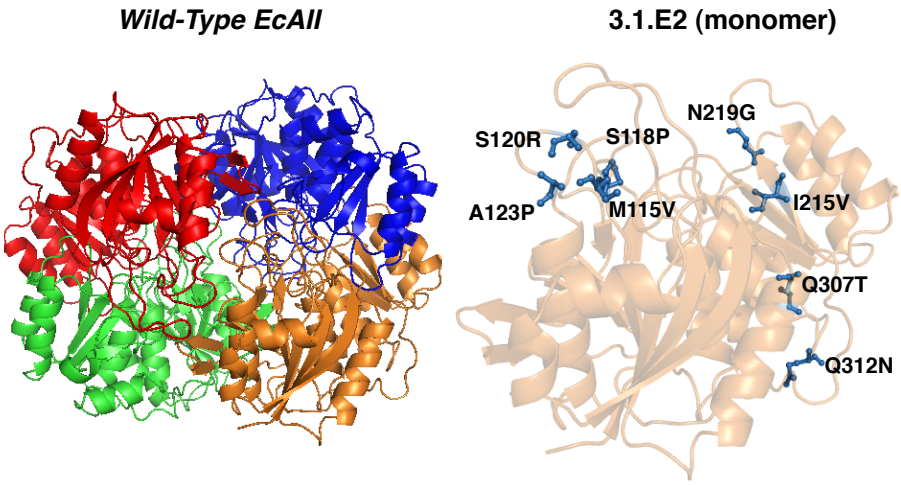


Figure 4.5 Location of the 8 amino acid mutations differentiating WT EcAII and 3.1.E2. Images generated by PyMol [105].

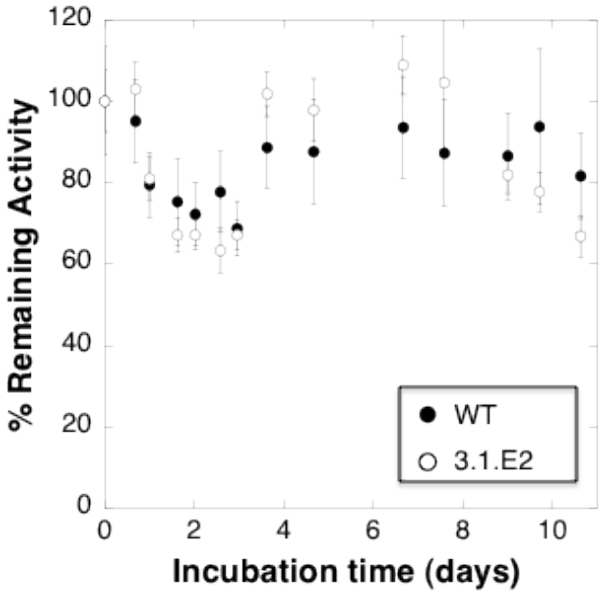


Figure 4.6 In serum stability of WT EcAII and 3.1.E2 measured quantitatively by percent remaining activity over time. Error bars shown are for S.D.

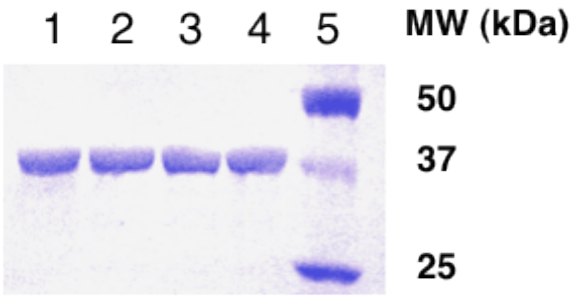


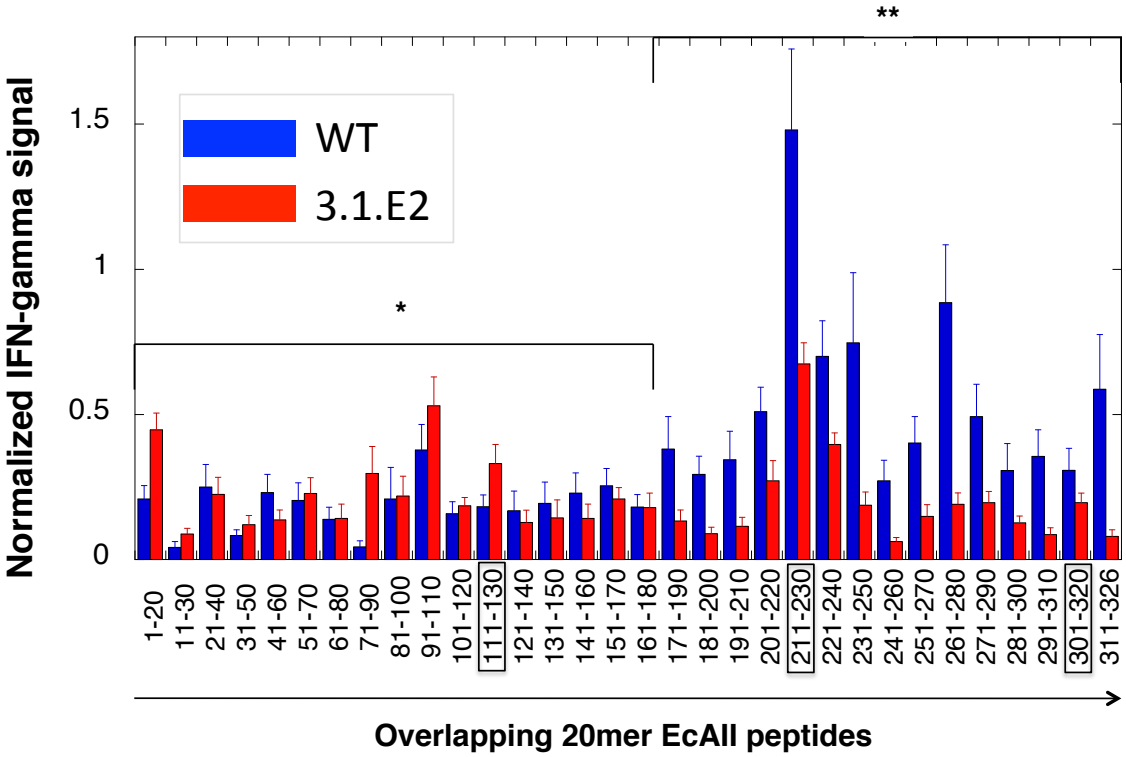
Figure 4.7 Reducing SDS-PAGE showing the purity of purified WT EcAII and engineered EcAII variants. Lane 1: WT EcAII; Lane 2: 1.1.C4; Lane 3: 2.2.G10; Lane 4: 3.1.E2; Lane 5: M.W. standards.

Evaluation of EcAII T-cell responses and immunogenicity using HLA-transgenic mice

The immunogenicity of purified, low endotoxin preparations of WT EcAII and 3.1.E2 were evaluated in transgenic mice expressing human HLA-DRB1*0401 under the mouse MHC-II promoter and deficient in the endogenous murine MHC-II locus. As a stringent test of the potential for immunogenicity, mice were immunized with a strong adjuvant (Complete Freund's Adjuvant) to induce robust CD4+ T-cell responses. The HLA transgenic mice were immunized with either WT EcAII or 3.1.E2 and T-cell responses were measured in draining lymph node cells by cytokine ELISPOT assays for IFN-γ levels following recall with either the initial enzyme itself or with overlapping 20-mer synthetic peptides corresponding to the sequence of the enzyme used in the initial immunization (**Figure 4.8A**). For WT EcAII, the highest level of T-cell activation was observed in response to 20-mer WTp211-230, which contained the predicted core region I_{216}VYNYANAS. Deimmunization resulted in a significant decrease in T-cell activation

by peptides containing the mutated sequences relative to I_{216}VYNYANAS and V_{304}LLQLALTYQ in the parental enzyme. In contrast, although mutagenesis of the M_{115}RPSTSMSA region resulted in a sequence with improved CPR score with for DRB1*0401, no statistical difference in cytokine stimulation could be observed for the mutant peptide. This was probably a consequence of the complex relationship between antigen processing, MHC-II binding, and TCR recognition and signaling [231]. Interestingly, while the cytokine responses induced by the 17 N-terminal overlapping peptides were essentially indistinguishable ($p = 0.182$) regardless of whether the mice had been immunized with WT or mutant enzyme, a significantly reduced response ($p < 0.0001$) was observed for the 3.1.E2-immunized population across the 15 C-terminal overlapping peptides. One possible explanation for this result is that the mutations in 3.1.E2 may have affected antigen processing, thus altering MHC-II loading[232]. Importantly, mice immunized with 3.1.E2 also displayed a statistically significant ($p = 0.02$) 10-fold reduction in anti-EcAII IgG titer relative to mice receiving the WT enzyme (**Figure 4.8B**). Given that the activation of CD4+ T-cells is in most cases required for the longevity and proliferation of B-cells and for antibody isotype-switching[24], this result strongly implicated that the removal of EcAII T-cell epitopes in 3.1.E2 resulted in reduced T-cell help and thus led to lower antibody titers relative to the WT enzyme.

(A)


**
*
WT
3.1.E2
Normalized IFN-gamma signal
1.5
1
0.5
0
1-20
11-30
21-40
31-50
41-60
51-70
61-80
71-90
81-100
91-110
101-120
111-130
121-140
131-150
141-160
151-170
161-180
171-190
181-200
191-210
201-220
211-230
221-240
231-250
241-260
251-270
261-280
271-290
281-300
291-310
301-320
311-326
Overlapping 20mer EcAll peptides

(B)

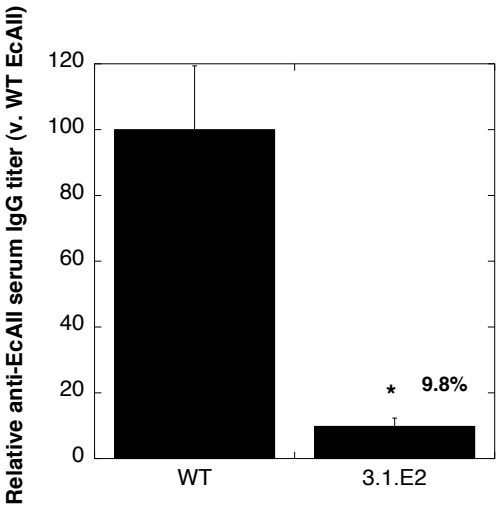

Relative anti-EcAll serum IgG titer (v. WT EcAll)
120
100
80
60
40
20
0
*
9.8%
WT
3.1.E2

Figure 4.8 T-cell activation and antibody responses in HLA-DRB1*0401 transgenic mice: (A) IFN-γ production by lymph node T-cells from HLA-DRB1*0401 transgenic mice immunized with either EcAII (n = 10 mice) or 3.1.E2 (n = 8 mice) and challenged with overlapping 20-mer peptides as shown. For each mouse, cytokine signals to each peptide were normalized using the cytokine signal generated to whole antigen in order to account for response variability across each sample population. Error bars shown are S.E.M. *:p = 0.182; **: p < 0.0001: paired Student's *t*-test, 2-tailed, comparing recall responses. (B) Relative anti-EcAII serum IgG titers induced by EcAII (n = 6 mice) and 3.1.E2 (n = 6 mice) in HLA-DRB1*0401 transgenic mice. Error bars shown are S.E.M. *:p = 0.02: unpaired Student's *t*-test, 2-tailed, comparing antibody titers.

4.3 DISCUSSION

Human or humanized protein deimmunization has so far relied on the introduction of one or at most, a very limited number of conservative amino acid substitutions that attempt to remove immunogenic epitopes without perturbing therapeutic function. However, more drastic re-engineering of the polypeptide sequence is often required for the deimmunization of heterologous enzymes that have not undergone tolerance induction. Introducing substantial changes in the primary sequence of enzymes without affecting stability of function poses a significant challenge. We showed that the use of combinatorial mutagenesis and neutral drift screens that directly interrogate protein function can be exploited to take large leaps in sequence space and thus generate variant polypeptides with reduced propensity to bind to MHC-II and elicit T-dependent antibody responses. The EcAII 3.1.E2 mutant contained 8 amino acid substitutions, 3 of which are not observed in any of the nearly 500 bacterial type II asparaginases in the database, yet retained near WT catalytic efficiency and stability. EcAII 3.1.E2 exhibited substantially reduced immunogenicity in HLA-transgenic mice and thus constitutes a very promising candidate for alleviating adverse responses in the treatment of childhood ALL. Further,

the development of an asparaginase displaying reduced immunogenicity could prove critical for longer term treatment in adult ALL or for relapsing patients.

The neutral drift screen we developed may be readily applied for the combinatorial deimmunization of a number of other heterologous therapeutic enzymes used in cancer treatment that function by systemic amino acid depletion, such as L-methioninase or arginine deiminase. Likewise, different neutral drift screens may be readily designed for the deimmunization of heterologous binding proteins, e.g. enzyme inhibitors[233], using established methods such as phage or microbial display. While in this work we employed a computational approach for monitoring immunogenicity, experimental methods for identifying T-cell epitopes can also be applied to monitor the immunogenicity propensity of mutant proteins.

Chapter 5: Conclusions and Future Perspectives

The role of *E. coli* L-Asparaginase II (EcAII) as an effective antileukemic agent for the treatment of ALL has been well established for over 40 years. Despite its proven clinical efficacy however, the use of EcAII has been limited by some toxicities, and critically, by the development of adverse antibody responses which in turn may mediate hypersensitivity reactions and/or neutralize the enzyme itself in a large fraction of patients. To date, the primary strategy for managing patients with adverse responses to EcAII (either native or PEGylated formulations) is treatment with *Er. caratovora* L-Asparaginase II (ErA), which while non-cross reactive with anti-EcAII antibodies [118], is also highly immunogenic and clinically inferior to EcAII with respect to both event-free survival and overall survival rates at 6 years [114]. The work detailed in this dissertation has focused on the investigation and development of novel strategies towards the engineering of a second-generation therapeutic asparaginase with reduced immunogenicity for the safer and more effective treatment of ALL. The results of Chapter 4 highlight the power of exploiting the evolutionary biology concept of neutral drift to modify immunogenic determinants of EcAII in order to generate a less immunogenic variant of the benchmark therapeutic. Additionally, as part of this work we explored of the biochemical properties and enzymology of human enzymes that possess some degree of functional similarity to asparaginase.

First, we provided experimental proof that the human asparaginase-like protein 1 (hASRGL1) belongs to the N-terminal nucleophile (Ntn) hydrolase family, and that Thr168 serves as its critical residue for both intramolecular processing and catalytic activity. Accordingly, kinetic analysis revealed that hASRGL1 displays both L-

Asparaginase ($k_{cat}/K_M \sim 10^3$ $M^{-1}s^{-1}$) and β-aspartyl peptidase activities, however fails to hydrolyze either L-Gln or GlcNAc-L-Asn, consistent with other characterized plant-type asparaginases [96, 97]. Moreover, the k_{cat}/K_M values exhibited for β-aspartyl peptides containing a hydrophobic amino acid were 2-4-fold higher relative to L-Asn, whereas substitution of a basic amino acid into the β-aspartyl position resulted in a 2-fold reduction in k_{cat}/K_M, again relative to L-Asn. Our results constituted the first described recombinant expression and characterization of this enzyme. Given the elucidated hASRGL1 activity profile described, we further proposed a potential physiological synergism with the enzyme Protein L-isoaspartyl Methyltransferase (PIMT, 2.1.1.77) for the repair and/or degradation of isoaspartyl-damaged proteins, whose accumulation is otherwise detrimental. Interestingly, Evtimova et al. [136] demonstrated that hASRGL1 is significantly upregulated in certain cancers (e.g. ovarian). Therefore, one future direction in the study of hASRGL1 should be in trying to uncover an explanation for its upregulation in malignancies, and further, how this implication may be therapeutically targeted or exploited. Of course, the other future direction would be to utilize the recombinant expression system (Chapter 2) and some adaptation of the high-throughput asparaginase screen (Chapter 4) described in this dissertation in order to engineer hASRGL1 for improved asparaginase activity given that the wild-type enzyme exhibits an approximately 1000-fold lower k_{cat}/K_M (L-Asn) relative to EcAII and would thus be unlikely to be immediately therapeutically relevant.

The optimization of bacterial expression and subsequent biochemical characterization of human protein N-terminal asparagine amidohydrolase (hNTAN1) represented the first example of bacterial recombinant expression of a mammalian N-end rule enzyme. Following optimization of the (i) expression RBS, (ii) N-terminus (i.e.

deletion of a nonnative encoded Gly2 residue), and (iii) fusion of an appropriate C-terminal affinity tag, we were able to obtain essentially homogeneous hNTAN1. Kinetic analysis revealed that the enzyme exhibited activity against the angiotensin peptide NRVYVHPF ($k_{cat}/K_M \sim 10^5$ $M^{-1}s^{-1}$), however failed to deamidate L-Asn, L-Gln, an analogous peptide with an acetylated N-terminal asparagine, or an analogous peptide with an N-terminal glutamine (QRVYVHPF). These results indicate that the recombinant human enzyme has substrate specificity similar to that of NTAN1 purified from porcine liver [191]. Additionally, we demonstrated that the alkylating agent iodoacetamide was able to completely abolish hNTAN1 activity, suggesting the existence of a nucleophilic Cys. Site-specific mutagenesis suggested that Cys75 may be strongly implicated as the nucleophilic residue. Much like with hASRGL1, the future directions for hNTAN1 should be two-fold. In one aspect, our results have the laid the groundwork for potentially investigating the N-end rule *in vitro* in order to both identify substrates targeted by the pathway, and as importantly, to begin to gain an understanding of the biochemical and genetic implications of the phenotypic results observed in *Ntan1*$^{-/-}$ and other N-end rule-targeted knockout mice. The other future direction for hNTAN1 related studies would be to engineer the substrate specificity of the enzyme such that it acquires asparaginase activity with therapeutically relevant k_{cat}/K_M.

Importantly, the work in this dissertation resulted in the isolation of an EcAII variant containing 8 amino acid substitutions within computationally predicted T-cell epitopes – of which 4 were non-conservative – while still exhibiting $k_{cat}/K_M \sim 10^6$ $M^{-1}s^{-1}$ for L-Asn hydrolysis. Immunization of HLA transgenic mice expressing the ALL-associated DRB1*0401 allele with this engineered variant resulted in significantly reduced T-cell responses and a 10-fold reduction in anti-EcAII IgG titers relative to the

existing therapeutic. The significant reduction in the immunogenicity of EcAII described here may be clinically relevant for ALL treatment immediately. Further rounds of EcAII epitope removal to generate a 'third generation' EcAII with potentially even less immunogenic potential, may turn out to be the most optimal strategy in developing a safer and effective asparaginase. This study also illustrated the power of employing neutral drift screens to achieve large jumps in sequence space as may be required for the deimmunization of heterologous proteins. Thus, another future perspective stemming from this work should be the adaptation of the neutral drift screen described here towards the deimmunization of other heterologous enzymes used in cancer treatment that are of interest in our lab, specifically L-methioninase and arginine deiminase. Similarly, neutral drift screens may be designed for the deimmunizaiton of a number of other heterologous proteins (e.g. enzyme inhibitors) using already well-established screening methods such as phage or microbial display.

Appendix I – Cyclotide Scaffold for Protein Engineering

To date, monoclonoal antibodies (mAbs) have served as the primary biomolecular scaffold in various therapeutic and diagnostic applications, owing to their high affinity, target specificity, and therapeutic efficacy [234]. However, it has become clear that mAbs suffer from a number of fundamental disadvantages [234, 235], which has in turn stimulated the exploration into engineering alternative protein scaffolds with improved features. The 10th type III domain of human fibronectin (Fn3) has been shown to be one such effective scaffold alternative. Hackel et al. [236] have designed Fn3 libraries with variable loop lengths and used an affinity maturation scheme to isolate a fibronectin with picomolar affinity to lysozyme. Designed ankyrin repeat domains (DARPins) are yet another novel class of binding molecules that can be selected to recognize a wide variety of ligands. For example, Zahnd et al. [237] used ribosome display to affinity mature a DARPin with picomolar affinity to human epidermal growth factor receptor 2 (Her2).

Cyclotides are naturally occurring, plant-derived mini proteins (~ 30 amino acids) that have both a cyclic backbone and a knotted arrangement of three conserved disulfide bonds, referred to as a cystine knot (**Figure A.1)**. The combination of the cyclic backbone and cystine knot motif is referred to as a cyclic cystine knot (CCK) and it confers cyclotides with remarkable stability and high resistance to a range of chemical, enzymatic, and thermal treatments [238], making them ideal molecular scaffolds for the generation of stable therapeutic affinity reagents. Cyclotides have already been shown to display a diverse range of therapeutically useful intrinsic activities, including antimicrobial and anti-HIV activities [239], however, a number of studies have additionally demonstrated the plasticity of cystine-knotted peptide framework regions and their tolerance to residue substitutions [240, 241], suggesting the notion that the

cyclotide scaffold can potentially be engineered across such regions to acquire therapeutically relevant, non-native binding functions while maintaining their stable characteristic CCK motif. Kimura et al. [242] recently demonstrated the potential of employing knottin peptides as protein engineering scaffolds. In their work, libraries of the knottin peptide *Ecballium elaterium* trypsin inhibitor (EETI-II) were constructed in which the 6-amino acid trypsin-binding loop was substituted with 11-amino acid loops containing the Arg-Gly-Asp integrin binding motif and randomized flanking residues, and subsequently screened using yeast surface display to isolate high affinity binders to specific integrin subtypes.

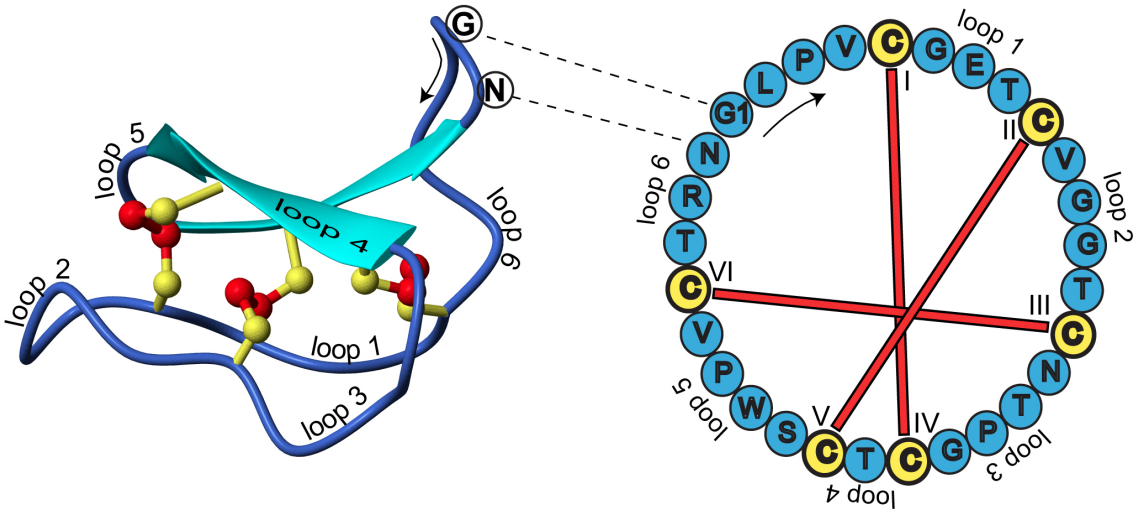


Figure A.1 Structure and sequence of the prototypic cyclotide kalataB1. This image was reproduced with permission from http://en.wikipedia.org/wiki/File:Cyclotide_structure.jpg

MCoTI-II (**Figure A.2**) is a powerful trypsin inhibitory cyclotide isolated from a plant of the squash family (*Momordica cochinchinensis*) that shares extensive structural features to the open-chain EETI-II described above. Avrutina et al. [243] revealed that the cyclic feature of MCoTI-II appears to impart the cyclotide with a nearly 3-fold greater trypsin-binding affinity (K_i = 0.03nM) relative to the open-chain EETI-II (K_i = 0.08nM), perhaps resulting from additional binding energy provided by the cyclization loop. Camarero et al. [244] have described the recombinant expression of a correctly folded and fully functional MCoTI-II in *E. coli* using intein-mediated native chemical ligation. However, this strategy constrains the localization of biosynthetically produced cyclotides to the bacterial cytoplasm. Nonetheless, an acyclic permutant of MCoTI-II, in which the proposed cyclization loop was removed, was synthesized and observed to retain picomolar affinity for trypsin (K_i = 0.3nM) [243]. Thus, using the open-chain MCoTI-II as a template, we anticipated the potential to construct libraries that could be screened for non-native binding activities with no constraint on bacterial localization, and subsequently, the cyclization loop could conceivably be re-introduced to isolated acyclic variants afterwards in order to generate novel binding cyclotides.

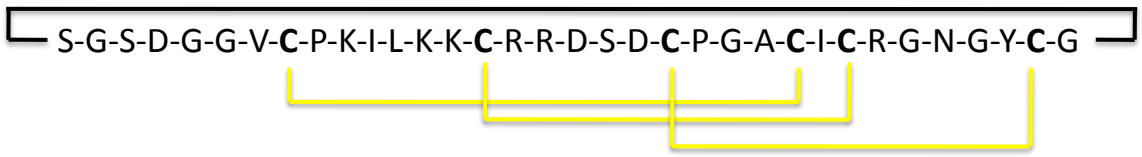


Figure A.2 Sequence and disulfide linkages of MCoTI-II. Disfulfide bonds are shown as yellow lines, while the black line connecting the terminal residues as shown represents the cyclization.

The two primary screening strategies pursued for this project to date were: (i) anchored periplasmic expression (APEx) [245] and (ii) phage display [246]. As such, it was first necessary to determine whether the circular permutation of open-chain MCoTI-II (oMCoTI-II), GV**C**PKILKK**C**RRDSD**C**PGA**CIC**RGNGY**C**G, could be (i) secreted to the bacterial periplasm (using a pelb leader sequence) with either an (ii) N- or C-terminal affinity tag, as both of these conditions would need to be met in carrying out either of the two screening strategies above. Using this particular permutation of oMCoTI-II, the recombinant protein displayed trypsin inhibitory activity when expressed with a C-terminal His_6 tag; however, no detectable activity was observed for oMCoTI-II with an N-terminal His_6 tag.

Thus, two approaches were ultimately developed for screening oMCoTI-II libraries:

I. In the anchored periplasmic expression (APEx) system, an oMCoTI-II-Fos fusion – where Fos was fused to the C-terminus of oMCoTI-II – was expressed and subsequently captured by an inner transmembrane NlpA-Jun fusion through the protein-protein interaction between Jun and Fos, thereby anchoring oMCoTI-II to the *E. coli* inner membrane. Following co-expression of the two fusion proteins, outer membrane permeabilization was used to create spheroplasts in which oMCoTI-II was accessible to exogenously added ligand. In this manner, controls were set up in which this system was evaluated by incubating spheroplasts with a biotin-trypsin conjugation and then labeling with streptavidin-phycoerythrin (PE) for subsequent FACS analysis. Theoretically, PE signal far above background levels could only be achieved when the following series of successive binding events occurred for a given spheroplast population (binding partners

in each event underlined): (i) oMCoTI-II-Fos bound to NlpA-Jun, (ii) oMCoTI-II-Fos bound to biotin-trypsin, and (iii) biotin-trypsin bound to streptavidin-PE (**Figure A.3**).

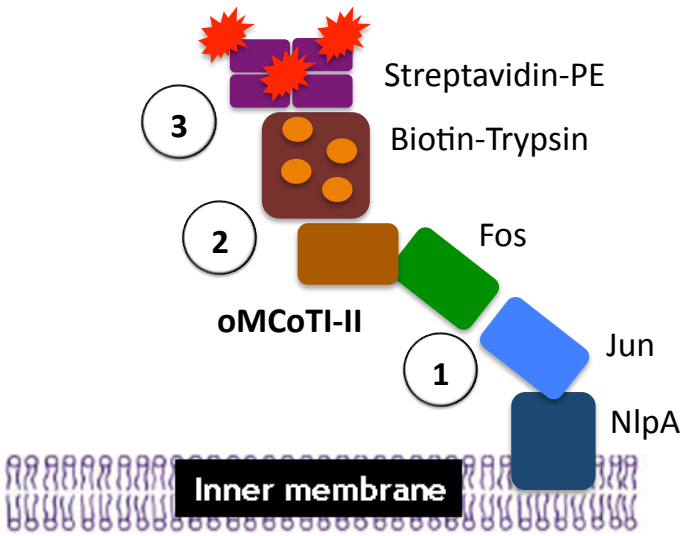


Figure A.3 Schematic of APEx approach for screening oMCoTI-II libraries. Enrichment of phycoerythrin (PE) signal is dependent upon 3 successive binding events as described in the text.

II. In the phage display system, oMCoTI-II was tethered to M13 phage via C-terminal fusion to the minor coat protein pIII. Multiple rounds of phage panning with varying concentrations of the ligand of interest (without need for conjugation to biotin) would thus allow for the isolation of high affinity acyclic cyclotide binders with novel binding activity.

To evaluate these screening approaches, a large collection of oMCoTI-II libraries were ultimately constructed in which the trypsin inhibitory loop, PKILKK [243], was replaced with completely randomized loops with length X, where X = 5, 6, 7, 8, 9, 10, or

11. Either of two randomization schemes were employed in order to generate minimalist libraries based on the success of this approach in the affinity maturation of other scaffold proteins [247, 248]. In one scheme, the degenerate codon NNC (where N = A, T, G, C) was encoded at each position within the loop, while in the other, a binary diversity of all Tyr and Ser residues was achieved using the degenerate codon TMT (where M = A or C). The substrate for which we hoped to generate affinity-matured binders against was the complement serine protease C1s.

Application of the APEx approach to isolate binders against biotin-C1s proved to be unsuccessful, as multiple attempts at optimization of the system failed to result in the enrichment of PE signal across iterative rounds of sorting, despite the modest validation of this system using controls whereby the PE signal for spheroplasts expressing the Jun/Fos capture machinery was > 4-fold greater relative to spheroplasts that did not express either fusion protein.

Application of the phage display approach also proved to be unsuccessful following multiple attempts at optimization. Despite very modest positive results obtained through polyclonal phage ELISAs comparing the binding affinity of initial phage library populations to those that had undergone multiple rounds of panning, monoclonal high binding variants against C1s were never isolated.

With fellow graduate student Mark Pogson, a new genetic selection strategy is currently in development to isolate oMCoTI-II variants evolved to bind various human proteases whose recombinant expression in *E. coli* would otherwise result in cell death. Thus, co-expression of oMCoTI-II could provide an effective means to link inhibitory activity to cell viability. Controls in which wild-type oMCoTI-II was co-expressed with human trypsin have provided promising early results, in which viability is dramatically

enhanced in cells that co-express oMCoTI-II relative to cells that express trypsin only. (**Figure A.4**).

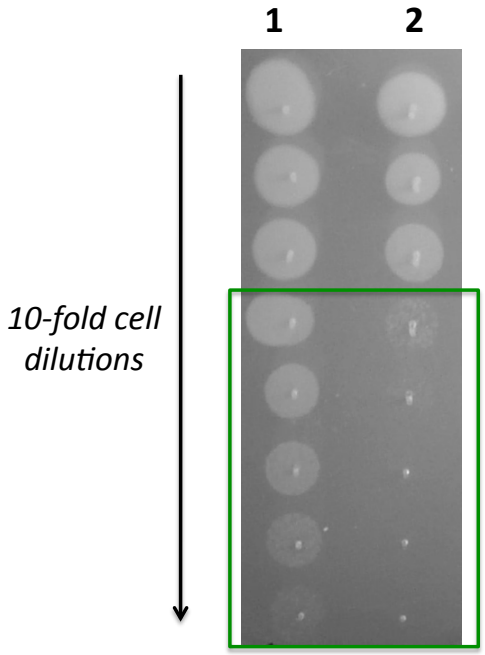


Figure A.4 Serial dilutions *E. coli* expressing recombinant human trypsin. (1) Cells co-expressing oMCoTI-II. (2) Cells expressing trypsin only. Cell viability in lane 1 is much greater as a result of the co-expression of trypsin inhibitory oMCoTI-II, whereby cell growth is observed over at least 4 additional log dilutions relative to cells expressing trypsin only (lane 2).

References

1. Leader, B., Q.J. Baca, and D.E. Golan, *Protein therapeutics: a summary and pharmacological classification*. Nat Rev Drug Discov, 2008. **7**(1): p. 21-39.
2. Strohl, W.R. and D.M. Knight, *Discovery and development of biopharmaceuticals: current issues*. Current Opinion in Biotechnology, 2009. **20**(6): p. 668-72.
3. Hamilton, S.R., et al., *Production of complex human glycoproteins in yeast*. Science, 2003. **301**(5637): p. 1244-6.
4. Mahmood, I. and M.D. Green, *Pharmacokinetic and pharmacodynamic considerations in the development of therapeutic proteins*. Clin Pharmacokinet, 2005. **44**(4): p. 331-47.
5. Wills, R.J., et al., *Interferon kinetics and adverse reactions after intravenous, intramuscular, and subcutaneous injection*. Clin Pharmacol Ther, 1984. **35**(5): p. 722-7.
6. Fishburn, C.S., *The pharmacology of PEGylation: balancing PD with PK to generate novel therapeutics*. J Pharm Sci, 2008. **97**(10): p. 4167-83.
7. Bailon, P., et al., *Rational design of a potent, long-lasting form of interferon: a 40 kDa branched polyethylene glycol-conjugated interferon alpha-2a for the treatment of hepatitis C*. Bioconjug Chem, 2001. **12**(2): p. 195-202.
8. Sroda, K., et al., *Repeated injections of PEG-PE liposomes generate anti-PEG antibodies*. Cell Mol Biol Lett, 2005. **10**(1): p. 37-47.
9. Armstrong, J.K., et al., *Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients*. Cancer, 2007. **110**(1): p. 103-11.
10. Millan, C.G., et al., *Drug, enzyme and peptide delivery using erythrocytes as carriers*. J Control Release, 2004. **95**(1): p. 27-49.
11. Schellenberger, V., et al., *A recombinant polypeptide extends the in vivo half-life of peptides and proteins in a tunable manner*. Nature Biotechnology, 2009. **27**(12): p. 1186-1190.
12. Fowler, S.B., et al., *Rational design of aggregation-resistant bioactive peptides: reengineering human calcitonin*. Proc Natl Acad Sci U S A, 2005. **102**(29): p. 10105-10.
13. Chennamsetty, N., et al., *Design of therapeutic proteins with enhanced stability*. Proc Natl Acad Sci USA, 2009. **106**(29): p. 11937-42.
14. Schellekens, H., *Bioequivalence and the immunogenicity of biopharmaceuticals*. Nat Rev Drug Discov, 2002. **1**(6): p. 457-62.
15. Onda, M., *Reducing the immunogenicity of protein therapeutics*. Curr Drug Targets, 2009. **10**(2): p. 131-9.
16. Hogquist, K.A., T.A. Baldwin, and S.C. Jameson, *Central tolerance: learning self-control in the thymus*. Nat Rev Immunol, 2005. **5**(10): p. 772-82.

17. Gilles, J.G., M.G. Jacquemin, and J.M. Saint-Remy, *Factor VIII inhibitors*. Thromb Haemost, 1997. **78**(1): p. 641-6.
18. Dummer, R., et al., *Formation of neutralizing antibodies against natural interferon-beta, but not against recombinant interferon-gamma during adjuvant therapy for high-risk malignant melanoma patients*. Cancer, 1991. **67**(9): p. 2300-4.
19. Zang, Y.C., et al., *Immunoregulation and blocking antibodies induced by interferon beta treatment in MS*. Neurology, 2000. **55**(3): p. 397-404.
20. Prabhakar, S.S. and T. Muhlfelder, *Antibodies to recombinant human erythropoietin causing pure red cell aplasia*. Clin Nephrol, 1997. **47**(5): p. 331-5.
21. Ritter, G., et al., *Serological analysis of human anti-human antibody responses in colon cancer patients treated with repeated doses of humanized monoclonal antibody A33*. Cancer Res, 2001. **61**(18): p. 6851-9.
22. Chirino, A.J., M.L. Ary, and S.A. Marshall, *Minimizing the immunogenicity of protein therapeutics*. Drug Discov Today, 2004. **9**(2): p. 82-90.
23. De Groot, A.S. and D.W. Scott, *Immunogenicity of protein therapeutics*. Trends Immunol, 2007. **28**(11): p. 482-90.
24. McHeyzer-Williams, L.J. and M.G. McHeyzer-Williams, *Antigen-specific memory B cell development*. Annu Rev Immunol, 2005. **23**: p. 487-513.
25. Schellekens, H., *Relationship between biopharmaceutical immunogenicity of epoetin alfa and pure red cell aplasia*. Curr Med Res Opin, 2003. **19**(5): p. 433-4.
26. De Groot, A.S. and L. Moise, *Prediction of immunogenicity for therapeutic proteins: state of the art*. Current opinion in drug discovery & development, 2007. **10**(3): p. 332-40.
27. Chester, K.A., M. Baker, and A. Mayer, *Overcoming the immunologic response to foreign enzymes in cancer therapy*. Expert Rev Clin Immunol, 2005. **1**(4): p. 549-59.
28. Jones, E.Y., et al., *MHC class II proteins and disease: a structural perspective*. Nat Rev Immunol, 2006. **6**(4): p. 271-82.
29. Wang, P., et al., *A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach*. PLoS Comput Biol, 2008. **4**(4): p. e1000048.
30. Huang, J., et al., *The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness*. Nature, 2010. **464**(7290): p. 932-936.
31. Koren, E., et al., *Clinical validation of the "in silico" prediction of immunogenicity of a human recombinant therapeutic protein*. Clin Immunol, 2007. **124**(1): p. 26-32.
32. Inaba, H., et al., *Thyrotropin receptor epitopes and their relation to histocompatibility leukocyte antigen-DR molecules in Graves' disease*. J Clin Endocrinol Metab, 2006. **91**(6): p. 2286-94.
33. Greenbaum, J.A., et al., *Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools*. J Mol Recognit, 2007. **20**(2): p. 75-82.

34. Steere, A.C., et al., *Antibiotic-refractory Lyme arthritis is associated with HLA-DR molecules that bind a Borrelia burgdorferi peptide*. J Exp Med, 2006. **203**(4): p. 961-71.
35. Baker, M.P. and T.D. Jones, *Identification and removal of immunogenicity in therapeutic proteins*. Curr Opin Drug Discov Devel, 2007. **10**(2): p. 219-27.
36. Jones, T.D., et al., *Identification and removal of a promiscuous CD4+ T cell epitope from the C1 domain of factor VIII*. J Thromb Haemost, 2005. **3**(5): p. 991-1000.
37. Tangri, S., et al., *Rationally engineered therapeutic proteins with reduced immunogenicity*. J Immunol, 2005. **174**(6): p. 3187-96.
38. Jones, T.D., et al., *The development of a modified human IFN-alpha2b linked to the Fc portion of human IgG1 as a novel potential therapeutic for the treatment of hepatitis C virus infection*. J Interferon Cytokine Res, 2004. **24**(9): p. 560-72.
39. Ueta, M., et al., *Immunosuppressive properties of human amniotic membrane for mixed lymphocyte reaction*. Clin Exp Immunol, 2002. **129**(3): p. 464-70.
40. Anthony, D.D. and P.V. Lehmann, *T-cell epitope mapping using the ELISPOT approach*. Methods, 2003. **29**(3): p. 260-9.
41. Nagata, S. and I. Pastan, *Removal of B cell epitopes as a practical approach for reducing the immunogenicity of foreign protein-based therapeutics*. Adv Drug Deliv Rev, 2009. **61**(11): p. 977-85.
42. Bontrop, R.E., et al., *Major histocompatibility complex class II polymorphisms in primates*. Immunol Rev, 1999. **167**: p. 339-50.
43. Klitgaard, J.L., et al., *Reduced susceptibility of recombinant polyclonal antibodies to inhibitory anti-variable domain antibody responses*. J Immunol, 2006. **177**(6): p. 3782-90.
44. Shirai, M., et al., *CTL responses of HLA-A2.1-transgenic mice specific for hepatitis C viral peptides predict epitopes for CTL of humans carrying HLA-A2.1*. J Immunol, 1995. **154**(6): p. 2733-42.
45. Man, S., et al., *Definition of a human T cell epitope from influenza A non-structural protein 1 using HLA-A2.1 transgenic mice*. Int Immunol, 1995. **7**(4): p. 597-605.
46. Cope, A.P., et al., *T cell responses to a human cartilage autoantigen in the context of rheumatoid arthritis-associated and nonassociated HLA-DR4 alleles*. Arthritis Rheum, 1999. **42**(7): p. 1497-507.
47. Khare, M., M. Rodriguez, and C.S. David, *HLA class II transgenic mice authenticate restriction of myelin oligodendrocyte glycoprotein-specific immune response implicated in multiple sclerosis pathogenesis*. Int Immunol, 2003. **15**(4): p. 535-46.
48. Yu, J.-J., et al., *Francisella tularensis T-Cell Antigen Identification Using Humanized HLA-DR4 Transgenic Mice*. Clinical and Vaccine Immunology, 2010. **17**(2): p. 215-222.
49. Forsthuber, T.G., et al., *T cell epitopes of human myelin oligodendrocyte glycoprotein identified in HLA-DR4 (DRB1*0401) transgenic mice are*

encephalitogenic and are presented by human B cells. J Immunol, 2001. **167**(12): p. 7119-25.
50. Sharma, B., *Immunogenicity of therapeutic proteins. Part 1: impact of product handling*. Biotechnol Adv, 2007. **25**(3): p. 310-7.
51. Sharma, B., *Immunogenicity of therapeutic proteins. Part 2: impact of container closures*. Biotechnol Adv, 2007. **25**(3): p. 318-24.
52. Sharma, B., *Immunogenicity of therapeutic proteins. Part 3: impact of manufacturing changes*. Biotechnol Adv, 2007. **25**(3): p. 325-31.
53. Banting, F.G., et al., *Pancreatic extracts in the treatment of diabetes mellitus: preliminary report. 1922*. CMAJ, 1991. **145**(10): p. 1281-6.
54. Richter, B., G. Neises, and K. Bergerhoff, *Human versus animal insulin in people with diabetes mellitus. A systematic review*. Endocrinol Metab Clin North Am, 2002. **31**(3): p. 723-49.
55. Presta, L.G., *Engineering of therapeutic antibodies to minimize immunogenicity and optimize function*. Adv Drug Deliv Rev, 2006. **58**(5-6): p. 640-56.
56. Pavlou, A.K. and M.J. Belsey, *The therapeutic antibodies market to 2008*. Eur J Pharm Biopharm, 2005. **59**(3): p. 389-96.
57. Jakobovits, A., *Production of fully human antibodies by transgenic mice*. Curr Opin Biotechnol, 1995. **6**(5): p. 561-6.
58. Lazar, H., et al., *Pharmacokinetics and safety profile of the human anti-Pseudomonas aeruginosa serotype O11 immunoglobulin M monoclonal antibody KBPA-101 in healthy volunteers*. Antimicrob Agents Chemother, 2009. **53**(8): p. 3442-6.
59. Stone, E.M., et al., *Replacing Mn(2+) with Co(2+) in human arginase i enhances cytotoxicity toward l-arginine auxotrophic cancer cell lines*. ACS Chem Biol. **5**(3): p. 333-42.
60. Schellekens, H., *Immunogenicity of therapeutic proteins: clinical implications and future prospects*. Clin Ther, 2002. **24**(11): p. 1720-40; discussion 1719.
61. Warmerdam, P.A.M., et al., *Elimination of a human T-cell region in staphylokinase by T-cell screening and computer modeling*. Thromb Haemost, 2002. **87**(4): p. 666-73.
62. Jones, T.D., et al., *The development of a modified human IFN-alpha2b linked to the Fc portion of human IgG1 as a novel potential therapeutic for the treatment of hepatitis C virus infection*. J Interferon Cytokine Res, 2004. **24**(9): p. 560-72.
63. Yeung, V.P., et al., *Elimination of an immunodominant CD4+ T cell epitope in human IFN-beta does not result in an in vivo response directed at the subdominant epitope*. J Immunol, 2004. **172**(11): p. 6658-65.
64. Harding, F.A., et al., *A beta-lactamase with reduced immunogenicity for the targeted delivery of chemotherapeutics using antibody-directed enzyme prodrug therapy*. Mol Cancer Ther, 2005. **4**(11): p. 1791-800.
65. Bander, N.H., et al., *Phase I trial of 177lutetium-labeled J591, a monoclonal antibody to prostate-specific membrane antigen, in patients with androgen-independent prostate cancer*. J Clin Oncol, 2005. **23**(21): p. 4591-601.

66. Macfarlane, D.J., et al., *Safety, pharmacokinetic and dosimetry evaluation of the proposed thrombus imaging agent 99mTc-DI-DD-3B6/22-80B3 Fab'*. Eur J Nucl Med Mol Imaging, 2006. **33**(6): p. 648-56.
67. Holgate, R.G. and M.P. Baker, *Circumventing immunogenicity in the development of therapeutic antibodies*. IDrugs, 2009. **12**(4): p. 233-7.
68. Broome, J.D., *Evidence that the L-asparaginase of guinea pig serum is responsible for its antilymphoma effects. II. Lymphoma 6C3HED cells cultured in a medium devoid of L-asparagine lose their susceptibility to the effects of guinea pig serum in vivo*. J Exp Med, 1963. **118**: p. 121-48.
69. Hill, J.M., et al., *L-asparaginase therapy for leukemia and other malignant neoplasms. Remission in human leukemia*. JAMA, 1967. **202**(9): p. 882-8.
70. Jones, B., et al., *Optimal use of L-asparaginase (NSC-109229) in acute lymphocytic leukemia*. Med Pediatr Oncol, 1977. **3**(4): p. 387-400.
71. Ertel, I.J., et al., *Effective dose of L-asparaginase for induction of remission in previously treated children with acute lymphocytic leukemia: a report from Childrens Cancer Study Group*. Cancer Res, 1979. **39**(10): p. 3893-6.
72. Clavell, L.A., et al., *Four-agent induction and intensive asparaginase therapy for treatment of childhood acute lymphoblastic leukemia*. N Engl J Med, 1986. **315**(11): p. 657-63.
73. Asselin, B.L., et al., *In vitro and in vivo killing of acute lymphoblastic leukemia cells by L-asparaginase*. Cancer Res, 1989. **49**(15): p. 4363-8.
74. Avramis, V.I., et al., *A randomized comparison of native Escherichia coli asparaginase and polyethylene glycol conjugated asparaginase for treatment of children with newly diagnosed standard-risk acute lymphoblastic leukemia: a Children's Cancer Group study*. Blood, 2002. **99**(6): p. 1986-94.
75. Pui, C.-H. and W.E. Evans, *Treatment of acute lymphoblastic leukemia*. N Engl J Med, 2006. **354**(2): p. 166-78.
76. Cooney, D.A. and R.E. Handschumacher, *L-asparaginase and L-asparagine metabolism*. Annu Rev Pharmacol, 1970. **10**: p. 421-40.
77. Horowitz, B., et al., *Asparagine synthetase activity of mouse leukemias*. Science, 1968. **160**(827): p. 533-5.
78. Haskell, C.M. and G.P. Canellos, *l-asparaginase resistance in human leukemia--asparagine synthetase*. Biochem Pharmacol, 1969. **18**(10): p. 2578-80.
79. Haskell, C.M., et al., *Biocmical and pharmacologic effects of L-asparaginase in man*. J Lab Clin Med, 1970. **75**(5): p. 763-70.
80. Haskell, C.M. and G.P. Canellos, *Asparaginase and leukaemia*. Lancet, 1969. **1**(7595): p. 626.
81. Ho, D.H., et al., *L-asparagine requirement and the effect of L-asparaginase on the normal and leukemic human bone marrow*. Cancer Res, 1970. **30**(2): p. 466-72.
82. Capizzi, R.L., et al., *L-asparaginase: clinical, biochemical, pharmacological, and immunological studies*. Ann Intern Med, 1971. **74**(6): p. 893-901.

83. Story, M.D., et al., *L-asparaginase kills lymphoma cells by apoptosis*. Cancer Chemother Pharmacol, 1993. **32**(2): p. 129-33.
84. Ueno, T., et al., *Cell cycle arrest and apoptosis of leukemia cells induced by L-asparaginase*. Leukemia, 1997. **11**(11): p. 1858-61.
85. Avramis, V.I. and E.H. Panosyan, *Pharmacokinetic/pharmacodynamic relationships of asparaginase formulations: the past, the present and recommendations for the future*. Clin Pharmacokinet, 2005. **44**(4): p. 367-93.
86. Kidd, J.G., *Regression of transplanted lymphomas induced in vivo by means of normal guinea pig serum. II. Studies on the nature of the active serum constituent: histological mechanism of the regression: tests for effects of guinea pig serum on lymphoma cells in vitro: discussion*. J Exp Med, 1953. **98**(6): p. 583-606.
87. Mashburn, L.T. and J.C. Wriston, Jr., *Tumor Inhibitory Effect of L-Asparaginase from Escherichia Coli*. Arch Biochem Biophys, 1964. **105**: p. 450-2.
88. Campbell, H.A. and L.T. Mashburn, *L-Asparaginase EC-2 from Escherichia coli. Some substrate specificity characteristics*. Biochemistry, 1969. **8**(9): p. 3768-75.
89. Wriston, J.C., *Asparaginase*. Meth Enzymol, 1985. **113**: p. 608-18.
90. Kurtzberg, J. Asparaginase, in *Cancer Medicine, 5th Edition* (Bast, R., Kufe, D., Pollock, R., Weichselbaum, R., Holland, J., Frei, E., Ed.), 2000. pp 699-705, BC Decker, Hamilton (ON).
91. Michalska, K. and M. Jaskolski, *Structural aspects of L-asparaginases, their friends and relations*. Acta Biochim Pol, 2006. **53**(4): p. 627-40.
92. Oinonen, C., et al., *Three-dimensional structure of human lysosomal aspartylglucosaminidase*. Nat Struct Biol, 1995. **2**(12): p. 1102-8.
93. Guo, H.C., et al., *Crystal structures of Flavobacterium glycosylasparaginase. An N-terminal nucleophile hydrolase activated by intramolecular proteolysis*. J Biol Chem, 1998. **273**(32): p. 20205-12.
94. Xuan, J., et al., *Crystal structure of glycosylasparaginase from Flavobacterium meningosepticum*. Protein Sci, 1998. **7**(3): p. 774-81.
95. Guan, C., et al., *Characterization and functional analysis of the cis-autoproteolysis active center of glycosylasparaginase*. J Biol Chem, 1998. **273**(16): p. 9695-702.
96. Borek, D., et al., *Expression, purification and catalytic activity of Lupinus luteus asparagine beta-amidohydrolase and its Escherichia coli homolog*. Eur J Biochem, 2004. **271**(15): p. 3215-26.
97. Hejazi, M., et al., *Isoaspartyl dipeptidase activity of plant-type asparaginases*. Biochem J, 2002. **364** (Pt 1): p. 129-36.
98. Cantor, J.R., et al., *The human asparaginase-like protein 1 hASRGL1 is an Ntn hydrolase with beta-aspartyl peptidase activity*. Biochemistry, 2009. **48**(46): p. 11026-31.
99. Jerlstrom, P.G., et al., *Structure and expression in Escherichia coli K-12 of the L-asparaginase I-encoding ansA gene and its flanking regions*. Gene, 1989. **78**(1): p. 37-46.

100. Distasio, J.A., et al., *Purification and characterization of L-asparaginase with anti-lymphoma activity from Vibrio succinogenes*. J Biol Chem, 1976. **251**(22): p. 6929-33.
101. Wade, H.E., et al., *A new L-asparaginase with antitumour activity?* Lancet, 1968. **2**(7571): p. 776-7.
102. Dinndorf, P.A., et al., *FDA Drug Approval Summary: Pegaspargase (Oncaspar(R)) for the First-Line Treatment of Children with Acute Lymphoblastic Leukemia (ALL)*. The Oncologist, 2007. **12**(8): p. 991-998.
103. Swain, A.L., et al., *Crystal structure of Escherichia coli L-asparaginase, an enzyme used in cancer therapy*. Proc Natl Acad Sci U S A, 1993. **90**(4): p. 1474-8.
104. Sanches, M., et al., *Structural comparison of Escherichia coli L-asparaginase in two monoclinic space groups*. Acta Crystallogr D Biol Crystallogr, 2003. **59**(Pt 3): p. 416-22.
105. Delano, D.W., *The PyMol Graphics System* 2002: Delano Scientific, Palo Alto, CA, USA.
106. Sanches, M., S. Krauchenco, and I. Polikarpov, *Structure, Substrate Complexation and Reaction Mechanism of Bacterial Asparaginases*. Current Chemical Biology, 2007. **1**(1): p. 75-86.
107. Wehner, A., et al., *Site-specific mutagenesis of Escherichia coli asparaginase II. None of the three histidine residues is required for catalysis*. Eur J Biochem, 1992. **208**(2): p. 475-80.
108. Harms, E., et al., *A catalytic role for threonine-12 of E. coli asparaginase II as established by site-directed mutagenesis*. FEBS Lett, 1991. **285**(1): p. 55-8.
109. Palm, G.J., et al., *A covalently bound catalytic intermediate in Escherichia coli asparaginase: crystal structure of a Thr-89-Val mutant*. FEBS Lett, 1996. **390**(2): p. 211-6.
110. Panosyan, E.H., et al., *Deamination of glutamine is a prerequisite for optimal asparagine deamination by asparaginases in vivo (CCG-1961)*. Anticancer Res, 2004. **24**(2C): p. 1121-5.
111. Derst, C., J. Henseling, and K.H. Röhm, *Engineering the substrate specificity of Escherichia coli asparaginase. II. Selective reduction of glutaminase activity by amino acid replacements at position 248*. Protein Sci, 2000. **9**(10): p. 2009-17.
112. Apostolidou, E., et al., *Treatment of acute lymphoblastic leukaemia : a new era*. Drugs, 2007. **67**(15): p. 2153-71.
113. Nesbit, M.E., I. Ertel, and G.D. Hammond, *L-Asparaginase as a single agent in acute lymphocytic leukemia: survey of studies form Childrens Cancer Study Group*. Cancer Treat Rep, 1981. **65 Suppl 4**: p. 101-7.
114. Duval, M., et al., *Comparison of Escherichia coli-asparaginase with Erwinia-asparaginase in the treatment of childhood lymphoid malignancies: results of a randomized European Organisation for Research and Treatment of Cancer-Children's Leukemia Group phase 3 trial*. Blood, 2002. **99**(8): p. 2734-9.

115. Masetti, R. and A. Pession, *First-line treatment of acute lymphoblastic leukemia with pegasparaginase*. Biologics, 2009. **3**: p. 359-68.
116. Zeidan, A., E.S. Wang, and M. Wetzler, *Pegasparaginase: where do we stand?* Expert Opin Biol Ther, 2009. **9**(1): p. 111-9.
117. Asselin, B.L., et al., *Comparative pharmacokinetic studies of three asparaginase preparations*. J Clin Oncol, 1993. **11**(9): p. 1780-6.
118. Avramis, V.I. and P.N. Tiwari, *Asparaginase (native ASNase or pegylated ASNase) in the treatment of acute lymphoblastic leukemia*. International journal of nanomedicine, 2006. **1**(3): p. 241-54.
119. Kwon, Y.M., et al., *L-Asparaginase encapsulated intact erythrocytes for treatment of acute lymphoblastic leukemia (ALL)*. Journal of Controlled Release, 2009. **139**(3): p. 182-189.
120. Cherry, J.R. and A.L. Fidantsef, *Directed evolution of industrial enzymes: an update*. Curr Opin Biotechnol, 2003. **14**(4): p. 438-43.
121. Johannes, T.W. and H. Zhao, *Directed evolution of enzymes and biosynthetic pathways*. Curr Opin Microbiol, 2006. **9**(3): p. 261-7.
122. Bershtein, S. and D.S. Tawfik, *Advances in laboratory evolution of enzymes*. Curr Opin Chem Biol, 2008. **12**(2): p. 151-8.
123. Kazlauskas, R.J. and U.T. Bornscheuer, *Finding better protein engineering strategies*. Nat Chem Biol, 2009. **5**(8): p. 526-9.
124. Bershtein, S., K. Goldin, and D.S. Tawfik, *Intense neutral drifts yield robust and evolvable consensus proteins*. J Mol Biol, 2008. **379**(5): p. 1029-44.
125. Bloom, J.D., et al., *Neutral genetic drift can alter promiscuous protein functions, potentially aiding functional evolution*. Biol Direct, 2007. **2**: p. 17.
126. Amitai, G., R.D. Gupta, and D.S. Tawfik, *Latent evolutionary potentials under the neutral mutational drift of an enzyme*. HFSP J, 2007. **1**(1): p. 67-78.
127. Peisajovich, S.G. and D.S. Tawfik, *Protein engineers turned evolutionists*. Nature Methods, 2007. **4**(12): p. 991-4.
128. van Nimwegen, E., J.P. Crutchfield, and M. Huynen, *Neutral evolution of mutational robustness*. Proc Natl Acad Sci U S A, 1999. **96**(17): p. 9716-20.
129. Broome, J.D., *Evidence that the L-asparaginase Activity of Guinea Pig Serum is responsible for its Antilymphoma Effects*. Nature, 1961. **191**: p. 1114-1115.
130. Borek, D., et al., *Expression, purification and catalytic activity of Lupinus luteus asparagine ß-amidohydrolase and its Escherichia coli homolog*. European Journal of Biochemistry, 2004. **271**(15): p. 3215-3226.
131. Brannigan, J.A., et al., *A protein catalytic framework with an N-terminal nucleophile is capable of self-activation*. Nature, 1995. **378**(6555): p. 416-9.
132. Saarela, J., et al., *Autoproteolytic activation of human aspartylglucosaminidase*. Biochem J, 2004. **378**(Pt 2): p. 363-71.
133. Michalska, K., A. Hernandez-Santoyo, and M. Jaskolski, *The mechanism of autocatalytic activation of plant-type L-asparaginases*. J Biol Chem, 2008. **283**(19): p. 13388-97.

134. Larsen, R.A., T.M. Knox, and C.G. Miller, *Aspartic peptide hydrolases in Salmonella enterica serovar typhimurium*. J Bacteriol, 2001. **183**(10): p. 3089-97.
135. Aswad, D.W., M.V. Paranandi, and B.T. Schurter, *Isoaspartate in peptides and proteins: formation, significance, and analysis*. J Pharm Biomed Anal, 2000. **21**(6): p. 1129-36.
136. Evtimova, V., et al., *Identification of CRASH, a gene deregulated in gynecological tumors*. Int J Oncol, 2004. **24**(1): p. 33-41.
137. Weidle, U.H., et al., *Cell growth stimulation by CRASH, an asparaginase-like protein overexpressed in human tumors and metastatic breast cancers*. Anticancer Res, 2009. **29**(4): p. 951-63.
138. Dieterich, D.C., et al., *Gliap--a novel untypical L-asparaginase localized to rat brain astrocytes*. J Neurochem, 2003. **85**(5): p. 1117-25.
139. Bush, L.A., et al., *A novel asparaginase-like protein is a sperm autoantigen in rats*. Mol Reprod Dev, 2002. **62**(2): p. 233-47.
140. Hoover, D.M. and J. Lubkowski, *DNAWorks: an automated method for designing oligonucleotides for PCR-based gene synthesis*. Nucleic Acids Res, 2002. **30**(10): p. e43.
141. Gill, S.C. and P.H. von Hippel, *Calculation of protein extinction coefficients from amino acid sequence data*. Anal Biochem, 1989. **182**(2): p. 319-26.
142. Griswold, K.E., et al., *Effects of codon usage versus putative 5'-mRNA structure on the expression of Fusarium solani cutinase in the Escherichia coli cytoplasm*. Protein Expr Purif, 2003. **27**(1): p. 134-42.
143. Shtraizent, N., et al., *Autoproteolytic cleavage and activation of human acid ceramidase*. J Biol Chem, 2008. **283**(17): p. 11253-9.
144. Suzuki, H. and H. Kumagai, *Autocatalytic processing of gamma-glutamyltranspeptidase*. J Biol Chem, 2002. **277**(45): p. 43536-43.
145. Boanca, G., A. Sand, and J.J. Barycki, *Uncoupling the enzymatic and autoprocessing activities of Helicobacter pylori gamma-glutamyltranspeptidase*. J Biol Chem, 2006. **281**(28): p. 19029-37.
146. Boanca, G., et al., *Autoprocessing of Helicobacter pylori gamma-glutamyltranspeptidase leads to the formation of a threonine-threonine catalytic dyad*. J Biol Chem, 2007. **282**(1): p. 534-41.
147. http://www.chem.agilent.com/Library/datasheets/Public/5980-3088.pdf.
148. Kane, J.F., *Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli*. Curr Opin Biotechnol, 1995. **6**(5): p. 494-500.
149. Kelo, E., et al., *Beta-aspartylpeptides as substrates of L-asparaginases from Escherichia coli and Erwinia chrysanthemi*. FEBS Lett, 2002. **528**(1-3): p. 130-2.
150. Johnson, B.A. and D.W. Aswad, *Fragmentation of isoaspartyl peptides and proteins by carboxypeptidase Y: release of isoaspartyl dipeptides as a result of internal and external cleavage*. Biochemistry, 1990. **29**(18): p. 4373-80.
151. Noronkoski, T., et al., *Glycosylasparaginase-catalyzed synthesis and hydrolysis of beta-aspartyl peptides*. J Biol Chem, 1998. **273**(41): p. 26295-7.

152. Kelley, L.A. and M.J. Sternberg, *Protein structure prediction on the Web: a case study using the Phyre server*. Nat Protoc, 2009. **4**(3): p. 363-71.
153. Aswad, D.W., *Stoichiometric methylation of porcine adrenocorticotropin by protein carboxyl methyltransferase requires deamidation of asparagine 25. Evidence for methylation at the alpha-carboxyl group of atypical L-isoaspartyl residues*. J Biol Chem, 1984. **259**(17): p. 10714-21.
154. O'Connor, C.M., D.W. Aswad, and S. Clarke, *Mammalian brain and erythrocyte carboxyl methyltransferases are similar enzymes that recognize both D-aspartyl and L-isoaspartyl residues in structurally altered protein substrates*. Proc Natl Acad Sci U S A, 1984. **81**(24): p. 7757-61.
155. Lowenson, J.D. and S. Clarke, *Identification of isoaspartyl-containing sequences in peptides and proteins that are usually poor substrates for the class II protein carboxyl methyltransferase*. J Biol Chem, 1990. **265**(6): p. 3106-10.
156. Lowenson, J.D. and S. Clarke, *Structural elements affecting the recognition of L-isoaspartyl residues by the L-isoaspartyl/D-aspartyl protein methyltransferase. Implications for the repair hypothesis*. J Biol Chem, 1991. **266**(29): p. 19396-406.
157. Reissner, K.J. and D.W. Aswad, *Deamidation and isoaspartate formation in proteins: unwanted alterations or surreptitious signals?* Cell Mol Life Sci, 2003. **60**(7): p. 1281-95.
158. Reissner, K.J., et al., *Synapsin I is a major endogenous substrate for protein L-isoaspartyl methyltransferase in mammalian brain*. J Biol Chem, 2006. **281**(13): p. 8389-98.
159. Kim, E., et al., *Phenotypic analysis of seizure-prone mice lacking L-isoaspartate (D-aspartate) O-methyltransferase*. J Biol Chem, 1999. **274**(29): p. 20671-8.
160. Ikegaya, Y., et al., *Aberrant synaptic transmission in the hippocampal CA3 region and cognitive deterioration in protein-repair enzyme-deficient mice*. Hippocampus, 2001. **11**(3): p. 287-98.
161. Lowenson, J.D., et al., *Limited accumulation of damaged proteins in l-isoaspartyl (D-aspartyl) O-methyltransferase-deficient mice*. J Biol Chem, 2001. **276**(23): p. 20695-702.
162. McClellan, A.J., et al., *Protein quality control: chaperones culling corrupt conformations*. Nat Cell Biol, 2005. **7**(8): p. 736-41.
163. Varshavsky, A., *The N-end rule: functions, mysteries, uses*. Proc Natl Acad Sci U S A, 1996. **93**(22): p. 12142-9.
164. Mogk, A., R. Schmidt, and B. Bukau, *The N-end rule pathway for regulated proteolysis: prokaryotic and eukaryotic strategies*. Trends Cell Biol, 2007. **17**(4): p. 165-72.
165. Ravid, T. and M. Hochstrasser, *Diversity of degradation signals in the ubiquitin-proteasome system*. Nat Rev Mol Cell Biol, 2008. **9**(9): p. 679-90.
166. Varshavsky, A., *The N-end rule pathway of protein degradation*. Genes Cells, 1997. **2**(1): p. 13-28.
167. Bachmair, A. and A. Varshavsky, *The degradation signal in a short-lived protein*. Cell, 1989. **56**(6): p. 1019-32.

168. Prakash, S., et al., *Substrate selection by the proteasome during degradation of protein complexes*. Nat Chem Biol, 2009. **5**(1): p. 29-36.
169. Tasaki, T. and Y.T. Kwon, *The mammalian N-end rule pathway: new insights into its components and physiological roles*. Trends Biochem Sci, 2007. **32**(11): p. 520-8.
170. Baker, R.T. and A. Varshavsky, *Yeast N-terminal amidase. A new enzyme and component of the N-end rule pathway*. J Biol Chem, 1995. **270**(20): p. 12065-74.
171. Kwon, Y.T., et al., *Altered activity, social behavior, and spatial memory in mice lacking the NTAN1p amidase and the asparagine branch of the N-end rule pathway*. Mol Cell Biol, 2000. **20**(11): p. 4135-48.
172. Wang, H., et al., *Glutamine-specific N-terminal amidase, a component of the N-end rule pathway*. Molecular Cell, 2009. **34**(6): p. 686-95.
173. Kwon, Y.T., et al., *An essential role of N-terminal arginylation in cardiovascular development*. Science, 2002. **297**(5578): p. 96-9.
174. Hu, R.G., et al., *The N-end rule pathway as a nitric oxide sensor controlling the levels of multiple regulators*. Nature, 2005. **437**(7061): p. 981-6.
175. Hu, R.G., et al., *The N-end rule pathway is a sensor of heme*. Proc Natl Acad Sci U S A, 2008. **105**(1): p. 76-81.
176. Kwon, Y.T., et al., *The mouse and human genes encoding the recognition component of the N-end rule pathway*. Proc Natl Acad Sci U S A, 1998. **95**(14): p. 7898-903.
177. An, J.Y., et al., *Impaired neurogenesis and cardiovascular development in mice lacking the E3 ubiquitin ligases UBR1 and UBR2 of the N-end rule pathway*. Proc Natl Acad Sci U S A, 2006. **103**(16): p. 6212-7.
178. Kwon, Y.T., et al., *Female lethality and apoptosis of spermatocytes in mice lacking the UBR2 ubiquitin ligase of the N-end rule pathway*. Mol Cell Biol, 2003. **23**(22): p. 8255-71.
179. Ouyang, Y., et al., *Loss of Ubr2, an E3 ubiquitin ligase, leads to chromosome fragility and impaired homologous recombinational repair*. Mutat Res, 2006. **596**(1-2): p. 64-75.
180. Zenker, M., et al., *Deficiency of UBR1, a ubiquitin ligase of the N-end rule pathway, causes pancreatic dysfunction, malformations and mental retardation (Johanson-Blizzard syndrome)*. Nat Genet, 2005. **37**(12): p. 1345-50.
181. Balogh, S.A., Y.T. Kwon, and V.H. Denenberg, *Varying intertrial interval reveals temporally defined memory deficits and enhancements in NTAN1-deficient mice*. Learn Mem, 2000. **7**(5): p. 279-86.
182. Lecker, S.H., et al., *Ubiquitin conjugation by the N-end rule pathway and mRNAs for its components increase in muscles of diabetic rats*. J Clin Invest, 1999. **104**(10): p. 1411-20.
183. Tasaki, T., et al., *Biochemical and genetic studies of UBR3, a ubiquitin ligase with a function in olfactory and other sensory systems*. J Biol Chem, 2007. **282**(25): p. 18510-20.

184. Eisele, F. and D.H. Wolf, *Degradation of misfolded protein in the cytoplasm is mediated by the ubiquitin ligase Ubr1*. FEBS Lett, 2008. **582**(30): p. 4143-6.
185. Lee, M.J., et al., *RGS4 and RGS5 are in vivo substrates of the N-end rule pathway*. Proc Natl Acad Sci U S A, 2005. **102**(42): p. 15030-5.
186. Du, F., et al., *Pairs of dipeptides synergistically activate the binding of substrate by ubiquitin ligase through dissociation of its autoinhibitory domain*. Proc Natl Acad Sci U S A, 2002. **99**(22): p. 14110-5.
187. Turner, G.C., F. Du, and A. Varshavsky, *Peptides accelerate their uptake by activating a ubiquitin-dependent proteolytic pathway*. Nature, 2000. **405**(6786): p. 579-83.
188. Rao, H., et al., *Degradation of a cohesin subunit by the N-end rule pathway is essential for chromosome stability*. Nature, 2001. **410**(6831): p. 955-9.
189. Ditzel, M., et al., *Degradation of DIAP1 by the N-end rule pathway is essential for regulating apoptosis*. Nat Cell Biol, 2003. **5**(5): p. 467-73.
190. Yoshida, S., et al., *A delayed leaf senescence mutant is defective in arginyl-tRNA:protein arginyltransferase, a component of the N-end rule pathway in Arabidopsis*. Plant J, 2002. **32**(1): p. 129-37.
191. Stewart, A.E., S.M. Arfin, and R.A. Bradshaw, *Protein NH2-terminal asparagine deamidase. Isolation and characterization of a new enzyme*. J Biol Chem, 1994. **269**(38): p. 23509-17.
192. Grigoryev, S., et al., *A mouse amidase specific for N-terminal asparagine. The gene, the enzyme, and their function in the N-end rule pathway*. J Biol Chem, 1996. **271**(45): p. 28521-32.
193. Schmidt, T.G. and A. Skerra, *The Strep-tag system for one-step purification and high-affinity detection or capturing of proteins*. Nat Protoc, 2007. **2**(6): p. 1528-35.
194. Kane, J.F., *Effects of rare codon clusters on high-level expression of heterologous proteins in Escherichia coli*. Current Opinion in Biotechnology, 1995. **6**(5): p. 494-500.
195. Salis, H.M., E.A. Mirsky, and C.A. Voigt, *Automated design of synthetic ribosome binding sites to control protein expression*. Nature Biotechnology, 2009. **27**(10): p. 946-950.
196. Frottin, F., et al., *The proteomics of N-terminal methionine cleavage*. Mol Cell Proteomics, 2006. **5**(12): p. 2336-49.
197. Kim, W., et al., *Trends in enzyme therapy for phenylketonuria*. Mol Ther, 2004. **10**(2): p. 220-4.
198. Chohan, S. and M.A. Becker, *Update on emerging urate-lowering therapies*. Current Opinion in Rheumatology, 2009. **21**(2): p. 143-149.
199. Ni, Y., U. Schwaneberg, and Z.H. Sun, *Arginine deiminase, a potential anti-tumor drug*. Cancer Letters, 2008. **261**(1): p. 1-11.
200. Tan, Y., M. Xu, and R.M. Hoffman, *Broad selective efficacy of recombinant methioninase and polyethylene glycol-modified recombinant methioninase on cancer cells In Vitro*. Anticancer Res. **30**(4): p. 1041-6.

201. Rytting, M., *Peg-asparaginase for acute lymphoblastic leukemia*, in *Expert Opin Biol Ther*. 2010. p. 833-9.
202. Schellekens, H., *Immunogenicity of therapeutic proteins: Clinical implications and future prospects*. Clinical Therapeutics, 2002. **24**(11): p. 1720-1740.
203. Jevsevar, S., M. Kunstelj, and V.G. Porekar, *PEGylation of therapeutic proteins*. Biotechnology Journal, 2010. **5**(1): p. 113-128.
204. Armstrong, J.K., et al., *Antibody against poly(ethylene glycol) adversely affects PEG-asparaginase therapy in acute lymphoblastic leukemia patients*, in *Cancer*. 2007. p. 103-11.
205. Sherman, M.R., M.G.P. Saifer, and F. Perez-Ruiz, *PEG-uricase in the management Of treatment-resistant gout and hyperuricemia*. Advanced Drug Delivery Reviews, 2008. **60**(1): p. 59-68.
206. Jones, E.Y., et al., *MHC class II proteins and disease: a structural perspective*, in *Nat Rev Immunol*. 2006. p. 271-82.
207. Wang, P., et al., *A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach*, in *PLoS Comput Biol*. 2008. p. e1000048.
208. Benkovic, S.J. and S. Hammes-Schiffer, *A perspective on enzyme catalysis*. Science, 2003. **301**(5637): p. 1196-1202.
209. Bershtein, S. and D.S. Tawfik, *Advances in laboratory evolution of enzymes*, in *Curr Opin Chem Biol*. 2008. p. 151-8.
210. Gupta, R.D. and D.S. Tawfik, *Directed enzyme evolution via small and effective neutral drift libraries*. Nature Methods, 2008. **5**(11): p. 939-942.
211. Link, A.J., K.J. Jeong, and G. Georgiou, *Beyond toothpicks: new methods for isolating mutant bacteria*. Nature Reviews Microbiology, 2007. **5**(9): p. 680-688.
212. Breaker, R.R., A. Banerji, and G.F. Joyce, *Continuous in-Vitro Evolution of Bacteriophage Rna-Polymerase Promoters*. Biochemistry, 1994. **33**(39): p. 11980-11986.
213. Gates, C.M., et al., *Affinity selective isolation of ligands from peptide libraries through display on a lac repressor "headpiece dimer"*. J Mol Biol, 1996. **255**(3): p. 373-86.
214. Yano, T. and H. Kagamiyama, *Directed evolution of ampicillin-resistant activity from a functionally unrelated DNA fragment: A laboratory model of molecular evolution*. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(3): p. 903-907.
215. Broome, J.D., *Evidence that the L-asparaginase of guinea pig serum is responsible for its antilymphoma effects. I. Properties of the L-asparaginase of guinea pig serum in relation to those of the antilymphoma substance*, in *J Exp Med*. 1963. p. 99-120.
216. Pui, C.H. and W.E. Evans, *Treatment of acute lymphoblastic leukemia*. N Engl J Med, 2006. **354**(2): p. 166-78.

217. Avramis, V.I. and P.N. Tiwari, *Asparaginase (native ASNase or pegylated ASNase) in the treatment of acute lymphoblastic leukemia*, in *International journal of nanomedicine*. 2006. p. 241-54.
218. Duval, M., et al., *Comparison of Escherichia coli-asparaginase with Erwinia-asparaginase in the treatment of childhood lymphoid malignancies: results of a randomized European Organisation for Research and Treatment of Cancer-Children's Leukemia Group phase 3 trial*, in *Blood*. 2002. p. 2734-9.
219. Datsenko, K.A. and B.L. Wanner, *One-step inactivation of chromosomal genes in Escherichia coli K-12 using PCR products*. Proc Natl Acad Sci U S A, 2000. **97**(12): p. 6640-5.
220. Skerra, A., *Use of the tetracycline promoter for the tightly regulated production of a murine antibody fragment in Escherichia coli*. Gene, 1994. **151**(1-2): p. 131-5.
221. Yoo, T.H., A.J. Link, and D.A. Tirrell, *Evolution of a fluorinated green fluorescent protein*. Proc Natl Acad Sci U S A, 2007. **104**(35): p. 13887-90.
222. Sturniolo, T., et al., *Generation of tissue-specific and promiscuous HLA ligand databases using DNA microarrays and virtual HLA class II matrices*. Nat Biotechnol, 1999. **17**(6): p. 555-61.
223. Singh, H. and G.P. Raghava, *ProPred: prediction of HLA-DR binding sites*. Bioinformatics, 2001. **17**(12): p. 1236-7.
224. Southwood, S., et al., *Several common HLA-DR types share largely overlapping peptide binding repertoires*, in *J Immunol*. 1998. p. 3363-73.
225. Yoo, T.H. and D.A. Tirrell, *High-throughput screening for methionyl-tRNA synthetases that enable residue-specific incorporation of noncanonical amino acids into recombinant proteins in bacterial cells*. Angew Chem Int Ed Engl, 2007. **46**(28): p. 5340-3.
226. Aida, Y. and M.J. Pabst, *Removal of endotoxin from protein solutions by phase separation using Triton X-114*. J Immunol Methods, 1990. **132**(2): p. 191-5.
227. Ito, K., et al., *HLA-DR4-IE chimeric class II transgenic, murine class II-deficient mice are susceptible to experimental allergic encephalomyelitis*. J Exp Med, 1996. **183**(6): p. 2635-44.
228. Dorak, M.T., et al., *Unravelling an HLA-DR association in childhood acute lymphoblastic leukemia*, in *Blood*. 1999. p. 694-700.
229. Wehner, A., et al., *Site-specific mutagenesis of Escherichia coli asparaginase II. None of the three histidine residues is required for catalysis*, in *Eur J Biochem*. 1992. p. 475-80.
230. Stecher, A.L., et al., *Stability of L-asparaginase: an enzyme used in leukemia treatment*, in *Pharm Acta Helv*. 1999. p. 1-9.
231. Huang, J., et al., *The kinetics of two-dimensional TCR and pMHC interactions determine T-cell responsiveness*, in *Nature*. 2010. p. 932-936.
232. Jensen, P.E., *Recent advances in antigen processing and presentation*, in *Nat Immunol*. 2007. p. 1041-8.

233. Stoop, A.A. and C.S. Craik, *Engineering of a macromolecular scaffold to develop specific protease inhibitors*. Nat Biotechnol, 2003. **21**(9): p. 1063-8.
234. Gill, D. and N. Damle, *Biopharmaceutical drug discovery using novel protein scaffolds*. Current opinion in biotechnology, 2006. **17**(6): p. 653-658.
235. Skerra, A., *Alternative non-antibody scaffolds for molecular recognition*. Current Opinion in Biotechnology, 2007. **18**(4): p. 295-304.
236. Hackel, B., A. Kapila, and K. Dane Wittrup, *Picomolar affinity fibronectin domains engineered utilizing loop length diversity, recursive mutagenesis, and loop shuffling*. Journal of molecular biology, 2008. **381**(5): p. 1238-1252.
237. Zahnd, C., et al., *A designed ankyrin repeat protein evolved to picomolar affinity to Her2*. J Mol Biol, 2007. **369**(4): p. 1015-28.
238. Craik, D.J., M. Cemazar, and N.L. Daly, *The cyclotides and related macrocyclic peptides as scaffolds in drug design*. Curr Opin Drug Discov Devel, 2006. **9**(2): p. 251-60.
239. Craik, D.J., R.J. Clark, and N.L. Daly, *Potential therapeutic applications of the cyclotides and related cystine knot mini-proteins*. Expert Opin Investig Drugs, 2007. **16**(5): p. 595-604.
240. Clark, R.J., N.L. Daly, and D.J. Craik, *Structural plasticity of the cyclic-cystine-knot framework: implications for biological activity and drug design*. Biochem J, 2006. **394**(Pt 1): p. 85-93.
241. Reiss, S., et al., *Inhibition of platelet aggregation by grafting RGD and KGD sequences on the structural scaffold of small disulfide-rich proteins*. Platelets, 2006. **17**(3): p. 153-7.
242. Kimura, R.H., et al., *Engineered cystine knot peptides that bind alphavbeta3, alphavbeta5, and alpha5beta1 integrins with low-nanomolar affinity*. Proteins, 2009. **77**(2): p. 359-69.
243. Avrutina, O., et al., *Trypsin inhibition by macrocyclic and open-chain variants of the squash inhibitor MCoTI-II*. Biol Chem, 2005. **386**(12): p. 1301-6.
244. Camarero, J.A., et al., *Biosynthesis of a fully functional cyclotide inside living bacterial cells*. Chembiochem, 2007. **8**(12): p. 1363-6.
245. Harvey, B.R., et al., *Anchored periplasmic expression, a versatile technology for the isolation of high-affinity antibodies from Escherichia coli-expressed libraries*. Proc Natl Acad Sci USA, 2004. **101**(25): p. 9193-8.
246. Smith, G.P. and V.A. Petrenko, *Phage Display*. Chem Rev, 1997. **97**(2): p. 391-410.
247. Sidhu, S.S. and A.A. Kossiakoff, *Exploring and designing protein function with restricted diversity*. Curr Opin Chem Biol, 2007. **11**(3): p. 347-54.
248. Koide, A., et al., *High-affinity single-domain binding proteins with a binary-code interface*. Proc Natl Acad Sci USA, 2007. **104**(16): p. 6632-7.

Vita

Jason Robert Cantor was born in Norwalk, Connecticut on November 12, 1982, the son of Sheldon and Marlene Cantor. After graduating from Stoughton High School in Stoughton, Massachusetts in May 2000, he entered the chemical engineering program at Cornell University in Ithaca, NY in August 2000. He graduated Magna Cum Laude with a Bachelor of Science degree from Cornell University in May 2004. In August 2004, he entered the Graduate School at The University of Texas at Austin. During this time, he also completed a 10-week internship at Lawrence Livermore National Laboratories in compliance with his acceptance of a United States Department of Homeland Security Graduate Fellowship.

Permanent address: 145 Robinette Road, Stoughton, MA 02072

This dissertation was typed by Jason Robert Cantor.