Design and analysis of immune-evading enzymes for ADEPT therapy

Daniel C. Osipovitch1†, Andrew S. Parker2†, Christabel D. Makokha3,6, Joseph Desrosiers4, Warren C. Kett3, Leonard Moise4, Chris Bailey-Kellogg2,7 and Karl E. Griswold3,5,7

1Program in Experimental and Molecular Medicine, Geisel School of Medicine, Dartmouth, Hanover NH 03784, USA, 2Department of Computer Science, Dartmouth, Hanover NH 03755, USA, 3Thayer School of Engineering, Dartmouth, Hanover NH 03755, USA, 4Institute for Immunology and Informatics, University of Rhode Island, Providence RI 02903, USA, 5Program in Molecular and Cellular Biology, Dartmouth, Hanover NH 03755, USA and 6Current address: Trinity Partners, LLC, 230 Third Avenue, Waltham, MA 02451–7528, USA

†These authors contributed equally to this work.

To whom correspondence should be addressed.
E-mail: cbk@cs.dartmouth.edu (C.B.-K.); karl.e.griswold@dartmouth.edu (K.E.G.)

Received March 31, 2012; revised June 7, 2012; accepted June 30, 2012

Edited by Bruce Tidor

The unparalleled specificity and activity of therapeutic proteins has reshaped many aspects of modern clinical practice, and aggressive development of new protein drugs promises a continued revolution in disease therapy. As a result of their biological origins, however, therapeutic proteins present unique design challenges for the biomolecular engineer. For example, protein drugs are subject to immune surveillance within the patient’s body; this anti-drug immune response can compromise therapeutic efficacy and even threaten patient safety. Thus, there is a growing demand for broadly applicable protein deimmunization strategies. We have recently developed optimization algorithms that integrate computational prediction of T-cell epitope/therapeutic protein

Introduction

Therapeutic proteins have proven to be exceptionally powerful drugs with the capacity to reshape disease therapy and clinical practice. The biopharmaceutical industry has capitalized on the molecular specificity of antibodies, the catalytic efficiency and selectivity of enzymes, and the precise regulatory properties of cytokines and hormones to produce a range of blockbuster biotherapeutics (Aggarwal, 2011). The success of numerous approved protein drugs has spurred further research and development efforts, yet proteins remain an underutilized reservoir of prospective therapeutic agents. One widely recognized hurdle to developing efficacious biotherapies is the fact that proteins, unlike small molecule drugs, are subject to immune surveillance within the patient’s body. An immune response directed towards a protein therapeutic may manifest a spectrum of clinical complications including accelerated drug clearance, neutralization and loss of efficacy or allergic responses ranging from rare anaphylactic reactions to more common delayed, infusion-like reactions (Schellekens, 2002a,b; De Groot and Scott, 2007). Growing awareness of the risks stemming from biotherapeutic immunogenicity has prompted increased scrutiny from regulatory agencies and, as a result, greater interest on the part of biopharma in pre-emptively addressing this issue (Giezen et al., 2008; De Groot and Martin, 2009). While grafting-based humanization techniques are routinely implemented with therapeutic antibodies (Almagro and Fransson, 2008), more general deimmunization strategies are needed to effectively tap the full diversity of protein families that exhibit therapeutic potential.

At a fundamental level, an anti-protein immune response is governed by molecular recognition of immunogenic peptides, or epitopes, which are found within the protein’s primary sequence (for a review see Trombetta and Mellman, 2005). Immune surveillance initiates with a patient’s antigen-presenting cells internalizing a therapeutic protein and proteolytically processing it into small peptide fragments. Fragments that represent potential epitopes are loaded into the groove of cognate type II major histocompatibility complex (MHC II) proteins, and these complexes are trafficked to the cell surface for display to the extracellular environment. Once there, the MHC II-peptide complexes are free to interact with receptor proteins on the surface of T-cell lymphocytes, and true immunogenic epitopes are distinguished by their capacity to form ternary MHC II-peptide-T-cell receptor complexes. These peptide-mediated cell–cell interactions drive a complex signaling cascade resulting in maturation and proliferation of B-cell lymphocytes that ultimately secrete immunoglobulin molecules able to bind the offending therapeutic protein. This well-characterized pathway affords the opportunity to re-engineer a protein’s immunogenic peptide sequences so as to disrupt interaction with MHC II and/or T-cell receptors. By successfully implementing such an ‘epitope deletion’ strategy, one can generate a deimmunized biotherapy.

Keywords: ADEPT/computational protein design/deimmunization/T-cell epitope/therapeutic protein
Deimmunization by T-cell epitope deletion is, conceptually, a straightforward process, but in practice the path to a deimmunized protein is technically challenging and subject to practical limitations. The fundamental tasks are identification of target epitopes and subsequent selection of epitope-deleting mutations. Experimental identification of a protein’s complete T-cell epitope profile requires specialized techniques, materials and an immense investment of time and resources (Malherbe, 2009). Therefore, a promising alternative leverages predictive computational tools to expedite T-cell epitope mapping (Wang et al., 2008; De Groot and Martin, 2009), and employs these same methods to assess, in silico, the extent to which a given mutation mitigates peptide immunogenicity (De Groot et al., 2005). But while the identification of key immunogenic epitopes and corresponding epitope-deleting mutations are requisite elements of protein deimmunization, they are not themselves sufficient to achieve this goal. By definition, an efficacious protein drug must retain its therapeutic function. This fact constrains deimmunizing mutations to those compatible with high level protein stability and activity. Thus, deimmunization by T-cell epitope deletion is an inherently complex, multi-objective optimization problem.

We have recently described a new computationally driven approach for protein deimmunization (Parker et al., 2010). Our method, Dynamic Programming for Deimmunizing Proteins (DP2), optimizes a target protein so as to reduce its immunogenicity, as assessed by T-cell epitope prediction, while simultaneously maintaining stability and activity, as assessed by sequence conservation or structure-based $\Delta \Delta G^\circ$ prediction. DP2 integrates validated prediction tools into a powerful computational framework for therapeutic protein design and, in doing so, is the first mathematically formalized method for optimally solving the combined immunogenicity and stability/activity problem. The DP2 algorithm’s rigorous mathematical structure guarantees identification of the best-scoring protein designs among the combinatorial set of possibilities. Our prior retrospective studies showed that DP2 can efficiently identify optimal sets of deimmunizing mutations that compare favorably with those found by extensive experimental efforts. Here, we leverage a therapeutically relevant protein to demonstrate the practical utility of the DP2 algorithm.

Antibody-directed enzyme prodrug therapies (ADEPT) are highly engineered chemotherapeutic regimens requiring carefully timed administration of multiple components (Schellmann et al., 2010). First, patients are given an antibody–enzyme fusion protein that localizes preferentially to cancer cells. Upon achieving high tumor-to-plasma ratios of the fusion protein, a relatively non-toxic prodrug is administered systemically. Subsequently, the prodrug is converted to a toxic agent via action of the fusion protein’s enzymatic domain. Thus, ADEPT therapies are designed to generate toxic chemotherapeutics at the site of malignancy, potentially improving efficacy and reducing side effects. To ensure site-specific prodrug conversion, ADEPT has almost uniformly employed enzymes of non-human origin. The bacterial enzymes of most ADEPT systems have activities orthogonal to the human proteome, thereby minimizing nonspecific conversion of prodrug by the patient’s endogenous biocatalysts. The use of bacterial enzymes, however, has manifested undesirable consequences stemming from the fact that proteins of non-human origin are predisposed towards high-level immunogenicity (Schellekens, 2002a,b). As with many other non-human therapeutic proteins, the development and testing of ADEPT technology has been plagued by immunogenicity-related complications (Bagshawe, 2009).

Here, we focus on the enzyme Enterobacter cloacae $\beta$-lactamase (P99BL), which has applications in ADEPT (Alderson et al., 2006). $\beta$-Lactamases are bacterial drug-resistance proteins that hydrolyze $\beta$-lactam antibiotics and have been co-opted as activators of various lactam-derivatized prodrugs. To mitigate immunogenicity-related risks, P99BL was previously subjected to a particularly rigorous experimental deimmunization effort (Harding et al., 2005). First, the full-length P99BL amino acid sequence was divided into a panel of 117 overlapping peptides, and each of these peptides was chemically synthesized. The immunogenicity of each peptide was then tested via proliferation assays using immune cells from 65 different human donors. The four native peptides that yielded the greatest immunostimulatory effect were then subjected to scanning alanine mutagenesis, which entailed the chemical synthesis of 55 new peptides, each containing a single alanine substitution. These 55 mutant peptides were analyzed for immunogenicity against a second panel of human immune cells from 66 donors. Those alanine substitutions that reduced the immunogenicity of the cognate peptide were subsequently engineered back into the full-length protein, and the expression, stability and activity of more than 16 protein variants were analyzed. Based on this exhaustive study, a single protein variant (the K21A/S324A double mutant) was deemed to have sufficient activity and stability to merit further study. Harding et al.’s isolation of the functionally enhanced K21A/S324A P99BL is a brilliant example of experimental T-cell epitope deletion, but the fact that this tour de force yielded only a single therapeutic candidate highlights the risks and uncertainty of such a monumental undertaking.

Using the P99BL as a model enzyme, we demonstrate the first experimental validation of our DP2 protein deimmunization algorithm. In stark contrast to conventional deimmunization strategies, our algorithm rapidly (in a matter of minutes) generated a multitude of protein designs predicted to have both reduced T-cell epitope content and high level activity and stability. Two of these designs were constructed and characterized, and the performance of these engineered proteins was benchmarked against both the wild-type P99BL and the previously deimmunized K21A/S324A double mutant. The results demonstrate that the DP2 algorithm generates high-performance biotherapeutic candidates with unprecedented speed and accuracy. While this experimental validation focuses on a single protein target, our methodology is general in nature and should be applicable to virtually any therapeutic protein of interest.

Methods

Computational methods

Given a wild-type amino acid sequence, our goal is to identify $m$-mutation variants that delete the greatest number of epitopes (reducing immunogenicity) while employing only conservative substitutions (maintaining stability and activity). This is a combinatorial optimization problem with \( \binom{n}{m} \) possible sets of mutated positions in an $n$ amino acid protein and
Conservative substitutions. We employ conservation statistics to determine position-specific mutations for consideration. Briefly, a multiple sequence alignment of homologs to the protein target is used to determine the amino acid frequencies at each position. The frequency of any given residue at a particular position is indicative of the substitution’s evolutionary acceptability and is thus considered predictive of structural and functional significance. In order to account for over- or under-sampling of evolutionary diversity in the extant sequence record, we weight sequences according to similarity within a phylogenetic tree (Gerstein et al., 1994). We allow substitutions with a weighted frequency above a specified background level. To avoid possible non-additive effects from coupled amino acids, we filter the resulting variants to ensure structural separation of substitutions. As a complementary assessment of structural conservation, we also calculate FoldX ΔΔG’s (Guerois et al., 2002) for putative protein designs.

Epitope prediction. We use the well-established ProPred epitope predictor (Singh and Raghava, 2001) to identify putative epitopes in the wild-type protein and prospective designs (Fig. 1). ProPred has proven effective in several applications (Mustafa and Shaban, 2006; Dinglasan et al., 2007; Klyushnenkova et al., 2007) and was one of the two most accurate predictors in a recent comparative study (Wang et al., 2008). ProPred is derived from the ‘pocket profile’ method TEPTPOE (Sturniolo et al., 1999), which assesses peptides with position-specific scoring matrices based on binding affinity measurements of all 20 amino acids against each of the 9 pockets in the binding grooves of representative MHC II alleles. We consider the eight most common human alleles, deemed representative of a large majority of the population (Southwood et al., 1998): DRB1*0101, 0301, 0401, 0701, 0801, 1101, 1301 and 1501. Given a peptide and allele, we sum the frame-specific scores and compare the result to a threshold value, essentially establishing an expected false-positive rate (5% for the plans presented here). As a complementary assessment of immunogenicity, we retrospectively examined predicted binding of resulting variants according to SMM-align (Nielsen et al., 2007) and ensured that the predictions were consistent (data not shown).

Optimization algorithm. DP$^2$ employs a dynamic programming algorithm that is both computationally efficient and optimal with respect to the given predictive models (i.e., guaranteed to find the minimum-epitope variant using a specified number of mutations from a given set). Briefly, DP$^2$ proceeds from N- to C-terminus, assessing the reduction in immunogenicity for each allowed substitution at the current position. New substitutions are compared against the best possible partial variants so far, relying on optimal substructure to consider only the nine residues at the tail of the partial variants. This approach assesses the entire complement of possible epitopes for a variant, thereby ensuring that a mutation to delete one epitope does not inadvertently introduce another. The dynamic programming approach also supports efficient identification of near-optimal variants, determined in order of epitope score by a tree search algorithm. To select variants for experimental study, we examine all variants with near-optimal epitope scores, assess them by the complementary prediction methods mentioned above, and ensure separation of mutations in both sequence (to target separate immunodominant regions) and structure (to avoid non-additive effects).

Experimental methods
Materials. Oligonucleotides (25 nmol scale, standard desalting) were purchased from Integrated DNA Technology (San Diego, CA, USA). Nitrocefin was purchased from Oxoid (Cambridge, UK). Human serum, human lysozyme and SYPRO Orange 5000× Protein Stain were purchased from...
Sigma (St Louis, MO, USA). MicroAmp® Fast Optical 0.1 ml 96-Well Plates and MicroAmp® Optical Adhesive Film were from Applied Biosystems (Bedford, MA, USA). Restriction enzymes and polymerase chain reaction (PCR) reagents were purchased from New England BioLabs (Ipswich, MA, USA). Growth media was purchased from Becton Dickinson (Franklin Lakes, NJ, USA). Plasmid purification kits and Ni-NTA resin were purchased from Qiagen (Valencia, CA, USA). PCR cleanup and gel extraction kits were from Zymo Research (Irvine, CA, USA). Peptides derived from P99βL were ordered from GenScript (Piscataway, NJ, USA), and DELFIA Eu-labeled Streptavidin was from PerkinElmer (Boston, MA, USA). Unless noted, all other chemicals and reagents were from Fisher Scientific (Pittsburgh, PA, USA).

Cloning, expression and purification. The wild-type P99βL gene was amplified from E. cloacae chromosomal DNA in whole cell lysates. The gene encoding the mature P99βL protein was appended with a 5'-sequence for the OmpA leader peptide and 3'-sequence encoding a GGGSAETVEHHHHHHH affinity tag (Amin et al., 2004). The PCR amplicon was digested with Ndel and HindIII, purified with a PCR cleanup kit, ligated into similarly digested pET26b, and transformed into BL21(DE3) [F--ompT hsdSB (ompR mB gal dcm)ompT hsdSB (DE3)]. HLA DR molecules were purchased from Benaroya Research Institute (Seattle, WA, USA), anti-HLA-DR antibody from BD Biosciences (San Diego, CA, USA) and Delfia Eu-labeled Streptavidin was from PerkinElmer (Boston, MA, USA). Unless noted, all other chemicals and reagents were from Fisher Scientific (Pittsburgh, PA, USA).

Thermostability. Differential scanning fluorimetry was performed essentially as described (Niesen et al., 2007) using an ABI 7500 Fast Real-Time PCR System from Applied Biosystems. Peptides and SYPRO Orange were diluted in PBS. Final protein concentrations were 100 µg/ml and final dye concentrations were 5X. Twenty microliter reactions were performed in 12 replicates. The PCR gradient was run from 25–94°C with a 1 min equilibration at each degree centigrade. Fluorescence was quantified using the preset TAMRA parameters. Melting temperatures were determined by data analysis with the DSF Analysis v3.0.xlsx Excel sheet (ftp://ftp.sgc.ox.ac.uk/pub/biophysics/) and GraphPad Prism v4.0 software.

Plasma stability. Enzymes were diluted into human serum to a final concentration of 100 µg/ml. The serum–enzyme mixture was incubated at 37°C and nitrocefin kinetics were analyzed at various time points. For the kinetic assays, enzyme in serum was diluted 1000-fold in PBS and combined 1:1 with 400 µM nitrocefin to yield a final enzyme concentration of 50 ng/ml and a final nitrocefin concentration of 200 µM. Plates were read as described above, and specific activities were calculated from initial slopes.

MHC binding assays. MHC II binding competition assays were performed as described (Steere et al., 2006). Briefly, in 96-well plates, non-biotinylated test peptides over a range of concentrations (0.1–100 µM) competed for binding to purified human MHC II molecules (50 nM) against a biotinylated tracer peptide at a fixed concentration (0.1 µM) for 24 h at 37°C. The sequences of the tracer peptides are provided as supplemental materials. Peptide-MHC II complexes were then captured on 96-well plates coated with pan anti-HLA-DR antibody (L243). Plates were washed and incubated with Europium-labeled streptavidin for 1 h at room temperature. Europium activation buffer was added to each well, and plates were developed for 15–20 min at room temperature before reading on a time resolved fluorescence plate reader. All measurements were performed in triplicate. Binding assays were performed for four alleles: DRB1*0101, 0401, 0701 and 1501. Percent maximal tracer binding was plotted against test peptide concentration and IC50 values were...
Results

Computational protein design
We applied DP\textsuperscript{2} to identify optimal and near-optimal variants of the wild-type P99BL (UniProtKB locus AMPC\_ENTCL/32–381), which we describe with positions numbered according to the crystal structure (pdb: 1XX2, chain A). For epitope analysis, we employed ProPred (Singh and Raghava, 2001) at a 5\% threshold. The initial epitope analysis predicts pervasive epitopes in P99BL: on the surface, in the core, and around the active site (Fig. 1). Conservative mutations were identified by analysis of Pfam 00144 homologs filtered for 35–90\% sequence identity to wild-type. Position-specific weighted frequencies were derived from a phylogenetic tree generated with ClustalW (Thompson et al., 1994), and target sites were substituted with only those amino acids appearing at >5\% weighted frequency.

DP\textsuperscript{2} identified A13D/I104T as the optimal two-mutation variant. A13D is predicted to eliminate binding by all five alleles that recognize wild-type 9mers overlapping residue 13 (Fig. 2a). I104T is predicted to eliminate all seven alleles binding one 9mer in the window around position 104 and three of the five alleles binding another overlapping 9mer (Fig. 2b). Categorization by allele reveals that A13D leaves no epitope on any allele and I104T deletes the epitopes for six of the eight alleles considered. DP\textsuperscript{2} selected this variant from a large and diverse set of allowable mutations. Interestingly, A13D deletes all epitopes without directly mutating any of the three predicted P1 anchors (L8, V11, V12). P1 anchor residues are hydrophobic amino acids that bind in the primary sequence, but no predicted epitope overlaps both residues (Fig. 2a and e).

Cloning, expression and purification
Wild-type P99BL was cloned from E. cloacae, expressed in E. coli, and secreted to the periplasmic space using an N-terminal OmpA signal sequence. The final construct included a C-terminal GGGS\_ETVEHHHHHH affinity tag, which facilitated purification by immobilized metal affinity chromatography. Mutant P99BL constructs were created by splice overlap extension PCR using the wild-type construct as template. The constructs tested included K21A/S324A (Harding et al., 2005), A13D/I104T, T42D/L149N and the cognate single mutants of both DP\textsuperscript{2} designs.

Induction of protein expression was found to be optimal at a high OD\textsubscript{600} (~3.5) and at a low temperature (16\°C). Initial studies with mid-log induction at 37\°C yielded protein which had undergone his-tag degradation despite the use of protease inhibitors during purification (data not shown). Protein purification from optimized conditions yielded homogenous preparations of >95\% purity (data not shown). This system allowed for robust and reproducible expression of P99BL enzymes, generally yielding between 8 and 40 mg protein per liter of culture after a 16 h induction.

Activity
The kinetics of the purified enzymes were analyzed using nitrocefin, a colorimetric β-lactam substrate. Results of the Michaelis–Menten kinetics analysis are provided in Table I. The parameters for the wild-type enzyme were in good agreement with prior reports (Bell and Pratt, 2002). The kinetics of the K21A/S324A double mutant benchmark, the only high-functioning construct detailed in prior deimmunization efforts (Harding et al., 2005), were statistically indistinguishable from the wild-type enzyme. The optimal DP\textsuperscript{2} design A13D/I104T also possessed wild-type kinetic parameters, and each of its component mutations yielded enzymes with wild-type activity. In contrast, the secondary T42D/L149N design exhibited a 3-fold decrease in \(k_{\text{cat}}\) and
2-fold decrease in $K_m$. Analysis of the corresponding individual point mutants suggested that the effects of the L149N mutation are dominant in the double mutant. In particular, both isolated mutations decreased $k_{cat}$, but L149N yielded an 80% reduction as compared with only a 30% reduction for T42D. Likewise, the favorable $K_m$ reduction of the T42D/L149N double mutant was attributable solely to L149N, as the isolated T42D mutation manifested a disadvantageous 50% increase in $K_m$ relative to wild-type. These tradeoffs between $k_{cat}$ and $K_m$ resulted in similar nitrocefin specificity constants ($k_{cat}/K_m$) for T42D, L149N and the corresponding double mutant.

Fig. 2. Design of P99BL deimmunizing mutations. For the wild-type P99BL (amino acids on the top row along the x-axis), epitopes for each of the eight common MHC II alleles are indicated as solid black lines spanning the 9mers. Allowed substitutions, as determined by evolutionary conservation statistics, are listed below each of the wild-type residues. Mutations selected with the DP$^2$ algorithm are boxed and highlighted in cyan: (A) A13D, (B) I104T, (C) T42D, and (D) L149N. Those from Harding et al. are boxed and highlighted in green: (E) K21A, and (F) S324A. Epitopes that remain after the specified mutation are shown as cyan or green hatched lines, respectively. Note that the K21A mutation from Harding et al. is predicted to introduce two new epitopes: one for DRB1*0401 and one for 0101.

### Table 1. P99BL kinetic parameters and apparent melting temperatures

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_m$ (µM)</th>
<th>$k_{cat}/K_m$ (s$^{-1}$ µM$^{-1}$)</th>
<th>$T_m$ (°C)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wild type</td>
<td>190 ± 10</td>
<td>33 ± 5</td>
<td>5.7 ± 0.2</td>
<td>56.04 ± 0.02</td>
</tr>
<tr>
<td>A13D</td>
<td>190 ± 10</td>
<td>32 ± 4</td>
<td>5.8 ± 0.1</td>
<td>48.74 ± 0.03</td>
</tr>
<tr>
<td>A13D</td>
<td>200 ± 10</td>
<td>34 ± 4</td>
<td>5.9 ± 0.1</td>
<td>56.31 ± 0.02</td>
</tr>
<tr>
<td>I104T</td>
<td>190 ± 10</td>
<td>32 ± 4</td>
<td>5.9 ± 0.2</td>
<td>49.23 ± 0.03</td>
</tr>
<tr>
<td>T42D, L149N</td>
<td>58 ± 1</td>
<td>19 ± 2</td>
<td>3.1 ± 0.1</td>
<td>50.48 ± 0.03</td>
</tr>
<tr>
<td>T42D</td>
<td>130 ± 10</td>
<td>48 ± 3</td>
<td>2.7 ± 0.1</td>
<td>55.35 ± 0.03</td>
</tr>
<tr>
<td>L149N</td>
<td>36 ± 9</td>
<td>15 ± 1</td>
<td>2.3 ± 0.3</td>
<td>51.87 ± 0.02</td>
</tr>
<tr>
<td>K21A, S324A</td>
<td>220 ± 20</td>
<td>38 ± 4</td>
<td>5.7 ± 0.1</td>
<td>56.06 ± 0.03</td>
</tr>
</tbody>
</table>
**Structural stability**

The relative thermostabilities of the various protein designs were assessed by differential scanning fluorimetry (Niesen et al., 2007), and apparent \( T_m \) data are provided in Table I. As with the kinetic analysis, the stability of the K21A/S324A benchmark was identical to that of the wild-type enzyme. The optimal DP2 plan A13D/I104T exhibited a 7°C decrease in \( T_m \), and analysis of the isolated point mutants traced the reduced stability to the I104T substitution. The secondary T42D/L149N design showed a similar 6°C decrease in stability, and analysis of the corresponding isolated point mutants indicated that the L149N substitution was the primary driver of this effect. While the decreased \( T_m \) values of the DP2 designs are notable compared with those of the wild-type and K21A/S324A proteins, inspection of the corresponding melting curves suggests that the two computationally designed variants exhibit negligibly small unfolded fractions at human body temperature (37°C, Supplementary Fig. S1), a point of particular relevance to therapeutic applications.

To obtain a more clinically relevant measure of the engineered proteins’ stability, variants were diluted into human plasma at 37°C, and specific activity towards the nitrocefin substrate was monitored as a function of time (Fig. 3). The wild-type enzyme and the optimal A13D/I104T variant each retained \(~70\%\) of their activity during extended incubation under biologically relevant conditions, although early time points indicate that the DP2 design is inactivated somewhat more quickly. The observed stability of the wild-type enzyme and the K21A/S324A variant are consistent with earlier studies (Harding et al., 2005), although notably, in our hands, the K21A/S324A variant was the single most stable protein. In contrast, the secondary DP2 design T42D/L149N exhibited a 60% decrease in activity by day 11. This compromised plasma stability appeared to result from a non-additive effect of the two mutations, as each of the corresponding point mutants was equally or more stable than the wild-type enzyme.

**MHC II binding**

MHC II binding assays were performed as a primary experimental measure of immunogenic potential for the wild-type and variant sequences. These assays represent a widely recognized means of validating computational predictions (Woodfolk, 2007; De Groot and Martin, 2009), and they are used here as a proxy for the success of epitope deletion. Peptides encompassing the predicted epitope sequences were synthesized and evaluated for their capacity to bind soluble DRB1*0101, 0401, 0701 and 1501 in a competition assay format. Peptide sequences and the IC50 values for each peptide/allele combination are provided in Fig. 4a and the fold-change in IC50 of engineered sequences versus...
wild-type are illustrated in Fig. 5. Peptides containing the wild-type A13 and I104 residues bound all four alleles with either moderate or high affinity, with the exception of I104 failing to bind DRB1*0701. Likewise, the L149 wild-type peptide bound with high affinity to all alleles. In contrast, the peptide containing wild-type T42 was a weak/non-binder of all alleles except DRB1*0101.

In general, DP2 mutations A13D and I104T significantly (α = 0.05) lowered MHC II binding affinities, with the exceptions of no change for DRB1*1501 and increased affinity of the I104T peptide for DRB1*0701. Peptides from the secondary DP2, containing mutations T42D and L149N, significantly lowered affinities across all alleles, with the exception of no change in binding of L149N to DRB1*0401 and 0701. Despite the substantial 18- and 48-fold decreased binding of the L149N peptide to alleles DRB1*0101 and 1501, respectively, its classification as a high-affinity binder across all alleles remained unchanged relative to the wild-type sequence. The two wild-type peptides containing either K21 or S324 are moderate to high-affinity ligands across all four MHC alleles tested. Notably, neither the K21A nor the S324A mutations from Harding et al. decreased MHC affinity for any allele, and in fact, these mutations generally increased binding affinity by a significant degree (α = 0.05). It should be noted that these mutations were selected based on reduced proliferation of human immune cells, which might be considered a gold standard for preclinical immunogenicity testing. That said, the blood donors used in the experimental mapping of the P99BL protein were not typed for MHC II allele status (Harding et al., 2005). Therefore it remains unknown exactly which alleles might have been represented, which might have been overrepresented, and which might have been completely absent from the donor cohorts.

To assess concordance of ProPred predictions with binding data, each peptide/allele combination was classified as either a true or false positive or as a true or false negative (Fig. 4b). With ProPred predictions set at a 5% threshold and binding set at IC50< 100 μM, the overall false-positive rate is 12.5% and false-negative rate 66%. The A13 and I104 peptides bound all alleles as predicted, with the exception of I104 binding to DRB1*0401. Binding predictions for mutations at these positions were borne out, with the exception of the A13D peptide, which bound DRB1*0701 and 1501 despite negative predictions. Notably, however, the IC50 for A13D binding DRB1*0701 was increased 60-fold to a value of 70 μM, and thus prediction of a deimmunizing effect may not be entirely unfounded. In general, predictions were confirmed for the wild-type T42 and L149 peptides; unexpectedly, however, the T42 peptide did not bind DRB1*0701 nor 1501. The T42D and L149N mutations were predicted to disrupt MHC II binding across all alleles tested. This correlated with the engineered T42D peptide evading binding to DRB1*0401, 0701 and 1501 in practice, but it was found to bind DRB1*0101 with high affinity. As noted above, the engineered L149N peptide proved to be a high-affinity binder with all alleles tested. With regard to the Harding et al. targets, though the K21 and S324 peptides and their alanine variants are not predicted to bind either DRB1*0701 or 1501, all bound. Thus, binding results, not ProPred predictions, suggest that peptides encompassing K21 and S324 are good targets for deletion of DRB1*0701 and 1501 epitopes, and it is possible that adjacent mutations predicted to disrupt binding to other alleles (Fig. 2e and f) might fortuitously mitigate DRB1*0701 and 1501 binding. Regardless, it is apparent that alanine substitutions do not adequately disrupt interaction with these two MHC II alleles. Finally, predictions of S324 binding to DRB1*0101 and 0401 and K21 binding to 0401 were positively confirmed, and as predicted, both of the respective alanine mutants bound DRB1*0101 and 0401.

**Discussion**

The literature provides numerous examples of protein deimmunization via T-cell epitope deletion. For example, experimental identification and subsequent disruption of T-cell epitopes has produced less immunogenic variants of staphylokokine (Warmerdam et al., 2002), erythropoietin (Tangri et al., 2005), factor VIII (Jones et al., 2005) and P99BL
(Harding et al., 2005). Experimental characterization of a protein’s T-cell epitope profile is, however, expensive in terms of both time and resources. Furthermore, identification of mutations that simultaneously deimmunize and preserve function is inherently complex. While both computational epitope prediction and stability evaluation tools are routinely employed by protein engineers (Polizzi et al., 2006; Koren et al., 2007; Jochens et al., 2010; Cantor et al., 2011; Moise et al., 2012), to date such tools have not been integrated to solve the multi-objective immunogenicity and stability/activity problem.

Here, we describe the first experimental application of our recently developed DP2 deimmunization methodology (Parker et al., 2010). In the interest of benchmarking our results against previous experimental efforts (Harding et al., 2005), we chose P99BL as a model system. Our objective was the efficient design and construction of enzyme variants exhibiting good catalytic proficiency, high-level stability and reduced epitope content. We applied the DP2 algorithm to the P99BL target and selected two protein designs: the globally optimal double mutant A13D/I104T and the secondary T42D/L149N design. To more clearly define the resulting structure–activity relationships, all four corresponding single mutants were also produced.

In each of the double mutant DP2 designs, a hydrophobic T-cell epitope anchor residue was targeted for mutation (I104T and L149N, respectively). The I104T substitution caused a 7°C decrease in apparent melting temperature, and the L149N substitution a 4°C decrease. In the case of the T42D/L149N design, the loss of thermostability correlated with a significant reduction in stability during long-term incubation in human plasma. The native L149 is close-packed (~10% solvent accessible surface area) at the center of a relatively hydrophobic pocket ringed by G116, P118, L119, W142, P295 and L293. Thus, the decreased stability of the double mutant might be rationalized in part by unfavorable energetics resulting from substitution of a highly polar paragine into this nonpolar environment. It should be emphasized, however, that decreased serum stability was only observed in the context of the T42D/L149N double mutant, suggesting a cooperative destabilizing effect of the two mutations. In contrast, the globally optimal DP2 design A13D/I104T exhibited 11-day plasma stability similar to wild-type P99BL, although its initial rate of inactivation may be somewhat faster than wild-type.

Notably, the K21A/S324A benchmark variant exhibited the highest level stability of any protein tested. In retrospect, this outcome is not unexpected as the DP2 algorithm predicts that both mutations are allowed by conservation analysis. More importantly, and as described in the introduction, this particular variant was the single best protein generated during an extensive and costly deimmunization effort that experimentally tested a large number of designs prior to focusing efforts on the most promising enzyme, K21A/S324A. Given this context, it bears noting that the DP2 designs tested here were the only two computational designs that were experimentally evaluated. Thus, we conclude that there is likely a beneficial tradeoff between moderately decreased protein stability, as observed with our designs, yet radically reduced time/effort to generate deimmunized candidates, via DP2. Certainly, it would be desirable to incorporate more accurate models of protein stability into the deimmunization algorithm, and ongoing efforts are indeed driving the continued evolution of our computational methods (Parker et al., 2011, 2012).

In contrast to the relatively low functional ‘hit rate’ of trial and error deimmunization methodologies, both DP2 designs retained comparatively high catalytic activity. In particular, the activity of the optimal A13D/I104T design was indistinguishable from wild-type, as were the activities of the corresponding individual mutants. Our secondary T42D/L149N design exhibited a 3-fold reduction in $k_{cat}$, and experiments with the constituent point mutants traced this effect to the L149N substitution. Inspection of the enzyme’s crystal structure shows that L149 is located immediately adjacent to Y150, a central residue in the charge relay system that deprotonates the S64 nucleophile (Goldberg et al., 2003) (Fig. 1). In general, active sites and buried hydrophobic residues have a low tolerance for mutation (Guo et al., 2004). Therefore, the reduced activity of T42D/L149N is not entirely surprising. Nonetheless, this variant accelerates nitrocefin hydrolysis by more than 10²-fold relative to the uncatalyzed reaction (uncatalyzed pseudo-first order rate constant $=4 \times 10^{-5}$ s⁻¹, Osipovitch unpublished results), and from a purely biochemical stand point, T42D/L149N is a very respectable catalyst. Overall, the fact that designs containing the hydrophobic-to-polar substitutions L149N, adjacent to a key active site residue, and I104T, a completely buried residue, retained reasonable levels of activity underscores our algorithm’s capacity for identifying deimmunizing mutations that are non-intuitive yet high-functioning. Such an outcome can only be guaranteed by integrated optimization of immunogenicity and stability/activity.

While catalytic proficiency and structural stability were important design constraints for our engineered enzymes, we ultimately sought variants having reduced T-cell epitope content. As an immunological measure of performance, we employed quantitative MHC II-peptide binding assays, and we set a relatively stringent requirement that peptide-MHC II IC₅₀ values should be increased above 100 μM for a mutation to qualify as ‘fully deimmunizing’. By this metric, our DP2 designs significantly outperformed both the wild-type enzyme and the previously generated K21A/S324A double mutant (Harding et al., 2005). In particular, our optimal A13D/I104T mutations successfully eliminated binding of the cognate peptides to the DRB1*0101 and 0401 alleles, and we speculate that the full-length variant protein would benefit from reduced immunogenicity in human patients.

It is worth noting that while the A13D mutation significantly decreased the binding affinity of its cognate peptide for the DRB1*0701 allele, the I104T mutation increased DRB1*0701 binding affinity. In the latter case, we emphasize that the wild-type I104 peptide itself is a moderate binder of this allele, and therefore the I104T mutation did not introduce a new epitope, but rather simply failed to eliminate a putative epitope that was already present in the wild-type sequence. On a global scale, failure to delete a single epitope for one MHC II allele does not preclude a successful deimmunization outcome, i.e. reduction or prevention of anti-drug antibodies, provided deletion of other epitopes lowers immunogenicity below a biologically relevant threshold (Yeung et al., 2004; Harding et al., 2005). Moreover, and importantly, the DP2 algorithm does not include a quantitative predictor of peptide–MHC II binding affinity, but rather
the algorithm assesses binding in a binary yes/no fashion based on a user-supplied false positive rate (5% in this case). Thus, the algorithm accurately predicted the fact that the I104T mutation would not eliminate binding to the DRB1*0701 allele (Fig. 4b), and consequently the relative affinities of the wild-type and mutant 104 peptides were irrelevant to the optimization calculations. In the future, quantitative predictors of peptide–MHC II binding affinity could be incorporated into the deimmunization algorithm, thereby accounting for relative changes in binding and rendering the methods less sensitive to threshold effects.

In examining the immunological performance of the secondary DP² design, it was found that the T42D and L149N mutations failed to completely eliminate MHC II binding as per our stringent threshold of IC₅₀ > 100 μM. However, the T42D mutation decreased the affinity of its cognate peptides for all four MHC II alleles, and the L149N mutation significantly decreased peptide binding affinity for two of four alleles. Notably, the L149 containing wild-type peptide exhibited the highest binding affinity of any tested peptide for alleles DRB1*0101 and 1501, and the second highest binding affinity towards alleles DRB1*0401 and 0701. The L149-containing peptide may therefore represent a highly immunogenic epitope, and mitigating the binding affinity of this peptide could ultimately prove important for clinical applications. Indeed, Harding et al. had also identified L149 as a potentially productive target site for deimmunization, but their stability analyses on crude enzyme preparations suggested that alanine substitution at this site resulted in unacceptable low stability (80% reduction by their measures) (Harding et al., 2005). To effectively target this immunogenic region while avoiding inadvertent introduction of new epitopes, DP² selected the less conservative but functional L149N substitution, which, while failing to eliminate MHC II binding of cognate peptide, proved to reduce binding affinity to the DRB1*0101 and 1501 alleles by 18-fold and 50-fold, respectively. It remains to be seen whether or not these significantly decreased affinities are sufficiently disruptive to inhibit T-cell stimulation.

**Conclusion**

The underlying premise of the DP² algorithm represents a paradigm shift for biotherapeutic deimmunization strategies. It formalizes the deimmunization problem and combines epitope prediction and stability estimation within a powerful computational framework for therapeutic protein design. DP² and more advanced implementations (Parker et al., 2011, 2012) are the first and, currently, only computational tools with the capacity to simultaneously optimize protein sequences for both low immunogenicity and high stability and activity. In contrast to serial analysis of immunogenicity then stability, the integrated nature and mathematical correctness of our dynamic programming algorithm guarantees identification of optimal protein designs with respect to the predictors. Moreover, DP² finds secondary variants in rank order, allowing for the elucidation of alternative molecular solutions and trends among high-scoring plans. In stark contrast to laborious trial-and-error deimmunization strategies, our methodology benefits from unprecedented speed and robust efficacy, as shown by the enhanced performance of both P99PL designs tested in this work. Importantly, the modular nature of DP² allows easy extensibility to new state-of-the-art predictors, and the continuing evolution of prediction tools will only enhance the algorithm’s functionality as opposed to driving its obsolescence. Because our algorithms are applicable to virtually any protein sequence, we anticipate that they will find broad utility in the growing field of therapeutic protein development.

**Supplementary data**

Supplementary data are available at PEDS online.

**Acknowledgements**

The authors would like to thank Dr Yingda Xu for assistance with mass spectral analysis of protein samples.

**Funding**

This work was supported by the National Institutes of Health [1 R01 GM 989771–01 to C.B.K. and K.E.G. and P20 R01 8787 to K.E.G.], the National Science Foundation [IIS-1017231 to C.B.K.] and the Neukom Institute for Computational Science [CompX Faculty Award to C.B.K. and K.E.G.].

**References**

Engineering immune-evading enzymes for ADEPT


