Published in final edited form as:

Cell Mol Life Sci. 2014 December ; 71(24): 4869–4880. doi:10.1007/s00018-014-1652-x.

Computationally driven deletion of broadly distributed T cell epitopes in a biotherapeutic candidate

Regina S. Salvat,

Thayer School of Engineering, Dartmouth College, 14 Engineering Dr., Hanover, NH 03755, USA

Andrew S Parker,

Department of Computer Science, Dartmouth College, 6211 Sudikoff Laboratory, Hanover, NH 03755, USA

Andrew Guillems,

Thayer School of Engineering, Dartmouth College, 14 Engineering Dr., Hanover, NH 03755, USA

Yoonjoo Choi,

Department of Computer Science, Dartmouth College, 6211 Sudikoff Laboratory, Hanover, NH 03755, USA

Chris Bailey-Kellogg, and

Department of Computer Science, Dartmouth College, 6211 Sudikoff Laboratory, Hanover, NH 03755, USA

Karl E. Griswold

Thayer School of Engineering, Dartmouth College, 14 Engineering Dr., Hanover, NH 03755, USA; Program in Molecular and Cellular Biology, Dartmouth College, Hanover, NH, USA

Chris Bailey-Kellogg: cbk@cs.dartmouth.edu; Karl E. Griswold: karl.e.griswold@dartmouth.edu

Abstract

Biotherapeutics are subject to immune surveillance within the body, and anti-biotherapeutic immune responses can compromise drug efficacy and patient safety. Initial development of targeted antidrug immune memory is coordinated by T cell recognition of immunogenic subsequences, termed “T cell epitopes.” Biotherapeutics may therefore be deimmunized by mutating key residues within cognate epitopes, but there exist complex trade-offs between immunogenicity, mutational load, and protein structure–function. Here, a protein deimmunization

© Springer Basel 2014

Correspondence to: Chris Bailey-Kellogg, cbk@cs.dartmouth.edu; Karl E. Griswold, karl.e.griswold@dartmouth.edu.

Present Address: A. S. Parker, Zettics, 150 Baker Ave. Ext., Ste. 201, Concord, MA, USA

Present Address: A. Guillems, Mascoma Corp., 67 Old Etna Road #300, Lebanon, NH, USA

Electronic supplementary material The online version of this article (doi:10.1007/s00018-014-1652-x) contains supplementary material, which is available to authorized users.

Conflict of interest Karl E. Griswold and Chris Bailey-Kellogg are Dartmouth faculty and co-members of Stealth Biologics, LLC, a Delaware biotechnology company. They acknowledge that there is a potential conflict of interest related to their association with this company, and they hereby affirm that the data presented in this paper is free of any bias. This work has been reviewed and approved as specified in these faculty members' Dartmouth conflict of interest management plans. The remaining authors declare no conflict of interest.

algorithm has been applied to P99 beta-lactamase, a component of antibody-directed enzyme prodrug therapies. The algorithm, integer programming for immunogenic proteins, seamlessly integrates computational prediction of T cell epitopes with both 1- and 2-body sequence potentials that assess protein tolerance to epitope-deleting mutations. Compared to previously deimmunized P99 variants, which bore only one or two mutations, the enzymes designed here contain 4–5 widely distributed substitutions. As a result, they exhibit broad reductions in major histocompatibility complex recognition. Despite their high mutational loads and markedly reduced immunoreactivity, all eight engineered variants possessed wild-type or better catalytic activity. Thus, the protein design algorithm is able to disrupt broadly distributed epitopes while maintaining protein function. As a result, this computational tool may prove useful in expanding the repertoire of next-generation biotherapeutics.

Keywords

Biotherapeutics; Protein engineering; Immunogenicity; Deimmunization; Epitope deletion; Computational protein design

Introduction

The development of increasingly diverse and powerful therapeutic proteins presents the biotechnologist with unprecedented opportunities as well as unique design challenges. As a class of macromolecules, proteins exhibit intrinsically complex sequence-structure-function relationships. From a drug development perspective, this complexity opens the door to a vast molecular design space that includes bioactive agents able to interface with virtually any disease target or pathway. At the same time, the biological origins of therapeutic proteins impart undesirable risk factors, such as the potential to induce an antidrug immune response in human patients [1, 2]. Immune reactions directed against biotherapeutics can compromise drug efficacy or cause more serious adverse events, such as anaphylaxis, infusion reactions, or induced autoimmunity [3, 4]. To fully capitalize on the emerging panacea of protein therapeutics, efficient and effective deimmunization strategies are needed.

While immunogenicity is a complex multifaceted issue, a primary driver of antidrug immune responses is recognition of non-self-proteins as foreign antigens [5]. In a healthy and functioning human immune system, extracellular proteins are constantly sampled by antigen presenting cells (APCs) [6]. Once internalized by APCs, a protein is cleaved into small peptide fragments, and putative immunogenic segments are loaded into the groove of class II major histocompatibility complex proteins (MHC II). The peptide-MHC II complexes are then displayed on the APC surface, where bona fide immunogenic peptides, termed “T cell epitopes,” form ternary complexes with surface receptors of cognate CD4⁺ T cells. This molecular recognition event initiates a signaling cascade that leads to T cell stimulation, B cell maturation, and ultimately production of circulating antibodies able to bind and clear the offending exogenous protein. Thus, inherently immunogenic proteins might be deimmunized by identifying constituent T cell epitopes and mutating residues responsible for MHC II complex formation. Initially, T cell epitope deletion strategies were purely experimental in nature [7–10]. These molecular engineering programs, which

integrated peptide-level epitope mapping with alanine scanning or similar site-directed mutagenesis, were time-, labor-, and resource-intensive. Moreover, they examined only a small segment of the entire protein design space, typically focusing on just one or two immunogenic regions. To more fully explore the sequence space associated with deimmunized proteins, higher-throughput search and design tools are needed.

Advances in bioinformatics have produced a variety of programs geared toward prediction of immunogenic T cell epitopes. Examples include ProPred [11], MHCpred [12], SVRMHC [13], aRB [14], SMM-align [15], NetM-HCIIpan [16], the IEDB consensus method [17], as well as proprietary tools such as Epimatrix [18] and iTope [19]. Many of these methods have demonstrated good predictive power [17, 20], and combined with appropriate experimental validation [21–23], they have the potential to accelerate biotherapeutic design and development. Indeed, such predictive tools have been successfully applied to numerous therapeutic candidates [1, 18, 24–26]. These prediction-based epitope deletion projects sometimes leverage bioinformatics (e.g., substitution matrices or evolutionary sequence conservation in homologous proteins) to identify epitope-deleting yet function-preserving mutations. However, at best they do so in a post hoc fashion that is typically focused only on one or two immunogenic peptides. While such sequential application of computational tools is not unreasonable, it precludes a global approach that simultaneously accounts for the combined effects of all mutations on protein structure, function, and immunogenicity.

Next-generation protein deimmunization tools have now seamlessly integrated T cell epitope prediction with global modeling of every mutation's functional consequences [27–30]. Dynamic programming for deimmunizing proteins (DP²) was the earliest of these optimization algorithms [27]. DP² employs either BLOSUM or position-specific substitution matrices as a prefilter to select a subset of deimmunizing mutations that are likely to preserve function. It then selects optimal sets of these mutations so as to minimize epitope content. The method has been experimentally evaluated using the *Enterobacter chloacae* P99 beta-lactamase (P99βL), a component of antibody-directed enzyme prodrug therapies. As an important point of reference, P99βL has previously undergone conventional, experimentally driven T cell epitope deletion. These benchmark experimental studies yielded a functional 2-mutation variant (K21A/S324A) that exhibited reduced immunogenicity in model studies [9]. More recently, application of DP² to P99βL generated a globally optimal 2-mutation design (A13D/I104T) that, upon construction and analysis, was found to possess catalytic activity equal to that of the wild-type enzyme [31]. Importantly, peptide fragments of the DP² design generally showed large reductions in binding affinity with human MHC II proteins, a measure of immunogenic potential. Thus, in contrast to arduous epitope mapping and scanning alanine mutagenesis, DP² required only minutes to design a P99βL variant possessing reduced immunoreactivity and high-level functionality.

While these initial studies demonstrated the speed and utility of integrated deimmunization algorithms, the DP² methodology has important limitations. In essence, DP²'s conservation-based mutation filter assumes that each residue position is independent of all other positions, and it therefore fails to account for deleterious (or beneficial) residue–residue interactions. Accounting for dependence between pairs of amino acids (“coupling”) becomes increasingly

important for proteins containing large numbers of distributed epitopes; there is a greater risk of disrupting coupled interactions at higher mutational loads. To accommodate this more complex design challenge, a revised methodology incorporating a 2-body sequence potential has been developed. This more advanced deimmunization algorithm, integer programming for immunogenic proteins (IP²), accounts for critical residue–residue interactions by quantifying covariation in the evolutionary sequence record. The capacity to quantify residue–residue coupling during molecular evolution has proven to be a powerful tool in bioinformatics [32], but incorporating such 2-body terms renders the protein design problem NP-hard. By formulating the dual objective optimization as an integer program, IP² has solved numerous real case studies using modest computational resources in practical time [28].

As an experimental validation of IP², a more aggressive deimmunization of P99βL has been pursued here. The results suggest that numerous epitopes, broadly distributed throughout a protein, can be quickly, efficiently, and functionally deleted using the IP² protein design tool.

Materials and methods

Materials

Oligonucleotides for sequencing and standard PCR methods (25 nmol scale, standard desalting) and oligonucleotides for gene synthesis (100-nmol scale, PAGE Purified) were purchased from Integrated DNA technology (San Diego, CA). Nitrocefin was purchased from Oxoid (Cambridge, UK). Human lysozyme and SYPRO Orange 5,000 × Protein Stain were purchased from Sigma (St. Louis, MO). MicroAmp[®] Fast Optical 0.1 ml 96-Well Plates and MicroAmp[®] Optical Adhesive Film were from Applied Biosystems (Bedford, MA). Restriction enzymes and PCR reagents were purchased from New England Bio-Labs (Ipswich, MA). Growth media were purchased from Becton–Dickinson (Franklin Lakes, NJ). Plasmid purification kits and Ni–NtA resin were purchased from Qiagen (Valencia, CA). PCR cleanup and gel extraction kits were from Zymo Research (Irvine, CA). Peptides derived from P99βL were ordered from GenScript (Piscataway, NJ) and were greater than 85 % pure. Biotinylated tracer peptides were purchased from twenty-first century Biochemicals (Marlborough, MA). MHC-II-DR molecules were purchased from Benaroya Research Institute (Seattle, Wa), anti-MHC-II-DR antibody from Biolegend (San Diego, CA), and DELFIA Eu-labeled streptavidin was from PerkinElmer (Boston, MA). Unless noted, all other chemicals and reagents were from VWR (Radnor, PA).

Computational deimmunization

To identify and assess mutations likely to be functionally acceptable, an IP² sequence potential was developed as described [28]. A multiple sequence alignment (MSA) of 94 homologs from Pfam 00144, including the wild type, was obtained by filtering for at least 30 % sequence identity to wild type, at most 90 % sequence identity to each other, and at most 25 % gaps. Position-specific one-body terms $\phi_i(a)$ were computed as the negative log frequency of each amino acid a at each position i . Only amino acids appearing at or above background frequencies [33] were considered as possible substitutions. Two-body terms $\phi_{i,j}$

(a,b) for pairs of amino acids (a,b) at coupled positions (i,j) were computed as the negative log amino acid frequency of the pair, minus the corresponding one-body terms, which avoids double counting. Only pairs of positions with significant coupling according to a χ^2 -based test were included in the sequence potential.

In addition to mutational constraints based on the evolutionary sequence record, residues in close proximity to the active site were held invariant. In particular, all positions with $C\beta$ (or Ca in the case of glycine) less than 7 Å from $C\beta$ of S64, Y150, K315, T316, or Ca of G317 were constrained to their wild-type residues. This active site lockdown was prompted by previous P99βL deimmunization efforts in which mutations proximal to key catalytic residues were found to compromise enzyme function (e.g., mutations to L149, see [9, 31]). Finally, prolines and cysteines were neither mutated out of nor substituted into the engineered enzyme variants.

To assess the effects of the functionally acceptable mutations on T cell epitope content, the ProPred epitope predictor [11] was applied at a 5 % threshold. The analysis considered MHC II alleles DRB1*0101, 0301, 0401, 0701, 0801, 1101, 1301, and 1501, which are known to be broadly representative of global populations [34]. Each non-amer peptide X considered in the optimization (i.e., incorporating a contiguous combination of wild-type residues and allowed substitutions) was classified as either a binder or non-binder of the eight different MHC II alleles. The number of binders was summed to generate the non-amer's epitope score $e(X)$.

The IP² objective function incorporates both the sequence potential (one-body ϕ_i and two-body $\phi_{i,j}$) and the epitope score (e), with relative contributions weighted by α and $(1 - \alpha)$, respectively

$$\alpha \left(\sum_{i,a} s_{i,a} * \phi_i(a) + \sum_{i,j,a,b} p_{i,j,a,b} * \phi_{i,j}(a,b) \right) + (1 - \alpha) \sum_{i,X} w_{i,X} * e(X) \quad (1)$$

Here, we define singleton binary variable $s_{i,a}$ to indicate whether or not the residue at position i is of amino acid type a . We define pairwise binary variable $p_{i,j,a,b}$, derived from $s_{i,a}$ and $s_{j,b}$, to indicate whether or not the residues at positions i and j are of amino acid types a and b , respectively. Finally, we define window binary variable $w_{i,X}$, derived from $s_{i,X[1]}$ through $s_{i+8,X[9]}$, to indicate whether or not the sequence of residues in the 9-residue window starting at position i is of the sequence of amino acid types in X . The algorithm chooses (sets to 1 instead of 0) some of the s , p , and w variables so as to minimize the objective function (sequence potential and epitope score), subject to constraints on the desired mutational load and consistency of the chosen variables (one amino acid at each position; pair and epitope variables derived from the corresponding single variables).

Cloning, expression, and purification

Gene synthesis was performed with an assembly reaction followed by an amplification reaction [35] using oligonucleotides that were 41 base pairs in length. All designs encoded an N-terminal OmpA leader sequence for expression in the periplasm and a C-terminal

hexa-His tag (sequence GGGSAETVEHHHHHH) for simplified purification. Genes were cloned into pET26b and expressed in BL21(DE) *E. coli* cells [F^- *ompT* *hsdS_B* ($r_B^- m_B^- gal dcm$ (DE3))].

Production of wild-type P99 β L and variants A, B, C, D, E, F, G, and H was performed as follows. Starter cultures were grown with aeration in LB medium containing 30 μ g/ml kanamycin (LB-Kan) at 37 °C overnight, subcultured 1:100 into 200–500 ml of fresh medium, and grown at 37 °C using 2 L baffed flasks until reaching an OD₆₀₀ of ~3.5. Cultures were then transferred to a 16 °C shaker, equilibrated for 15 min, and induced with 1 mM IPTG. Following 16 h of induction, protein was purified using the protocol described in the Epicentre® PeriPreps Periplasting Kit with slight modifications. Briefly, cells were pelleted at 6,000g for 10 min and resuspended in PeriPreps Periplasting Buffer containing 1.5 μ g/ml human lysozyme. After 5 min, cells were quenched with ice-cold water, recovered on ice for 10 min, spun at 14,000 g for 10 min, and the supernatant was reserved as the periplasmic fraction. A 400- μ l bed volume of Ni-NTA resin was washed with water and equilibrated with PBS (137 mM NaCl, 2.6 mM KCl, 10 mM Na₂HPO₄, 1.7 mM KH₂PO₄, pH 7.4). Clarified periplasmic fraction was flowed through the resin by gravity, the column was washed two times with 1 ml of PBS containing 20 mM imidazole, and the enzyme was eluted with 2 ml of PBS containing 200 mM imidazole. The elution fraction was either dialyzed against three changes of 4 L PBS or subjected to spin column ultrafiltration against three washes of 15 ml PBS and concentrated to 0.5–2 mg/ml. Purified protein was stored at 4 °C prior to further analysis. The purities of all proteins were determined by reverse-phase HPLC analysis (Agilent 1200 Series HPLC) on a Vydac 214TP 180 mm C4 column, eluted at 65 °C with a 25–45 % gradient of [90 % acetonitrile/9.9 % water/0.1 % trifluoroacetic acid] in [99.9 % water/0.1 % trifluoroacetic acid] at a flow rate of 1 ml/min. All protein preparations were >95 % pure.

Kinetic studies

Nitrocefin substrate stock was prepared immediately prior to the experiments by dissolving nitrocefin powder in DMSO to a concentration of 20 mM. Triplicate assays were run in 96-well plate format at 30 °C measuring absorbance at 490 nm (Molecular Devices SpectraMax 190 plate reader). Absorbance measurements were converted to micromolar product concentrations using the appropriate molar absorptivity ($\epsilon_M = 20,500 \text{ M}^{-1} \text{ cm}^{-1}$). The assay buffer was PBS, and each well contained a final enzyme concentration of 50 ng/ μ l, 0.04 % BSA, and nitrocefin at concentrations ranging from 10 to 500 μ M. Initial reaction rates were plotted against substrate concentration, and Michaelis–Menten kinetic parameters were determined by nonlinear regression using GraphPad Prism v.5 software (La Jolla, Ca). Measurements were made in triplicate, and enzymes were purified and assayed in biological duplicate.

Thermostability

Differential scanning fluorimetry was performed essentially as described (Niesen, Berglund et al. 2007) using an ABI 7500 Fast Real-Time PCR System from Applied Bio-systems (Bedford, MA). Proteins and SYPRO Orange were diluted in PBS. Final protein

concentrations were 100 µg/ml, and final dye concentrations were 5× (from the supplier's 5000× concentrate). Twenty microliter reactions were performed in 12 replicates. The PCR gradient was run from 25 to 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 v.4 software.

MHC Binding assays

MHC II competition binding assays were performed using a 384-well high throughput assay as described in detail elsewhere [23]. Binding assays were performed for the four alleles: DRB1*0101, 0401, 0701, and 1501, which provide broad representation of class II MHC binding pockets [34]. The data were fit to the one-site log(IC₅₀) model (Eq. 2) by nonlinear regression in GraphPad Prism v.5 software.

$$Y = \text{Bottom} + (\text{Top} - \text{Bottom}) / (1 + 10^{(X - \text{LogIC}_{50})}) \quad (2)$$

Result

Computational design

The majority of predicted epitopes in wild-type P99βL occur in clusters that are broadly distributed throughout the P99βL sequence and structure (Fig. 1). The fundamental objective of the current work was deletion of broadly dispersed epitopes using function-preserving mutations. IP² was implemented to generate optimal, deimmunized P99βL designs at mutational loads of both 4 and 5 substitutions per variant (four different plans at each mutational load). The α term in the IP² objective function (see materials and methods) provides a relative weight for sequence score versus epitope score. At α = 1, IP² will generate designs with the best possible sequence score irrespective of the epitope score (conservative designs). Conversely, at α = 0, designs will have a minimal epitope score irrespective of the sequence score (aggressive designs). As a preliminary test of the trade-offs inherent to the two objective functions, plans were designed over a range of α values (Table 1). Interestingly, only seven sites were mutagenized among the entire panel of eight enzyme variants (Fig. 1). Accessing greater diversity of substitutions at the 4- and 5-mutation loads would therefore require lower values of α, which would in turn skew designs away from the more conservative mutations.

Although the deimmunized enzyme plans showed a high degree of relatedness, there existed systematic trends that correlated with the design parameters and related objectives (Table 1; Fig. 2). The K21E substitution was shared by all eight designs, as it was attractive from the perspective of both conservation (frequency of 14 % in the MSA) and immunogenicity (deleting 6 out of eight associated epitopes). (Note: we report here only the easy-to-see one-body conservation frequency as an indicator of conservativeness, but optimization accounts for two-body coupling as well.) Similarly, R105S was present in all but the most conservative 4-mutation design, as it possesses a frequency of 32 % and deletes 6 of 12 associated epitopes. The most aggressive plans (α = 0.5 and 0.7) included mutations at both A13 and I262. Both aspartic acid and glutamic acid were substituted at the former. A13D,

associated with $\alpha = 0.5$, is less desirable with respect to the sequence score (9 % frequency) yet deletes all seven predicted epitopes. Conversely A13E, associated with $\alpha = 0.7$, is predicted to be more conservative (55 % frequency) but fails to disrupt one of two predicted DRB1*0401 binders. Threonine was the only alternative residue at I262, appearing at a frequency of 15 % and deleting five of eight epitopes. Mutations at A13 and I262 were conspicuously absent in the most conservative plans, whereas R210H (deletes 3/3 epitopes, 23 % frequency) was selected at $\alpha = 0.9$ and R133H (deletes 3/4 epitopes, 6 % frequency) at $\alpha = 0.97$. Mutations at M235 were the final distinguishing feature of aggressive versus conservative designs. Specifically, aggressive 4-mutation plans avoided residue M235, but conservative plans ($\alpha = 0.96$) included the M235Q mutation (15 % frequency), which is predicted to delete 4/5 wild-type epitopes while introducing one new DRB1*1301 epitope (Fig. 2). Similarly, the most conservative 5-mutation plan ($\alpha = 0.97$) also incorporated M235Q, whereas the remaining 5-mutation designs employed M235D (10 % frequency). The more aggressive M235D substitution deletes 4/5 wild-type epitopes while avoiding introduction of any new epitopes.

Cloning, expression, and purification

Genes encoding IP² designs with C-terminal hexa-histidine tags were assembled from overlapping synthetic oligonucleotides, cloned into the pET26b vector, and expressed in BL21(DE3) *E. coli* hosts. Enzymes were purified from osmotic shockates by immobilized metal affinity chromatography, yielding >95 % purity as analyzed by HPLC (data not shown). The wild-type protein yielded 30 μ g per milliliter culture volume, whereas all other designs yielded between 1 and 9 μ g per milliliter culture volume. The differences in yields may be due to the fact that mutant genes were optimized for PCR assembly, yielding distinctly different codon usage compared to the wild-type P99 β L construct [31]. Purified enzymes were dialyzed into phosphate-buffered saline (PBS) and stored at 4 °C prior to analysis.

Thermostability analysis

To assess the structural integrity of the deimmunized designs, their relative thermostability was quantified using differential scanning fluorimetry (table 1). The apparent T_m of the wild-type enzyme, 56.37 °C, closely matched that reported previously [31]. In comparison, each of the IP² designs experienced a moderate decrease in stability, although none showed any detectable unfolding at the clinically relevant temperature of 37 °C (Supplemental Fig. 1). The 4-mutation designs exhibited an average $T_m = 50.8$ °C, with the most conservative design D ($\alpha = 0.98$) having the highest T_m of any variant tested (52.81 °C). In comparison, the 5-mutation designs exhibited a lower average $T_m = 49.6$ °C, although the difference was not statistically significant ($P = 0.1727$, two-tailed t test). In aggregate, the eight designs generated by IP² (overall average $T_m = 50.2$ °C) compared favorably with prior DP² designs (average $T_m = 52.0$ °C, across six variants [31]) despite the fact that the proteins analyzed here bore up to fivefold higher mutational loads.

Kinetic analysis

The functionality of the P99 β L enzymes was quantified by kinetic analysis using the colorimetric substrate nitrocefin (Table 1). Michaelis–Menten parameters for the wild-type enzyme were found to correlate well those from the literature [9, 31]. The IP² variants' parameters were compared to those of wild-type P99 β L (extra sum of squares F test), and deimmunizing mutations were found to have no significant effect on K_m , with the exception of variant E, which was improved by 23 % ($P = 0.0208$). In contrast, seven of the eight deimmunized variants exhibited maximum reaction rates that were significantly different from that of wild-type P99 β L ($P < 0.05$); variant B was equivalent to wild-type, variant E exhibited a marginally decreased k_{cat} (6 % reduction), and the remaining six variants exhibited enhanced k_{cat} . Of particular interest was variant G, a 5-mutation design that showed a 62.5 % increase in k_{cat} . Similarly, variants A, C, and H each exhibited 30 % or greater enhanced maximum rates, while variants D and F were found to have more modest gains (13–16 %). Seven of the eight variants possessed mild to moderately improved catalytic efficiency (10–35 % higher k_{cat}/K_m).

Interestingly, the average performance of the 4-mutation designs was not significantly different from that of the 5-mutation designs ($P = 0.705, 0.983, 0.804$ for k_{cat} , K_m , and k_{cat}/K_m , respectively; unpaired two-tailed t tests). Catalytic performance did not correlate closely with α , but this is likely due to the fact that all eight IP² variants exhibited uniformly high activity (essentially wild type or better). Thus, the 1- and 2-body sequence potential implemented in IP² was able to identify highly functional mutations even under conditions where equal weighting was placed on the deimmunizing objective function.

Peptide-MHC II

The immunoreactive potential of P99 β L peptide fragments was evaluated via quantitative binding studies with recombinant human MHC II. For the wild-type sequence, synthetic peptides were designed so as to span all epitopes associated with each of the seven sites that were mutagenized (A13, K21, R105, R133, R210, M235, and I262) (Fig. 2). Cognate peptides were synthesized for each of the nine mutations that comprised the deimmunized variants. The affinities of wild-type and engineered peptides for human MHC II alleles DRB1*0101, 0401, 0701, and 1501 were then quantified and compared (Fig. 3). Of the 36 pairwise comparisons, mutations decreased immunoreactivity in 21 cases, had no effect in 12 cases, and increased MHC II binding in only three cases. Mutations designed by IP² can therefore be considered generally disruptive with respect to MHC II immunoreactivity.

Based on IC₅₀ values, each peptide was classified as either a strong (IC₅₀ < 1 μ M), moderate (1 μ M < IC₅₀ < 10 μ M), weak (10 μ M < IC₅₀ < 100 μ M), or non-binder (IC₅₀ > 100 μ M). Using the 100- μ M cutoff to delineate binders and non-binders, the experimental data for alleles DRB1*0101, 0401, 0701, and 1501 were correlated with the ProPred predictions (Supplemental table 1). The false-positive rate (ProPred predicted binders that were observed as non-binders) was 9 %, and the false-negative rate (ProPred predicted non-binders that were observed as binders) was 31 %. While other studies suggest a more stringent cutoff for experimental MHC II binding [17], the 100- μ M threshold was chosen so as to maintain consistency with prior work on P99 β L [31]. In that study, similar false-

positive and false-negative rates were observed (6 and 33 %, respectively). Below is a detailed description of the peptide binding results.

Analysis of immunoreactivity—epitope cluster 1

The A13 epitope was predicted to bind all four MHC II alleles tested (Fig. 2), and experimentally it was found to bind all but DRB1*0101. Both mutations at residue A13 resulted in broad disruption of MHC II binding, but A13D proved more potent against DRB1*0401 and A13E more potent against 1501 (Fig. 3). This result is consistent with the IP² design output, as A13D is predicted to delete both 0401 epitopes whereas A13E is predicted to delete only 1 of the 2 (Fig. 2). With respect to allele 0701, both engineered peptides remained below the 100- μ M threshold that classified experimental binders, but consistent with predictions, each mutation reduced 0701 binding by nearly two orders of magnitude. As a result, both mutations converted a strong 0701 binder to a weak binder. It remains to be seen whether or not such an effect is sufficient to mitigate an immune response in vivo.

Analysis of immunoreactivity—epitope cluster 2

Among the four alleles tested, wild-type peptide K21 was predicted to bind only DRB1*0401 (Fig. 2). Consistent with prior analyses [31], however, it was found to bind all four alleles. The K21E mutation decreased binding threefold to fivefold for alleles 0101, 0401, and 1501. Thus, it transformed three moderate binders into weak binders (Fig. 3). As predicted, binding of the engineered peptide to allele 0701 remained unchanged relative to the wild-type sequence.

Analysis of immunoreactivity—epitope cluster 3

In the case of R105, binding was predicted for all four alleles, with multiple epitopes for 0101, 0401, and 1501 (Fig. 2). Experimentally, binding was observed with only alleles 0701 and 1501. The R105S mutation ablated binding to 0701, as predicted, and reduced binding to 1501 by more than fivefold (Fig. 3). The fact that the engineered peptide remained classified as a weak binder of 1501 is consistent with the prediction that it would disrupt only one of two 1501 epitopes in this peptide (Fig. 2). While the R105S mutation had no impact on 0401 binding, it increased affinity for the 0101 allele, thereby changing the wild-type non-binder to a weak binder of this MHC II protein.

Analysis of immunoreactivity—epitope cluster 4

R133 was predicted to bind only 1501, and indeed, it was found to be an especially high affinity binder of this allele. The wild-type peptide also bound 0701, albeit with low affinity. The R133H mutation was not predicted to influence binding among the four alleles tested here, but was instead designed to delete the untested DRB1*0801 and 1301 epitopes (Fig. 2). Unexpectedly, however, R133H was found to increase binding to alleles 0701 (twofold) and 0101 (sixfold). In the former instance, both the wild-type and engineered peptides were classified as weak binders, but in the latter instance, the wild-type non-binder was transformed into a weak binder (Fig. 3).

Analysis of immunoreactivity—epitope cluster 5

R210 was a predicted binder for 0701 and 1501, contributing to two distinct epitopes for the latter (Fig. 2). Experimentally, the wild-type peptide was found to bind 1501 weakly but showed no binding to 0701. The R210H mutation successfully decreased 1501 binding, reducing the IC_{50} to 91.5 μ M, a value just below the threshold for experimental binders (Fig. 3). With respect to the other three non-binding alleles, the R210H mutation did not influence MHC II interactions.

Analysis of immunoreactivity—epitope cluster 6

M235 was predicted bind 0101 and 0401, and experiments showed it to be a weak binder of both alleles, as well as weak binder of 0701. Consistent with predictions, the M235D mutation ablated binding to both 0101 and 0401. Additionally, it decreased binding to 0701, albeit by less than twofold (Fig. 3). The alternative M235Q mutation was less effective than M235D, yielding no change for 0101 and only marginally decreased binding to 0401 and 0701. These observations reflect the IP² design parameters to some extent, as M235D appears only in more aggressive plans with $\alpha < 0.9$, whereas M235Q appears exclusively in conservative plans with $\alpha > 0.9$ (Table 1).

Analysis of immunoreactivity—epitope cluster 7

The wild-type I262 peptide contained predicted epitopes for all four tested alleles (Fig. 2), and experiments confirmed it to be the most highly immunoreactive peptide tested. The I262T mutation ablated binding to 0101, as predicted, and it reduced the variant peptide's affinity for 0401, 0701, and 1501 by as much as 17-fold (Fig. 3).

Analysis of immunoreactivity—aggregate protein scores

To enable direct experimental comparisons between full-length proteins, a semiquantitative aggregate epitope score was calculated by summing the strong, moderate, and weak binding peptides derived from each design (see Supplemental Table 1 for annotation of individual peptide binding categories). For example, wild-type P99 β L peptides exhibited a total of 28 binding interactions: 15 weak (green bars), 6 moderate (blue bars), and seven strong (red bars) (Fig. 4). Six of the eight designs from the current study yielded a net reduction in total epitope counts, with the two most aggressive 5-mutation plans E and F exhibiting only 24 binding interactions. Importantly, E and F both yielded a net reduction of three strong binders and one moderate binder (weak binding counts remained at 15, equivalent to wild type). In contrast, the two most conservative designs D and H were found to have 29 binding interactions, one more than wild type. However, in both cases, the additional epitope was a weak binder, and interestingly, both mutants encoded the R210H mutation, which rendered DRB1*1501 a borderline weak binder of the cognate peptide (i.e., $IC_{50} = 91.5 \mu$ M, just below the cutoff for non-binding). Perhaps more importantly, variants D and H each possessed only 10 moderate to high binders, whereas the wild-type P99 β L and all three previously deimmunized variants exhibited 13 such epitopes. Thus, for D and H, three higher affinity binding interactions ($IC_{50} < 10 \mu$ M) were deleted at the expense of one added weak interaction. In a similar fashion, all eight of the IP² designs benefited from a net 3–4 count reduction in strong and/or moderate binders. Finally, it bears noting that, as separate

groups, the 4-mutation designs and the 5-mutation designs showed a strong positive correlation between α and the experimental epitope score, namely for a given mutational load, lower values of α yielded greater reductions in MHC II binding, as expected. This observation is in contrast to the lack of correlation with kinetic parameters, and it provides evidence of the algorithm's predictive power with respect to disruption of peptide-MHC II binding.

Discussion

The biotherapeutics literature boasts numerous successful T cell epitope deletion studies, but with very few exceptions, these programs have focused on only one or two immunogenic regions of a given protein target [7, 9, 10, 31, 36–38]. Indeed, it has been suggested that broad T cell epitope deletion may be an intractable problem [39]. Perhaps the most aggressive studies to date are the deimmunization of *E. coli* asparaginase [25] and the reengineering of Factor VIII peptide fragments [24]. The former project modified three separate immunogenic regions with a total of eight mutations. The deletion of these epitopes relied on iterative stochastic substitution of classical anchor residues, functional analysis of the resultant libraries using a somewhat sophisticated high throughput screen, and post hoc determination of which mutations might prove deimmunizing. Ultimately the project was quite successful in producing a deimmunized asparaginase, but the overall process yielded a low hit rate of both (i) functional variants from the library population and (ii) deimmunized variants within the functional subset. In the case of the factor VIII project, the authors sought to deimmunize six separate immunogenic regions; however, this work was pursued exclusively in the context of synthetic peptide fragments. No effort was made to incorporate the broad panel of deimmunizing mutations into a full-length protein, and it remains to be seen whether or not such a heavily mutated factor VIII enzyme would be stable and active. Given the necessity that deimmunized biotherapeutics maintain high-level function, the current literature indicates that deletion of numerous dispersed epitopes is a particularly challenging task. While there is speculation that highly aggressive T cell epitope deletion might be possible [40], to date, there has been no report of simultaneous disruption among four or more immunogenic regions in any full-length protein target.

The current study was designed to evaluate the utility of a new deimmunization algorithm using P99 β L as a model biotherapeutic. In one prior report, deimmunized P99 β L variants containing 1–2 mutations were designed with a less sophisticated algorithm, DP² [31]. Another study had previously generated a 2-mutation P99 β L variant using a conventional trial and error experimental approach [9]. In comparison, the P99 β L designs described here contained 4–5 mutations and targeted 2–5 times as many immunogenic regions. In an effort to accommodate these high mutational loads without compromising enzyme function, active site residues were held invariant and deimmunized designs were generated with the more advanced algorithm, IP², which optimizes for both sequence conservation and critical residue–residue interactions [28]. Compared to the earlier 2-mutation DP² designs, on average, the current IP² variants exhibited faster overall maximum reaction rates, higher overall catalytic efficiency, and equivalent average melting temperatures (Fig. 5). Thus, the current designs accommodated higher mutational loads while manifesting better catalytic proficiency and comparable structural integrity.

Functional deletion of broadly distributed P99 β L epitopes was the fundamental motivation behind the current work. As a metric of success, quantitative peptide-MHC II binding studies were performed on both wild-type and engineered P99 β L peptide fragments. Importantly, analogous experimental measures have been made previously for the DP² designed P99 β L variants and the conventionally deimmunized K21A, S324A double mutant [31]. These data enable direct comparisons with the IP² designs of the current study. The K21A, S324A double mutant was found to exhibit reduced immunogenicity in ex vivo cellular assays [9], but its cognate peptide fragments were, surprisingly, found to have somewhat increased affinity for MHC II alleles DRB1*0101, 0401, 0701, and 1501 (Fig. 4). In contrast, the earlier DP² variants each converted one or more high affinity epitopes into moderate affinity epitopes, and the globally optimal A13D, I104T design had a net deletion of one epitope (Fig. 4). Importantly, however, neither the earlier DP² designs nor the conventionally deimmunized variant yielded a net reduction in combined moderate and strong binding peptides (see horizontal blue hatched line, Fig. 4). This observation has direct implications for immunogenicity, as multiple lines of evidence indicate that the stability of peptide-MHC II complexes is a critical determinant of immunodominant T cell epitopes [41, 42]. It is encouraging, therefore, that all eight of the IP² designs yielded a net reduction of 3–4 moderate to high affinity peptide epitopes. In this context, one might reasonably speculate that even designs D and H would have reduced immunogenicity, as they each showed a net decrease of three moderate to strong binders at the cost of a total epitope count one greater than wild type. The more aggressively deimmunized plans E and F benefited from a net reduction of four epitopes, three strong binders, and one moderate binder. As evidenced by the results of the previous deimmunization studies, such broad reduction of P99 β L immunoreactivity is not possible at lower mutational loads.

Combined with the results of prior analyses [31], the data presented here highlight the utility of integrated protein deimmunization algorithms that simultaneously incorporate T cell epitope prediction and computational optimization of protein function. An important next step will be validating these engineered proteins' reduced immunogenicity with more advanced immunological techniques, such as ex vivo human immune cell assays or in vivo analysis in humanized mice. Likewise, demonstrating the algorithm's broad utility will require application to additional therapeutic protein targets. Pending completion of these important extensions of the current work, we anticipate that IP² and related design tools might significantly accelerate the development of safe and efficacious next-generation biotherapies.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by NIH grant R01-GM-098977 to CBK and KEG. RSS was supported in part by a Luce Foundation Fellowship and in part by a Thayer Innovation Program Fellowship from the Thayer School of Engineering. The authors would like to thank Thomas Scanlon, Warren Kett, and Deeptak Verma for their insights and support.

References

1. Baker MP, Reynolds HM, Lumicisi B, Bryson CJ. Immunogenicity of protein therapeutics: the key causes, consequences and challenges. *Self Nonself*. 2010; 1(4):314–322.10.4161/self.1.4.13904 [PubMed: 21487506]
2. Barbosa MD. Immunogenicity of biotherapeutics in the context of developing biosimilars and biobetters. *Drug Discov Today*. 2011; 16(7–8):345–353.10.1016/j.drudis.2011.01.011 [PubMed: 21300174]
3. De Groot AS, Scott DW. Immunogenicity of protein therapeutics. *Trends Immunol*. 2007; 28(11): 482–490.10.1016/j.it.2007.07.011 [PubMed: 17964218]
4. Schellekens H. Immunogenicity of protein therapeutics, or how to make antibodies without T-cells. *Infamm Res*. 2007; 56:S351–S352.
5. Schellekens H. Factors influencing the immunogenicity of therapeutic proteins. *Nephrol Dial Transplant*. 2005; 20(Suppl 6):vi3–vi9.10.1093/ndt/gfh1092 [PubMed: 15958824]
6. Trombetta ES, Mellman I. Cell biology of antigen processing in vitro and in vivo. *Annu Rev Immunol*. 2005; 23:975–1028.10.1146/annurev.immunol.22.012703.104538 [PubMed: 15771591]
7. Warmerdam PA, Plaisance S, Vanderlick K, Vandervoort P, Brepoels K, Collen D, De Maeyer M. Elimination of a human T-cell region in staphylokinase by T-cell screening and computer modeling. *Thromb Haemost*. 2002; 87(4):666–673. [PubMed: 12008950]
8. Jones TB, Hart RP, Popovich PG. Molecular control of physiological and pathological T-cell recruitment after mouse spinal cord injury. *J Neurosci*. 2005; 25(28):6576–6583.10.1523/JNEUROSCI.0305-05.2005 [PubMed: 16014718]
9. Harding FA, Liu AD, Stickler M, Razo OJ, Chin R, Faravashi N, Viola W, Graycar T, Yeung VP, Aehle W, Meijer D, Wong S, Rashid MH, Valdes AM, Schellenberger V. A beta-lactamase with reduced immunogenicity for the targeted delivery of chemotherapeutics using antibody-directed enzyme prodrug therapy. *Mol Cancer Ther*. 2005; 4(11):1791–1800.10.1158/1535-7163.MCT-05-0189 [PubMed: 16276001]
10. Cizeau J, Grenkow DM, Brown JG, Entwistle J, MacDonald GC. Engineering and biological characterization of VB6-845, an anti-EpCAM immunotoxin containing a T-cell epitope-depleted variant of the plant toxin bouganin. *J Immunother*. 2009; 32(6):574–584.10.1097/CJI.0b013e3181a6981c [PubMed: 19483652]
11. Singh H, Raghava GP. ProPred: prediction of HLA-DR binding sites. *Bioinformatics*. 2001; 17(12):1236–1237. [PubMed: 11751237]
12. Guan P, Doytchinova IA, Zygouri C, Flower DR. MHCpred: bringing a quantitative dimension to the online prediction of MHC binding. *Appl Bioinform*. 2003; 2(1):63–66.
13. Wan J, Liu W, Xu Q, Ren Y, Flower DR, Li T. SVRMHC prediction server for MHC-binding peptides. *BMC Bioinform*. 2006; 7:463.10.1186/1471-2105-7-463
14. Bui HH, Sidney J, Peters B, Sathiamurthy M, Sinichi A, Putton KA, Mothe BR, Chisari FV, Watkins DI, Sette A. Automated generation and evaluation of specific MHC binding predictive tools: ARB matrix applications. *Immunogenetics*. 2005; 57(5):304–314.10.1007/s00251-005-0798-y [PubMed: 15868141]
15. Nielsen M, Lundegaard C, Lund O. Prediction of MHC class II binding affinity using SMM-align, a novel stabilization matrix alignment method. *BMC Bioinform*. 2007; 8:238.10.1186/1471-2105-8-238
16. Nielsen M, Justesen S, Lund O, Lundegaard C, Buus S. NetMHCIIpan-2.0—Improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure. *Immunome Res*. 2010; 6:9.10.1186/1745-7580-6-9 [PubMed: 21073747]
17. Wang P, Sidney J, Dow C, Mothe B, Sette A, Peters B. A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach. *PLoS Comput Biol*. 2008; 4(4):e1000048.10.1371/journal.pcbi.1000048 [PubMed: 18389056]
18. De Groot AS, Knopp PM, Martin W. De-immunization of therapeutic proteins by T-cell epitope modification. *Dev Biol (Basel)*. 2005; 122:171–194. [PubMed: 16375261]

19. Perry LC, Jones TD, Baker MP. New approaches to prediction of immune responses to therapeutic proteins during preclinical development. *Drugs R D*. 2008; 9(6):385–396.10.2165/0126839-200809060-00004 [PubMed: 18989990]
20. De Groot AS, Martin W. Reducing risk, improving outcomes: bioengineering less immunogenic protein therapeutics. *Clin Immunol*. 2009; 131(2):189–201.10.1016/j.clim.2009.01.009 [PubMed: 19269256]
21. Kern F, LiPira G, Gratama JW, Manca F, Roederer M. Measuring Ag-specific immune responses: understanding immunopathogenesis and improving diagnostics in infectious disease, autoimmunity and cancer. *Trends Immunol*. 2005; 26(9):477–484.10.1016/j.it.2005.07.005 [PubMed: 16039158]
22. Li Pira G, Ivaldi F, Moretti P, Manca F. High throughput T epitope mapping and vaccine development. *J Biomed Biotechnol*. 2010; 2010:325720.10.1155/2010/325720 [PubMed: 20617148]
23. Salvat RS, Moise L, Bailey-Kellogg C, Griswold KE. A high throughput MHC II binding assay for quantitative analysis of peptide epitopes. *J Vis Exp* (in press). 201410.1093/protein/gzs044
24. Moise L, Song C, Martin WD, Tassone R, De Groot AS, Scott DW. Effect of HLA DR epitope deimmunization of Factor VIII in vitro and in vivo. *Clin Immunol*. 2012; 142(3):320–331.10.1016/j.clim.2011.11.010 [PubMed: 22222093]
25. Cantor JR, Yoo TH, Dixit A, Iverson BL, Forsthuber TG, Georgiou G. Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. *Proc Natl Acad Sci USA*. 2011; 108(4):1272–1277.10.1073/pnas.1014739108 [PubMed: 21209329]
26. Koren E, De Groot AS, Jawa V, Beck KD, Boone T, Rivera D, Li L, Mytych D, Koscec M, Weeraratne D, Swanson S, Martin W. Clinical validation of the “in silico” prediction of immunogenicity of a human recombinant therapeutic protein. *Clin Immunol*. 2007; 124(1):26–32.10.1016/j.clim.2007.03.544 [PubMed: 17490912]
27. Parker AS, Zheng W, Griswold KE, Bailey-Kellogg C. Optimization algorithms for functional deimmunization of therapeutic proteins. *BMC Bioinform*. 2010; 11:180.10.1186/1471-2105-11-180
28. Parker AS, Griswold KE, Bailey-Kellogg C. Optimization of therapeutic proteins to delete T-cell epitopes while maintaining beneficial residue interactions. *J Bioinform Comput Biol*. 2011; 9(2):207–229. [PubMed: 21523929]
29. Parker AS, Choi Y, Griswold KE, Bailey-Kellogg C. Structure-guided deimmunization of therapeutic proteins. *J Comput Biol*. 2013; 20(2):152–165.10.1089/cmb.2012.0251 [PubMed: 23384000]
30. Choi Y, Griswold Ke, Bailey-Kellogg C. Structure-based redesign of proteins for minimal T-cell epitope content. 1096-987× (electronic).
31. Osipovitch DC, Parker AS, Makokha CD, Desrosiers J, Kett WC, Moise L, Bailey-Kellogg C, Griswold KE. Design and analysis of immune-evading enzymes for ADePT therapy. *Protein Eng Des Sel*. 2012; 25(10):613–623.10.1093/protein/gzs044 [PubMed: 22898588]
32. Sandler I, Zigdon N, Levy E, Aharoni A. The functional importance of co-evolving residues in proteins. *Cell Mol Life Sci*. 2014; 71:673–682. [PubMed: 23995987]
33. McCaldon P, Argos P. Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide-sequences. *Proteins Struct Funct Genet*. 1988; 4(2):99–122.10.1002/prot.340040204 [PubMed: 3227018]
34. Southwood S, Sidney J, Kondo A, del Guercio MF, Appella E, Hoffman S, Kubo RT, Chesnut RW, Grey HM, Sette A. Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol*. 1998; 160(7):3363–3373. [PubMed: 9531296]
35. Stemmer WP, Cramer A, Ha KD, Brennan TM, Heyneker HL. Single-step assembly of a gene and entire plasmid from large numbers of oligodeoxyribonucleotides. *Gene*. 1995; 164(1):49–53. [PubMed: 7590320]
36. Mazor R, Vassall AN, Eberle JA, Beers R, Weldon JE, Venzon DJ, Tsang KY, Benhar I, Pastan I. Identification and elimination of an immunodominant T-cell epitope in recombinant immunotoxins based on *Pseudomonas* exotoxin A. *Proc Natl Acad Sci*. 2012; 109(51):E3597–E3603.10.1073/pnas.1218138109 [PubMed: 23213206]

37. Yeung VP, Chang J, Miller J, Barnett C, Stickler M, Harding FA. Elimination of an immunodominant CD4+ T cell epitope in human IFN- β does not result in an in vivo response directed at the subdominant epitope. *J Immunol.* 2004; 172(11):6658–6665. [PubMed: 15153481]
38. Tangri S, Mothe BR, Eisenbraun J, Sidney J, Southwood S, Briggs K, Zinckgraf J, Bilsel P, Newman M, Chesnut R, LiCalsi C, Sette A. Rationally engineered therapeutic proteins with reduced immunogenicity. *J Immunol.* 2005; 174(6):3187–3196. [PubMed: 15749848]
39. Onda M. Reducing the immunogenicity of protein therapeutics. *Curr Drug Targets.* 2009; 10(2): 131–139. [PubMed: 19199909]
40. Lee S. Implications of available design space for identification of non-immunogenic protein therapeutics. *Biomed Microdevices.* 2010; 12(2):283–286.10.1007/s10544-009-9383-8 [PubMed: 20012559]
41. Hill JA, Wang DQ, Jevnikar AM, Cairns E, Bell DA. The relationship between predicted peptide-MHC class II affinity and T-cell activation in a HLA-DR beta 1*0401 transgenic mouse model. *Arthrit Res Therapy.* 2003; 5(1):R40–R48.10.1186/ar605
42. Weaver JM, Sant AJ. Understanding the focused CD4 T cell response to antigen and pathogenic organisms. *Immunol Res.* 2009; 45(2–3):123–143.10.1007/s12026-009-8095-8 [PubMed: 19198764]
43. Schrödinger L. The PyMOL Molecular Graphics System (1.5.0.4).

Abbreviations

APCs	Antigen-presenting cells
MHC II	Class II major histocompatibility complex proteins
DP²	Dynamic programming for deimmunizing proteins
P99βL	<i>Enterobacter chloacae</i> P99 beta-lactamase
IP²	Integer programming for immunogenic proteins

Fig. 1.

Epitope Map of P99βL. The total number of predicted epitopes (y-axis) that start at a given amino acid is mapped onto the primary sequence of P99βL (x-axis). Using a Propred threshold of 5 %, epitopes were predicted for MHC II alleles DRB1*0101, 0301, 0401, 0701, 0801, 1101, 1301, and 1501 (i.e., max score for any position is 8). Sites of IP² mutations are indicated with arrows and residue numbers. Inset: Predicted epitopes mapped onto the P99βL peptide backbone (PDB 1XX2 chain A). The density of predicted epitopes is indicated by size and color gradients. *Thick tubes* indicate a high density of overlapping epitopes, while thin white tubes indicate no epitopes in that protein segment. Residues targeted with mutations are rendered as *pink ball* and *sticks*. Rendered with PyMOL [43]

Fig. 2.

Allele-specific epitope predictions for P99βL peptides. Names of synthetic peptides are shown above each plot, the corresponding amino acid sequences are indicated on the x-axis, and the relevant human MHC II alleles are shown on the y-axis. For the wild-type P99βL sequences, predicted epitopes for each allele are indicated as solid black lines spanning the 9mer peptide. Epitopes that remain after the specified mutation are shown as overlaid *hatched lines*

Fig. 3.

Peptide binding affinities for human MHC II proteins. IC_{50} values are plotted as cognate wild-type and variant pairs, where lower IC_{50} values correspond to higher affinity binding with human MHC II. The slope of the connecting lines is a relative measure of deimmunizing efficacy, where larger positive slopes indicate a greater fold decrease in affinity relative to wild type. Lines with negative slopes indicate a mutation that enhanced MHC II binding. Shading indicates binding strength by category (strong = dark gray; moderate = medium gray; weak = light gray; non-binding = white)

Fig. 4.

Aggregated immunoreactivity scores for full-length protein designs. The binding strength of individual peptides for MHC II alleles DRB1*0101, 0401, 0701, and 1501 were binned as strong ($IC_{50} < 1 \mu M$), moderate ($1 \mu M \leq IC_{50} < 10 \mu M$), weak ($10 \mu M \leq IC_{50} < 100 \mu M$), or non-binding ($IC_{50} \geq 100 \mu M$, not shown). The counts for each enzyme's constituent peptides were summed and plotted by semiquantitative category (y-axis). The horizontal hatched lines are visual guides for the strong and moderate counts of wild-type P99 β L. Asterisk indicates values from a previously published study [31]. Annotation of binding category for individual peptides is provided in Supplemental Table 1

Fig. 5.

Relative activity and stability of various deimmunized P99 β L variants. Relative scores were calculated by dividing a given variant's parameter by the corresponding wild-type value from the cited study. Hatched horizontal line indicates wild-type values. IP² variants are shown in blue, DP² variants in green, and the conventionally deimmunized variant in red. Error bars indicate SEM. The horizontal hatched line is a visual guide to the normalized wild-type value of each parameter. Asterisk indicates values from a previously published study [31]

Table 1

Thermostability and activity of deimmunized variants

Design	A13	K21	R105	R133	R210	M235	I262	No. Mut.	α	k_{cat} (s ⁻¹)	K_m (μ M)	k_{cat}/K_m (s ⁻¹ μ M ⁻¹)	T_m (°C)
Wild type								0	1.000	320 $\pm$ 10	43 $\pm$ 3	7.5 $\pm$ 0.6	56.37 $\pm$ 0.05
A	D	E	S				T	4	0.500	420 $\pm$ 20	43 $\pm$ 5	10 $\pm$ 1	50.41 $\pm$ 0.07
B	E	E	S				T	4	0.701	300 $\pm$ 40	40 $\pm$ 20	7 $\pm$ 3	50.32 $\pm$ 0.09
C		E	S		H	Q		4	0.963	440 $\pm$ 20	50 $\pm$ 5	8.8 $\pm$ 0.9	49.65 $\pm$ 0.07
D		E		H	H	Q		4	0.975	370 $\pm$ 10	42 $\pm$ 4	8.8 $\pm$ 0.9	52.81 $\pm$ 0.01
E	D	E	S			D	T	5	0.500	300 $\pm$ 10	33 $\pm$ 3	9.2 $\pm$ 0.9	50.1 $\pm$ 0.1
F	E	E	S			D	T	5	0.700	360 $\pm$ 10	39 $\pm$ 3	9.1 $\pm$ 0.8	50.16 $\pm$ 0.06
G	E	E	S		H	D		5	0.904	520 $\pm$ 20	51 $\pm$ 5	10 $\pm$ 1	49.53 $\pm$ 0.02
H		E	S	H	H	Q		5	0.972	420 $\pm$ 10	51 $\pm$ 4	8.3 $\pm$ 0.7	48.5 $\pm$ 0.1

Values provided as mean $\pm$ SEM