Removing T-cell epitopes with computational protein design

Chris King[a,1], Esteban N. Garza[b], Ronit Mazor[c], Jonathan L. Linehan[d], Ira Pastan[c], Marion Pepper[b], and David Baker[a]

[a]Institute for Protein Design, Department of Biochemistry and [b]Department of Immunology, University of Washington, Seattle, WA 98195; and [c]National Cancer Institute and [d]National Institute of Allergy and Infectious Disease, National Institutes of Health, Bethesda, MD 20892

Edited by Barry Honig, Howard Hughes Medical Institute, Columbia University, New York, NY, and approved April 24, 2014 (received for review December 4, 2013)

Immune responses can make protein therapeutics ineffective or even dangerous. We describe a general computational protein design method for reducing immunogenicity by eliminating known and predicted T-cell epitopes and maximizing the content of human peptide sequences without disrupting protein structure and function. We show that the method recapitulates previous experimental results on immunogenicity reduction, and we use it to disrupt T-cell epitopes in GFP and *Pseudomonas* exotoxin A without disrupting function.

deimmunization | machine learning | biotherapeutics | Rosetta | immunotoxin

Immunogenicity is a major problem in the development of protein therapeutics. Repeated administration of a protein therapeutic can lead to B-cell activation and production of antibodies, rendering the therapeutic clinically ineffective or cross-reacting with host proteins (1). Affinity maturation of antibody-producing memory B cells is initiated by T-cell recognition of peptide epitopes displayed on major histocompatibility complex class II (MHCII) proteins on the surface of mature antigen-presenting cells. Immunogenicity may be reduced by eliminating known T-cell epitopes from the protein sequence and/or increasing the prevalence of sequences already found in the host genome to which T cells would already be tolerant, an approach that has met with substantial clinical success in the humanization of recombinant antibodies (2). However, unlike antibodies, which have been extensively characterized, the mutational tolerance of most proteins is generally not known, and hence, the extension of this approach to proteins of arbitrary structure and function remains a major challenge. Deimmunization efforts have relied, for the most part, on experimental characterization of a large number of point mutants followed by a combination of individual mutations (3, 4).

To reduce or eliminate immunogenicity, it would be desirable to have a method that eliminates MHCII-binding epitopes and increases host sequence content without disrupting interactions essential for proper folding and function. The peptide-binding repertoire of many MHCII alleles has been extensively characterized (5), and a number of methods has been developed for predicting the affinity of novel peptides for a given MHCII (6). Coupling of epitope prediction methods with methods for predicting the structural and functional consequences of mutations offers the possibility of reducing the immunogenicity of a target protein without disrupting structure and function. Epitope prediction methods, homolog substitution matrices, and structural stability calculations have been combined to predict optimal epitope-eliminating mutations (7, 8). Epitope prediction methods have been integrated with structure-based protein design (9) by combining the 9mer epitope PROPRED matrices with protein design of all residues in a flexible backbone method that allows substantial redesign of protein cores. The combined method was able to eliminate epitope-like sequences while maintaining native-like values for a number of predicted protein stability metrics, but folding, function, and immunogenicity were not evaluated experimentally.

Here, we describe the integration of the Rosetta computational protein design method with experimental immunogenic epitope data, MHC epitope prediction tools, and host genomic data to enable the design of proteins with reduced immunogenicity while retaining function and stability. Our approach goes beyond PROPRED by implementing a more accurate machine learning-based epitope prediction method for 28 different H-2, HLA-DR, and HLA-DQ alleles, restricts design to select 15mer epitope regions, and uses a greedy stepwise protein design algorithm (10) to eliminate the most immunogenic epitopes with as few mutations as possible, avoiding disruptive core mutations likely to destabilize the protein. We compare the performance of our epitope predictor with PROPRED and another leading epitope prediction method for 13 different human and mouse MHC alleles, show the effectiveness and generality of the method with in silico tests on previously characterized deimmunized protein targets, and show experimentally for GFP in mice and *Pseudomonas* exotoxin A (PE38) in humans that the method eliminates T-cell epitopes without disrupting function.

Results

Overview of Computational Method. For a given target protein and set of host MHC alleles, potential T-cell epitopes are first identified using a support vector machine (SVM). These regions are then optimized to eliminate the T-cell epitopes while retaining structure and function using an extension of the Rosetta all-atom protein design methodology with modifications to both the energy function used in the design calculations and the optimization procedure.

The energy function used in the sequence optimization is supplemented with two terms that incorporate immunologically relevant data. The first term calculates predicted epitope content using SVMs trained with experimentally determined peptide–MHC binding data. Scores from SVMs for each MHC allele in each 15mer sequence frame are averaged and then summed over

Significance

Proteins represent the fastest-growing class of pharmaceuticals for a diverse range of clinical applications. Computational protein design has the potential to create a novel class of therapeutics with tunable biophysical properties. However, the immune system reacts to T-cell epitope sequences in non-human proteins, leading to neutralization and elimination by the immune system. Here, we combine machine learning with structure-based protein design to identify and redesign T-cell epitopes without disrupting function of the target protein. We test the method experimentally, removing T-cell epitopes from GFP and *Pseudomonas* exotoxin A while maintaining function.

Author contributions: C.K., E.N.G., R.M., I.P., M.P., and D.B. designed research; C.K., E.N.G., and R.M. performed research; C.K., E.N.G., R.M., and J.L.L. contributed new reagents/analytic tools; C.K., E.N.G., R.M., I.P., M.P., and D.B. analyzed data; and C.K. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

[1]To whom correspondence should be addressed. E-mail: chrisk1@uw.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1321126111/-/DCSupplemental.

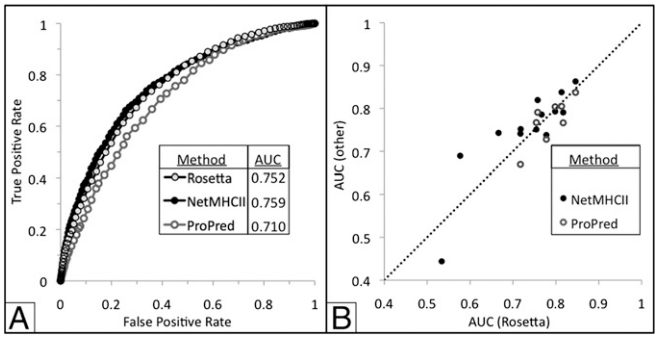

Fig. 1. Performance of Rosetta SVM T-cell epitope prediction. (*A*) ROC curve true-positive rate vs. false-positive rate for all testing data and comparison with current methods. Total AUC is listed for each method. (*B*) Predictive performance over each allele test set. *x* Axis, Rosetta AUC; *y* axis, other method AUC. Points below the 1:1 dotted line indicate where Rosetta performs better than other methods.

each frame. The second term uses known host genome 9mer data and known epitope data, rewarding each host 9mer in proportion to its frequency of occurrence in the host genome and penalizing known epitopes. Both deimmunization scores favor negatively charged residues on the surface of the protein; hence, we also introduce a net charge constraint into the total objective function, penalizing deviations from the input protein formal charge. By weighting this term appropriately against the deimmunization terms, negatively charged residues may be placed at critical epitope-disrupting surface positions while compensatory positively charged residues are introduced at other positions.

Sequence optimization is carried out using a protocol that focuses on solutions that reduce the energy with a relatively small number of mutations and allows use of computationally expensive objective functions in design calculations that otherwise might be unacceptably slow. The energies of all point mutants at each design position are first computed and then sorted. At each position, all point mutants within a certain threshold of the lowest energy mutant are saved for subsequent combination. These lowest energy point mutations are then combined using a greedy stepwise, steepest descent heuristic (*Methods*), allowing for structural relaxation at each step, until mutations have been attempted at all design positions. Multiple diverse near-optimal designs can be generated in parallel by stochastically accepting the placement of near-optimal mutations during the combination process.

T-Cell Epitope Prediction. We initially trained our SVMs on experimentally measured MHC binding affinities for 26 allelic variants (11). For each MHC allele type represented in the training set, we collected all epitope sequences from the Immune Epitope Database (IEDB) known to elicit T-cell activation (5). We restricted additional analysis to alleles with at least 25 known T-cell epitopes in the IEDB. In the absence of sufficient data on nonbinding peptides for each allele, we assumed that most 15mers in the host genome would not constitute strong MHC binders and generated negative datasets by randomly choosing 1,000 15mers from the human genome (12). We compared our SVM-based epitope predictions with the predictions of previous methods using a subset of T-cell epitopes withheld from initial training. Sensitivity and specificity were evaluated over all alleles for our method, NetMHCII-v2.2 (13), and PROPRED (14), and predictive performance was evaluated by calculating the area under the receiving operator characteristic (ROC) curve (AUC) for each allele type both independently and over the entire set. Because PROPRED contains only matrices for human leukocyte antigen DR beta (HLA-DRB) alleles, we combined testing data for the eight DRB alleles covered by all three methods and generated a standard ROC plot (Fig. 1*A*). The highest AUC was achieved by NetMHCII (AUC = 0.759). Rosetta performed comparably but slightly worse (AUC = 0.752), whereas PROPRED achieved significantly lower prediction accuracy (AUC = 0.710). Because more T-cell epitopes have been characterized for some alleles in the test set, combining all testing data weights performance analysis with alleles with more data points. If we average AUCs with equal weight over the shared DRB allele set, again, NetMHCII performs best (AUC = 0.792), with Rosetta slightly lower (AUC = 0.785) and PROPRED lower still (AUC = 0.771) (Fig. 1*B*).

Large-Scale Benchmarking and Calibration of Design Method. We first tested the ability of the method to eliminate putative human T-cell epitopes from eight proteins from pathogenic organisms that contain known MHC-binding epitopes. The computational design protocol was used to simultaneously eliminate all predicted epitopes for eight representative human DRB1 alleles (*SI Methods*), collectively covering almost 95% of the human population (15). To evaluate the tradeoff between epitope removal and protein stability, multiple design simulations were carried out, with an increasing weight on the SVM-based epitope scoring term. Increasing the weight on this term decreases the number of MHCII predicted epitopes and increases the Rosetta energy (Fig. 2*A*).

Because amino acid substitutions predicted to disrupt peptide–MHC binding might destabilize the overall protein, a balance between the stability of the protein and disruption of possible MHC binding must be sought. A weight on the epitope scoring term of 2.0 eliminates 79% of the predicted epitopes and 84% of the known epitopes without increasing Rosetta energy above that of the native protein (Table S1). Because sequence changes are permitted only at critical predicted epitope positions, the number of mutations is minimized, thus allowing for substantial reduction in predicted immunogenicity while maintaining average sequence identity at 66%. Similar calculations were carried out varying the weight of the term favoring 9mer sequences found in the human genome (Fig. 2*B*). As the weight increases, the average number of human genome 9mers over the designed regions increases, and the Rosetta energy becomes more unfavorable. At a weight of 3.5, human 9mer sequences increase from 0% to 4.3% of redesigned epitopes, whereas the average Rosetta energy increases only 10% over baseline (Table S2). These weights were used for the remainder of this work unless otherwise noted in *Methods*.

Recapitulation of Previous Immunogenicity Reduction Data. We next attempted to recapitulate the results of previous experimentally validated protein deimmunization efforts where immunogenicity was reduced without disrupting biological function. Cantor et al. (16) sought to remove T-cell epitopes from *Escherichia coli* L-asparaginase II, an enzyme approved for treatment of acute lymphoblastic leukemia, while maintaining native-like enzymatic activity and protein stability. Cantor et al. (16) first used a neutral drift selection scheme to identify allowable mutations in each of three predicted HLA-DRB1*04:01 epitopes and then combined

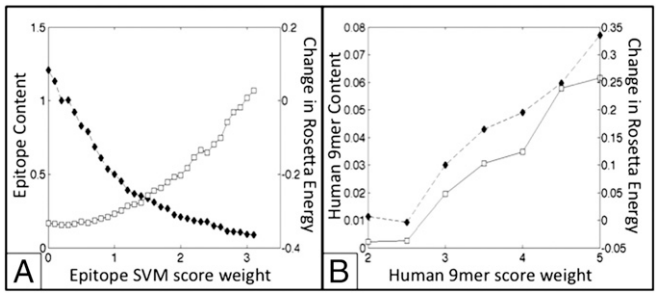

Fig. 2. Tradeoffs between Rosetta energy and extent of deimmunization. Rosetta energy (□) of redesigned proteins increases, whereas (*A*) epitope content decreases and (*B*) human 9mer count increases (♦) as the weights on the associated score terms are increased.

mutations in each epitope to produce a fully functional enzyme with reduced immunogenicity in vivo. We applied our protocol to the biologically active tetrameric state of *E. coli* L-asparaginase II, constraining design to the residue positions chosen for randomization in the previous study, and compared our predictions against the mutations identified in the selection screen and the predicted immunogenicity of the resultant design sequence. After Rosetta redesign for all 24 epitope residues, two of five mutations assumed residue identities found in the sequences of experimentally isolated activity-preserving variants, with one mutation occurring at a position not tested in the above study. In addition, two of three redesigned epitopes have predicted HLA-DRB1*04:01 binding affinity lower than the peptides tested in the previous study (Table 1).

Tangri et al. (4) attempted the deimmunization of erythropoietin (Epo), a growth hormone used in treatment of anemia and myelodysplasia. Tangri et al. (4) screened a set of Epo-derived peptides for binding to a set of 15 different HLA alleles, targeted two potential epitopes for deimmunization by screening a small set of point mutants in each epitope for reduction of T-cell activation, combined point mutants, and screened for both immunogenicity and biological activity. We applied our deimmunization design protocol to this protein using the structure of Epo bound to its native receptor and targeting the same epitope positions experimentally explored by Tangri et al. (4). Because the HLA allele set of the human donors in the study is unknown, we designed simultaneously against all eight alleles in the HLA-DRB1 allele set. In the first epitope region, Rosetta introduces only one substitution (the same mutation found to be optimal in ref. 4). Rosetta similarly recovers the point mutation position in the second epitope (Table 1) but substitutes an alanine to preserve the protein's net charge and makes three additional mutations to further disrupt MHC binding. In both cases, the Rosetta design minimizes the number of predicted MHC-binding peptides while minimizing the energy of the protein in the epitope regions.

Epitope Removal in Superfolder GFP. As an experimental proof of concept, we chose the fluorescent reporter protein superfolder GFP (sfGFP) (17). GFP is used to identify and track genetically modified stem cells in vivo for numerous applications (18) but concern about the immunogenicity of cells expressing GFP remains (19). We sought to redesign sfGFP to eliminate T-cell epitopes without disrupting fluorescence. To do so, we targeted the top four predicted H-2-IAb epitopes in the sfGFP sequence (Fig. 3*A*). None of these epitopes were present in our known epitope database, although epitope 84 had been previously identified as immunodominant in WT GFP. The design algorithm is nearly deterministic; to generate multiple candidates for testing, we stochastically sampled alternative sequences by random inclusion of locally near-optimal mutations at each design position. Eight designs were chosen for testing based on sequence diversity and predicted stability (Table 2 and *SI Methods, sfGFP Deimmunization Design Sequences Alignment*). sfGFP and all eight designs were expressed in *E. coli* and purified as soluble protein. To determine whether fluorescence was affected by the design mutations, emission and absorbance spectra were obtained for sfGFP and all eight deimmunized proteins. All eight designs showed fluorescence absorbance and emission spectra comparable with sfGFP, with fluorescence excitation peaks at 485 nm (Fig. 4*A* and Fig. S1).

We then investigated whether existing epitopes were correctly identified and removed and whether new epitopes were not introduced by the design mutations. sfGFP and the deimmunized variant (sfGFP.di.v3.2) were chosen for immunological testing. This variant was selected as the most aggressive design, because it had the highest Rosetta energy but still maintained function. GFP:I-Ab tetramer reagents were generated using both native and design peptide sequences for three of the predicted epitope regions. For both constructs, five mice were injected with the protein in complete Freund's adjuvant (CFA). After 6 d, spleens from all 10 mice were stained with a multicombinatorial panel of GFP:I-Ab tetramers corresponding to both the native and design sequences of all three predicted epitope regions (Table 2), and tetramer-positive cells were magnetically bead-enriched. For mice challenged with WT sfGFP, flow cytometry confirmed epitope 84 as the immunodominant epitope, with epitope 223/224 recognized by a smaller number of T cells (Fig. 4*B*). For mice challenged with the deimmunized protein, all three tetramers corresponding to the three redesigned epitope regions failed to isolate T cells above background levels, except for one mouse that responded weakly to

Table 1. Recapitulation of previous *E. coli* L-asparaginase II deimmunization efforts (16) against one HLA allele and Epo (4) against eight HLA alleles

	Sequence	Predicted IC_{50} (nM)	Rosetta energy
EcAII: Epitope 1 (115–123) (rank 6)			
Native	MRPSTSMSA	194.3	−12.0
Cantor et al. (16)	***V***RP***PTR***MS***P***	339.9	77.4
Rosetta	MRPQT***F***MSA	87.2	−9.4
EcAII: Epitope 2 (216–224) (rank 11)			
Native	IVYNYANAS	217.3	−15.4
Cantor et al. (16)	***V***VY***G***YANAS	195.8	−13.8
Rosetta	IVYNY***S***NA**M**	197.1	−11.2
EcAII: Epitope 3 (304–312) (rank 3)			
Native	VLLQLALTQ	135.3	−17.4
Cantor et al. (16)	VLL***T***LALT***N***	122.4	−11.3
Rosetta	VLLQLAL**W**Q	193.1	−13.4
Epo: Epitope 1 (101–115) (rank 7)			
Native	GLRSLTTLLRALGAQ	7.7	−20.0
Tangri et al. (4)	GLRSLT***D***LLRALGAQ	12.1	−20.3
Rosetta	GLRSLT***D***LLRALGAQ	12.1	−20.3
Epo: Epitope 2 (136–150) (rank 1)			
Native	DTFRKLFRVYSNFLR	5.0	−23.5
Tangri et al. (4)	DTFRKLFRVY***D***NFLR	24.0	−19.9
Rosetta	DTFRK**EFFDY***A***NFLR	70.2	−13.7

MHC IC_{50} values are predicted by Rosetta SVM, and Rosetta energies are the sum of total residue energies over the epitope region. IC_{50} values for Epo are listed as the lowest predicted across the allele set. Epitope ranks as a function of predicted immunogenicity are listed next to the residue ranges. Mutations are highlighted in bold, and known activity-preserving mutations are italicized. EcAII, *E. coli* L-asparaginase II.

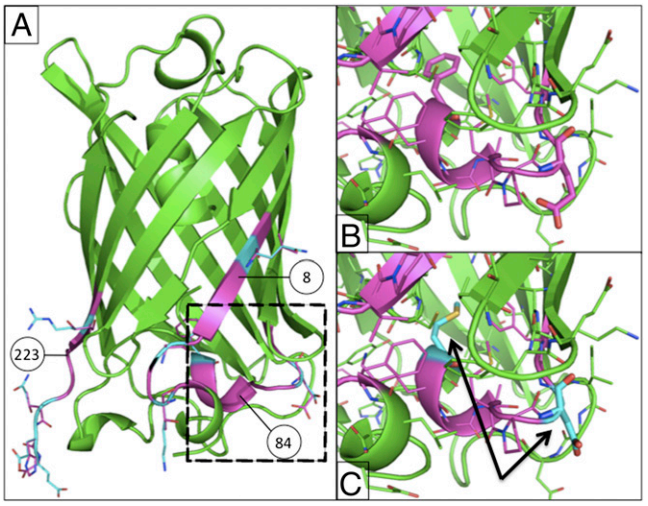

Fig. 3. Rosetta design model for sfGFP deimmunization. (*A*) Published coordinates of sfGFP crystal structure. Both known and predicted epitopes were targeted for design. Epitope indices from Table 2 are labeled in circles. (*B*) Close-up view of immunodominant epitope. (*C*) Rosetta deimmunization design of *B*. Cyan, design mutations; green, sfGFP; magenta, predicted epitopes.

the mutant epitope 223/224 (Fig. 4*C*). These two experiments confirmed that the design mutations effectively eliminated or greatly reduced the antigenicity of WT T-cell epitopes without creating new immunodominant epitopes or disrupting fluorescence.

Epitope Removal in Domain III of PE38. We next sought to remove T-cell epitopes from the toxin domain of the cancer therapeutic HA22, a recombinant immunotoxin containing a 38-kDa fragment of PE38 (20), while maintaining cytotoxic activity. HA22 has been used successfully to treat refractory hairy cell leukemia in a recent phase 1 clinical trial (21), and it has produced complete remission in several patients with acute lymphoblastic leukemia (22). However, HA22 has shown limited effectiveness in treating patients who are not immune-compromised, because the presence of the bacterial PE38 moiety leads to host immune response and the production of neutralizing antidrug antibodies (23). To address this issue, we targeted three epitope regions in PE38 previously identified in humans (24) and designed five mutations predicted to eliminate binding to a diverse set of 14 human HLA alleles while maintaining favorable interactions with the toxin substrate, eukaryotic elongation factor 2 (Fig. S2). The five mutants were expressed and purified for subsequent testing of cytotoxicity in two Burkitt lymphoma cell lines. Two mutants in the region 466–480, A476D and A476D+D474Y, had cytotoxic activity reduced by ~80%, but three mutants in the region 547–564 displayed equal or greater cytotoxicity than the WT toxin (Fig. 5*A*). The two most active mutants (L552N and L552E) were chosen for additional characterization of immunogenicity. Peripheral blood mononuclear cells (PBMCs) derived from two patients and one naïve donor were stimulated with mutant antigen, and IL-2 response was measured after restimulation with the WT and mutant epitope peptides. Both mutants caused a significant reduction in T-cell response for both epitope peptides in all three samples ($P > 0.01$ in Student t test) (Fig. 5*B*).

Discussion

We have developed a computational protein design method that incorporates host genome information and MHC-binding prediction tools to reduce the immunogenicity of arbitrary protein targets. The method removes MHC epitopes and increases human sequence content while maintaining protein stability and interactions with binding partners. Mutations predicted by the method partially recapitulate the mutations of previous successful deimmunization efforts. The effectiveness of the method was verified experimentally by successfully predicting and eliminating an immunodominant T-cell epitope from sfGFP and eliminating a known T-cell epitope from PE38. We redesigned these proteins to mitigate T-cell immune responses in both mice and humans, respectively, showing that the method presented is generally applicable to humans or other species for which sufficient MHC data exist.

Because of the possibility of disrupting folding and function of the target protein, most deimmunization efforts rely on experimental testing of a limited number of point mutants followed by conservative attempts to combine a small subset of these mutations into a functional product. Although varied in their approaches, proprietary methods used by companies, such as EpiVax (25) and Antitope (26), typically combine matrix-based epitope scoring, experimental characterization of MHC binding or T-cell epitope mapping, stepwise mutation of antigenic amino acids, and introduction of tolerizing epitopes where possible. Using structure-based design simulations, we introduced nine mutations into sfGFP simultaneously without disrupting function. One redesigned epitope is largely buried in the core of the protein (Fig. 3 *B* and *C*), requiring the selection of mutations that simultaneously eliminates MHC-binding propensity without disrupting the folding free energy of the protein. Eliminating such epitopes would be difficult using the above proprietary approaches, because multiple substitutions may be required to compensate for packing defects caused by mutations in the epitope region, and nonconservative mutations near the protein-binding interface of exotoxin A would not have been predicted without a structure-based design calculation.

This work has focused only on eliminating the most immunoreactive epitopes for a given set of MHC alleles. Immunological testing of a larger number of human patients would be required to fully cover the breadth of HLA allotype diversity, and there may exist highly conserved T-cell reactive amino acids in protein sequences for which no immune silencing mutations are possible, necessitating the prediction and removal of discontinuous B-cell epitopes from the protein surface. Methods exist for prediction of B-cell epitopes but suffer from lower predictive power because of the difficulty of obtaining sufficient structural data for the entire repertoire of discontinuous 3D epitopes (27). Nevertheless, B-cell epitope removal methods have proven successful for a number of clinical targets (28, 29). Such methods could be incorporated into a comprehensive deimmunization pipeline, and additional testing of systemic antibody response would be required to show complete immune evasion.

Table 2. sfGFP epitopes targeted for redesign

	sfGFP			sfGFP.di.v3.2		
Index	Native sequence	Predicted IC_{50} (nM)	Rosetta energy	Design sequence	Predicted IC_{50} (nM)	Rosetta energy
8	FTGVVPILV	548	−12.9	**FKGR**VPI**Q**V	998	−10.6
84	FKSAMPEGY	784	−8.4	**M**KSAMP**D**GY	4,465	−8.9
223	FVTAAGITH	542	−8.4	FV**R**AAGI**QE**	3,312	−7.9
224	VTAAGITHG	954	−7.15	V**R**AAGI**QEE**	2,354	−6.5

Mutations are highlighted in bold.

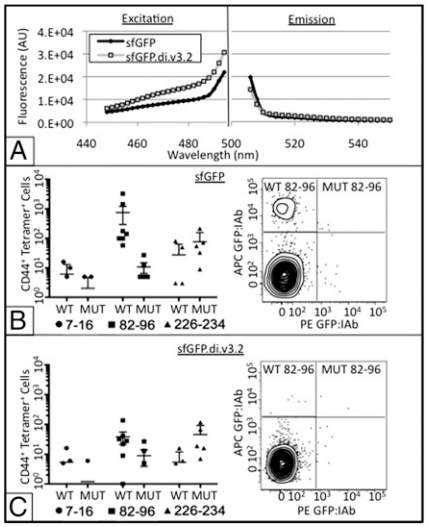

Fig. 4. Redesign of sfGFP reduces T-cell reactivity without disrupting fluorescence. (*A*) Deimmunized sfGFP excitation and emission spectra in arbitrary units (AU). Excitation spectrum measured at 510-nm emission. Emission spectra measured at 488-nm excitation. (*B*) Flow cytometry analysis of tetramer-enriched populations of $CD3^+$ $CD4^+$ $CD44^+$ GFP:I-Ab^+ lymphocytes labeled with phycoerythrin (PE) and/or A-phycocyanin (APC). Total $CD44^+$ $CD4^+$ tetramer-positive cells for each of six GFP:I-Ab tetramers in mice immunized with WT sfGFP. Immunization with the native sfGFP leads to the expansion and activation of $CD4^+$ T cells responding to epitopes 82–96. (*C*) Total $CD44^+$ $CD4^+$ tetramer-positive cells for each of six GFP:I-Ab tetramers in mice immunized with the designed sfGFP 3.2. Mice immunized with the designed sfGFP 3.2 no longer respond to the native sfGFP 82–96 epitopes or the designed epitopes 82–96 in sfGFP 3.2. MUT, mutant.

Here, we have focused on application to protein therapeutics delivered extracellularly, where MHCII mediates the primary immunological pathway. The method is readily extensible to both MHC class I- and MHCII-based deimmunization, and it is, thus, applicable to therapeutics expressed intracellularly, such as gene therapy products. Because epitope scores are averaged over all MHC allele SVMs, degenerate binding epitopes are penalized more strongly and thus, are the first to be targeted in our greedy design algorithm. This scoring scheme is critical, because mounting evidence points to the correlation between the number of host MHC alleles that a given epitope binds and its propensity to initiate an immune reaction in vivo (4, 30). Our epitope prediction method was trained on large-scale peptide-binding affinity data, although predictions can be made using other sources. When experimental binding constants are not available, MHC–peptide structure simulations have shown promise in calculating accurate sequence specificities based on MHC–peptide energetics (31). As improvements in energy functions lead to improvement in the prediction of the effects of mutations on stability and function and high-throughput experimental MHC–peptide-binding data become increasingly available, computational protein design will play an increasingly prominent role in development of next generation protein therapeutics.

Methods

Penalizing Predicted Epitopes. ***Epitope SVM construction.*** A detailed description of SVM construction is provided in *SI Methods*. Briefly, one SVM model was trained for each MHC using publicly available peptide–MHC binding constants for 29 human and mouse MHC alleles; 15mer peptide sequences were first aligned and then encoded into numerical feature vectors as described in *SI Methods*. After encoded, SVM regression models were trained using libSVM to recapitulate peptide–MHC binding IC_{50} values by minimizing mean-squared error in fivefold cross-validation tests.

Epitope SVM scoring. The total epitope prediction score of a protein during Rosetta design is calculated by sliding a 15-residue window across the protein sequence, calculating the average SVM binding score over all allele SVM models in the user-defined allele list, and summing the contributions from each overlapping sequence frame.

Rewarding 9mer Sequences That Occur in the Host Genome and Penalizing Known Epitopes. ***Host 9mer database construction.*** Protein translations of Ensembl gene predictions for *Homo sapiens* and *Mus musculus* genomes were downloaded from ftp://ftp.ensembl.org/pub/current_fasta/ on February 28, 2012. The number of occurrences of every unique contiguous 9mer peptide found in all hypothetical translation products was first calculated. Each 9mer was assigned a score that rewards common sequences; 9mers that occur between 1 and 10 times were assigned a score of $-\log(n) - 1$, where n is the total count of the 9mer, and 9mers that occur more than 10 times were given a constant score of −2.0 to prevent domination of scoring by widespread repeat sequences. Thus, each unique sequence was given a score in the range [−2, −1].

Known epitope database construction. All epitope sequences known to elicit T-cell activation through the MHCII pathway in either humans or mice were downloaded from the IEDB on February 28, 2012; 9mer core sequences were predicted as the epitope subsequence with the highest predicted MHC binding affinity as scored by Rosetta SVMs. All 9mer epitope sequences were assigned a constant score of 10.0.

9mer Database scoring. 9mer Subsequences with associated scores (genomic 9mers and known epitope sequences) are loaded from a user-supplied table at runtime and stored in a hash table for quick look-up during design. The total score is the sum of the scores for each 9mer subsequence.

Rosetta Design Calculations. ***Structure preprocessing and simulation parameters.*** Before design calculations, all protein structures were subject to multiple cycles of backbone minimization and rotamer optimization with position restraints on side-chain heavy atoms, allowing for small structural changes to bring the structure to the local energy function minimum. For all design calculations, talaris2013 Rosetta score weights were used; native residues were given a constant energy bonus of −0.2 Rosetta energy units (REU), nonnative residues were given a penalty of 0.2 REU, and nonnative cysteine and histidine residues were disallowed in design.

Rosetta greedy optimization design. All design simulations were implemented using Rosetta Scripts (32). Greedy sequence design and rotamer optimization were carried out using the Rosetta greedy descent optimization algorithm as previously described (11). Details are provided in *SI Methods*. Briefly, every amino acid point mutant is sampled independently, and the total energy is stored. Optimal mutations are sorted by energy before combinatorial optimization is attempted in a rank-ordered, steepest descent fashion. Multiple diverse solutions can be generated by stochastically attempting combination of near-optimal substitutions at each position.

Deimmunization design simulations. Details of each design calculation are provided in *SI Methods*. Briefly, all simulations followed a similar workflow as follows. Crystal structures were downloaded from the Protein Data Bank

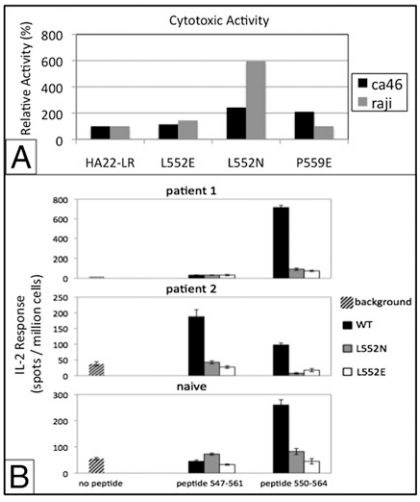

Fig. 5. Redesign of exotoxin A reduces T-cell reactivity without loss of function. (*A*) Relative cytotoxicity for the original HA22-LR toxin and three computationally designed variants in two cell types. (*B*) ELISpot IL-2 response was measured for PBMCs derived from two patients and one naïve donor after restimulation with two WT peptides and four mutant peptides.

and preminimized with Rosetta as described above. Designable residues were selected on the basis of average predicted MHC binding affinity; all residues in any epitope frame that score above a certain threshold are selected for design. For comparison with previous experimental works, residues were selected so as to provide a meaningful comparison with the published data. Deimmunization design was then carried out as described above. For design targets undergoing experimental characterization, multiple design sequences were generated by randomly sampling near-optimal mutations at each design position and combining these mutations in a stochastic manner.

sfGFP Activity and Immunogenicity Assays. ***Identification of eGFP$_{82–96}$:I-A^b epitope.*** To begin to narrow down a region in the EGFP protein containing a CD4$^+$ T-cell epitope in C57BL/6 mice, the EGFP nucleotide sequence (Clontech) was parsed into five equally sized fragments, and each fragment was inserted into the TOPO cloning site of the pTrcHis2-TOPO vector containing an isopropyl beta-D-1 thiogalactopyranoside-inducible promoter and a 6× His C-terminal epitope tag (Invitrogen). EGFP protein fragments were expressed in Rosetta 2 competent *E. coli* (EMD Millipore). Bacteria were lysed with BugBuster protein extraction reagent (Novagen) and sonicated, and EGFP fragments were purified with His-Bind resin columns (Novagen). Next, individual mice were immunized s.c. with 25 μg purified whole EGFP protein (Biovision, Inc.) in CFA (Sigma). After 10 d, draining lymph node cells were negatively selected for CD4$^+$ T cells (Miltenyi Biotech), and an anti–IFN-γ ELISPOT assay was done by interrogating with the above purified EGFP protein fragments (antibodies were from eBioscience, and 96-well Multiscreen filter plates were from Millipore). On identifying the EGFP protein fragment that gave the best IFN-γ–producing CD4$^+$ T-cell response, a 15mer overlapping peptide library (each offset by 2 aa) was constructed (Mimotopes). This library was then used for interrogation in another ELISPOT assay, as described above, to narrow down the 15mer epitope to amino acids 82–96.

Immunizations. C57BL/6 mice were injected s.c. at the base of the tail with either 50 μg sfGFP or 50 μg sfGFP.di.v3.2 emulsified in 50 μL CFA (Sigma-Aldrich). C57BL/6 mice were injected i.p. with 100 μg sfGFP or 100 μg sfGFP.di.v3.2 in aluminum hydroxide adjuvant (Brenntag).

Tetramer production. Biotin-labeled soluble I-A^b molecules containing EGFP peptide (FKSAMPEGY) covalently attached to the I-A^b β-chain were produced in *Drosophila melanogaster* S2 cells and then purified and made into tetramers with streptavidin-phycoerythrin or streptavidin-allophycocyanin (Prozyme) as described (33).

Cell enrichment and flow cytometry. All antibodies were from eBioscience unless noted. Single-cell suspensions of spleens and lymph nodes were stained for 1 h at room temperature with eGFP:I-A^b allophycocyanin tetramer. Samples were then enriched for bead-bound cells on magnetized columns, and a portion was removed for counting as described (33). For identification of surface phenotype, the rest of the sample underwent surface staining on ice with a mixture of antibodies specific for B220 (RA3-6B2), CD11b (MI-70), CD11c (N418), CD44 (IM7; BD), CD4 (RM4-5; BD), CD3 (145-2C11), and CD8 (5H10; BioLegend), each conjugated with a different fluorochrome. Cells were then analyzed on a Canto (BD) flow cytometer. Data were analyzed with FlowJo software (TreeStar).

Fluorescence spectra. sfGFP samples were diluted to a uniform concentration of 9.8 μM in PBS, and fluorescence spectra were measured on a SpectraMax plate reader. Excitation was measured from 448 to 500 nm (2-nm intervals) at 510-nm emission. Emission was measured from 498 to 550 nm (2-nm intervals) at 488-nm excitation. Fluorescence spectra were normalized by subtracting the signal obtained from pure PBS buffer.

Exotoxin A activity and immunogenicity assays. ***Construction, expression, purification, and cytotoxic activity of recombinant immunotoxin.*** The mutations L552E, L552N, and P559E were introduced into a plasmid expressing HA22 VH-PE38 using PCR overlap extension. The mutant RIT were purified as previously described (34). Cytotoxicity assays were performed on CD22$^+$ human Burkitt lymphoma cell lines (CA46 and Raji). The assay was performed as previously described (24).

In vitro expansion and ELISpot. PBMCs from two patients who were previously treated with PE38 RIT and one naïve donor were obtained after informed consent. The PBMCs were stimulated with parent RIT, HA22-L552E, or HA22-L552N and cultured in 37 °C with 5% (vol/vol) CO_2 for 14 d. IL-2 (10 U/mL; Millipore) was added every 3 d. On day 14, cells were harvested and restimulated with either WT peptides 93 and 94 (GPEEEGGRLETILGW and EEGGRLETILGWPLA) or mutant peptides (GPEEEGGR**E**ETILGW, EEGGR**E**ETILGWPLA, GPEEEGGR**N**ETILGW, and EEGGR**N**ETILGWPLA). IL-2 ELISpot was used to detect T-cell activation according to the manufacturer's instructions.

Rosetta command line demonstration. Command line examples, input files, and instructions for running all protein design simulations are included in a freely available archived demonstration at https://zenodo.org/record/8436.

ACKNOWLEDGMENTS. This research was supported by the Defense Threat Reduction Agency and the Intramural Research Program of the National Institutes of Health, the National Cancer Institute, Center for Cancer Research.

1. Zubler RH (2001) *Naive and Memory B Cells in T-Cell-Dependent and T-Independent Responses. Springer Seminars in Immunopathology* (Springer, Berlin), pp 405–419.
2. Goldenberg MM (1999) Trastuzumab, a recombinant DNA-derived humanized monoclonal antibody, a novel agent for the treatment of metastatic breast cancer. *Clin Ther* 21(2):309–318.
3. Harding FA, et al. (2005) A beta-lactamase with reduced immunogenicity for the targeted delivery of chemotherapeutics using antibody-directed enzyme prodrug therapy. *Mol Cancer Ther* 4(11):1791–1800.
4. Tangri S, et al. (2005) Rationally engineered therapeutic proteins with reduced immunogenicity. *J Immunol* 174(6):3187–3196.
5. Vita R, et al. (2010) The immune epitope database 2.0. *Nucleic Acids Res* 38(Database issue):D854–D862.
6. Nielsen M, Lund O, Buus S, Lundegaard C (2010) MHC class II epitope predictive algorithms. *Immunology* 130(3):319–328.
7. Parker AS, Griswold KE, Bailey-Kellogg C (2011) Optimization of therapeutic proteins to delete T-cell epitopes while maintaining beneficial residue interactions. *J Bioinform Comput Biol* 9(2):207–229.
8. Parker AS, Zheng W, Griswold KE, Bailey-Kellogg C (2010) Optimization algorithms for functional deimmunization of therapeutic proteins. *BMC Bioinformatics* 11(1):180.
9. Choi Y, Griswold KE, Bailey-Kellogg C (2013) Structure-based redesign of proteins for minimal T-cell epitope content. *J Comput Chem* 34(10):879–891.
10. Nivón LG, Bjelic S, King C, Baker D (2014) Automating human intuition for protein design. *Proteins* 82(5):858–866.
11. Wang P, et al. (2008) Immune epitope database: MHC-II binding dataset. Available at http://tools.iedb.org/mhcii/download/. Accessed February 14, 2013.
12. Flicek P, et al. (2012) Ensembl 2012. *Nucleic Acids Res* 40(Database issue):D84–D90.
13. Nielsen M, Justesen S, Lund O, Lundegaard C, Buus S (2010) NetMHCIIpan-2.0—Improved pan-specific HLA-DR predictions using a novel concurrent alignment and weight optimization training procedure. *Immunome Res* 6(1):9.
14. Singh H, Raghava GPS (2001) ProPred: Prediction of HLA-DR binding sites. *Bioinformatics* 17(12):1236–1237.
15. Southwood S, et al. (1998) Several common HLA-DR types share largely overlapping peptide binding repertoires. *J Immunol* 160(7):3363–3373.
16. Cantor JR, et al. (2011) Therapeutic enzyme deimmunization by combinatorial T-cell epitope removal using neutral drift. *Proc Natl Acad Sci USA* 108(4):1272–1277.
17. Pedelacq JD, Cabantous S, Tran T, Terwilliger TC, Waldo GS (2006) Engineering and characterization of a superfolder green fluorescent protein. *Nat Biotechnol* 24(1):79–88.
18. Brazelton TR, Blau HM (2005) Optimizing techniques for tracking transplanted stem cells in vivo. *Stem Cells* 23(9):1251–1265.
19. Persons DA, et al. (1998) Use of the green fluorescent protein as a marker to identify and track genetically modified hematopoietic cells. *Nat Med* 4(10):1201–1205.
20. FitzGerald DJ, Wayne AS, Kreitman RJ, Pastan I (2011) Treatment of hematologic malignancies with immunotoxins and antibody-drug conjugates. *Cancer Res* 71(20):6300–6309.
21. Kreitman RJ, et al. (2012) Phase I trial of anti-CD22 recombinant immunotoxin moxetumomab pasudotox (CAT-8015 or HA22) in patients with hairy cell leukemia. *J Clin Oncol* 30(15):1822–1828.
22. Wayne AS, et al. (2011) A novel anti-CD22 immunotoxin, moxetumomab pasudotox: Phase I study in pediatric acute lymphoblastic leukemia (ALL). *ASH Annual Meeting Abstracts* 118(21):248.
23. Hassan R, et al. (2007) Phase I study of SS1P, a recombinant anti-mesothelin immunotoxin given as a bolus I.V. infusion to patients with mesothelin-expressing mesothelioma, ovarian, and pancreatic cancers. *Clin Cancer Res* 13(17):5144–5149.
24. Mazor R, et al. (2012) Identification and elimination of an immunodominant T-cell epitope in recombinant immunotoxins based on Pseudomonas exotoxin A. *Proc Natl Acad Sci USA* 109(51):E3597–E3603.
25. De Groot AS, Terry F, Cousens L, Martin W (2013) Beyond humanization and deimmunization: Tolerization as a method for reducing the immunogenicity of biologics. *Expert Rev Clin Pharmacol* 6(6):651–662.
26. Baker M, Carr F (2010) Pre-clinical considerations in the assessment of immunogenicity for protein therapeutics. *Curr Drug Saf* 5(4):308–313.
27. Greenbaum JA, et al. (2007) Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. *J Mol Recognit* 20(2):75–82.
28. Nagata S, Pastan I (2009) Removal of B cell epitopes as a practical approach for reducing the immunogenicity of foreign protein-based therapeutics. *Adv Drug Deliv Rev* 61(11):977–985.
29. Onda M, et al. (2008) An immunotoxin with greatly reduced immunogenicity by identification and removal of B cell epitopes. *Proc Natl Acad Sci USA* 105(32):11311–11316.
30. Hammer J, et al. (1993) Promiscuous and allele-specific anchors in HLA-DR-binding peptides. *Cell* 74(1):197–203.
31. Yanover C, Bradley P (2011) Large-scale characterization of peptide-MHC binding landscapes with structural simulations. *Proc Natl Acad Sci USA* 108(17):6981–6986.
32. Fleishman SJ, et al. (2011) RosettaScripts: A scripting language interface to the Rosetta macromolecular modeling suite. *PLoS ONE* 6(6):e20161.
33. Moon JJ, et al. (2007) Naive CD4$^+$ T cell frequency varies for different epitopes and predicts repertoire diversity and response magnitude. *Immunity* 27(2):203–213.
34. Pastan I, Beers R, Bera TK (2004) *Recombinant Immunotoxins in the Treatment of Cancer. Antibody Engineering* (Springer, Berlin), pp 503–518.