

An *In Vitro* Human Cell-Based Assay to Rank the Relative Immunogenicity of Proteins

Marcia Stickler,* Narapon Rochanayon,* O. Jennifer Razo,* Jeanette Mucha,* Wendy Gebel,* Nargol Faravashi,* Regina Chin,* Susan Holmes,† and Fiona A. Harding,*¹

*Genencor International, Palo Alto, California 94304, and †Department of Statistics, Stanford University, Stanford, California 94305

Received on July 3, 2003; accepted on October 13, 2003

A method to rank proteins based on their relative immunogenicity has been devised. A statistical analysis of peptide-specific responses in large human donor pools provides a structure index value metric that ranked four industrial enzymes in the order determined by both mouse and guinea pig exposure models. The ranking method also compared favorably with human sensitization rates measured in occupationally exposed workers. Structure index values for other proteins known to cause immune responses in humans were also determined and found to be higher than the value determined for human β 2-microglobulin. Using values from known immunogenic and putative nonimmunogenic proteins, a cut-off value was established. The structure index value calculation provides a comparative method to predict subsequent immunogenicity on a human population basis without the need to use animal models. Information provided by this assay can be used in the early development of protein therapies and other protein-based applications to select or create reduced immunogenicity variants.

Key Words: immunogenicity; cellular activation; human; antigen/peptide/epitopes.

Proteins have the capacity to induce potentially life-threatening immune responses. This limitation has hindered their widespread use in consumer end-use applications. Recently, numerous highly publicized adverse events have come to the attention of the FDA, resulting in the requirement for immunogenicity testing both prior to and after approval of new protein therapeutics (Casadevall *et al.*, 2002; Chamberlain and Mire-Sluis, 2003; Kore *et al.*, 2002; Li *et al.*, 2001; Nordlee *et al.*, 1996). The potential immunogenicity of transgenic foods has been a concern for some time, and this field has struggled to develop predictive tools for food allergens (Hileman *et al.*,

2002; Oehlschlager *et al.*, 2001). In spite of the effort, there is no current standard test for potential immunogenicity.

Predictive methods for determining the immunogenicity of potential commercial proteins are hampered by the complex nature of the human population. Responsiveness of an individual to a particular protein is controlled by many interacting parameters, including the stability (Hall *et al.*, 2002) and proteolytic activity (Gough *et al.*, 2001) among other properties (Hall *et al.*, 2002), and additional extrinsic factors such as the route and dose of exposure (Braun *et al.*, 1997; Ge *et al.*, 2001), the presence of adjuvants including the presence or absence of endotoxin contamination (Alving, 2002; Brimnes *et al.*, 2003), underlying immune responsiveness (for example, tolerance to prevalent self proteins or pre-existing immunity), the presence of protein aggregates (Braun *et al.*, 1997), and the HLA molecules present. All of these parameters on an individual basis contribute to the population-based response. To analyze the collective response of a representative sample of human donors, a large enough sample set must be tested to ensure the appropriate genetic mix, and both recall and primary responses must be tabulated.

In addition to the ongoing research in the area of food allergens, the immunogenicity of proteins has long been a concern in the enzyme manufacturing industry (Bernstein *et al.*, 1999; Johnsen *et al.*, 1997; Kimber *et al.*, 1996; Pepys *et al.*, 1985; Vanhanen *et al.*, 1997, 2000, 2001). Occupational exposure to proteins has been documented to result in immune responses in industrial and laboratory workers. Conversion to skin prick test positivity (SPT+) can be controlled by reduction of the level of airborne protein (Sarlo and Kirchner, 2002; Schweigert *et al.*, 2000). When a new protein is to be manufactured, an occupational exposure guideline must be established. A commonly accepted method to determine these guidelines is the guinea pig intratracheal test (GPIT) (Sarlo *et al.*, 1997). The GPIT test, while useful, is time consuming and expensive. Recently, a mouse-based test (mouse intranasal test, MINT) has been established that reproduces the results seen in the GPIT (Robinson *et al.*, 1998).

Animal models have their limitations (Bussiere, 2003). The use of partially outbred guinea pigs in the GPIT necessitates

All authors except S.H. are employed by Genencor International. Genencor International has applied for a patent on the disclosed method.

¹ To whom correspondence should be addressed at Genencor International, 925 Page Mill Road, Palo Alto, CA 94304. Fax: (650) 845-6509. E-mail: fharding@genencor.com.

Toxicological Sciences vol. 77 no. 2 © Society of Toxicology 2004; all rights reserved.

the use of large numbers of animals to achieve statistical significance when comparing responses between groups. Inter-experiment variation in control animal responses is very high, which makes potency determinations based on a single set of control responses less convincing. The MINT assay does not suffer from as much variability in antibody responses; the mice used are typically BDF1 mice, a cross between two highly inbred mouse strains. While this additional level of control allows for more robust data analyses, different strains of mice will return very different potency rankings for similar enzymes (Blaikie and Basketter, 1999a,b). This is likely due to the specificity of the immune response in a mouse line that has been inbred to express very limited MHC molecules (Sarlo *et al.*, 2000). Finally, while the data from an individual laboratory using the MINT assay is robust, the MINT assay is also plagued by interlaboratory differences (Blaikie and Basketter, 1999a).

Finally, any animal test will suffer from the inability to provide a mechanistic description of the immune response to a given protein in humans. First and foremost, inbred strains of mice will present peptide molecules with the specificity conferred by their murine MHC molecules. Human HLA molecules, while highly related to mouse MHC molecules, do not have identical peptide specificities (Bono and Strominger, 1982; Schwaiger *et al.*, 1993). Human HLA transgenic mice have become available for application to the mechanistic study of human immune responses (Black *et al.*, 2002; Boyton and Altmann, 2002; Chen *et al.*, 2002, 2003; Das *et al.*, 2000; Ito *et al.*, 1996; Raju *et al.*, 2002; Sonderstrup *et al.*, 1999). HLA transgenic mice suffer from their species-specific immune system complexities (Farrar *et al.*, 2000; Kim and Jang, 1992). HLA transgenic mice are often used for mapping studies when expressing a single HLA molecule, a situation not found in humans. This is especially of note for HLA-DQ transgenic mice where cross-pairing between different HLA-DQ alleles has been shown to create new peptide presentation specificities (Krco *et al.*, 1999). While the number of available HLA class II transgenic mice expressing common class II alleles is increasing, there are not enough different strains to represent the complexity of human HLA class II frequencies.

To avoid the issues arising from immunogenicity analyses in animals other than humans and to incorporate the complexities of the human target population for commercial proteins, we have developed a method to rank the immunogenicity of proteins using human peripheral blood monocytes (PBMC) as the test subject. The method is based on data gathered using a previously described epitope mapping technique, which relies on the population-based determination of CD4⁺ T cell responses (Stickler *et al.*, 2000, 2003a,b). Because large replicates of human samples are used, the information provided is applicable to general populations. The information is gathered by testing CD4⁺ T cell responses to peptides presented by dendritic cells and, therefore, includes response data, depending on the protein analyzed, representing both recall and pri-

mary immune responses (Harding, 2003). The data do not suffer from the specificity issues surrounding the use of inbred mice. This method can rank proteins based on their overall immunogenicity but cannot provide relative potency information unless the data are compared to pre-existing animal data. Four well-characterized industrial allergens were placed in the order determined by the GPIT, the BDF1 MINT, and by comparison with human sensitization in occupationally exposed workers (Robinson *et al.*, 1998). Proteins known to be either immunogenic or presumably tolerizing in humans were tested as positive and negative controls for the method.

MATERIALS AND METHODS

Peptides. All peptides were obtained from a commercial source (Mimotopes, San Diego, CA); 15-mer peptides offset by three amino acids that described the entire sequence of the proteins of interest were synthesized in a Multipin format (Maeji *et al.*, 1990). Peptides were resuspended in DMSO at approximately 1 to 2 mg/ml and stored at -70°C until use.

Protein sequences. Amino acid sequences from the following industrial proteases were tested in the assay: *Bacillus lentus* subtilisin (Swissprot accession P29600), *Bacillus amyloliquefaciens* subtilisin (BPN' Y217L, Swissprot accession P00782), *Bacillus licheniformis* subtilisin (Alcalase[®], Swissprot accession P00780), and *Bacillus licheniformis* α -amylase (Swissprot accession P06278). The following other proteins were tested: human interferon- β (Swissprot accession P01674), human β 2-microglobulin (accession AAH32589), human erythropoietin (Swissprot accession P01588), human thrombopoietin (pir accession G02729), Bertholletia excelsa 2S storage protein (Ber e1, Swissprot accession P04403), and mouse Vh36-60 gene family member (similar to Swissprot accession P01823).

Human donor blood samples. Buffy coat samples were obtained from two commercial sources (Stanford Blood Center, Palo Alto, CA, and BloodSource, Sacramento, CA). Buffy coat samples were further purified by density separation. Each sample was HLA typed for HLA-DR β and HLA-DQ β using a commercial PCR-based kit (Bio-Synthesis, Lewisville, TX).

Preparation of dendritic cells and CD4⁺ T cells. The preparation of monocyte-derived dendritic cells and CD4⁺ T cells has been described previously (Stickler *et al.*, 2000; Zhou and Tedder, 1996). Briefly, monocytes were purified by adherence to plastic in AIM V medium (Gibco/Life Technologies, Baltimore, MD). Adherent cells were cultured in AIM V media containing 500 units/ml of recombinant human IL-4 (Endogen, Woburn, MA) and 800 units/ml recombinant human GM-CSF (Endogen) for 5 days. On day 5, recombinant human IL-1 α (Endogen) and recombinant human TNF- α (Endogen) were added at 50 units/ml and 0.2 units/ml, respectively. On day 7, the fully matured dendritic cells were treated with 50 $\mu\text{g}/\text{ml}$ mitomycin c (Sigma Chemical Co., St. Louis, MO) for 1 h at 37°C . Treated dendritic cells were dislodged with 50 mM EDTA in PBS, washed in AIM V media, counted, and resuspended in AIM V media at 2×10^5 cells/ml.

CD4⁺ T cells were purified by negative selection from frozen aliquots of PBMC using Collect CD4 columns (Cedarlane, Toronto, Ontario, Canada) or Dynabeads[®] (Dynal Biotech, Oslo, Norway). CD4⁺ T cell populations were routinely >80% pure and >95% viable as judged by Trypan blue (Sigma Chemical Co.) exclusion. CD4⁺ T cells were resuspended in AIM V media at 2×10^6 cells/ml.

Assay conditions. CD4⁺ T cells and dendritic cells were plated in round-bottomed 96-well format plates at 100 μl of each cell mix per well. The final cell number per well was 2×10^4 DC and 2×10^5 CD4⁺ T cells. Peptide was added to a final concentration of ~ 5 $\mu\text{g}/\text{ml}$ in 0.25–0.5% DMSO. Control wells contained DMSO without added peptide. Each peptide was tested in duplicate. Cultures were incubated at 37°C in 5% CO_2 for 5 days. On day 5, 0.5 μCi of tritiated thymidine (NEN/DuPont, Boston, MA) was added to each

well. On day 6, the cultures were harvested onto glass fiber mats using a TomTec manual harvester (TomTec, Hamden, CT) and then processed for scintillation counting. Proliferation was assessed by determining the average CPM value for each set of duplicate wells (TriLux Beta, Wallac, Finland).

Data analysis. For each individual, average CPM values for all the peptides were determined. The average CPM values for each peptide were divided by the average CPM value of the control (DMSO only) wells to calculate a stimulation index (SI). A positive response was recorded if the SI value was equal to or larger than 2.95. The 2.95 value was determined empirically, tested by receiver-operator curve analysis, and shown to result in a highly accurate and efficient value for the analysis of I-mune assay data (Stickler *et al.*, 2003a). For each protein assessed, positive responses to individual peptides by individual donors were compiled. Donor blood samples were tested with each peptide set to yield an average of at least two responses (response = SI of 2.95 or greater) per peptide. For example, since each donor responded to an average of three peptides per 100 peptides tested (average background of ~3.15% for 11 industrial enzymes [Stickler *et al.*, 2003a]), this would result in the need to test 67 donors to get an expected two responses per peptide for the 100 peptides tested. Data for each protein tested is graphed as the percentage of responders within the population tested to each peptide in the set.

To determine the background response for a given protein, the percentage of responses for each peptide in the set was averaged and a standard deviation was calculated.

Statistical methods. The total variation distance between the empirical frequencies and the uniform distribution (Kotz and Johnson, 1988), the structure index value, was calculated based on the following equation:

$$\text{Structure Index} = \sum \left| f(i) - \frac{1}{p} \right|$$

where Σ is the sum over all peptides in the peptide set of the absolute value of the proportion of responses to each peptide minus the frequency of that peptide in the set; $f(i)$ is defined as the frequency of responses for an individual peptide divided by the total number of responses accumulated; and p is the number of peptides in the peptide set. For more information on this, see Results. For an example of the calculation for BPN' Y217L, see the Supplementary Material online.

Statistical significance of peptide response frequency was calculated based on Poisson statistics (Stickler *et al.*, 2003a). The average frequency of responders to all the peptides was used to calculate a Poisson distribution based on the total number of responses and the number of peptides in the set. A frequency of response to a peptide is considered significant if $p < 0.05$.

RESULTS

HLA Types Within the Donor Pool

The HLA-DR β 1 allelic expression was determined for 184 random individuals to compare allelic frequencies in our donor pool with the U.S. general population. It is important that our data be representative of the larger population if our results are to be considered generally accurate. HLA typing was performed using low-stringency PCR determinations. The data compiled for the Stanford and Sacramento samples are compared to the Caucasian HLA-DR β 1 frequencies published previously (Marsh *et al.*, 2000; Fig. 1). Our donor population is enriched for HLA-DR4 and HLA-DR15. However, our frequencies for these alleles are well within the reported ranges (5.2–24.8% for HLA-DR4 and 5.7–25.6% for HLA-DR15). Similarly, for HLA-DR3, -DR7, and -DR11, our frequencies are lower than the average Caucasian frequency but within the

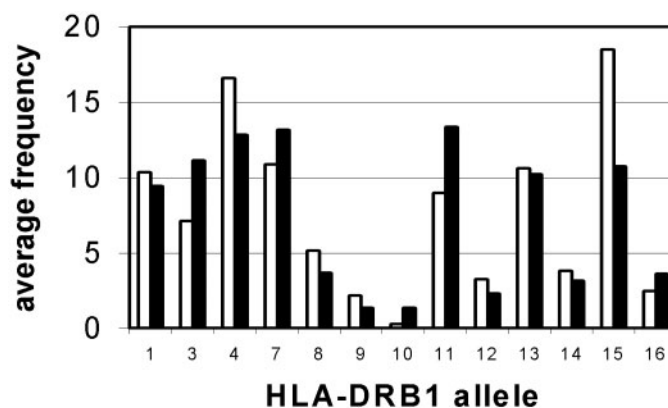


FIG. 1. HLA-DR β 1 allelic frequency in the community donor pool (white bars) compared with published frequencies (black bars). HLA-DR β 1 allelic frequency in the community donor pool was compared with the frequency of HLA-DR β 1 alleles in the U.S. general Caucasian population (Marsh *et al.*, 2000).

reported ranges for those alleles. Also of note, HLA-DR15 is found at a higher frequency in ethnic populations (persons of Asian and Hispanic descent; see ethnicity data at the U.S. Census web site, www.census.gov), which are heavily represented in the San Francisco Bay area.

Epitope Mapping Data for Four Known Respiratory Allergens

Epitope mapping was performed on a set of four industrial enzymes to provide the data used in the ranking analysis. These enzymes were selected for this analysis because they are known to cause immune responses in occupationally exposed workers (Sarlo *et al.*, 1997). However, these enzymes have not been found to induce widespread allergic type responses in the general population (Pepys *et al.*, 1973, 1985). Recently, a retrospective study of atopic women found that only 0.15% of all donors tested exhibited any evidence of exposure to a *Bacillus protease* included in common laundry detergent formulations (Sarlo *et al.*, 2003). Therefore, while these proteins are known to be capable of inducing immune responses in exposed workers, the majority of the community donors tested in our assays are not expected to carry significant *a priori* exposure to these enzymes.

α -amylase. Individuals ($n = 82$) were tested with peptides derived from the α -amylase sequence. The average percentage of response to all peptides in this set was $2.80 \pm 3.69\%$ (average $\pm$ standard deviation), well within our overall average for 11 industrial enzymes of 3.15 ± 1.57 (Stickler *et al.*, 2003a). Prominent responses were noted to amino acids 34–48 (peptide 12), 160–174 (peptide 54), and 442–456 (peptide 148; Fig. 2). All three of these responses were highly significant above the background response ($p < 0.0001$).

B. lentus subtilisin. Individuals ($n = 65$) were tested with two replicate peptide sets for this protein. The results were

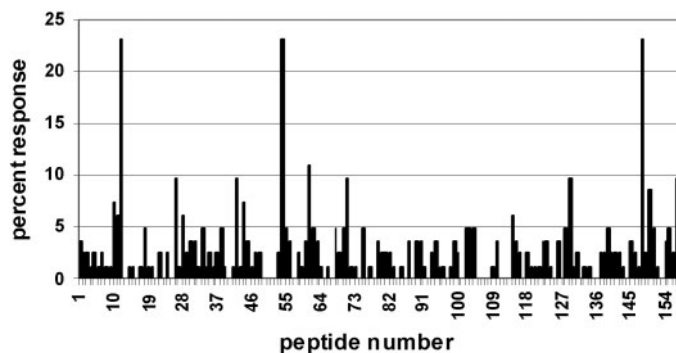


FIG. 2. I-mune assay results for *B. licheniformis* α -amylase. Community donors ($n = 82$) were tested with peptides derived from the sequence of α -amylase. The consecutive 15-mer peptides offset by three amino acids are listed on the x axis, and the percentage of the donors who responded to each peptide is shown on the y axis.

compiled. The average percentage of response for this peptide set was found to be $3.45 \pm 2.90\%$. A prominent response was noted at amino acids 160–174 (peptide 55; $p = 0.0004$; Fig. 3).

BPN' Y217L. Individuals ($n = 113$) were tested with two peptide sets. The compiled average percentage of response for this data set was 3.24% . Prominent responses were noted at amino acids 70–84 (peptide 24) and 109–123 (peptide 37; Fig. 4). Both regions accumulated responses at a statistically significant level ($p < 0.001$). A region of responses was also noted around amino acid 154, but the number of responses there was not significant.

Alcalase. Individuals ($n = 92$) were tested with peptides derived from this sequence. The background response to this protein was low at 2.35% . The same peptide set was tested in two temporally spaced analyses and the data were compiled. Additionally, there were significantly more peptides returning no response within the set for this protein. A prominent response was noted at amino acids 19–33 (peptide 7; $p < 0.0001$; Fig. 5).

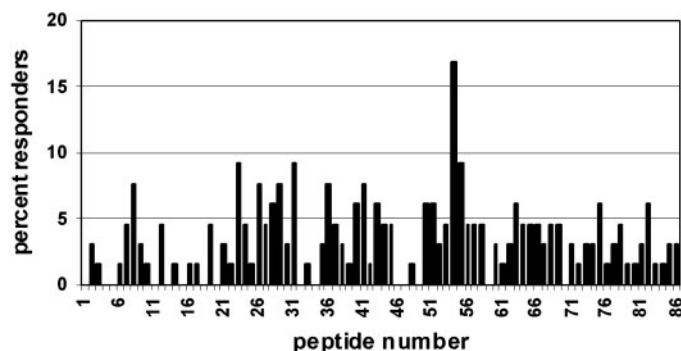


FIG. 3. I-mune assay results for *B. lentus* subtilisin. Community donors ($n = 65$) were tested with peptides derived from the sequence of *B. lentus* subtilisin. The consecutive 15-mer peptides offset by three amino acids are listed on the x axis, and the percentage of the donors who responded to each peptide is shown on the y axis.

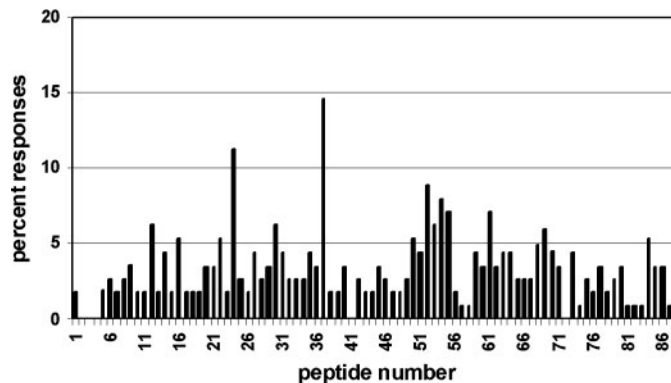


FIG. 4. I-mune assay results for BPN' Y217L. Community donors ($n = 113$) were tested with peptides derived from the sequence of BPN' Y217L. The consecutive 15-mer peptides offset by three amino acids are listed on the x axis, and the percentage of the donors who responded to each peptide is shown on the y axis.

Structure Index Calculations

We have created a new statistic to measure the structuration of the sample. We could have used any distance from our proportions to the uniform distribution that puts equal weight in all the categories (peptides). We could have also used entropy (Kotz and Johnson, 1988), which would have resulted in an index in the form $H(\text{prop}) = -\sum \text{prop}(i) \log(\text{prop}(i))$, where prop is the vector of proportions of the samples that fall into each category. The larger the entropy, the further from certainty and the less structured that sample. We instead chose the total variation distance, or L1 distance (Kotz and Johnson, 1988), that simply takes the sum of the absolute values of the differences between the proportions and the uniform distribution that puts equal probability of falling into any category. We call this value, calculated using data generated in the complete data set (all of the donor responses to all of the peptides tested)

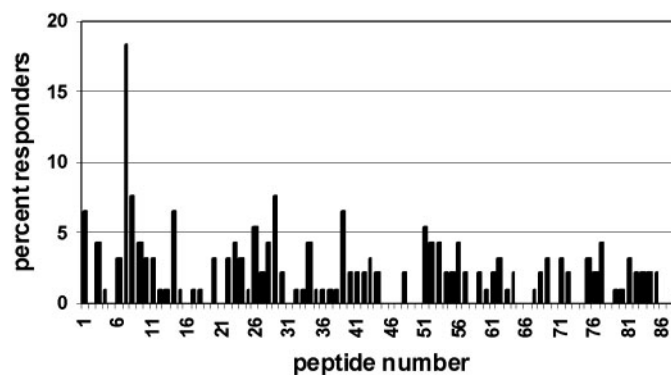


FIG. 5. I-mune assay results for Alcalase. Community donors ($n = 92$) were tested with peptides derived from the sequence of Alcalase. The consecutive 15-mer peptides offset by three amino acids are listed on the x axis, and the percentage of the donors who responded to each peptide is shown on the y axis.

the structure index value. (See supplemental data online for an example of the calculation.)

In theory, if every peptide in the data set had the same number of responses, $f(i)-1/p$ would equal zero. In other words, the proportion of the responses at each peptide would equal the proportion of the data set represented by one peptide, and the difference between these values would equal zero. The absolute value of the sum of the data for all the peptides (zero at each one) would equal zero. On the other hand, if all the accumulated responses were at one peptide, the value would approach 2.0.

To ensure comparability of the structure index values, a stable response pattern must be achieved within the data set. The number of donors necessary was determined empirically for BPN' Y217L, as shown in Figure 6. After about two responses per peptide, the structure index value reaches a plateau level. For the BPN' Y217L peptide set, this occurred after testing approximately 50 donors. For all additional peptide set testing, enough donors to achieve at least two responses per peptide were used.

For each of the industrial enzymes, the compiled responses were used to calculate the structure index values. The structure index values were as follows: 0.81 for amylase, 0.72 for Alcalase, 0.64 for B. lentus subtilisin, and 0.52 for BPN' Y217L (see Table 1). This result indicates that there is more activity induced by the amylase peptide set, when CD4+ T cell activation is measured by a level of proliferation resulting in an SI of 2.95 or greater, as compared with activity measured using the other peptide sets. The result for BPN' Y217L indicates that the peptide set derived from the sequence of this protein was the least active, with the lowest amount of structure. The structure

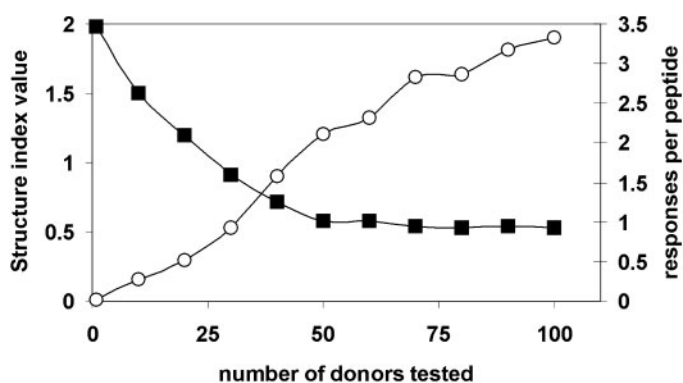


FIG. 6. Calculated structure values decrease with increasing number of responses per peptide. Structure index value (left y axis) and the average number of responses per peptide (right y axis) were calculated for the BPN' Y217L peptide set. Results are graphed versus the accumulating number of donors (x axis). The structure index values are shown as black squares, and the average response per peptide is shown as open circles. The structure index value reaches a plateau level at approximately two responses per peptide, which occurs after testing 50 donors.

TABLE 1
Structure Index Values for Four Respiratory Allergens

Allergen	Peptides	n	Responses per peptide	Epitope regions	Structure value
Amylase	157	82	2.29	3	0.81
B. lentus subtilisin	86	65	2.24	1	0.64
Alcalase	88	92	2.16	1	0.72
BPN' Y217L	88	113	3.61	2	0.52

index values rank the four tested proteins as follows: amylase > Alcalase > B. lentus subtilisin > BPN' Y217L.

Comparison to Animal Models and Human Exposure Rates

Two animal models have been used for the prediction of allergenicity and immunogenicity of industrial proteins. The GPIT and MINT assays have been validated by comparing the potency differences determined experimentally with skin prick test (SPT) positivity rates in occupationally exposed industrial workers (Sarlo *et al.*, 1997). In humans, the percentage of SPT+ workers was compared to the level of enzyme in the formulation under manufacture, and potency values were determined that ranked the proteins as follows: amylase >> savinase (B. lentus subtilisin) = Alcalase > BPN' Y217L. Both the guinea pig (GPIT) and BDF1 mouse (MINT) models ranked the proteins in the same order: amylase > Alcalase > B. lentus subtilisin > BPN' Y217L. However, the relative immunogenicity values differed in the two models. Figure 7 shows the calculated structure values for the four industrial enzymes graphed versus the published GPIT (Fig. 7A) and MINT (Fig. 7B) potency values. Human cell-based structure data presented here correlates well with both methods (R^2 values of 0.86 and 0.84, respectively) for the four tested proteins.

Structure Index Values of Additional Proteins

To provide more tests of the statistical method, the structure index value for a putative nonimmunogenic protein was tested. Human β 2-microglobulin was selected since this protein has been shown to induce both central and peripheral tolerance (Guery *et al.*, 1995). Peptides were tested and the results were compiled for 87 community donors. The structure value was calculated and found to be 0.39 (see Table 2). Due to the tolerance induction to this ubiquitous protein, a value of 0.39 likely defines the functional (rather than theoretical) lower limit of the structure value calculations.

Structure index values were then determined for several proteins known to cause immune responses at different frequencies in the human population. These included Ber e1, the major allergen found in Brazil nuts, human interferon- β (IFN- β), human thrombopoietin (Tpo), and a mouse VH 36-60

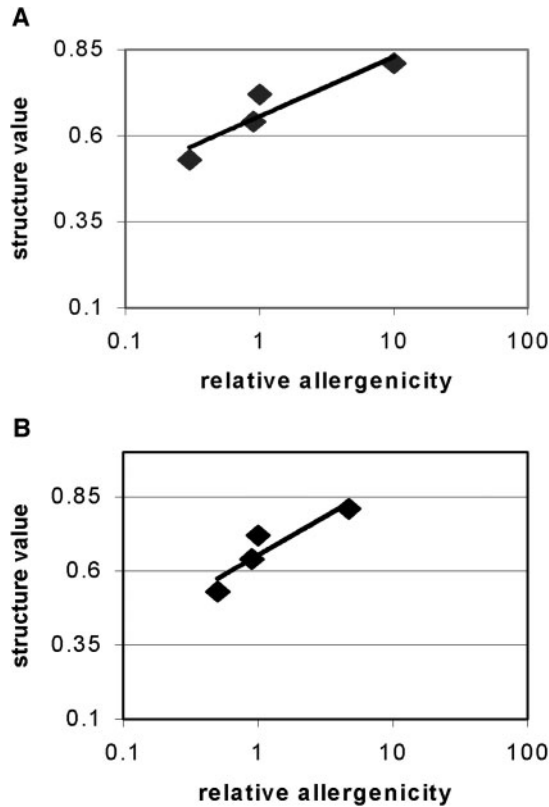


FIG. 7. A comparison between published (A) GPIT and (B) MINT ranking data and the structure index values for four industrial enzymes. The relative allergenicity of α -amylase, Alcalase, BPN' Y217L, and B. lentus subtilisin as determined in guinea pig- (GPIT-) and mouse (MINT)-based assays (x axis) is compared with the structure index values (y axis).

family member V region gene (Table 2). Human IFN- β , Tpo, and Ber e1 are all known to induce immune responses in humans (Basser *et al.*, 2002; Li *et al.*, 2001; Scagnolari *et al.*, 2002; Sicherer and Sampson, 2000). The structure values for IFN- β , Tpo, and Ber e1 are all comparatively high. The value for the mouse VH region is comparatively low, suggesting that this protein is comparatively nonimmunogenic. This result is consistent with a structural analysis of potential immunogenicity of the mouse heavy chain families (Olsson *et al.*, 1991). The structure value for human erythropoietin (Epo) was found to be 0.45. Erythropoietin was used to treat chronic anemia without immunological incident for over a decade. Recently, numerous cases of pure red cell anemia (PRCA) have been linked to the use of this molecule (Casadevall *et al.*, 2002). The frequency of these events is very small, and the appearance of these cases coincides with some changes in the manufacturing and use of the molecule.

Determination of a Cut-off Value for Reduced Immunological Risk

The structure index values for the four respiratory allergens amylase, B. lentus subtilisin, BPN'Y217L, and Alcalase were

TABLE 2
Structure Index Values for Selected Additional Proteins

Protein	Peptides	n	Average response per peptide (%)	Responses per peptide	Epitope regions	Structure value
hTpo	52	99	2.56	2.54	1	0.65
hIFN- β	52	88	3.17	2.79	1	0.75
Ber e1	27	92	4.27	3.92	2	0.66
hEpo	52	178	4.32	7.69	1	0.45
Mouse Vh 36-60 family	35	74	7.0	5.23	0	0.38
β 2-microglobulin	36	87	3.90	3.39	0	0.39

combined with the values for human IFN- β , Tpo, and Ber e1 to determine a cut-off value for a putatively immunogenic protein. The average of these seven values was found to be 0.68 with a standard deviation of 0.09 (Table 3). Two standard deviations below the mean is a value of 0.50. To determine the limit using the comparatively less immunogenic proteins, the same analysis was performed using values for β 2-microglobulin, the mouse VH36-60, and Epo. The average structure index value for these three proteins is 0.41 with a standard deviation of 0.04. Two standard deviations above the average are a value of 0.49. Therefore, a protein with a structure index value of less than 0.50 could be categorized by this analysis using these proteins as benchmarks of a comparatively less immunogenic protein.

Application of Structure Index Values for the Creation of Reduced Immunogenicity Proteins

As an example of how the structure analysis could be used to guide the design of reduced immunogenic proteins, structure

TABLE 3
Calculation of a Suggested Cut-off Value for Comparatively Less Immunogenic Proteins

More immunogenic		Less immunogenic	
Protein	SIV	Protein	SIV
Amylase	0.81	β 2-microglobulin	0.39
B. lentus	0.64	ScFv	0.38
BPN' Y217L	0.52	Epo	0.45
Alcalase	0.72		
hIFN-beta	0.75		
Ber e1	0.66		
Tpo	0.65		
Mean	0.68	Mean	0.41
95% CI	0.59-0.77	95% CI	0.31-0.50
Variance	0.009	Variance	0.001
SD	0.09	SD	0.04
SE	0.04	SE	0.02
Mean - (2*SD)	0.50	Mean + (2*SD)	0.49

values were recalculated for all epitope-containing proteins where the epitope peptide response rates were artificially reduced to the background response level. The complete data sets containing all the responses for each donor to each peptide were numerically modified for structure recalculations by reducing the responses to epitope peptides to average background levels. Since positive responses were removed from the calculation, an equivalent number of responses were scattered randomly through the data set to maintain the same overall rate of response (Fig. 8). This analysis shows that by modifying the epitope responses within the tested proteins a variant could be created that would demonstrate a reduced structure index value. Some of the more immunogenic proteins, such as amylase and Alcalase, could not be modified to reduce their structure index values to below the cut-off value of 0.50. In these cases, it might not be possible to create reduced immunogenicity variants. However, effective modifications of Tpo, Ber e1, and BPN'Y217L might be possible.

DISCUSSION

We have devised a method to assess the overall immunogenic potential of any protein by an analysis of the response rate of individual donors to a set of peptides describing the protein of interest. The immunogenicity ranking is comparative and is based on an *in vitro* correlation of exposure of large human donor pools to the proteins under study. Therefore, proteins selected or constructed using this method as a guide will not be nonimmunogenic for every individual member of the population. Data provided from this analysis is best thought of as a population-based risk reduction tool. The method uses dendritic cells as antigen-presenting cells, 15-mer peptides offset by three amino acids that encompass the entire sequence

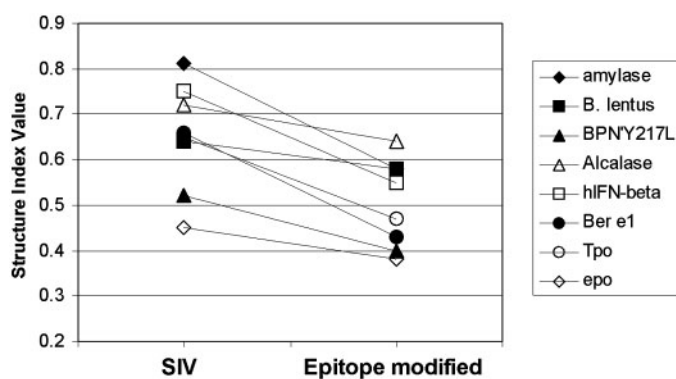


FIG. 8. Calculated structure index values for epitope-containing proteins, and theoretical epitope-modified variants. The data sets for each protein indicated (amylase, Alcalase, BPN'Y217L, B. lentus subtilisin, Tpo, Epo, human IFN- β , and Ber e1) were modified by reducing the number of responses to the epitope peptides to the background level. The overall response rate was kept the same by adding back responses randomly over the data sets. The resultant structure index values are shown for each of the potential modified variants.

of the protein, and CD4+ T cells from the dendritic cell donors. A response is tallied if the average CPM of tritiated thymidine incorporation for a particular peptide is greater than or equal to 2.95 times the background CPM. Results per peptide are tabulated for a large donor set that should reflect general HLA allele frequencies, with some variations. A statistical calculation based on the determination of “difference from linearity” is performed, and this structure value can be used to rank the relative immunogenicity of these proteins. We found that the ranking returned by this method closely reflects immunogenicity determinations (by the MID assay [Sarlo *et al.*, 1997]) and allergenicity of these proteins as respiratory allergens when determined in occupationally exposed workers (Sarlo *et al.*, 1997) or in the GPIT or MINT (Robinson *et al.*, 1998) assays.

In addition to correctly ranking the allergenicity of four known respiratory allergens, the structure index values for a set of known human immunogens (hTPO [Li *et al.*, 2001], hIFN- β [PRISMS Study Group, 2001], and Ber e1 [Nordlee *et al.*, 1996]) were found to be comparatively high, indicating that these proteins might be capable of inducing immune responses in a significant number of exposed people. The average percentage of response values of the two human-derived protein immunogens were within the range defined by 11 industrial enzymes: 3.15 ± 1.6 (for example, see Table 2). Background values within this range are assumed to represent responses to novel proteins, as the industrial enzymes tested are not widely encountered by community donors (Pepys *et al.*, 1973, 1985; Sarlo *et al.*, 2003). A high structure value in the context of a low background rate suggests that humans are largely immunologically “naïve” to these human sequence-derived proteins, and that the immune system “ignores” the proteins until non-physiological doses are introduced in the presence of an adjuvant signal. The structure value for Ber e1, the major protein allergen in Brazil nuts, was also found to be high. Our community donor pool is likely exposed to Brazil nuts in food, and, therefore, the value reflects contributions from allergic donors, exposed yet unallergic donors, and donors who are truly unexposed. Conversely, the structure value for a mouse VH 36-60 gene family member was low, commensurate with its predicted immunogenicity (Olsson *et al.*, 1991). The structure value determined for β 2-microglobulin was also low, as would be expected given that this molecule is presumed to be subject to both peripheral and central tolerance mechanisms (Guery *et al.*, 1995). Finally, the structure value for human erythropoietin was also found to be low, consistent with its safe use and low frequency of adverse affects (Casadevall *et al.*, 2002).

The range of structure values was low. A putative nonimmunogenic protein, β 2-microglobulin, had a value of 0.39, while the most immunogenic protein verified by human exposure, amylase, had a value of 0.81. The correlation between these values and potency in animal models is logarithmic. Small differences in the value can indicate large differences in immunogenicity. Therefore, it is critical to test a large enough

donor set (greater than two responses per peptide) to make accurate comparisons.

The comparative ranking of proteins tested in this assay assumes that the immunogenicity of whole protein molecules would be compared *in vivo* at the same dose, in the same formulation, in a matched set of donors, and over the same dose course. This analysis also precludes any processing and/or presentation differences in the proteins, as well as general physical and structural properties (i.e., stability, activity, and multimerization).

The method described also allows for the localization of T cell epitopes in any protein of interest. CD4⁺ T cell epitopes can be determined in the absence of individuals exposed to the test protein (Stickler *et al.*, 2000, 2003a,b). Modification of peptide epitope sequences to select variants less likely to induce immune responses can be performed using unexposed community donors. An analysis of donor responses to the modified peptide variants can be used to calculate structure values for the new protein. Note that testing of protein variants designed to be less immunogenic, by virtue of provoking fewer responses *in vitro* with large replicates of human donors, cannot be rationally tested in guinea pigs or mice, as rodents often have different CD4⁺ T cell epitopes than the human population (see [Mucha *et al.*, 2002] for B. lentus and BPN⁺ Y217L epitopes in guinea pigs). Transgenic mice are limited in their utility, since they typically do not express more than one HLA allele and enough strains to represent the complexity of the human population HLA allele frequencies are not currently available.

Ranking of proteins does not imply any fold potency differences. Potency determination can be extrapolated from an alignment of our data with animal data. However, the values determined would be subject to the same inherent inaccuracies as the animal data used. It follows from the data presented that the selection of a protein with the lowest structure value would minimize the risk of inducing immune responses in human subjects on a population basis. As a general rule for the selection of lead candidates, we calculated that proteins should have structure values less than 0.5. We currently have no human *in vivo* data to support the use of this method, a criticism that can be levied at almost all predictive and functional methods to modify the immunogenicity of proteins. Full validation of this ranking method will require testing of known immunogenic and allergenic proteins and their modified variants and will not be complete until human donor data become available.

The method described here is an assay to determine the relative immunogenicity of proteins in human subjects that does not expose the donor to the protein of interest. Proteins can be ranked relative to one another. The method encompasses a previously described technique to identify immunodominant peptide epitopes. Taken together, this information may allow for the selection of reduced immunogenicity proteins and can direct the rational modification of proteins to

create and test hypo-immunogenic variants appropriate for use in humans.

REFERENCES

- Alving, C. R. (2002). Design and selection of vaccine adjuvants: Animal models and human trials. *Vaccine* **20**(Suppl. 3), S56–S64.
- Basser, R. L., O'Flaherty, E., Green, M., Edmonds, M., Nichol, J., Menchaca, D. M., Cohen, B., and Begley, C. G. (2002). Development of pancytopenia with neutralizing antibodies to thrombopoietin after multicycle chemotherapy supported by megakaryocyte growth and development factor. *Blood* **99**, 2599–2602.
- Bernstein, I. L., Bernstein, J. A., Miller, M., Tierzieva, S., Bernstein, D. I., Lummus, Z., Selgrade, M. K., Doerfler, D. L., and Seligy, V. L. (1999). Immune responses in farm workers after exposure to *Bacillus thuringiensis* pesticides. *Environ. Health Perspect.* **107**, 575–582.
- Black, K. E., Murray, J. A., and David, C. S. (2002). HLA-DQ determines the response to exogenous wheat proteins: A model of gluten sensitivity in transgenic knockout mice. *J. Immunol.* **169**, 5595–5600.
- Blaikie, L., and Basketter, D. A. (1999a). Experience with a mouse intranasal test for the predictive identification of respiratory sensitization potential of proteins. *Food Chem. Toxicol.* **37**, 889–896.
- Blaikie, L., and Basketter, D. A. (1999b). Strain variation in the IgG1 antibody response to proteins administered intranasally in the mouse. *Food Chem. Toxicol.* **37**, 897–904.
- Bono, M. R., and Strominger, J. L. (1982). Direct evidence of homology between human DC-1 antigen and murine I-A molecules. *Nature* **299**, 836–840.
- Boyton, R. J., and Altmann, D. M. (2002). Transgenic models of autoimmune disease. *Clin. Exp. Immunol.* **127**, 4–11.
- Braun, A., Kwee, L., Labow, M. A., and Alsenz, J. (1997). Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon α (IFN- α) in normal and transgenic mice. *Pharm. Res.* **14**, 1472–1478.
- Brimmes, M. K., Bonifaz, L., Steinman, R. M., and Moran, T. M. (2003). Influenza virus-induced dendritic cell maturation is associated with the induction of strong T cell immunity to a coadministered, normally nonimmunogenic protein. *J. Exp. Med.* **198**, 133–144.
- Bussiere, J. L. (2003). Animal models as indicators of immunogenicity of therapeutic proteins in humans. *Dev. Biol.* **112**, 135–139.
- Casadevall, N., Nataf, J., Viron, B., Kolta, A., Kiladjian, J. J., Martin-Dupont, P., Michaud, P., Papo, T., Ugo, V., Teyssandier, I., *et al.* (2002). Pure red cell aplasia and antierythropoietin antibodies in patients treated with recombinant erythropoietin. *N. Engl. J. Med.* **346**, 469–475.
- Chamberlain, P., and Mire-Sluis, A. R. (2003). An overview of scientific and regulatory issues for the immunogenicity of biological products. *Dev. Biol.* **112**, 3–11.
- Chen, D., Ueda, R., Harding, F., Patil, N., Mao, Y., Kurahara, C., Platenburg, G., and Huang, M. (2003). Characterization of HLA DR3/DQ2 transgenic mice: A potential humanized animal model for autoimmune disease studies. *Eur. J. Immunol.* **33**, 172–182.
- Chen, Z., Dudek, N., Wijburg, O., Strugnell, R., Brown, L., Deliyannis, G., Jackson, D., Koentgen, F., Gordon, T., and McCluskey, J. (2002). A 320-kilobase artificial chromosome encoding the human HLA DR3-DQ2 MHC haplotype confers HLA restriction in transgenic mice. *J. Immunol.* **168**, 3050–3056.
- Das, P., Abraham, R., and David, C. (2000). HLA transgenic mice as models of human autoimmune diseases. *Rev. Immunogenet.* **2**, 105–114.

- Farrar, J. D., Smith, J. D., Murphy, T. L., Leung, S., Stark, G. R., and Murphy, K. M. (2000). Selective loss of type I interferon-induced STAT4 activation caused by a minisatellite insertion in mouse Stat2. *Nat. Immunol.* **1**, 65–69.
- Ge, Y., Powell, S., Van Roey, M., and McArthur, J. G. (2001). Factors influencing the development of an antifactor IX (FIX) immune response following administration of adeno-associated virus-FIX. *Blood* **97**, 3733–3737.
- Gough, L., Sewell, H. F., and Shakib, F. (2001). The proteolytic activity of the major dust mite allergen Der p1 enhances the IgE antibody response to a bystander antigen. *Clin. Exp. Allergy* **31**, 1594–1598.
- Guery, J. C., Sette, A., Appella, E., and Adorini, L. (1995). Constitutive presentation of dominant epitopes from endogenous naturally processed self- β 2-microglobulin to class II-restricted T cells leads to self-tolerance. *J. Immunol.* **154**, 545–554.
- Hall, F. C., Rabinowitz, J. D., Busch, R., Visconti, K. C., Belmares, M., Patil, N. S., Cope, A. P., Patel, S., McConnell, H. M., Mellins, E. D., et al. (2002). Relationship between kinetic stability and immunogenicity of HLA-DR4/peptide complexes. *Eur. J. Immunol.* **32**, 662–670.
- Harding, F. (2003). CD4+ T cell epitope identification: Applications to allergy. *Clin. Exp. Allergy* **33**, 557–565.
- Hileman, R. E., Silvanovich, A., Goodman, R. E., Rice, E. A., Holleschak, G., Astwood, J. D., and Hefle, S. L. (2002). Bioinformatic methods for allergenicity assessment using a comprehensive allergen database. *Int. Arch. Allergy Immunol.* **128**, 280–291.
- Ito, K., Bian, H. J., Molina, M., Han, J., Magram, J., Saar, E., Belunis, C., Bolin, D. R., Arceo, R., Campbell, R., et al. (1996). HLA-DR4-IE chimeric class II transgenic, murine class II-deficient mice are susceptible to experimental allergic encephalomyelitis. *J. Exp. Med.* **183**, 2635–2644.
- Johnsen, C., Sorensen, T., Larson, A., Secher, A., Andreassen, E., Kofoed, G., Nielsen, L., and Gyntelberg, F. (1997). Allergy risk in an enzyme-producing plant: A retrospective follow-up study. *Occup. Environ. Med.* **54**, 671–675.
- Kim, B. S., and Jang, Y. S. (1992). Constraints in antigen processing result in unresponsiveness to a T cell epitope of hen egg lysozyme in C57BL/6 mice. *Eur. J. Immunol.* **22**, 775–782.
- Kimber, I., Bernstein, I. L., Karol, M. H., Robinson, M. K., Sarlo, K., and Selgrade, M. K. (1996). Workshop overview. Identification of respiratory allergens. *Fundam. Appl. Toxicol.* **33**, 1–10.
- Kore, E., Zuckerman, L. A., and Mire-Sluis, A. R. (2002). Immune responses to therapeutic proteins in humans—clinical significance, assessment and prediction. *Curr. Pharm. Biotechnol.* **3**, 349–360.
- Kotz, S., and Johnson, N., Eds. (1988). *Encyclopedia of Statistical Sciences*, pp. 300, 512. John Wiley and Sons, New York.
- Krco, C. J., Watanabe, S., Harders, J., Griffiths, M. M., Luthra, H., and David, C. S. (1999). Identification of T cell determinants on human type II collagen recognized by HLA-DQ8 and HLA-DQ6 transgenic mice. *J. Immunol.* **163**, 1661–1665.
- Li, J., Yang, C., Xia, Y., Bertino, A., Glaspy, J., Roberts, M., and Kuter, D. J. (2001). Thrombocytopenia caused by the development of antibodies to thrombopoietin. *Blood* **98**, 3241–3248.
- Maeji, N. J., Bray, A. M., and Geysen, H. M. (1990). Multi-pin peptide synthesis strategy for T cell determinant analysis. *J. Immunol. Methods* **134**, 23–33.
- Marsh, S. G. E., Parham, P., and Barber, L. D. (2000). *The HLA Facts Book*. Academic Press, San Diego.
- Mucha, J. M., Stickler, M. M., Poulouse, A., Ganshaw, G., Saldajeno, M., Collier, K., Huang, M. T., and Harding, F. A. (2002). Enhanced immunogenicity of a functional enzyme by T cell epitope modification. *BMC Immunol.* **3**, 2.
- Nordlee, J. A., Taylor, S. L., Townsend, J. A., Thomas, L. A., and Bush, R. K. (1996). Identification of a Brazil nut allergen in transgenic soybeans. *N. Engl. J. Med.* **334**, 688–692.
- Oehlschlager, S., Reece, P., Brown, A., Hughson, E., Hird, H., Chisholm, J., Atkinson, H., Meredith, C., Pumphrey, R., Wilson, P., et al. (2001). Food allergy—towards predictive testing for novel foods. *Food Addit. Contam.* **18**, 1099–1107.
- Olsson, P. G., Hammarstrom, L., and Smith, C. I. (1991). Antigenicity of mouse monoclonal antibodies. A study on the variable region of the heavy chain. *J. Theor. Biol.* **151**, 111–122.
- Pepys, J., Mitchell, J., Hawkins, R., and Malo, J. L. (1985). A longitudinal study of possible allergy to enzyme detergents. *Clin. Allergy* **15**, 101–115.
- Pepys, J., Wells, I. D., D'Souza, M. F., and Greenberg, M. (1973). Clinical and immunological responses to enzymes of *Bacillus subtilis* in factory workers and consumers. *Clin. Allergy* **3**, 143–160.
- PRISMS Study Group. (2001). PRISMS-4: Long-term efficacy of interferon-beta-1a in relapsing MS. *Neurology* **56**, 1628–1636.
- Raju, R., Marietta, E., Vinasco, J., Conti-Fine, B. M., Infante, A. J., and David, C. S. (2002). Cryptic determinants and promiscuous sequences on human acetylcholine receptor: HLA-dependent dichotomy in T cell function. *Hum. Immunol.* **63**, 237–247.
- Robinson, M. K., Horn, P. A., Kawabata, T. T., Babcock, L. S., Fletcher, E. R., and Sarlo, K. (1998). Use of the mouse intranasal test (MINT) to determine the allergenic potency of detergent enzymes: Comparison to the guinea pig intratracheal (GPIT) test. *Toxicol. Sci.* **43**, 39–46.
- Sarlo, K., Adamson, G. M., and Stachlewitz, R. (2003). Evidence of allergic antibody to a *Bacillus* serine protease in atopic women: A retrospective study. *Toxicol. Sci.* **72**, 229.
- Sarlo, K., Fletcher, R., Gaines, W., and Ritz, H. L. (1997). Respiratory allergenicity of detergent enzymes in the guinea pig intratracheal test: Association with sensitization of occupationally exposed individuals. *Fundam. Appl. Toxicol.* **39**, 44–52.
- Sarlo, K., and Kirchner, D. B. (2002). Occupational asthma and allergy in the detergent industry: New developments. *Curr. Opin. Allergy Clin. Immunol.* **2**, 97–101.
- Sarlo, K., Parris, J. S., Clark, E. D., Horn, P. A., Robinson, M. K., McCay, J. A., Peachee, V. L., Veloso, Y. L., and White, K. L. Jr. (2000). Influence of MHC background on the antibody response to detergent enzymes in the mouse intranasal test [in process citation]. *Toxicol. Sci.* **58**, 299–305.
- Scagnolari, C., Bellomi, F., Turriziani, O., Bagnato, F., Tomassini, V., Lavolpe, V., Ruggieri, M., Bruschi, F., Meucci, G., Dicuonzo, G., et al. (2002). Neutralizing and binding antibodies to IFN- β : Relative frequency in relapsing-remitting multiple sclerosis patients treated with different IFN- β preparations. *J. Interferon Cytokine Res.* **22**, 207–213.
- Schwaiger, F. W., Weyers, E., Epplen, C., Brun, J., Ruff, G., Crawford, A., and Epplen, J. T. (1993). The paradox of MHC-DRB exon/intron evolution: α -helix and β -sheet encoding regions diverge while hypervariable intronic simple repeats coevolve with β -sheet codons. *J. Mol. Evol.* **37**, 260–272.
- Schweigert, M. K., Mackenzie, D. P., and Sarlo, K. (2000). Occupational asthma and allergy associated with the use of enzymes in the detergent industry: A review of the epidemiology, toxicology, and methods of prevention. *Clin. Exp. Allergy* **30**, 1511–1518.
- Sicherer, S. H., and Sampson, H. A. (2000). Peanut and tree nut allergy. *Curr. Opin. Pediatr.* **12**, 567–573.
- Sonderstrup, G., Cope, A. P., Patel, S., Congia, M., Hain, N., Hall, F. C., Parry, S. L., Fugger, L. H., Michie, S., and McDevitt, H. O. (1999). HLA class II transgenic mice: Models of the human CD4+ T cell immune response. *Immunol. Rev.* **172**, 335–343.
- Stickler, M., Chin, R., Faravashi, N., Gebel, W., Razo, O. J., Rochanayon, N., Power, S., Valdes, A. M., Holmes, S., and Harding, F. A. (2003a). Human population-based identification of CD4(+) T cell peptide epitope determinants. *J. Immunol. Methods* **281**, 95–108.
- Stickler, M., Mucha, J., Estell, D., Power, S., and Harding, F. (2003b). A human dendritic cell-based method to identify CD4(+) T cell

- epitopes in potential protein allergens. *Environ. Health Perspect.* **111**, 251–254.
- Stickler, M. M., Estell, D. A., and Harding, F. A. (2000). CD4+ T cell epitope determination using unexposed human donor peripheral blood mononuclear cells. *J. Immunother.* **23**, 654–660.
- Vanhanen, M., Tuomi, T., Nordman, H., Tupasela, O., Holmberg, P. C., Miettinen, M., Mutanen, P., and Leisola, M. (1997). Sensitization to industrial enzymes in enzyme research and production. *Scand. J. Work Environ. Health* **23**, 385–391.
- Vanhanen, M., Tuomi, T., Tiikkainen, U., Tupasela, O., Tuomainen, A., Luukkonen, R., and Nordman, H. (2001). Sensitization to enzymes in the animal feed industry. *Occup. Environ. Med.* **58**, 119–123.
- Vanhanen, M., Tuomi, T., Tiikkainen, U., Tupasela, O., Voutilainen, R., and Nordman, H. (2000). Risk of enzyme allergy in the detergent industry. *Occup. Environ. Med.* **57**, 121–125.
- Zhou, L.-J., and Tedder, T. F. (1996). CD14+ blood monocytes can differentiate into functionally mature CD83+ dendritic cells. *Proc. Natl. Acad. Sci. USA* **93**, 2588–2592.