Protein Engineering, Design & Selection vol. 24 no. 11 pp. 845–853, 2011
Published online September 9, 2011 doi:10.1093/protein/gzr046

Exploiting bias in a non-immune human antibody library to predict antigenicity

Margaret E.Ackerman[1,2,6], Jennifer I.Lai[3], Ira Pastan[5] and K. Dane Wittrup[3,4]

[1]Thayer School of Engineering, Dartmouth College, Hanover, NH 03766, USA, [2]Departments of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [3]Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA, [4]Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA and [5]Laboratory of Molecular Biology, National Cancer Institute, National Institutes of Health, Bethesda, MD 20892, USA

[6]To whom correspondence should be addressed.
E-mail: Margaret.E.Ackerman@Dartmouth.edu

Received May 26, 2011; revised August 9, 2011; accepted August 14, 2011

Edited by Dennis Burton

Non-immune human antibody fragment libraries have generated antigen-binding proteins useful as prospective research, imaging, diagnostic and therapeutic agents. However, because the generation of such libraries relies on cloning antibody sequences from the circulating immune repertoire rather than truly naïve, germline sequences, their composition may reflect the deletion of autoreactive sequences, making them less suited for isolating binding clones to human antigens, but perhaps useful in applications where an in vitro handle on representative circulating antibody diversity is desired. Here we demonstrate that a large non-immune human scFv library is relatively depleted of sequences capable of recognizing human antigens as compared with orthologs antigens. Additionally, because this non-naïve, non-immune library may capture a representative section of antibody diversity, we explore its possible utility in conducting early pre-screens to predict the antigenicity of prospective therapeutics and find a correlation between the clinical immunogenicity of a small panel of protein therapeutics with their propensity for interacting with the library.

Keywords: antibody/antibody fragment/antigenicity/immunogenicity/protein library/protein selection/immune library/scFv

Introduction

Immune repertoire libraries have long been used to isolate antibody fragments capable of recognizing antigens useful in diagnosis and therapy, and have more recently been utilized to characterize the antibody response to natural infection with great success. Non-immune libraries, generated from the naturally present diversity of the circulating B cell repertoire, differ from naïve libraries in which germline sequences are present, in that they reflect the repertoire biases imparted during B cell development. One of these biases is selection against self-binding sequences, and functions as a control against autoimmunity. However, the targets of many protein engineering efforts are human antigens, and while large non-immune libraries have been constructed in almost every protein display format (Sheets *et al.*, 1998; de Haard *et al.*, 1999; Feldhaus *et al.*, 2003), it has been observed that they are perhaps not the most conducive to successful isolation of high-affinity binding reagents against human antigens, and the field has consequently moved toward the use of completely synthetic germline libraries (Hanes *et al.*, 2000).

Here we first test whether a bias against recognition of self antigens can be observed in a non-immune library, and then posit that the influence of positive and negative selection factors in B cell development, may make non-immune, but non-naïve, libraries a useful platform for study of the initial generation of the antibody response. Accordingly, we investigate whether a non-immune library has the capacity to be used to study the antigenicity or immunogenicity of therapeutic proteins prospectively. The use of immune libraries to study the adaptive immune response retrospectively, even allowing germline lineage antibody diversification studies to be performed is well established (Hicar *et al.*, 2010; Bonsignori *et al.*, 2011; Scheid *et al.*, 2011).

Biologics that induce an adaptive immune response can cause significant toxicities, including inflammation, anaphylaxis, complement-mediated cell lysis and result in altered pharmacokinetics leading to both limited efficacy and toxicity (Schellekens, 2002; Subramanyam, 2006). While 'foreign-ness' is a primary factor in immunogenicity (van den Berg and Rand, 2004; Kanduc, 2008), the numerous cases in which even completely human proteins have led to immune stimulation (Peces *et al.*, 1996; Zang *et al.*, 2000) indicate that immunogenicity is a complex phenomenon with many contributing factors. With the introduction of new therapies consisting of novel recombinant and completely non-human proteins, immunogenicity, i.e. the tendency of a protein drug to be recognized as foreign by the immune system, has become a significant concern in the drug development, approval and monitoring process (Shankar *et al.*, 2006; Ponce *et al.*, 2009). Accordingly, strategies to reduce immunogenicity have been pursued with great interest.

As both arms of the adaptive immune system function in determining the immunogenicity of a therapeutic protein, significant effort has been devoted to developing tools for evaluating both B and T cell immunogenicity. Although there is sometimes significant overlap between the epitopes recognized by both T and B cells, in most cases the reactive epitopes are completely distinct (Brons *et al.*, 1996; Rosenberg and Atassi, 1997), and must therefore be addressed independently. Similarly, while T cells contribute significantly to the production of a strong B cell response, anti-drug antibodies are primarily responsible for the altered pharmacokinetics and complications resulting from administering an immunogenic drug (Ponce *et al.*, 2009). Therefore, both classes of epitopes are significant and the most complete means to assess clinical immunogenicity would involve the study of both.

Despite having very different mechanisms of action, both T and B cells have a functional diversity that is generated by high levels of genetic recombination, and tempered by deletion of variants that are self-reactive or fail certain checks of quality control (Goodnow, 1996; Nemazee, 2000; Ferry *et al.*, 2006). Several aspects of T cell biology have made assessments of T cell immunogenicity a tractable problem for both in silico and *in vitro* prediction. First, T cell epitopes are linear. Second, major histocompatibility complex (MHC) structures are known, and computational tools exist for evaluating the fit of various peptide fragments in silico (Brusic *et al.*, 2004). Finally, the pool of MHCs has a diversity limited enough to be sufficiently covered for testing *in vitro*. Together, these factors have allowed the development of in vitro T cell stimulation assays that are predictive of clinical immunogenicity (Walden, 1996; Stickler *et al.*, 2000; Warmerdam *et al.*, 2002).

In contrast, satisfactory methods for comparable in vitro determination of B cell epitopes are lacking, largely due to the significant technical difficulties associated with the prevalence of conformational, i.e. non-linear, discontinuous epitopes, the vast diversity of antibody sequences, and their initially low-affinity interactions with the target antigen. Despite more recent results with linear epitopes (Kanduc, 2009), these problems impede study by either experimental or computational methods.

The combined effect of these hurdles has been to effectively limit B cell epitope determination to *in vivo* study. Current methods entail generating immune libraries by injecting the biologic, waiting for an anti-drug immune response and then harvesting drug-reactive antibodies, or synthesizing antibody libraries, mapping the epitopes they recognize and then rationally altering these specific epitopes (Keil and Wagner, 1989; Ritter *et al.*, 2001; Spencer *et al.*, 2002; Mayer *et al.*, 2004; Onda *et al.*, 2006). This strategy has frequently been performed in animal models, under the supposition that the same epitopes will be immunogenic in multiple species, but has also been performed either with human samples (Laroche *et al.*, 2000; Gustafsson *et al.*, 2009) or mice with humanized immune systems (Wierda, 2001). While these studies have been quite useful when comparing variants of a particular protein (Keil and Wagner, 1989, Zwickl *et al.*, 1991), their limitations are fairly obvious in that they are conducted at a late stage and are time and resource intensive. Better tools, particularly those that can be implemented early in development are clearly needed.

Here, we investigate whether a simple *in vitro* assay utilizing protein engineering tools could be used to score the clinical immunogenicity of protein therapeutics prospectively, in a manner similar to the well-established use of molecular display techniques in studying immune responses retrospectively. We propose that by utilizing a large non-immune antibody sequence repertoire, we can capture a sequence-function landscape reflective of the positive and negative selection factors in B cell development and the initial immune repertoire available for the generation of an antibody response to a therapeutic protein. We hypothesize that by panning a non-immune library against a prospective therapeutic in a system capable of capturing even low-affinity interactions via avidity, that we may be able to provide a means to assess immunogenicity prior to clinical study.

Importantly, while T cell epitopes in proteins provide strong predictive power with respect to immunogenicity, no satisfactory in vitro assay for B cell, antibody-mediated immunogenicity has been previously described. The success of immune libraries in developing useful antibodies and in studying the development and history of the immune response to natural infection demonstrates that they are a powerful tool and indicates that it is possible that they may provide insight into evaluating the B cell immunogenicity of candidate protein drugs.

Materials and methods

Protein preparation

Nine species of albumin, and two pooled polyclonal immunoglobulin (Ig)G samples were purchased in high purity from Sigma. Carcinoembryonic antigen (CEA)-binding scFv bearing His_6 tags were cloned into pCT-based vectors, produced in YVH10 yeast as described (Graff *et al.*, 2004), and purified using affinity chromatography (Clontech Talon resin). A33 antibodies were a gift of Gerd Ritter at Memorial Sloan Kettering and the Ludwig Institute. The chimeric EGFR antibody Ab806 was a gift of Jamie Spangler, Massachusetts Institute of Technology. Immunotoxins were produced as described (Kreitman *et al.*, 2000). Briefly, HA22-based immunotoxins were produced in BL21 *Escherichia coli* in a pEM15-based plasmid, refolded from inclusion bodies, and purified utilizing ion exchange and size exclusion chromatography.

Proteins were biotinylated essentially according to the manufacturer's instructions (Pierce sulfo-NHS-LC-biotin), except for tuning the relative molar amounts of reagent: protein to minimize excessive biotinylation. Typically, this required decreasing the concentration of biotinylation reagent in half-log steps. Free biotin was removed by extensive serial buffer exchange using Amicon spin columns of appropriate molecular weight cutoffs. Successful biotinylation was confirmed by western blot, and quantified using a biotin quantification kit (Pierce). Target biotinylation levels were between one and two biotin groups per molecule.

Yeast handling

The scFv library was grown and induced as described (Feldhaus *et al*. 2003; Chao *et al.*, 2006). Briefly, yeast were thawed and subcultured twice before induction at 20°C for 24 h. Induction quality was assessed by labeling for the C-terminal c-myc tag expression utilizing an anti-c-myc antibody and secondary (Invitrogen A21281, A11309) and analysis on a Coulter EPICS flow cytometer. Populations were always maintained at least at a 10-fold excess of their theoretical size in order to minimize loss of diversity.

Bead coating

Streptavidin-coated magnetic beads (Invitrogen 11047) were coated with protein as follows. Approximately 1×10^7 beads (25 μl) were washed twice with phosphate-buffered saline (PBS) with 0.1% lysozyme or albumin carrier protein, then incubated with 50 μl of biotinylated target protein at a concentration of 100 μg/ml, diluted with 500 μl PBS and rotated overnight at 4°C. Complete coating was confirmed by incubating beads with differing amounts of target protein, then

labeling beads with streptavidin-PE and analyzing on a flow cytometer, as described (Ackerman *et al.*, 2009). Following coating, beads were washed twice in PBS to remove excess protein.

Assay

For each protein panel, 4×10^9 yeast cells were spun down, and in order to deplete the population of streptavidin binding scFvs, were then incubated with 100 μl streptavidin beads at 4°C for 1 h. The yeast:bead slurry was then placed on a magnet for 5 min, and unbound yeast were removed using a pipette. The negatively selected yeast were then incubated with pooled beads, each coated with a single protein of interest from the panel being studied. These incubations were carried out for 1–3 h at 4°C on a rotator. The yeast:bead slurry was again placed on a magnet for 5 min, and unbound yeast were removed using a pipette. The remaining beads and yeast were resuspended in 1 ml of PBS, added to 50 ml of growth media for regrowth, and then were subsequently induced. This process was repeated a second time. Following the second pooled bead selection, truncation mutants were removed by labeling for c-myc expression and collecting all full-length (c-myc positive) clones on a FACSAria sorter. The sizes of all populations were determined by plating.

Pooled selections utilizing panels of proteins of interest were performed due to the need to regrow and induce cultures between steps. Display level can vary dramatically between cultures and strongly impacts selection, and so in order to evaluate colony counts fairly, all growth and induction steps were performed in pooled batches so as to keep induction conditions and display levels constant. Pooled selections furthermore reduce possible differences due to handling in early steps.

Following these initial pooled selections, yeast were induced and for each protein in the panel tested, 10^7 yeast, representing at least a 10-fold oversampling as determined by the plating of an aliquot of each population, were negatively selected as described above, then incubated with 10–25 μl of beads coated with each antigen individually for 1–3 h at 4 °C on a rotator. The slurry was placed on a magnet for 5 min, after which unbound yeast were removed and discarded. The remaining yeast:bead mixture was resuspended in 1 ml of PBS, and a 50 μl aliquot designated Wash 1 was removed for serial dilution and plating to determine the size of the initially bound population. The resuspended yeast and beads were then incubated at 4°C for 15 min on a rotator, and then placed on the magnet a second time. Unbound yeast were removed, bound yeast and beads were resuspended and an aliquot-designated Wash 2 was removed for plating. This process was then repeated a third time, and the final bead-bound population (Wash 3) was retained for further analysis.

Plated dilutions were incubated for 2 days at 30°C, and then colonies were counted. Generally, two separate dilutions had countable populations (between 5 and 200 colonies), and these counts were averaged.

Data analysis

Data for individual selections were collected in triplicate from three separate inductions of the populations isolated in pooled selection rounds. As total colony counts are greatly dependent on the expression level of the yeast (Ackerman *et al.*, 2009), and expression level can vary significantly from one induction to another, counts between replicates were not compared directly, and it was necessary to standardize the replicates to account for differences in display. Thus, the average number of binders retrieved after each wash was calculated, and a ratio of individual counts to this average was calculated. This ratio to the mean was determined for each replicate, prior to averaging the triplicate data sets. This method of analysis prevents any one replicate from dominating the analysis due to higher expression, and resulting higher colony counts. Similarly, this analysis around the mean further allowed the tendency of single high or low counts within a replicate to greatly skew the entire data set to be minimized.

Follow-up experiments and sequencing

For some data sets, the Wash 3 yeast populations isolated against specific proteins were grown, induced and subjected to follow-up study. For the albumin panel, individual clones from the Wash 3 populations were subsequently bead-selected against their specific target antigen and binding counts were determined as described previously (Ackerman *et al.*, 2009). The specificity of the Wash 3 population was further analyzed by flow cytometry. Target protein was multiplexed by pre-incubating with streptavidin-PE, incubated with the Wash 3 yeast population, pelleted, washed and then analyzed by flow cytometry to observe specific binding. The percent of yeast binding to target antigen specifically was calculated by determining the percentage of yeast with PE signal above background (secondary alone).

Plasmids were collected from selected yeast populations by zymoprep (Zymo Research D2004), and then transformed into XL-1 blue cells (Stratagene 200130). Colonies were picked, grown for 10 h in LB Amp with 10% glycerol, and submitted to Agencourt Biosciences for sequencing using the atermmid primer (Chao *et al.*, 2006).

Results

Assay tools and process

The method described here combines two protein engineering tools: yeast surface display of a non-immune human scFv library (Feldhaus *et al.*, 2003), and highly avid magnetic beads that are capable of isolating even very weak specific interactions (Ackerman *et al.*, 2009), to investigate the intrinsic reactivity of a given protein with a non-immune antibody repertoire (Fig. 1). The yeast display platform, in which proteins of interest are secreted as Aga2p fusions and captured on the surface of yeast through covalent bonding to the Aga1p mating receptor, has proven a robust technology for the interrogation of biological libraries (Gai and Wittrup, 2007), and is commonly used for the isolation and engineering of antibody fragments.

The non-immune human scFv library used was generated by cloning heavy and light chain variable domains from B cells in the lymph nodes and spleens of 58 adults, resulting in a library of 10^9 sequences (Feldhaus *et al.*, 2003). As many of these B cells have undergone negative selection against self-proteins, we anticipated that the library would be depleted for variable domains reactive toward human proteins. Furthermore, we hypothesized that the diversity of scFv sequences in this library may capture a useful statistical

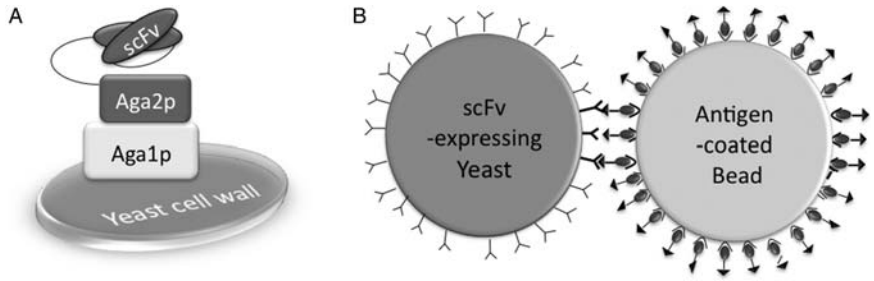

Fig. 1. Protein engineering tools used in the immunogenicity assay (**A**) Yeast display platform, in which a library of human antibody fragments (scFv) are fused to the Aga2p mating receptor and captured on the yeast cell wall through a covalent interaction with the Aga1p receptor. A c-myc tag is incorporated at the C-terminus of the displayed scFv, providing a handle to isolate full-length clones. (**B**) Highly avid magnetic bead-based selection of yeast expressing scFvs that recognize the target protein via multivalent interactions.

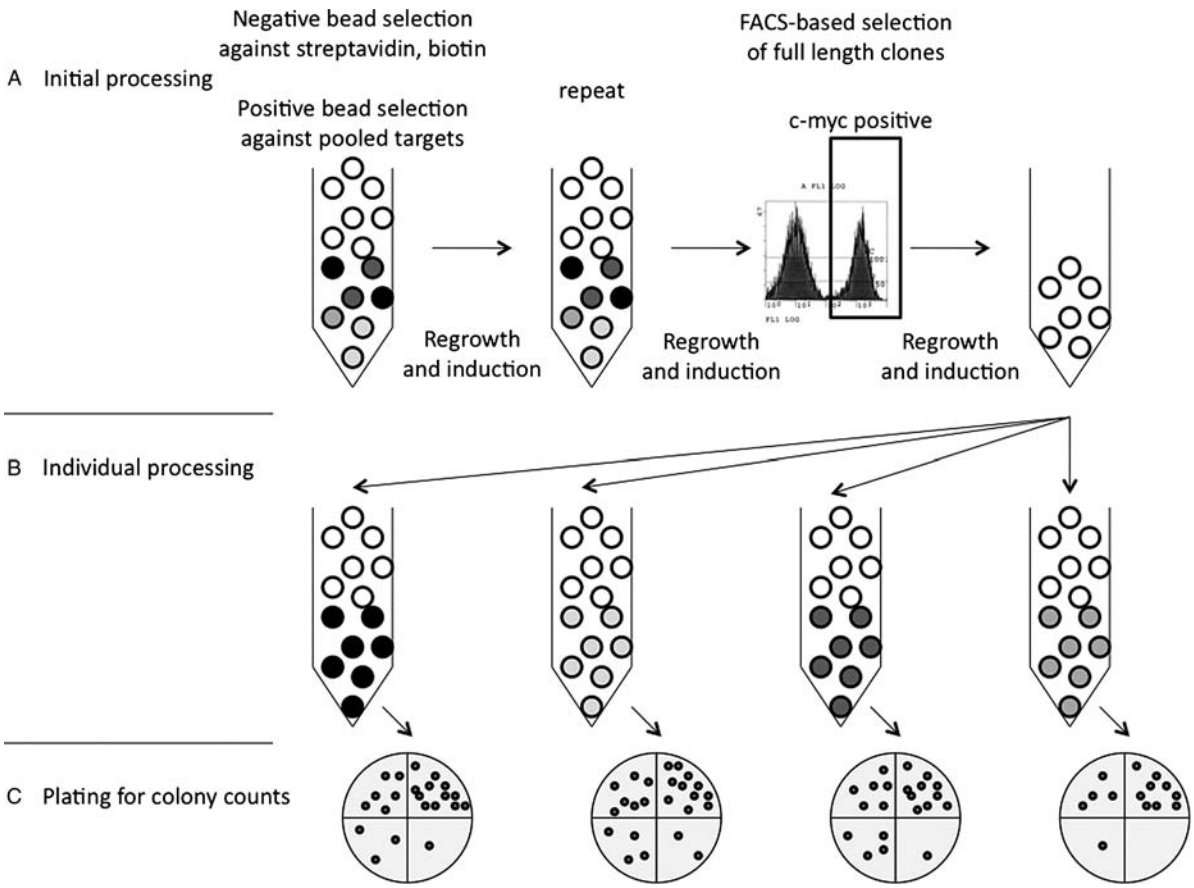

Fig. 2. Assay process (**A**) Initial processing: the scFv library is first depleted of reagent binders by removing yeast that bind magnetic beads, then enriched for yeast which bind the panel of proteins of interest by isolating yeast which bind to pooled, protein-coated beads. This process is repeated, and then full-length clones are isolated by FACS. (**B**) In triplicate, the resulting yeast population is then depleted of reagent binders again, and then divided into aliquots that are incubated with beads coated with each protein of interest individually. (**C**) The bound yeast and bead slurry is washed three times, and each wash population is serially diluted and then plated to determine the number of yeast expressing scFvs that interact with each protein.

sample of the diversity of the antibody sequence-function landscape *in vivo*. Therefore, we sought to determine whether this library could be used to predict the relative initial B cell mediated immune response, and thereby theoretically allow the identification of antigenic or immunogenic proteins *in vitro*. To make this assessment, yeast capable of binding to the protein-coated beads were selected, isolated and counted. We hypothesized that this count provides a means to assess the initial pre-immune reactivity of a human antibody repertoire to a given protein.

The assay is schematized in Fig. 2. Briefly, the assay consists of two phases, bulk enrichment of scFv clones which recognize any protein in the panel being compared, followed by quantitative assessment of the number of scFv clones in the enriched population which bind to each protein in the panel.

Assay quality control

To avoid enriching scFv clones which recognize the reagents utilized in the assay rather than the proteins being evaluated,

negative selections were performed against uncoated and biotin-conjugated magnetic beads. These negative selections represent a critical step as the selections can otherwise isolate clones that recognize any available binding epitope, including these surfaces which may be presented unintentionally.

As truncated scFvs likely have exposed hydrophobic regions, their residual presence in the scFv library could potentially skew the binding counts via non-specific interactions. Accordingly, a single FACS sort gated on the c-myc tag located at the C-terminus of the displayed scFv was conducted to isolate full-length clones. Despite possible enrichment via non-specific mechanisms, the prevalence of truncated clones was not found to increase significantly in two rounds of pooled bead selections, and remained at roughly the same level as was initially present in the library (Fig. 3A). Nonetheless, the c-myc selection yielded a nearly pure population of full-length clones for subsequent analysis.

In the quantitative, single-protein selections, beads and bound yeast were resuspended and washed one to three times prior to plating, yielding three separate colony counts. In general, the count of the first wash (Wash 1) was high compared with subsequent counts (Fig. 3B), likely due to the high density of yeast cells present in the incubation mixture prior to washing (100 million/ml), which may have resulted in passive trapping of non-binding yeast by migrating beads. After the beads were resuspended once, density was significantly decreased, along with the probability of grossly inflating colony counts via passive trapping. Alternatively, it is possible that these initially high counts may indicate the presence of weakly interacting clones. Figure 3B presents colony counts for populations recognizing both a protein with no human ortholog, streptavidin, as well as human and bovine albumin, and human and murine IgG, and demonstrates a trend of decreased reactivity toward these two human proteins relative to their orthologs from other species. For this test, binders to streptavidin were not pre-depleted from the library, allowing comparison of the prevalence of scFvs recognizing this bacterial component versus fully human blood proteins.

Assay results for albumin and IgG

Next, in order to investigate whether decreased reactivity toward human antigens was a reproducible quality of the non-immune library, pilot experiments with albumin and IgG were conducted in triplicate, yielding the colony count ratios presented in Fig. 4A, in which human and bovine albumin, and human and murine IgG were compared for their reactivity with the non-immune library. For all graphs, the count for the human protein variant was assigned a score of one, and the ratio of binding to other targets was calculated. The average and standard deviation of three replicates are presented.

We then determined assay reproducibility by performing two separate replicates starting from the initial library and followed through final plate counts independently for both bovine and human albumin, and murine and human IgG (Fig. 4B). With good agreement across these independent assays, bovine albumin scored as ~3 times more reactive to the scFv library as human, and murine IgG scored as ~6 times more reactive than human IgG.

Extended panel of human and orthologs proteins

As a panel of orthologous proteins provides a more rigorous test than pairwise comparisons, we next tested albumin from eight different species as well as human albumin produced recombinantly in yeast; results are presented in Fig. 5A. As would be expected given the selection against recognition of

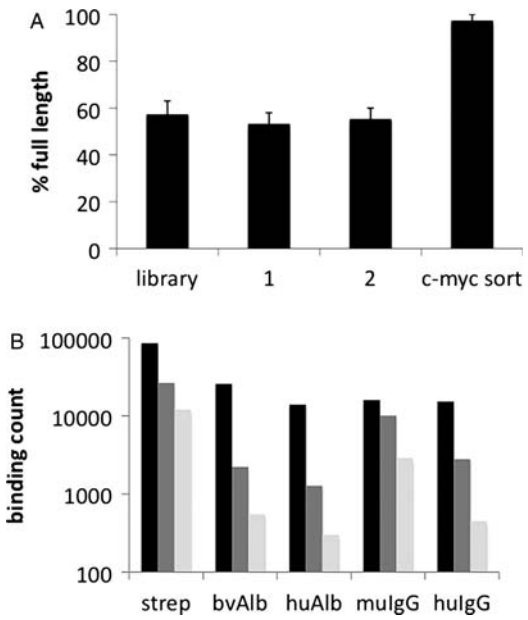

Fig. 3. Assay quality control (**A**) Population composition: the naïve library, populations singly and doubly selected against pooled beads, and following the FACS sort were labeled for c-myc expression, demonstrating the removal of truncation mutants following FACS. (**B**) Effect of washes on binding counts in a single replicate: Wash 1, 2 and 3 binding counts against five targets of interest, demonstrating the trends observed for binding count over the course of three washes.

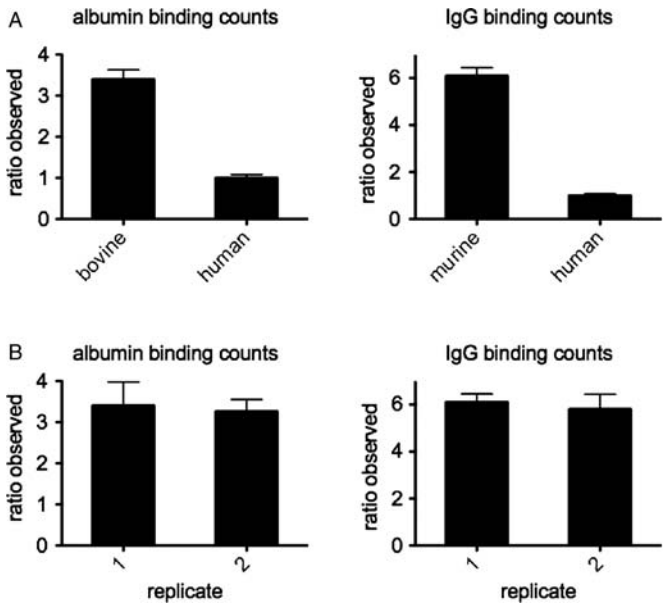

Fig. 4. Human ortholog point comparisons (**A**) Ratio of binding clones from the yeast displayed scFv library for bovine and human albumin, and murine and human IgG (**B**) Assay reproducibility: ratio of yeast binding to bovine:human albumin and murine:human IgG in two separate assays starting from the initial library and followed through triplicate final plate counts independently.

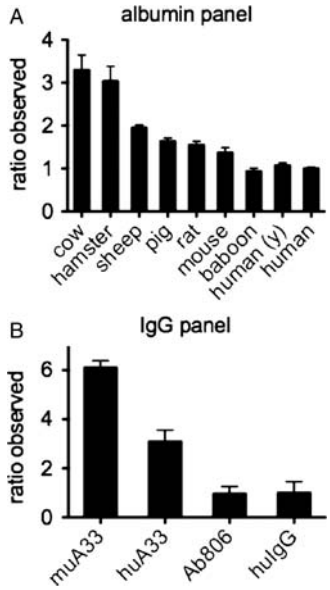

Fig. 5. Albumin and IgG panels (**A**) Ratio observed of yeast binding to each species of albumin relative to binding to human albumin. (**B**) Ratio observed of yeast binding to muA33, huA33, Ab806 and huIgG to huIgG. Rank of reactivity agrees with clinical data.

self-antigens in the immune repertoires used in its construction, the non-immune library reacted to the lowest extent against human albumin, yeast-produced recombinant human albumin and baboon albumin, predicted to have a high degree of sequence homology. All other species ranked as more immunogenic in this assay. However, these scores did not quantitatively agree with their phylogenetic similarity to human albumin. Among other possibilities, this may reflect differences in the surface-accessible fraction of non-conserved amino acids, as well as the type of residues substituted, such as bulky or aromatic side chains, which are widely considered to be more immunogenic. Additionally, because there is not a clinical data set regarding the immunogenicity of these albumin molecules, these scores are of limited value as a correlate beyond the finding that the human antigens were recognized to the lowest extent.

Next, to assess the relevance of colony counts as a possible correlate to clinical immunogenicity, a panel of antibodies was tested, including a human polyclonal sample and three antibodies that have been used in the clinic: muA33 and huA33, murine and humanized antibodies to the A33 antigen, a cell surface protein expressed in the colon; and Ab806, a chimeric antibody that recognizes the epidermal growth factor receptor.(Mishima *et al.*, 2001) Murine anti-A33 (muA33) was found to be highly immunogenic in patients (Welt *et al.*, 1996), leading to its humanization. Despite humanization, the resulting variant, huA33, retained a considerable degree of immunogenicity, and ~60% of patients treated with huA33 developed an immune response characterized by significant titers of anti-huA33 antibodies after 1–2 injections (Ritter *et al.*, 2001; Welt *et al.*, 2003). In contrast, Ab806, which, despite having a limited clinical history, has been safer (one of eight patients generated anti-drug antibodies), scored similarly to bulk human IgG. Figure 5B presents the relative number of yeast binding to muA33, huA33, Ab806 and huIgG, with notable agreement with clinical immunogenicity profiles.

Fine discrimination between protein variants

As the antibodies and albumin variants tested are rather dissimilar, we next looked at whether the non-immune library was able to capture differences between proteins with only a few amino acid point mutations. Reactivity differences between proteins with relatively few amino acid substitutions may imply an ability to precisely map reactive epitopes, and offer a means to identify and remove such sites in advance of clinical testing. A panel of CEA binding scFvs were produced, including a murine scFv (MFE) (Chester *et al.*, 1994), humanized MFE (hMFE) (Boehm *et al.*, 2000), a stabilized version of the humanized construct (shMFE) (Graff *et al.*, 2004) and a high-affinity mutant (sm3e) (Graff *et al.*, 2004). These scFvs diverge by 4 (shMFE), 6 (sm3e) and 28 (MFE) amino acids from the humanized scFv, hMFE. MFE, the murine scFv, scored highest in the immunogenicity assay, and hMFE, the humanized scFv scored lowest (Fig. 6A). The stability- and affinity-modified variants, shMFE and sm3e scored slightly higher, in accordance with the addition of mutations moving these sequences away from the humanized sequence.

Finally, a panel of immunotoxins, consisting of an anti-CD22 scFv fused to pseudomonas exotoxin A, HA22,(Salvatore *et al.*, 2002) and two variants with reduced immunogenicity: HA22-8x, (Onda *et al.*, 2006, 2008) which diverges by eight amino acids; and HA22-LR-8x,(Weldon *et al.*, 2009) in which the translocation domain of the toxin is also deleted, were tested. In mouse models, HA22-8x and HA22-LR-8x have significantly reduced immunogenicity

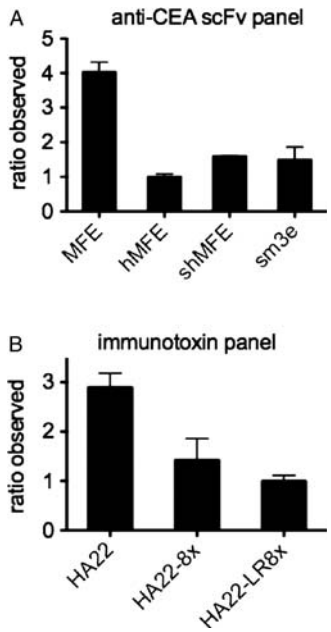

Fig. 6. Fine discrimination between sequence variants (**A**) Ratio observed of yeast binding to each scFv relative to hMFE in a CEA-scFv panel: four anti-CEA scFvs were screened, a murine scFv (MFE, 28 amino acids diverged from hMFE), humanized scFv (hMFE), stabilized humanized scFv (shMFE, four amino acids diverged from hMFE), and a stabilized, humanzed, high-affinity mutant (sm3e, six amino acids diverged from hMFE). (**B**) Immunotoxin panel: HA22 anti-CD22 scFv-pseudomonas exotoxin A fusion (HA22) and reduced immunogenicity variants HA22-8x (eight amino acids diverged from wild-type immunotoxin), and HA22-LR-8x (eight amino acids diverged and translocation domain deleted). Data are presented as the ratio of yeast binding to each immunotoxin relative to HA22-LR-8x.

relative to HA22, (Onda *et al*., 2008) (Hansen *et al*., 2010) and HA22-8x has been demonstrated to have decreased reactivity with immune sera from HA22-treated patients, suggesting the successful removal of immunogenic epitopes from this variant (Onda *et al*., 2006). In agreement with these studies, our assay reported that the native HA22 immunotoxin was twice as reactive toward human scFvs as HA22-8x, and three times as reactive as HA22-LR-8x (Fig. 6B).

Diversity and specificity of selected clones

To probe the validity of these colony count-based scores, clones isolated from the final wash of albumin selections were sequenced in order to assess the diversity of this population. No duplicate sequences were found when up to 20 clones isolated against a single target were sequenced, indicating that the assay protocol is not so stringent as to eliminate diversity. Furthermore, when 100 clones from the albumin selections (10–20 per albumin variant) were sequenced, no duplicated sequences were found. Importantly, however, while we expect our colony count-based scores may correlate with the number of immunogenic sites on a given therapeutic, it is unlikely that each binding clone represents a different epitope, as a polyclonal response can be generated against a single immunogenic epitope (Onda *et al*., 2006). Rather, each binding clone may represent a possible starting point for affinity maturation and the generation of an adaptive immune response.

The specificity of the populations isolated was then evaluated. First, for several of the albumin samples, FACS-based analysis was conducted on the Wash 3 populations for binding to biotinylated antigen detected via streptavidin-PE. Results of this analysis are presented in Fig. 7A, and demonstrated that the majority of each population (>65%) demonstrated antigen-specific binding of a high-enough affinity to be detected by FACS, whereas the naïve library demonstrated no detectable binding (<1%).

To expand on this population-based analysis, and to capture binding that may have been too weak to be detected by flow cytometry, we further analyzed the specificity of individual clones from the final wash by determining whether or not they could be enriched from a larger non-binding population via specific interactions with their presumed target antigen. The Wash 3 clones were mixed with non-binding yeast at a ratio of 1:1000, and then selected against antigen-coated beads. All clones tested demonstrated specific binding to protein-coated beads as determined by their enrichment relative to their prevalence before selection, while a negative control clone demonstrated no enrichment (Fig. 7B). Together, these results indicate that selectivity is achieved while maintaining diversity, and points toward the validity of colony count-based scores.

Collectively, these results indicate that non-immune libraries may reflect negative selection against self-antigens. This finding suggests that they may capture a useful fraction of the sequence-function landscape and therefore provide a useful means of evaluating the antigencity of therapeutic proteins. The *in vitro* assay presented here involving a simple count of binders from a non-immune library against a protein of interest consistently scored human proteins as having lower reactivity compared with their orthologs from other species, agreement with clinical immunogenicity for a limited panel of antibodies used as therapeutics, differentiation between protein variants that diverge by as few as two to eight amino acids, and finally scored the reactivity of an immunotoxin in good agreement with mouse and human clinical data. Accordingly, an *in vitro* test based on this methodology may provide a simple and rapid means to prescreen therapeutic proteins for B cell immunogenicity.

Discussion

Koshland wrote that ‘of all mysteries of modern science, the mechanism of self versus non-self recognition in the immune system ranks at or near the top’ (Koshland, 1990). Accordingly, significant efforts have been put into the development of *in vitro* and in silico immunogenicity prediction techniques, the success of which has depended strongly on the ability to circumvent technical limitations in adequately capturing these basic functions of the immune system.

As T cell epitopes are necessary for stimulating the B cell response, their removal can decrease immunogenicity in proteins even where B cell epitopes remain intact. The inverse case, i.e. a decrease in immunogenicity due to removal of B cell epitopes independent of T cell epitopes, has also been demonstrated and reviewed recently(Nagata and Pastan, 2009). As the anti-drug antibodies formed are responsible for the toxicity and altered pharmacokinetics of an immunogenic drug, B cell epitopes merit independent study. Additionally, although foreign proteins have a large potential number of B cell epitopes, experimental evidence suggests the presence of immunogenic hotspots that generate polyclonal antibody responses. A technique capable of reliably identifying these sites is likely to provide a good proxy for proteins, such as the immunotoxin evaluated here, which cannot intuitively or

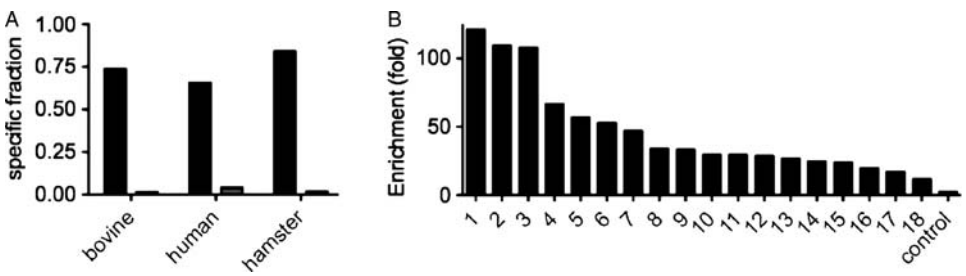

Fig. 7. Antigen specificity of Wash 3 populations and isolated clones (**A**) Percent of Wash 3 population (black bars) and naïve library (gray bars) binding to antigen as determined by flow cytometry. (**B**) Clones from the Wash 3 population were mixed with non-binding yeast and selected against antigen-coated beads. Enrichment relative to the initial prevalance of each clone was determined, demonstrating specific binding of all clones tested, while a lysozyme-binding control clone demonstrated no enrichment.

computationally be more comprehensively humanized early in development.

Despite the numerous factors responsible for the clinical immunogenicity profile of a protein drug, the assay reported here, which relies on the simple premise of counting interactions between a human antibody fragment library and candidate drugs, may effectively measure B cell antigenicity and provide a convenient *in vitro* means to assess the safety profile of protein therapeutics. We have demonstrated that this assay scores human proteins as less reactive than their orthologs, agreement with clinical immunogenicity for a limited panel of antibodies that included a humanized antibody with anomalously high clinical immunogenicity, and that the assay described here has the ability to make fine discriminations between protein variants that have relatively few amino acid substitutions. The latter result indicates that not only may this method capture global immunogenicity differences, but might be used to precisely identify specific immunogenic epitopes. Furthermore, the use of highly avid magnetic beads eliminates the need for affinity maturation and allows rapid screening. However, further validation including testing a larger panel of therapeutic antibodies and other approved protein biologics is merited. Likewise, test sets of proteins with known stabilities or degrees of humanness could be used to support the use of this method in evaluating immunogenicity.

The non-immune library used in these studies is a key variable, as it must adequately capture *in vivo* antibody diversity. The technique is predicated on the premise that the library used is representative of the antibody repertoire *in vivo*. As the library was constructed from adult B cells, it indeed ought to reflect the effect of negative selection against self-reactive antibody sequences. However, this library was not designed for this purpose, and was not controlled for differences in immune repertoires based on age or B cell subsets. These limitations are significant as self- and poly-reactive clones are common in some B cell subsets. Another shortcoming of this library is that it does not exhibit the same antibody subtype distribution as native diversity, e.g. Class 6 variable heavy chain domains are heavily over-represented.

Similarly, because one of the mechanisms by which autoreactive antibodies are resolved is by changing the pairing of heavy and light chains, and the means of construction for this library shuffled chain pairings, it is possible that the newly generated pairings restore auto-reactivity, leading to inflation of the binding counts against fully human proteins. The observation that human proteins scored as less reactive implied that this altered pairing does not typically result in self-reactivity may reflect the predominance of the Vh domain in driving antigen recognition. However, it remains to be determined whether the binders isolated to fully human proteins are the result of non-native Vh–Vl pairings or are reflective of an inherent and underlying capacity of the immune system to self-react that has not been matured or developed further due to the lack of sufficient affinity and/or co-stimulatory factors. The prevalence of autoimmune disorders with auto-reactive antibodies indicates that given the right signals, the existing B cell repertoire can generate antibodies to recognize virtually any antigen, whether self or foreign. Thus, immunogenicity is not simply determined by availability of an antibody sequence with some ability to recognize a foreign protein, but rather by a complex set of signals present in germinal centers, representing a tremendous hurdle to any technology aimed at predicting immunogenicity (Schwickert *et al.*, 2011).

Despite these limitations, the reactivity counts generated using this library correlate well with expectations based on sequence homology and human clinical studies. Perhaps this is not surprising, as immunogenicity reductions do tend to translate across species and between individuals—indicating that immunogenicity is perhaps more a matter of the immunogenic sites of the protein therapeutic than of the specific characteristics of the immune repertoire(Nagata and Pastan, 2009). That is, despite the sequence diversity of the antibody repertoire, the same sites seem to frequently act as immunogenic hotspots, and therefore a method to identify and remove these sites may translate into immunogenicity reductions across individuals.

Future study, perhaps using an scFv library expressly designed for immunogenicity prediction, in which native Vh–Vl pairings were maintained, and naturally occurring subclass usage and appropriate B cell subsets were accurately captured, might allow investigation into trends among self-reactive sequences and immunogenic epitopes, as well as permit epitope mapping, and feed into existing databases of immunogenic epitopes (Vita *et al.*, 2009), and thereby facilitate development of improved in silico methods. Similarly, library-based methods analogs to that presented here could be utilized to characterize differences in the immune repertoire between individuals with different genetic or disease backgrounds. Overall, a library designed for this purpose, with careful attention paid to the classes of B cells utilized could not only improve the method, but also perhaps allow study of the B cell lineages, complementarity determining region sequences, and antibody classes that might be more prone to the generation of autoimmune antibodies. The great utility of immune libraries in other efforts suggests that they are useful tools in studying the immune response and this study suggests that they may provide a means to evaluate the intrinsic B cell antigenicity of proteins. Furthermore, use of libraries in other display formats may also prove promising in efforts to prospectively analyze antibody diversity and characterize the B cell immunome. Finally, this study supports use of antibody libraries that have not been negatively selected against self antigens in protein engineering efforts with the objective of identifying binders to native human proteins.

Funding

This work was supported by funding from the MIT-Portugal Program.

References

Ackerman,M., Levary,D., Tobon,G., Hackel,B., Orcutt,K.D. and Wittrup,K.D. (2009) *Biotechnol Prog*, **25**, 774–783. First published on 2009/04/14, doi:10.1002/btpr.174.

Boehm,M.K., Corper,A.L., Wan,T., *et al*. (2000) *Biochem J*, **346**, 519–528. First published on 2000/03/24.

Bonsignori,M., Hwang,K.K., Chen,X., *et al*. (2011) *J Virol*. First published on 2011/07/29, doi:JVI.05045-11 [pii]; 10.1128/JVI.05045-11.

Brons,N.H., Blaich,A., Wiesmuller,K.H., Schneider,F., Jung,G. and Muller,C.P. (1996) *Scand J Immunol*, **44**, 478–484. First published on 1996/11/01.

Brusic,V., Bajic,V.B. and Petrovsky,N. (2004) *Methods*, **34**, 436–443. First published on 2004/11/16, doi:S1046-2023(04)00125-2 [pii]; 10.1016/j.ymeth.2004.06.006.

Chao,G., Lau,W.L., Hackel,B.J., Sazinsky,S.L., Lippow,S.M. and Wittrup,K.D. (2006) *Nat Protoc*, **1**, 755–768. First published on 2007/04/05, doi:nprot.2006.94 [pii]; 10.1038/nprot.2006.94.

Chester,K.A., Begent,R.H., Robson,L., *et al*. (1994) *Lancet*, **343**, 455–456. First published on 1994/02/19.

de Haard,H.J., van Neer,N., Reurs,A., Hufton,S.E., Roovers,R.C., Henderikx,P., de Bruine,A.P., Arends,J.W. and Hoogenboom,H.R. (1999) *J Biol Chem*, **274**, 18218–18230. First published on 1999/06/22.

Feldhaus,M.J., Siegel,R.W., Opresko,L.K., *et al*. (2003) *Nat Biotechnol*, **21**, 163–170. First published on 2003/01/22, doi:10.1038/nbt785; nbt785 [pii].

Ferry,H., Leung,J.C., Lewis,G., Nijnik,A., Silver,K., Lambe,T. and Cornall,R.J. (2006) *Transplantation*, **81**, 308–315. First published on 2006/02/16, doi:10.1097/01.tp.0000203830.79357.39; 00007890-20060 2150-00002 [pii].

Gai,S.A. and Wittrup,K.D. (2007) *Curr Opin Struct Biol*, **17**, 467–473. First published on 2007/09/18, doi:S0959-440X(07)00119-4 [pii]; 10.1016/j.sbi.2007.08.012.

Goodnow,C.C. (1996) *Proc Natl Acad Sci U S A*, **93**, 2264–2271. First published on 1996/03/19.

Graff,C.P., Chester,K., Begent,R. and Wittrup,K.D. (2004) *Protein Eng Des Sel*, **17**, 293–304. First published on 2004/04/30, doi:10.1093/protein/gzh038; gzh038 [pii].

Gustafsson,E., Haas,P.J., Walse,B., Hijnen,M., Furebring,C., Ohlin,M., van Strijp,J.A. and van Kessel,K.P. (2009) *BMC Immunol*, **10**, 13. First published on 2009/03/17, doi:1471-2172-10-13 [pii]; 10.1186/1471-2172-10-13.

Hanes,J., Schaffitzel,C., Knappik,A. and Pluckthun,A. (2000) *Nat Biotechnol*, **18**, 1287–1292. First published on 2000/12/02, doi:10.1038/82407.

Hansen,J.K., Weldon,J.E., Xiang,L., Beers,R., Onda,M. and Pastan,I. (2010) *J Immunother*, **33**, 297–304. First published on 2010/05/07, doi:10.1097/CJI.0b013e3181cd1164; 00002371-201004000-00008 [pii].

Hicar,M.D., Kalams,S.A., Spearman,P.W. and Crowe,J.E., Jr (2010) *Vaccine*, **28** (Suppl. 2), B18–B23. First published on 2010/06/01, doi:S0264-410X(10)00144-1 [pii]; 10.1016/j.vaccine.2010.02.002.

Kanduc,D. (2008) *Adv Exp Med Biol*, **640**, 198–207. First published on 2008/12/11, doi:10.1007/978-0-387-09789-3_15.

Kanduc,D. (2009) *Curr Pharm Des*, **15**, 3283–3289. First published on 2009/10/29.

Keil,W. and Wagner,R.R. (1989) *Virology*, **170**, 392–407. First published on 1989/06/01.

Koshland,D.E., Jr (1990) *Science*, **248**, 1273. First published on 1990/06/15.

Kreitman,R.J., Margulies,I., Stetler-Stevenson,M., Wang,Q.C., FitzGerald,D.J. and Pastan,I. (2000) *Clin Cancer Res*, **6**, 1476–1487. First published on 2000/04/25.

Laroche,Y., Heymans,S., Capaert,S., De Cock,F., Demarsin,E. and Collen,D. (2000) *Blood*, **96**, 1425–1432. First published on 2000/08/15.

Mayer,A., Sharma,S.K., Tolner,B., Minton,N.P., Purdy,D., Amlot,P., Tharakan,G., Begent,R.H. and Chester,K.A. (2004) *Br J Cancer*, **90**, 2402–2410. First published on 2004/05/27, doi:10.1038/sj.bjc.6601888; 6601888 [pii].

Mishima,K., Johns,T.G., Luwor,R.B., *et al*. (2001) *Cancer Res*, **61**, 5349–5354. First published on 2001/07/17.

Nagata,S. and Pastan,I. (2009) *Adv Drug Deliv Rev*, **61**, 977–985. First published on 2009/08/15, doi:S0169-409X(09)00223-3 [pii]; 10.1016/j.addr.2009.07.014.

Nemazee,D. (2000) *Annu Rev Immunol*, **18**, 19–51. First published on 2000/06/03, doi:18/1/19 [pii]; 10.1146/annurev.immunol.18.1.19.

Onda,M., Nagata,S., FitzGerald,D.J., *et al*. (2006) *J Immunol*, **177**, 8822–8834. First published on 2006/12/05, doi:177/12/8822 [pii].

Onda,M., Beers,R., Xiang,L., Nagata,S., Wang,Q.C. and Pastan,I. (2008) *Proc Natl Acad Sci USA*, **105**, 11311–11316. First published on 2008/08/06, doi:0804851105 [pii]; 10.1073/pnas.0804851105.

Peces,R., de la Torre,M., Alcazar,R. and Urra,J.M. (1996) *N Engl J Med*, **335**, 523–524. First published on 1996/08/15.

Ponce,R., Abad,L., Amaravadi,L., *et al*. (2009) *Regul Toxicol Pharmacol*, **54**, 164–182. First published on 2009/04/07, doi:S0273-2300(09)00059-2 [pii]; 10.1016/j.yrtph.2009.03.012.

Ritter,G., Cohen,L.S., Williams,C., Jr, Richards,E.C., Old,L.J. and Welt,S. (2001) *Cancer Res*, **61**, 6851–6859. First published on 2001/09/18.

Rosenberg,J.S. and Atassi,M.Z. (1997) *Immunol Invest*, **26**, 473–489. First published on 1997/06/01.

Salvatore,G., Beers,R., Margulies,I., Kreitman,R.J. and Pastan,I. (2002) *Clin Cancer Res*, **8**, 995–1002. First published on 2002/04/12.

Scheid,J.F., Mouquet,H., Ueberheide,B., *et al*. (2011) *Science*. First published on 2011/07/19, doi:science.1207227 [pii]; 10.1126/science.1207227.

Schellekens,H. (2002) *Clin Ther*, **24**, 1720–1740; discussion 1719. First published on 2002/12/28, doi:S0149291802800753 [pii].

Schwickert,T.A., Victora,G.D., Fooksman,D.R., Kamphorst,A.O., Mugnier,M.R., Gitlin,A.D., Dustin,M.L. and Nussenzweig,M.C. (2011) *J Exp Med*, **208**, 1243–1252. First published on 2011/05/18, doi:jem.20102477 [pii]; 10.1084/jem.20102477.

Shankar,G., Shores,E., Wagner,C. and Mire-Sluis,A. (2006) *Trends Biotechnol*, **24**, 274–280. First published on 2006/04/25, doi:S0167-7799(06)00082-5 [pii]; 10.1016/j.tibtech.2006.04.001.

Sheets,M.D., Amersdorfer,P., Finnern,R., *et al*. (1998) *Proc Natl Acad Sci USA*, **95**, 6157–6162. First published on 1998/05/30.

Spencer,D.I., Robson,L., Purdy,D., *et al*. (2002) *Proteomics*, **2**, 271–279. First published on 2002/03/29, doi:10.1002/1615-9861(200203)2:3<271::AID-PROT271>3.0.CO;2-W [pii].

Stickler,M.M., Estell,D.A. and Harding,F.A. (2000) *J Immunother*, **23**, 654–660. First published on 2001/02/24.

Subramanyam,M. (2006) *J Immunotoxicol*, **3**, 151–156. First published on 2008/10/30, doi:756649429 [pii]; 10.1080/15476910600845740.

van den Berg,H.A. and Rand,D.A. (2004) *J Theor Biol*, **231**, 535–548. First published on 2004/10/19, doi:S0022-5193(04)00329-7 [pii]; 10.1016/j.jtbi.2004.07.008.

Vita,R., Zarebski,L., Greenbaum,J.A., Emami,H., Hoof,I., Salimi,N., Damle,R., Sette,A. and Peters,B. (2009) *Nucleic Acids Res*. First published on 2009/11/13, doi:gkp1004 [pii]; 10.1093/nar/gkp1004.

Walden,P. (1996) *Curr Opin Immunol*, **8**, 68–74. First published on 1996/02/01, doi:S0952-7915(96)80107-5 [pii].

Warmerdam,P.A., Plaisance,S., Vanderlick,K., Vandervoort,P., Brepoels,K., Collen,D. and De Maeyer,M. (2002) *Thromb Haemost*, **87**, 666–673. First published on 2002/05/15.

Weldon,J.E., Xiang,L., Chertov,O., Margulies,I., Kreitman,R.J., FitzGerald,D.J. and Pastan,I. (2009) *Blood*, **113**, 3792–3800. First published on 2008/11/08, doi:blood-2008-08-173195 [pii]; 10.1182/blood-2008-08-173195.

Welt,S., Scott,A.M., Divgi,C.R., *et al*. (1996) *J Clin Oncol*, **14**, 1787–1797. First published on 1996/06/01.

Welt,S., Ritter,G., Williams,C., Jr, Cohen,L.S., John,M., Jungbluth,A., Richards,E.A., Old,L.J. and Kemeny,N.E. (2003) *Clin Cancer Res*, **9**, 1338–1346. First published on 2003/04/10.

Wierda,D.S.H. and Zwickl,C.M. (2001) *Toxicology*, **158**, 71–74. First published on.

Zang,Y.C., Yang,D., Hong,J., Tejada-Simon,M.V., Rivera,V.M. and Zhang,J.Z. (2000) *Neurology*, **55**, 397–404. First published on 2000/08/10.

Zwickl,C.M., Cocke,K.S., Tamura,R.N., Holzhausen,L.M., Brophy,G.T., Bick,P.H. and Wierda,D. (1991) *Fundam Appl Toxicol*, **16**, 275–287. First published on 1991/02/01.