Implications of Immunogenicity in Drug Development
Michele Gunsior, PhD, Covance Laboratories, Chantilly, VA

Introduction
Biotherapeutics emerged in the 1970s as viable alternatives to small molecule drugs when the advent of recombinant DNA technology allowed for cost-efficient, large-scale manufacture of proteins in cultured cells under controlled conditions. In 1982, recombinant human insulin, developed by Genentech and Eli Lilly, was the first biotechnology therapy to be approved by the FDA. Since then the biotechnology industry has reached more than 50 billion dollars in annual revenue, and continued growth is expected as pharmaceutical companies shift more resources into large molecule drug development, or make high-value acquisitions of biologics companies. There are more than 200 biotechnology products currently on the market with 400-plus additional therapies in clinical trials targeting diseases such as cancer, Alzheimer’s, rheumatoid arthritis, multiple sclerosis, and HIV/AIDS. In 2014, it is predicted that 50% of the top 100 drug sales will be biologics. This is a significant increase from 28% in 2008 and 11% in 2000 [EvaluatePharma 2009].

In contrast to chemically synthesized drugs, biotherapeutics such as proteins, peptides, and antibodies generally possess large and complex structures often modified by glycosylation or pegylation, and may be non-human in origin. These characteristics greatly increase the potential to induce an antibody-mediated immune response over their small-molecule counterparts, and in consequence may cause severe illness in humans. Real-world examples of therapeutics that have caused serious health consequences for healthy volunteers and patients include erythropoietin [Casadevall, 2002], factor VIII [Lavigne-Lissalde, 2005], and insulin therapy. Immunogenicity testing is thus an important part of developing large-molecule therapies.

Immune Response
The principal function of an immune system is to protect the body against infection. For higher organisms this is accomplished with the innate (immediate, non-specific) and adaptive (specific, antibody-mediated) immune response. The immune system must distinguish the organism’s own cells, tissues and proteins from non-self antigens, although in autoimmune disorders such as rheumatoid arthritis and systemic lupus erythematosus the immune system targets self tissues and proteins. Most non-self antigens will elicit an antibody-mediated immune response, termed immunogenicity, in animals and humans. This adaptive response to a particular antigen is heterogeneous; no two biological systems react in the same manner, and the response of a single system will change over time. The immune response is used advantageously in vaccinations, however, a primary concern in the development of non-vaccine therapeutics is understanding, measuring, and mitigating an unwanted, drug-induced antibody response.

The myriad of factors that cause immunogenicity within a particular system can be categorized as drug-related or patient-related. Some examples of drug-related factors are related to protein structure such as alterations due to pegylation, glycosylation, and oxidation of the molecule. While...
pegylation is often used to increase the bioavailability of drug and reduce immunogenicity, anti-PEG-antibodies have been detected and in some cases shown to be associated with changes to pharmacokinetics (PK) data and loss of efficacy [Armstrong, 2007]. Structural characteristics of a protein such as repetitive amino acid sequences (whether in secondary or tertiary structure) and the tendency to form aggregates can affect immunogenicity rates as the immune system is primed to respond to aggregates and other patterned, repetitive sequences. Additionally, a particular drug meant to mimic an endogenous counterpart may differ slightly in sequence, causing the immune system to recognize the therapeutic as “non-self.” A second drug-related factor is related to manufacturing process and drug formulation. For example, drug production in eukaryotic or prokaryotic systems may introduce immunogenic foreign proteins into the product. Furthermore, the formulation itself may increase the propensity towards aggregation as well as contain excipients or impurities that induce an immune response.

Other drug-related factors include the route of administration, dose level, and frequency of dosing. It has been generally accepted that subcutaneous (SC) exposure is more immunogenic than either intramuscular (IM) or intravenous (IV) dosing. However, anecdotal evidence presented at a recent conference on immunogenicity (Boris Gorovits, Immunogenicity Summit, October 2009) indicates that this type of ranking is not necessarily supported by the data. A paper exercise evaluating data on 20 biotherapeutics of various classes indicated that in many cases, the rate of immunogenicity by IV exposure was equal to SC exposure. Given that many exceptions exist, specific testing of individual products is encouraged [Koren, 2008; Ponce, 2009]. Dose and frequency of drug administration effects are assessed more generally, as these are not truly independent of other factors. A single dose is usually not expected to produce a robust or well-developed immune response compared with multiple doses over time. However, frequent and high concentration dosing regimens is one mechanism used to induce drug tolerance in patients.

Patient or host-related factors, such as genetic disposition, immune status, disease state, and the presence of co-medications play a role in the immunogenicity rate. For example, a healthy individual will have a different immune response to a given drug than that of the targeted disease state population. For example, Factor VIII drug therapy is unlikely to induce the same response in healthy subjects whose immune system is more likely to see the drug as “self” than the disease population who lacks Factor VIII protein, and thus will label the drug therapy “non-self” and xenobiotic in origin.

**Impact of Immune Response**

Since a critical part of the drug development process is to accurately assess drug exposure to determine bioavailability in the test species, it is important to understand how the presence of anti-drug antibodies (ADA, also anti-therapeutic antibodies, ATA) may impact this assessment. The antibodies produced as a result of drug challenge may be sustaining, clearing, or neutralizing, all of which can affect the pharmacokinetic and pharmacodynamic properties of a drug. Sustaining antibodies prolong the drug half-life by protecting it from normal clearance; in contrast, a clearing antibody more quickly decreases drug concentration below therapeutic levels or removes drug from circulation. Binding antibodies may not clear a drug, but can interfere with pharmacokinetic analysis by preventing measurement of the therapeutic. Neutralizing antibodies abrogate a drug's therapeutic effect and potentially worsen the patient’s disease if not also inducing serious side-effects. The extent, duration, and severity of the response can have a wide range of clinical consequences from little or no effect, to severe and life-threatening illness. The most well-known example of severe illness is pure red cell aplasia in which neutralizing antibodies targeted against erythropoietin cross react with the endogenous protein. Finally, the presence of ADA can lead to other clinically important adverse events such as hypersensitivity reactions, anaphylaxis, serum sickness, and severe infusion reactions.

The early immune response typically comprised of low affinity, low concentration IgM molecules, is of limited magnitude, and peaks within 7-14 days. As the immune response matures, often with repeated drug (antigen) exposure, isotype switching leads to a more robust response consisting largely of IgG antibodies. The presence of IgG4 indicates a well-entrenched immune response, with high affinity antibodies that are more likely to neutralize the effect of drug in higher concentrations.

Understanding the antibody response to drug exposure is necessary to evaluating the significance of that response. Determining the magnitude of response is important in understanding and ultimately determining the safety and efficacy of a particular drug which may include:

- Response duration - whether sustained or transient
- Occurrence of adverse events such as allergic reactions
- Evaluating the production of neutralizing antibodies
- Correlating change in PK/TK profiles
The challenge then becomes developing sensitive and specific immunogenicity assays to determine the presence, specificity, magnitude, and neutralizing potential of ADA.

**Methodology for Immunogenicity Testing and Analysis**

A typical immunogenicity analysis follows a tiered approach that includes three types of analysis (Figure 1): an initial screen, in which study samples are assessed for the presence of ADA, a confirmation assay, which determines whether putatively positive samples react specifically with free drug (versus nonspecific reactivity), and a titer evaluation that provides a relative level of ADA. When assessment of neutralizing antibody potential is required, a fourth type of assay, most often cell-based, is implemented to determine whether ADAs in confirmed positive samples specifically prevent drug from binding to its target.

An inadequately designed ADA assay can hinder the appropriate assessment of clinical findings and delay product development or marketing. The ability to correlate immunogenicity and clinical events therefore relies upon a properly developed and validated assay. The approach to immunogenicity testing has undergone significant changes in recent years. For example, Shankar et al., (2008) published a more concrete plan for developing and validating reliable testing methods than was previously considered.

Specifically they expanded upon previous industry opinions [Mire-Sluis, 2004] by recommending a more statistical and objective approach in evaluating certain method parameters such as cutpoint and confirmatory cut-off limits.

The screening cut point is defined as the level at and above which a sample is considered positive (or reactive) for ADA and below which a sample is considered negative. A risk-based approach requires a cutpoint selection that leads to a 5% false positive rate as it is more appropriate for a certain number of samples be considered false positive than false negative. The false positive samples can then be proven non-reactive in the confirmatory assay ensuring that the method can detect low levels of true positives. The screening cutpoint is determined by a statistical evaluation of a drug-naïve population representative of the study samples. Any false positives should then be distinguished from true positive (reactive) samples through a competitive assay, termed confirmation assay, with free drug. The specificity cut point is determined during evaluation of the screening cut point by incubating the same naïve samples (or, alternately, naïve samples spiked with low positive control) with an excess of drug and evaluating the reduction in signal of spiked versus unspiked samples. A detailed explanation for determining both the screening and confirmatory cutpoints can be found in Shankar, et al.

Immunogenicity assessments can be made using a number of analytical platforms such as:

- Enzyme-linked Immunosorbent Assay (ELISA)
- Radioimmunoassay (RIA)
- Electrochemiluminescence (ECL)
- Surface Plasmon Resonance (SPR)

Each of the platforms has advantages and disadvantages; a well-thought out development plan will take into consideration these practicalities in designing the immunoassay. For example, ECL assays are often the platform of choice as a homogenous bridging assay. ECL eliminates multiple wash steps and increases the ability to detect low affinity antibodies as well as various isotypes. For further characterization of the immune response, SPR, such as the Biacore system, may be used to determine the specific antibody isotype or binding characteristics.

ADA assays are considered qualitative (or sometimes quasi-quantitative) because the positive control does not accurately reflect the heterogeneity of immune system response, or, usually, originate from the species of interest. The positive control is often a monoclonal antibody or polyclonal sera derived from hyper-immunization of a particular species, and will vary from the study sample in terms of affinity, avidity, specificity, and isotype. Given the inherent dissimilarity between positive control and study sample, the use of mass units in defining the level of response is of limited value, and not recommended by the FDA [U.S. Dept. Health and Human Services, 2009]. The purpose of a positive control is therefore to demonstrate the method functionality and monitor assay performance. The ideal positive control is derived from the species of interest, which is usually not possible for human subjects. If the positive control is from a different species, it should be diluted into the matrix of study samples. In method design and execution, low, high, and negative controls are used to monitor assay performance. A high positive control with signal at the upper end of linear response is useful in monitoring method performance over time, reagent
qualifications, and troubleshooting, or when a method has a demonstrable hook effect. In contrast, it is the low positive control that ensures reliable assay performance and provides confidence that the assay can detect low levels of ADA.

Both drug quantitation and ADA determination assays are subject to interference from matrix components. For drug quantitation assays this includes the presence of ADA; ADA assays are similarly prone to interference from free drug. Antibody drugs in particular have long half-lives, and ADA may stabilize drugs thereby prolonging the half-life. Effort is made to minimize these interferences as much as possible: time points for ADA analysis are typically chosen when the drug is in the elimination phase, which may be months following drug challenge, or the assay format may include an acid-dissociation step to remove or decrease antibody:drug complexes. The affect of drug on antibody detection, e.g. tolerance, in an analytical method is evaluated by incubating increasing levels of drug with the low positive control. The lowest concentration of drug that prevents detection of ADA is termed the drug tolerance; drug tolerance depends, like sensitivity, upon the particular positive control chosen. It is therefore suggested that ADA method validations are designed to appropriately assess the drug tolerance over the range of possible drug concentrations that may be seen in test samples. Caution is required, however, when interpreting results of study samples with measurable quantities of drug; since drug tolerance is dependent upon the chosen positive control – which itself is not representative of all possible immune responses – the assessment of ADA (positive or negative) and the corresponding level (titer) may not be accurate. Similarly, since ADA may interfere with quantitation of drug, the concentration of drug obtained from a particular sample may not be a true reflection of the actual level.

The main purpose of evaluating immunogenicity in animals is to help interpret the PK and PD properties of the drug, as well as any toxicity findings:

\[
\text{ADA + PK + PD + toxicity = immunogenicity assessment.}
\]

Preclinical Implications of Immunogenicity

Until recently, the focus of regulatory agencies and opinion papers has largely been the design and application of immunogenicity studies on human subjects, but an important part of any drug development program are preclinical toxicology studies in which the effects of drug are evaluated in animal models. The particular choice of animal will depend on the drug characteristics and physiological function so as to better understand the mechanism of action and effect in vivo as well as to assess toxicity. However, animal models are generally poor predictors of immunogenicity in humans. For example, the recombinant human monoclonal antibody Humira produced a 100% induction of ADA in cynomolgus monkeys yet the clinical immunogenicity rate was only 12% [Ponce, 2009]. As many large molecule drugs are human in origin, it is expected that the animal species will recognize these as non-self proteins and mount an immune response. The development of ADA and magnitude of the response can have clear impacts on the ability to interpret study findings. As discussed earlier, antibodies can enhance the clearance of drug (reducing drug exposure and limiting efficacy, as well as altering PK and PD profiles), sustain bioavailability which may result in increased toxicity or increased efficacy, or neutralize the drug’s pharmacological effect which can decrease efficacy and toxicity. Thus the main purpose of evaluating immunogenicity in animals is to help interpret the PK and PD properties of the drug, as well as any toxicity findings, or more simply, ADA + PK + PD + toxicity = immunogenicity assessment. As a result, careful consideration should be made when selecting appropriate animal models.

Accurate assessment of ADA in study samples is impacted by several factors, including the concentration of drug in matrix and the choice of positive controls. The effect of high drug concentration can sometimes be minimized by the assay format, or by the use of techniques such as acid dissociation to decrease drug:ADA complexes. The other limitations of an ADA method are those imposed by the choice of positive control, which will not adequately represent the heterogeneity of the study samples. A biological system mounts a polyclonal immune response in which antibodies produced are of differing affinity and avidity. Excellent method drug tolerance and sensitivity may be illustrated with the use of a high affinity monoclonal antibody control, but this is not reflective of the biological response, which likely contains many high affinity antibodies and a host of low affinity/avidity antibodies. The drug tolerance and sensitivity obtained from method validation should not be applied to study samples, or used cautiously with careful consideration of the type of positive control and its similarity to the biological response; very different sensitivities and drug tolerance results will be obtained with a polyclonal preparation versus a single high-affinity monoclonal antibody control.

Immunogenicity assessment in animals is best used to interpret study findings and assess the validity of toxicology studies. During protocol development, blood samples should be collected for any study in which animals are exposed to drug for greater than seven days. In an effort to manage sample analysis, Ponce, et al. (2009) provided a decision tree that guides the investigator on whether ADA analyses will meaningfully contribute to data interpretation. For example, if suitable PK assays and PD biomarkers are available, and no change in profile is seen, then an ADA screen may not be warranted. If
Proper risk assessment considers the probability of generating an immune response and the consequences of such a response.

However, changes are observed, then at minimum the screening assay should be performed, followed by titer if shown positive for ADA. Though a tiered approach (screen, confirmation, titer) is recommended for clinical sample analysis, confirmation is generally not needed in animal studies, nor is the fourth tier of neutralizing antibody assessment.

Clinical Implications of Immunogenicity

It is widely understood that “the induction of immunogenicity in animal models is not predictive of a potential for antibody formation in humans” [ICH S6 Guidance] and thus the immunogenicity rate in a particular animal model does not necessarily preclude the initiation of clinical studies. An approach that takes into account a number of factors including the causes and consequences of immunogenicity is applied to any new biotherapeutic under consideration for marketing and licensure; proper risk assessment considers the probability of generating an immune response and the consequences of such a response [Rosenberg; Koren, 2008]. The risk is then categorized as low, medium, or high and the bioanalytical testing strategy is defined accordingly [Shankar, 2007].

To appraise the probability of an immune response, the causes of immunogenicity discussed under the Immune Response section above are considered along with other factors such as:

- Presence of drug conjugates (as to a toxin)
- Exposure of neoantigenic epitopes upon drug denaturation or fragmentation
- Natural tolerance or use of immunosuppressive drugs
- Length of drug treatment
- Homology of drug product to endogenous counterpart
- Redundancy of an endogenous counterpart
- Mechanism of action

Following assessment of immunogenic potential, consideration of the possible clinical sequelae, as well as whether the target indication is a life-threatening disease is needed to develop a risk-management plan. In mild-to-moderate responses to ADA, the patient may experience loss of drug efficacy, or allergic hypersensitivity reactions such as an infusion-site reaction. More serious consequences include overstimulation of endogenous mechanisms such as that following treatment with TGN1412 [Medicines and Healthcare Products Regulatory Agency, 2006], anaphylactic shock, or cross reactivity with, and neutralization of endogenous protein leading to an autoimmune deficient syndrome (Table 1).

The type of risk inherent to a particular drug dictates the testing and sampling strategy. Low and medium risk drugs should have ADA samples taken frequently in Phase I and Phase II trials to help understand the human immune response to the drug; testing may be less frequent in Phase III. By contrast, a high risk drug, one with potential to cross react with and neutralize an endogenous counterpart, might require frequent sampling and real-time ADA analysis (including neutralization potential) at all stages of clinical development, or sequential patient dosing (rather than a cohort design) to better mitigate risks. At a minimum, all immunogenicity assessments during clinical trials should include screen, confirmatory, and titer analysis. The timing for implementation of a neutralizing antibody assay, and type of format required, will vary with drug product and associated risk to the patient. A cell-based assay for a monoclonal antibody designated as low risk may be of limited value (while taking considerable time and effort to develop and validate).

Approaches to attenuate immunogenicity and minimize patient risk include drug redesign (to remove epitopes of
high immunogenic potential), change in dose and/or dosing procedure (IV administration versus inhalation), and exclusion of sensitive subjects (a patient with known allergies to a similar product could be prescreened with a skin test, or be administered drug in a clinical setting where severe allergic reactions can be readily treated). Once immunogenicity is observed, the clinician must decide the best path forward for a patient. If ADA inhibits the function of a life-saving therapy, it may be necessary to institute a tolerizing regimen to inure the immune system against the therapeutic. These decisions are made on a case-by-case basis with consideration of all available information.

**Follow-on Biologics**

Follow-on biologics, or biosimilars, are large-molecule therapeutics similar to a licensed drug, but manufactured and marketed by a third-party following patent expiration of the innovator molecule. While there is currently no mechanism for approval and licensure of biosimilars in the US, the EU and Japan both have a legal, regulated process that includes assessment of immunogenicity between the innovator drug and follow-on biologic at all stages of the development process to ensure comparable quality, safety, and efficacy.

Paramount in assessment of immunogenicity is that the methods show comparable sensitivity and specificity. Key questions to consider when developing a single method that shows equal sensitivity and specificity to both the innovator and biosimilar drug products [U.S. Dept. Health and Human Services, 2009] include the following:

- What is best the assay format and positive control to use?
- If a bridging assay is employed, should the capture and detection reagents be innovator molecule or biosimilar?
- Should the positive control be one generated from the reference product or follow-on biologic, or perhaps some combination of both?
- If a single method can not be developed for immunogenicity comparison, then do the separate methods show equivalent sensitivity and specificity?

Answers to these and other questions are important to distinguish between the effects of different manufacturing processes, purity of biosimilar versus originator therapeutic, or biosimilar itself from differences in assay performance.

**Conclusion**

In contrast to new chemical entities, biotherapeutic drugs are potentially immunogenic, and, in rare cases, treatment with such products can lead to severe and devastating illnesses in humans. Therefore, the importance in understanding how immunogenicity affects drug exposure, efficacy, and toxicity at all stages of the drug development process can not be overstated. The complexity of the immune response necessitates production of a sound risk-assessment strategy and a fit-for-purpose approach in testing of preclinical and clinical samples. Whatever the application, properly designed and validated immunoanalytical methods provide confidence in the data used for these purposes.

**References**


