Analysis and Immunogenic Potential of Aggregates and Particles

A Practical Approach, Part 2

Anthony Mire-Sluis, Barry Cherney, Russell Madsen, Alla Polozova, Amy Rosenberg, Holly Smith, Taruna Arora, and Linda Narhi

The conclusion of this CMC Forum continued to focus on the latest developments in detection and characterization of protein aggregates (**1**). Afternoon sessions detailed the most recent experiments probing the role of protein aggregates in immunogenicity, with discussions on the best models to use and initial results. Topics included potential thresholds for immunogenicity, linking laboratory and clinical data, and predicting and testing potential immunogenicity of products throughout a development lifecycle.

AFTERNOON SESSIONS

Amy Rosenberg (Division of Therapeutic Proteins, CDER, FDA) described a preponderance of data concerning the ability of aggregates to induce and enhance immune responses as well as their significance to the

WWW.PHOTOS.COM

immunogenicity of therapeutic proteins. Her data came from both animal and human studies.

Rosenberg showed robust animal data regarding the ability of protein aggregates to induce immune responses and break tolerance to endogenous proteins. She highlighted studies showing that the greater the degree of antigenic organizations — in which proteins are arrayed at defined intervals such as for viral like particles (VLPs) — the greater the probability of immune response induction. Whereas deliberate aggregation of proteins induced immune responses in animals, elimination of aggregates from therapeutic protein products (by

ultracentrifugation) or disaggregation of protein aggregates (by high-pressure techniques) reduced or eliminated immune responses. Those results further bolstered the contention that protein aggregates enhance immunogenicity. Moreover, murine studies showed that aggregate levels, and not the total amount of protein, correlated with the rapidity with which an antibody response was induced as well as the strength (titer) of ensuring antibody response.

Rosenberg showed data from human studies that also demonstrated a positive (though unwanted) role of aggregates in inducing immune responses. Novel particulate delivery systems for therapeutic proteins were tested. In that case, the presence of human growth hormone (HGH) on microspheres dramatically increased the incidence of HGH antibody response and correlative data in clinical studies of products that contained various types (denatured or native) and amounts of aggregates. In one such study on a highly aggregated IVIG product that induced anaphylaxis, skin testing confirmed that the immune response was directed only to the aggregated fraction and not to the native monomeric fraction of ultracentrifuged IVIG.

An additional study assessing human T-cell responses to a highly

aggregated human serum albumin (HSA) product found that the response was directed specifically to the aggregated fraction. Several human correlative studies further contributed to data implicating aggregates in immunogenicity.

• The sustained high-titer immune response to highly aggregated and denatured HGH, purified from pituitary glands, was diminished by near elimination of the prominent aggregate peak. Instead, it became a transient, lower-titer response.

• Oxidized human IFN-a2b forms high molecular weight (MW) native aggregates, which break tolerance when injected into tolerant human INF-a2b transgenic mice. In human studies, a close correlation was found between the amount of oxidized (hence aggregated) IFN-a2b and the incidence and titer of neutralizing antibodies to the product.

• Highly aggregated rhu IL-2 is highly immunogenic, whereas natural IL-2, which is monomeric, has low immunogenicity.

Why do aggregates, particularly those of high MW or those present in higher-order particulate structures so efficiently induce or enhance immune responses? The explanation can be viewed teleologically and on cellular and molecular levels. A human's immune system is designed to fight microbial attack. It does so by efficiently recognizing, processing, and generating effector responses to materials bearing microbial "signatures." Those signatures consist of particles of defined size bearing molecular patterns (e.g., protein–lipid arrays) and three-dimensional structures. Antigen-presenting cells are highly efficient at phagocytosing or endocystosing particles and high-MW protein arrays that trigger both innate (by means of Toll-like and other inflammatory receptors) and adaptive immune responses. Although monomeric proteins may be minimally immunogenic, arraying them to resemble microbes (such as on VLPs) induces robust immune responses. That is the basis of a successful HPV vaccine, in which the L1 pentamers are arrayed in VLPs.

Given the preponderance of data speaking to the capacity of large protein aggregates and particles to induce immune responses, aggregates should be considered a clear risk to product safety and efficacy and therefore a CQA. In accordance with this view, all aggregates (both soluble and particulate) should be characterized at lot release and on stability, with associated risk evaluations made and an appropriate risk-reduction strategy implemented — which may include additional specifications where appropriate.

Taruna Arora (Amgen Inc.) described three testing strategies for testing the immunogenic potential of protein aggregates: in silico, in vitro, and in vivo models. Each has its advantages and disadvantages. Arora described a new human IgG tolerant mouse model that has been generated through cross-breeding of human IgG2 kappa/lambda transgenic mice with wild type mice. This heterozygous mouse is tolerant of human IgG2/ kappa and IgG2/lambda antibodies and can elicit robust immune responses to foreign antigens. Use of such models early in development allows for the early management of potential immunogenic sites. Once validated, each model plays a role in the manufacturing assessment of candidate molecules.

Depending on how they were formed (e.g., pH, shaking, oxidation, and so on), aggregates and particles exhibit different properties (e.g., reversibility, size, shape, native or non-native protein content). Different types of protein particles induce immunogenicity in both in vitro and in vivo models.

Holly Smith (Eli Lilly and Co.) discussed the effect of immunogenicity on nonclinical studies. She described how antidrug antibodies can affect pharmacokinetic (PK) and/or pharmacodynamic (PD) studies; cause hypersensitivity, immune complex formation, and/or neutralization of biological activity; or have no influence at all. Neutralizing antibodies can prevent drugs from moving within a patient's body and affect PK/PD or directly inhibit the product activity. Any antibody that affects either PK/PD or efficacy can invalidate a toxicology study if animals are not exposed to drugs or their effects. Smith described a case in which antibodies to a product in a nonclinical toxicology study reduced PK by 50%, and no apparent toxicodynamics were observed. Further analysis illustrated that the PD (as assessed through a biomarker) was also reduced in correlation with the reduced PK, thus confirming that immunogenicity can influence the utility and interpretation of nonclinical studies.

AFTERNOON PANEL DISCUSSION

After the presentations, a panel consisting of Arora, Jack Ragheb (CDER, FDA), Barbara Rellahan, (CDER, FDA), Rosenberg, and Smith considered several questions and encouraged audience participation. The following is an overview of the questions and answers discussed.

Do we believe that subvisible particles of therapeutic proteins are immunogenic? If so, is this a generic property or specific to each protein?

Nonhuman studies have shown that aggregated materials can induce immune responses but can be confounded by the way aggregates are formed, which can lead to subtly different types of aggregates and elicit different responses. There appears to be clear potential for aggregated proteins to break tolerance and induce immune responses. In the vaccine arena, it is clear that aggregated proteins have a higher potential to induce an immune response. From our understanding of immune systems, it is not surprising that a more aggregated molecule would mimic a virus or bacteria in the way it gets recognized and processed by the immune system. Disaggregation studies in animals also show that an immune response can be reduced if aggregation is decreased.

We cannot ignore the work carried out with vaccines, where aggregation clearly increases immune responses in humans. Also, multiple cases have shown that reduction of aggregation in products (e.g., HGH, interferon, IVIg, HSA, factor VIII) caused a reduction of immunogenicity in humans. In addition, an analysis of the antibodies produced in response to certain products has shown to bind to the aggregated form and not the monomeric form.

The nature of drug products and the patient population can definitely influence their propensity to induce an immune response. Such factors include

- endogenous counterpart or not
- closeness to "natural" sequence
- product impurities (e.g., host-cell proteins)
- dosing regimen (e.g,. single use or chronic, SQ or IV)
- patient population (immuno-suppressed, immune dysfunction).

How might an aggregate profile affect immunogenicity (including subvisible and visible particles, changes in particle profile with time, etc.)? The type of aggregate is crucial to its propensity to induce an immune response. Aggregates can have very different properties according to how they are created: through feezing and thawing, agitation/vortexing, heat, solid-particle coating, oxidation

(whether through metal, peroxide, or other means), gluteraldehyde exposure, high concentration, or pH. This appears to also depend on the protein itself and can affect immunogenic potential. As described in the morning session, aggregates can be classified by several different characteristics, all of which can affect potential immunogenicity. Crystalline structure and the exposure of T- and B-cell epitopes can be especially relevant to immunogenicity

Protein-coated particles appear to be the most potent "aggregates," as demonstrated in animal models and hinted at with HGH in a depot formulation in humans. For mGH, aggregates created by vortexing appear more potent than those created by freeze–thaw. In innate immune responses, aggregates that were caused by stirring appeared to induce more cellular activation: Material with the highest number of particles and in a partially folded or native state appeared most immunogenic. In an interferon transgenic mouse model, metal catalyzed and heat/pH induced aggregates broke tolerance, but heat alone and crosslinked aggregates did not. Other in vivo studies, however, showed that aggregates failed to induce immunogenicity or were less immunogenic than monomers.

Do we believe that current assays can predict immunogenicity? How much trust should we put in such assays? In silico models help identify MHC class II binding linear peptide sequences. Whether a T cell in vivo will recognize that sequence is not an absolute. Such tools do not distinguish beetween conformational and nonconformational epitopes (e.g., B-cell epitopes). Overprediction of immunogenicity risk has been seen in these models. In silico prediction is very dependent on how well computational algorithms have been created and for what MHC families. Self-tolerance is not accounted for by those tools either. In silico prediction does not consider factors beyond sequence that can break tolerance or enhance immunogenicity.

T-cell assays provide a more functional readout regarding the

ability of proteins and peptides to bind to MHC and induce cellular activation. Dendritic cell assays can uptake and process epitopes and induce cytokine production. They can be limited, as for any predictive immunogenicity test, by the representation of MHC families and lack of physiological environment. In addition, they can be quite variable and are complex to carry out. You need to include dendritic cells and/or PBMCs with T cells for in vitro processing and presentation to represent intercellular processing and activation of the immune system in vivo.

In vivo assays should include all cellular interactions required to induce an immune response. The genetic background of murine models is crucial to their ability to predict what would happen in humans (e.g., tolerance, B- and T-cell receptor types, signal pathways), and peptides are still presented by murine molecules. How well human protein transgenic/tolerant models correlate with those of human immune system remains to

be completely understood. Such models are more relevant for high-homology proteins and/or when using a species-specific protein. Animal models are especially useful for investigating consequences of mounting an immune response. Relative and comparative immunogenicity may be a more relevant focus for animal models at present.

No model can predict what will happen in patients where the disease, route, or comedication can influence the induction of immunogenicity unless animal models can be designed to account for some of those factors (e.g., disease-state models, route). Ultimately, there is a paucity of data comparing two molecules — one predicted to be immunogenic and another to not be immunogenic — tested head-to-head in a clinical study. "Deimmunization," however, has resulted in molecules that did not have immunogenic results in the clinic.

Are these tests at a stage to influence formulations and protein sequence development, and/or function as part of candidate selection? Some companies now routinely use computational screening to select different amino acid sequences for a product at the earliest stage of development. Such models, however, are used in association with other attributes (e.g., potency, stability, solubility, potential to aggregate). The tests can reduce potential risk during candidate selection despite their limitations (better than nothing).

Generally, there appears to be a routine whereby high-risk clones from in silico prediction are then passed into in vitro studies to confirm or refute potential immunogenicity — which, depending on the candidate, would lead to "deimmunization" and retesting. Once the sequence has been altered, however, you should go back through attribute screening to ensure that you have not replaced one issue with another.

Would the purification of particles be possible, and would testing in these assays provide useful data? It is reasonably difficult to separate out protein particles because there are often so few in quantity and they tend to morph together if centrifuged. You can filter your product, but protein particles can pass through filters. Those that do not will often stick to the filter (and may represent the more strongly associated ones). Separating by size (e.g., gradients) may be more relevant. As has been described, the nature of particles very much depends on how they are created but can also be influenced by their separation (some may be more stable than others). So, you may end up with an imbalanced population of particles of a particular nature, or proteins themselves may be affected by the purification and not representative of actual product.

Can anything in the design or analysis of clinical studies provide better links between particle levels and their effects on patients? Actually, testing products in clinical studies for levels of aggregates, submicron particles, and subvisible and visible particles would help. A lot of discussion in industry and with regulators has focused on tracking material lots in clinical studies. Whether it is workable to intentionally create "more variable" lots in regard to any quality attribute during development remains to be seen. So intentionally creating lots with higher or lower levels of particles and tracking outcomes in clinical studies is not easy but could be attempted. But using a single lot per patient throughout a study is not trivial. Samples need to be collected

at appropriate time points during clinical studies to prevent drug interference/ . This also provides for long enough studies to observe late development of antibodies and better understand immunogenic potential of products and any protein particles they may contain. ⊕

*Corresponding author **Anthony Mire-Sluis**, PhD, is vice president of corporate, product, and device quality at Amgen Inc., MS-37-2-C, One Amgen Center Drive, Thousand Oaks, CA 91320-1799; 805-313-2415. **Taruna Arora**, PhD, is principal scientist of protein science at Amgen Inc. **Russell Madsen** is president of The Williamsburg Group, LLC. **Amy Rosenberg** is the director at the Division of Therapeutic Proteins, Office of Biotechnology Products, CDER, FDA. **Barry Cherney** is deputy director at the Division of Therapeutic Proteins, Office of Biotechnology Products, CDER, FDA. **Alla Polozova**, PhD, is senior scientist of analytical biochemistry at MedImmune. **Linda Narhi**, PhD, is scientific executive director of process and product aevelopment at Amgen Inc. **Holly W. Smith**, MA, is senior research scientist of investigative toxicology at Eli Lilly and Company.*