Pharmaceutical aspects of subvisible particles in protein formulations

Weinbuch, D.

CHAPTER 3

PRECLINICAL MODELS USED FOR IMMUNOGENICITY PREDICTION OF THERAPEUTIC PROTEINS

Vera Brinks[1], Daniel Weinbuch[2], Matthew Baker[3], Yann Dean[4], Philippe Stas[5], Stefan Kostense[6], Bonita Rup[7], Wim Jiskoot[2]

[1] Department of Pharmaceutics, Utrecht Institute for Pharmaceutical Sciences (UIPS), Utrecht University, Universiteitsweg 99 3584 CG Utrecht, The Netherlands

[2] Division of Drug Delivery Technology, Leiden Academic Centre for Drug Research (LACDR), Leiden University, P.O. Box 9502, 2300 RA Leiden, The Netherlands

[3] Antitope Ltd, Babraham Research Campus, Babraham Cambridge CB22 3AT, UK

[4] Eclosion SA, Chemin des Aulx, 14 1228 Plan-les-Ouates, Geneva, Switzerland

[5] BLA Consult bvba, Kortestraat 62 9473 Welle, Belgium

[6] Crucell Holland BV, Leiden, The Netherlands

[7] Protein Bioanalytics, Pfizer Inc., Andover Massachusetts, USA

The chapter was published in the Journal *Pharmaceutical Research*: *PharmRes 2013 30(7):1719-1728*

Abstract

All therapeutic proteins are potentially immunogenic. Antibodies formed against these drugs can decrease efficacy, leading to drastically increased therapeutic costs and in rare cases to serious and sometimes life threatening side-effects. Many efforts are therefore undertaken to develop therapeutic proteins with minimal immunogenicity. For this, immunogenicity prediction of candidate drugs during early drug development is essential. Several *in silico*, *in vitro,* and *in vivo* models are used to predict immunogenicity of drug leads, to modify potentially immunogenic properties, and to continue development of drug candidates with expected low immunogenicity. Despite the extensive use of these predictive models, their actual predictive value varies. Important reasons for this uncertainty are the limited/insufficient knowledge on the immune mechanisms underlying immunogenicity of therapeutic proteins, the fact that different predictive models explore different components of the immune system, and the lack of an integrated clinical validation. In this review, we discuss the predictive models in use, summarize aspects of immunogenicity that these models predict, and explore the merits and the limitations of each of the models.

Introduction

Therapeutic proteins are very successful in treating a wide variety of life-threatening diseases such as multiple sclerosis, diabetes, chronic kidney failure and a wide variety of cancers. In contrast to small molecule drugs, they do not possess intrinsic toxicity due to harmful metabolites or off-target effects, and their side effects are mainly caused by exaggerated pharmacodynamic effects (1). Because of their success and versatility, therapeutic proteins are the fastest growing class of drugs and make up about one third of the drug market.

One of the major attention points of therapeutic proteins is immunogenicity. Anti-drug antibodies (ADAs) induced by nearly all therapeutic proteins can interfere with drug-efficacy, alter PK/PD or induce severe, sometimes life-threatening, side-effects in a subset of the patients (2–5). The potential danger of immunogenicity of therapeutic proteins caught public attention around 2002 when an increased number of patients treated with Eprex® (epoetin alpha) were reported to form antibodies that cross-reacted with endogenous erythropoietin. As a result red blood cell production arrested and blood transfusions were vital for these patients' survival (5–7). Besides the apparent risk for patient safety, immunogenicity also poses a financial burden.

In order to minimize side effects caused by ADA formation, immunogenicity assessment of therapeutic proteins during drug development is critical. By identifying immunogenic properties at an early stage, and subsequently modifying those properties, immunogenicity in patients could be minimized. Many efforts have been undertaken to develop *in silico*, *in vitro,* and *in vivo* models that predict different aspects of immunogenicity of therapeutic proteins (8–10). However, current limited knowledge on the general principles that apply to the induction of antibodies by these drugs makes it very difficult to determine the risk factors for immunogenicity and predict the clinical consequences of immunogenicity of a new protein drug (11). Also, the clinically observed immunogenicity against specific drugs is variable depending on other factors such as the disease treated, concomitant treatment and patient background (12). In addition, direct clinical evidence showing that use of these predictive models to guide drug development actually helps to lower immunogenicity is largely missing, as few drug candidates are clinically tested and therefore direct comparisons of candidates showing predicted high and low risk are rarely obtained. Despite these limitations, several *in silico*, *in vitro,* and *in vivo* models are currently applied for different aspects of preclinical immunogenicity prediction (13–15) (Table 1).

Table 1: Main applications and limitations of current *in silico*, *in vitro* and *in vivo* predictive tools for protein immunogenicity.

Category	What does it predict?	Advantages	Disadvantages
In silico	Presence of potential CD4+ T cell epitopes	Fast, low cost	Focus on primary structure; no information on contribution of other factors (e.g., glycosylation, formulation, aggregation)
	Presence of neoepitopes	Fast, low cost	Does not address the actual T-cell activation
In vitro	Presence of CD4+ T cell epitopes	Relatively fast, low cost	Focus on activation of specific immune cells
	Presence of neoepitopes	Measures biological effects	Large donor sets needed
	Activation of T cells	Can be used to screen product-related factors other than primary structure	Assay variability
In vivo			
Conventional animals	Relative immunogenicity	*In vivo* correlate of immunogenicity	Per definition a classical immune response against therapeutic proteins Overestimation of immunogenicity Time consuming, expensive Non-human immune system
Non-human primates	Relative immunogenicity	Express similar proteins as humans, therefore may have similar immune mechanism underlying immunogenicity	Predictive value strongly depends on protein
	Likely the presence of neoepitopes	*In vivo* correlate of immunogenicity	Time consuming, expensive
	Breaking of tolerance (depends on protein)	*In vivo* correlate of immunogenicity	Needs clinical validation
Transgenic immune	Presence of neoepitopes	Express protein of interest similar to tolerant mice humans, can be used to study breaking of tolerance	Mice respond with murine immune system
	Relative immunogenicity	*In vivo* correlate of immunogenicty	Time consuming, expensive
	Breaking of tolerance	*In vivo* correlate of immunogenicty	Needs clinical validation
Human immune system xenograft models	Presence of neoepitopes	Express many human immune proteins that xenograft models are (potential) therapeutic targets, can be used to study breaking of tolerance.	For most models human T cells lack the ability to recognize antigens in a HLA-restricted manner (i.e. no value to predict T cell-de pendent ADA), BLT mice are exception.
	Relative immunogenicity	Mice respond with a (reconstituted) human immune system	Time consuming, expensive
	Breaking of tolerance (if human protein is expressed due to xenografting)	*In vivo* correlate of immunogenicty	Needs clinical validation

This review summarizes the models in use and potential future models to predict immunogenicity of therapeutic proteins. It gives insight into the rationale of each of the models and discusses the specific aspects of immunogenicity predicted by them. It ends with recommendations on future studies that need to be performed in order to improve predictability.

Models predicting CD4+ T cell epitopes and CD4+ T cell activation

Most of the *in silico* and *in vitro* models used to predict immunogenicity of therapeutic proteins focus on identifying CD4+ T helper cell epitopes and measuring activation of CD4+ T cells (Table 1). In an adaptive immune response against foreign proteins, CD4+ T cells and their epitopes are crucial for the induction of an immune response, which is characterized by isotype switched antibodies such as IgG, by affinity maturation, and the formation of immunological memory. The observation that some patients treated with therapeutic proteins produce high affinity, isotype switched antibodies, suggests that ADA immunogenicity in these cases is driven via a CD4+ T cell dependent mechanism (16), involving T cell epitopes present in the protein sequence. Presentation of these epitopes by major histocompatibility complex (MHC) class II molecules on antigen presenting cells (APC) can engage T cells to initiate a cascade of events resulting in an ADA response by B cells. Assuming that therapeutic proteins evoke an antibody response via this T cell-dependent mechanism, the prediction of T cell epitopes and a corresponding T cell response could be an effective way to identify immunogenic sequences and, by eliminating them, reduce the potential for immunogenicity. The main methods employed in detecting CD4+ T cell epitopes and CD4+ T cell responses to proteins are (i) *in silico* analysis of MHC class II binding peptides and (ii) *in vitro* T cell stimulation. Both techniques enable CD4+ T cell epitopes to be predicted in the context of human MHC class II.

In silico models

In silico models use the amino acid sequence of therapeutic proteins to predict the presence of peptides in these proteins that bind to MHC molecules. Several first-generation models are based on quantitative matrices. They use experimental data of the many peptides known to bind to specific HLA allotypes and in addition they score each of the amino acids depending on their position in the binding groove. Even though this approach has been mainly applied to MHC class I (which is driving cytotoxic responses), *in silico* tools for predicting the presence of MHC class II-binding epitopes such as Tepitope, MHCPred, Epimatrix and SVMHC are also developed (17–20). More recently, *in silico* models based on artificial neural networks (ANN) have been developed, which involve

data modeling tools able to "learn" which peptides could bind to MHC. The information needed during a learning phase is provided by a set of peptide sequences from both known MHC binders and confirmed non-binders, and is used by the ANN to find patterns for a prediction of new sequences. Examples of such models are ANNPrep and Comprep (21). Several neural network servers are available, such as NetCHOP, NetCTLpan and NetMHCpan (http://tools.immuneepitope.org). The ANN methods are very adaptive and have the ability to self-improve.

A drawback of the above mentioned methods is their reliance on extremely large data sets, which require intensive experimental work. To overcome this problem, some structure-based methods have been developed that also examine the three dimensional structures of the binding groove of the HLA molecules using force field analysis based on crystal structures and other structural approaches. Up until now, two models applying to HLA class II molecules have been described: Epibase and the methods developed by Davies and colleagues (22,23).

The latest *in silico* algorithms aim to combine T cell epitope identifications with predictors of proteasomal cleavage sites and transport efficiency of the peptides to the endoplasmic reticulum (where peptides are loaded onto MHC class I molecules). Combined prediction methods could indeed lead to a bridging between pure T cell epitope prediction and the actual T-helper cell stimulation by the loaded peptides, as these also take into account processes involved in antigen presentation (24). Unfortunately, the reliability of the currently available models (e.g. Fragpredict, PAProC) is still very low. Moreover, applications are mostly available for MHC class I prediction and more research on their applicability for MHC class II binding is needed.

In general, *in silico* methods allow a rapid and relatively low-cost analysis of protein sequences for peptides that bind to MHC class II. They are very useful in modeling interactions between known CD4+ T cell epitopes (identified from *in vitro* T cell assays) and MHC class II, and have shown similar epitopes as identified with *in vitro* methods (discussed later) (25–29). The use of *in silico* tools has enabled the generation of a number of therapeutic proteins in which the CD4+ T cell epitopes have been removed by mutations that disrupt binding to MHC class II. However, whereas good accuracy can be reached with some of these tools in generating a peptide map of the peptides that are capable to bind MHC class II receptors *in vitro* based on the primary protein structure, the application of the tools is limited. The major reason is that these tools are largely restricted to the prediction of interactions between peptide sequences and MHC molecules and therefore do not take into account other factors that affect immune responses, such as antigen uptake and processing by the APC, T cell activation through the

T cell receptor (TCR), tolerance of T cells to epitopes encountered during development in the thymus, and the involvement of other immune cells. Therefore, *in silico* methods are suitable to predict the presence of potential CD4+ T cell epitopes on a given protein sequence, however, information on subsequent activation of T cells and interactions among other immune components is lacking. Because these tools use primary amino acid sequence information, they do not take into account the effect of non-sequence related factors such as formulation, impurities, and aggregates on antibody response (Table I).

***In vitro* models**

The limitation of *in silico* methods in providing information on activation of CD4+ T cells can be partly overcome by *in vitro* T cell assays. *In vitro* T cell assays are used to assess the potential of whole proteins to activate CD4+ T cells, as well as map T cell epitopes using peptides spanning the sequence of interest. The assays typically involve large numbers of patient or healthy donors to represent a large proportion of human leucocyte antigen (HLA) allotypes in the world population and to reach sufficient statistical power (Table 1). There are many different methods in practice, but in general, peripheral blood mononuclear cells (PBMCs), including APCs and T cells from patients, naïve donors or antigen-exposed individuals are harvested and brought into contact with either the whole antigen or peptide fragments (mostly 15 residues per peptide overlapping 10 or 12 amino acids). Subsequently, the type and strength of the immune response can be determined by various intracellular and extracellular T cell markers. A method commonly used for the determination of T cell activation is ELISPOT. This ELISA based method detects cytokines secreted by activated T cells, such as IL-2, IL-4, or INF-γ, where INF-γ ELISPOT seems to be favored by many researchers (30,31). More recently, flow cytometry was implemented as a more direct method for detection of T cell activation by analyzing the expression of CD25 at the cell surface of CD4+ T cells (32). There is evidence that the repertoire of epitopes presented in patients is similar to the epitopes identified *in vitro*. Two independent research groups have identified CD4+ T cell epitopes in the C1 and A2 domains of Factor VIII using *in vitro* primed T cells from healthy donors (33,34). These observations have enabled the use of community donor blood for mapping T cell epitopes and determining the T cell activation potential by whole proteins (Table 1). Data from *in vitro* T cell assays, such as the number and potency (immunodominance) of individual T cell epitopes or proteins, are therefore used to predict the relative risk of activating a T cell dependent immune response between multiple variants of a therapeutic protein during pre-clinical development.

In general, the use of *in vitro* T cell assays allows the qualitative and quantitative measurement of T cell epitopes and their role in the activation of T-helper cells, which in

turn enables strategies such as T cell epitope removal to be employed. In addition, *in vitro* T cell assays can be used to monitor T cell activation and proliferation of differently formulated products or in presence of aggregates, which is not possible with *in silico* methods. However, as these assays are based on cell material from donors, batch-to-batch and donor variability makes standardization an issue (Table 1). The use of large donor pools is a requirement, making the tools relatively low-throughput at this time (20).

Combined use

Often, CD4+ T cell based *in silico* and *in vitro* tools are combined in preclinical immunogenicity prediction. Because these models simplify the complexity of the immune system and its responses, they are used to assess relative potential for immunogenicity due to the presence of CD4+ T cell epitopes and activation, between similar products directed against the same target. *In silico* tools are particularly used to screen early stage drug candidates or libraries, in order to exclude the protein variants or designs that show a significantly higher number of potential T cell epitopes compared to other variants. *In vitro* models in particular are also used to study the effect of product-related factors other than primary structure on T cell activation. They are, however, less suitable to predict aspects of immunogenicity that involve complex immune processes such as breaking of immune tolerance (discussed later on), incidence of antibody formation, and clinical consequences of ADAs.

Models Predicting B cell epitopes and B cell activation

B cell epitopes are important in eliciting an immune response against bacteria and viruses. A repeated array of B cell epitopes on the surface of these microorganisms can bind to multiple B cell receptors on the B cell surface, and by crosslinking them, directly activate B cells to give an antibody response. This type of immune response is different from the classical T cell dependent immune responses in many ways, but two of the most important features are that T cells are not necessarily involved in antibody formation and that immunological memory formation against foreign antigens is absent (35). Clinical studies on patients treated with therapeutic interferon beta and an anti-TNF antibody have shown that patients being antibody positive during first treatment did not show a fast increase in antibody titers when treatment was restarted. This indicates that no immunological memory was formed (36,37). Although a limited number of patients was included, the data suggest that repeated B cell epitopes and crosslinking of B cell receptors could be important in immunogenicity. It might be hypothesized that protein aggregates could

express repeated cell epitopes needed to crosslink B cell receptors and activate B cells. This has been suggested as a mechanism for breaking of immune tolerance (38).

Models predicting individual B cell epitopes are available, however, taking into account that the structure of proteins is highly dependent on production conditions, formulation, and handling and that, with changing structure, other B cell epitopes can form, it seems almost impossible with our current knowledge to accurately predict B cell epitopes using *in silico* models (39–41). Nonetheless, advances in their predictive value are made (42). Also it is questionable if these models would have any value for immunogenicity prediction, because individual epitopes are incapable of crosslinking B cell receptors; instead, repeated epitopes are needed for this. Models looking at repeated protein structure are therefore more likely suitable in predicting immunogenicity of therapeutic proteins. While current *in silico* methods are unsuitable for this, *in vitro* B cell models could be a solution. However, the maintenance of B cells *in vitro* is a highly complicated task and current assays using PBMCs in short term suspension or in monolayer format, do not represent *in vivo* behavior sufficiently (43).

So, for now no models are available that could predict immunogenicity of therapeutic proteins due to repeated structures or repeated B cell epitopes. Moreover, if such models would become available, they would likely encounter similar limitations as the T cell epitope models in that they would focus on a single component aspect of the immune response, and not take into account the biological complexity of the entire immune system.

In vivo models

In vivo models used to predict immunogenicity of therapeutic proteins have the advantage over *in silico* and *in vitro* tools that immunogenicity can be studied in an organism with an intact immune system. In contrast to the simplified nature of *in silico* and *in vitro* models, *in vivo* models allow the interplay between immune cells and complex processes underlying antibody formation against therapeutic proteins. However, because preclinical assessment of immunogenicity *in vivo* is expensive and time consuming, animal models are less suitable for large-scale screening. Moreover, care has to be taken that the animal models are representative for the immune processes taking place in humans. They are therefore mostly used after lead selection by *in silico* and *in vitro* models (Table 1). Similar to the models described before, the predictive value of animal models depends on the items that need prediction, on the type of therapeutic protein and on the similarity of the processes underlying immunogenicity compared to those in humans. These include the similarities or differences in pharmacokinetics, pharmacodynamics, and target binding

between humans and the species of animal model. Similar to the *in silico* and *in vitro* models, animal models cannot be used to predict incidence of immunogenicity in patients. Also the specificity of humoral antibody responses and therefore potential for clinical effect will be hard to predict. However, they might be used to assess relative immunogenicity, presence of 'neo-epitopes' and breaking of immune tolerance (Table 1). We assume that animal models with an immune system that is genetically most similar to the human immune system are most predictive. Therefore, conventional animal models such as rats and mice would have least predictive value, while transgenic animal models and non-human primates would have highest predictive value. Recently developed animal models such as the human xenograft mouse models are being investigated for immunogenicity prediction.

Conventional Animal Models

Animal models such as rats and mice have been often used in the early years of preclinical immunogenicity prediction. However, most human therapeutic proteins are foreign proteins (i.e. have limited sequence homology) for these animals and as a result they will usually develop an ADA response against a foreign protein. This may not be informative, as the exact mechanisms underlying immunogenicity might be different in humans (38). Even when the therapeutic protein is foreign in both humans and the animal model (e.g., plant derived or bacterial proteins), species differences in the immune system, and restriction in genetic diversity between animals (in the case of inbred strains) might introduce false results.

When assessing the predictive value of conventional animal models, it is expected that they overestimate immunogenicity development in patients since rats and mice are likely to form antibodies against all (recombinant human) therapeutic proteins (44). This also implies that the ADAs will mostly be neutralizing. Therefore these animals are insensitive to discriminate between binding and neutralizing antibody responses which both can occur in patients and are therefore unsuitable to predict clinical relevance of antibody formation. In addition, Katsutani et al. (45) have shown that wildtype mice seem unsuitable to assess the presence of neoepitopes. Using human tissue plasminogen activator as antigen, they have shown that site specific modification does not lead to increased recognition of epitopes in these mice. Because these animals already recognize multiple epitopes due to foreignness of the protein, the prediction of neo-epitopes is very difficult, especially when taking into account species differences in MHC class II. However, for some proteins, rats and mice might be of value to determine the relative immunogenicity between products of the same class. For example, Bellomi et al. (46) have used BALB/c mice to assess relative difference between interferon beta 1a formulations.

They found that a new formulation of interferon beta 1a was less immunogenic compared to commercially available formulations, Avonex and Rebif.

Mice Rendered Immune Tolerant to Human Proteins

In order to prevent therapeutic proteins from inducing an ADA response in mice due to their foreignness, transgenic mice that express a human protein have been developed. As a result these mice are, like humans, immune tolerant for the particular human protein they express. Studies in such mice have shown that the immunogenicity of clinical preparations of recombinant human interferon alpha, interferon beta and monoclonal antibodies (mAbs) is significantly enhanced by the presence of aggregates (47–50). In particular, aggregates induced by metal catalyzed oxidation and aggregates composed of monomers that still exhibit native structural elements appear most immunogenic. However, by using these models it is not possible to predict what level of aggregation is needed to induce an antibody response in patients. These models have shown to predict relative immunogenicity of interferon beta products (51), with the most immunogenic product in patients (Betaferon) being more immunogenic in these mice compared to other products such as Avonex and Rebif. However absolute incidences of antibody positive individuals differed between the immune tolerant mice and patients.

Transgenic mouse models also have been shown to predict neo-epitopes when given a modified form of human insulin and tissue plasminogen activator (52,53). So, these models can therefore be used to determine relative immunogenicity of protein variants and formulations. Moreover, studies conducted with immune tolerant mice have shown that although being tolerant for human growth hormone, an immune response could be induced when these mice were treated with a sustained-release formulation. This illustrates that – in addition to predicting immunogenicity due to aggregation, relative immunogenicity, and neo-epitopes – these models can be used to study breaking of immune tolerance (54). Immune tolerant murine models are, however, limited by their inability to predict the incidence of immunogenicity or clinical consequences of ADA formation (Table 1).

A major disadvantage of the immune tolerant mice is that they, like conventional animal models, respond against a therapeutic protein via a rodent immune system. If the mechanisms underlying immunogenicity are T cell (epitope) triggered, absence of human MHC class II in these mice likely limits the usefulness of such models. In turn, differences in B cell repertoire might affect prediction for B cell epitopes if these appear to be the trigger for immunogenicity (Table 1).

Non-human primates

Because proteins expressed in humans and non-human primates show a high degree of homology, non-human primates are expected to be immune tolerant for most human proteins. Also their immune system is more similar to the human immune system compared to rodent models and transgenic mice. Therefore, the mechanisms underlying the antibody response in non-human primates would, in theory, better reflect the human immune response against therapeutic proteins. Non-human primates such as chimpanzees and rhesus monkeys have been shown to predict the presence of neo-epitopes and relative immunogenicity of protein structural variants of various human proteins such as tissue plasminogen activator, growth hormone and insulin (45,55,56). In theory, they might also be suitable to study breaking of immune tolerance for therapeutic proteins, which are similar to their endogenous proteins. In one occasion non-human primates have also been shown to predict development of neutralizing (cross-reactive) antibodies to thrombopoietin that was also observed clinically (57). However, is questionable if this is generally applicable to other therapeutic proteins (Table 1). Despite their apparent superiority as predictive model, non- human primates are incapable of predicting incidence of immunogenicity in patients. Moreover, non-human primates cannot be used to predict immunogenicity of all therapeutic proteins; apparently their predictive value strongly depends on the protein in question. For example, for interleukin 3 they have shown very poor predictability (58). This implies that the predictive value of these models is only known for already tested proteins (Table 1).

HLA Transgenic Mice

Mice expressing specific human HLA allotypes (and lacking endogenous mouse MHC class II) have been developed and used for research to evaluate the involvement of human HLA alleles in indications such as allergy and autoimmune diseases (59,60). Applications of these models in predicting the immunogenicity of protein therapeutics are currently being developed. These models will be particularly valuable when immunogenicity is driven by CD4+ T cell epitopes. To improve the suitability of these mice, they should be crossed with mice where immune tolerance against a specific recombinant therapeutic or class of therapeutics is induced by either transgenic expression of the protein of interest or induction of tolerance during neonatal development (61,62). For example, in order to produce a model that might be suitable to predict the potential immunogenicity of mAbs, transgenic mice that express (monoclonal) human immunoglobulin could be bred with transgenicmice that express human HLA alleles (63–66). In order to avoid generating mice that are tolerant to both human and murine mAb variable region sequences, these mice should not express endogenous mouse MHC class II and mouse immunoglobulins.

However, obtaining HLA-diversity, which is comparable to that of the human population, will be a significant challenge.

Human Immune System Xenograft Models

As an alternative to transgenic mice, models based on immunodeficient NOD scid IL2Rγ/- or Rag2−/−γc−/− mice are being developed. These mice lack functional mouse T- and B cells, have no functional complement system, have diminished mouse NK functioning, and lack mouse macrophage activity. These mice have shown to be very successful for engraftment of human immune cells and therefore have a functional human-like immune system (67,68). Neonatal immunodeficient mice are used for engraftment of CD34+ human hematopoetic progenitor cells, which can be isolated from fetal human tissue. This engraftment leads to the reconstitution of 40–60% of human CD45+ mononuclear cells in peripheral blood and spleen, and gives sizable compartments of human B cells, T cells, natural killer cells, monocyte/macrophages, and dendritic cells. Since these mice express human MHC class II and should be tolerant to human immunoglobulins, they might be suitable for the prediction of the immunogenicity of therapeutic proteins including mAbs. In addition, as the biological activity of the therapeutic target probably plays an important role in immunogenicity (e.g., soluble versus membrane-bound), these models offer the advantage to express many therapeutic targets (e.g., TNF, BAFF, CD3, CD20) as human (69–71).

There are, however, limitations in using some of the currently available engraftment models. First, they are not tolerant against all human proteins. Second, there is no germline transfer of genes encoding human immune cells, so each mouse has to be generated on an individual basis. As shown in some studies, this may lead to considerable variability in immune responses to antigens that stimulate potent responses in humans. Furthermore, some strains of these mice do not express HLA molecules on thymic epithelial cells. Consequently, human T cells developing in these humanized mice lack the ability to recognize antigens in an HLA-restricted manner, precluding the investigation of human T cell responses against therapeutic proteins (72). However BLT mice, which are immunodeficient mice in which human liver and thymus fragments are implanted under the renal capsule and which are given additional haematopoietic stem cells intravenously, do have HLA restricted T cells (73). However B cell responses in these BLT mice appear to be generally limited to IgM, potentially due to immature lymph node architecture.

General Conclusions and Recommendations

The main limitations in predictive value of the models presented in this paper are (i) insufficient knowledge on the interplay of immune mechanisms underlying immunogenicity of therapeutic proteins and (ii) insufficient clinical validation. Future studies should therefore address these two topics.

The mechanisms underlying immunogenicity of therapeutic proteins are not well studied. For example, it is still uncertain whether the primary mechanism by which therapeutic proteins induce antibodies is driven via a T cell dependent mechanism, via repeated B cell epitopes, via another yet unknown mechanism, or whether immunogenicity is a combination of all of these. We also do not know whether there is a general immune mechanism explaining immunogenicity of all (recombinant human) therapeutics, or if this mechanism is product specific. Special attention should be taken when considering proteins that are non-human, vs. human proteins in patients with endogenous counterparts vs. human proteins used in replacement therapy for patients deficient in the endogenous counterpart. To answer these questions, more studies in animals, but also more in depth studies in patients are needed. One of the priorities should be to elucidate to what extent T- and B cell epitopes are triggering ADA formation, and if there is HLA restriction in this response. Also we should focus on understanding contributions of aggregates. These are considered one of the major risk factors of immunogenicity. However, despite numerous publications we still do not know which specific types of aggregates are immunogenic and why they can induce ADAs. Is this because of better uptake by APC or are they capable of directly activating B cells? It is also not clear whether low levels of aggregates found in many therapeutic proteins play a role in the protein immunogenicity. In addition, insight in treatment and patient-related factors affecting immunogenicity should be gained. For example, we do not know if a patient forming antibodies against a certain drug can be retreated with that same or a similar drug on a later occasion without having a memory response. For now this is (almost) not studied, although sparse clinical data suggests that this might be possible for some therapeutic proteins (36,37). We also need more data on why some individuals form ADAs and others do not, while being treated with the same drug.

Another focus should be on validating the current predictive models. Data from *in silico*, *in vitro,* and *in vivo* models should be combined with clinical data in order to answer questions like: Does the removal of T cell or B cell epitopes lower immunogenicity in patients? And to what extent are *in vivo* models capable of predicting immunogenicity in patients? Clinical immunogenicity data comparing the original and corresponding “deimmunized” variants of the same protein species should give insight into the effect of

predicted epitope removal on immunogenicity. Also comparing predicted CD4+ T cell epitopes, *in silico,* and *in vitro*, with actual peptides recognized by MHC II in patients would be needed to validate the suitability of epitope prediction by these models. The assessment of predictive value of animal models might be achieved by comparing antibody incidences of different products between animals and patients. As mentioned before, parameters such as antibody titer and clinical effect of ADAs might not be suitable in assessing predictive value. Foundations such as the European Immunogenicity Platform (www.e-i-p.eu) gather experts in the field to discuss these items and to start collaborations aiming to answer some of the questions mentioned above. However, the studies comparing *in silico*, *in vitro*, *in vivo* and clinical data encounter some challenges. *In silico*, *in vitro,* and *in vivo* models are used to predict relative potential for immunogenicity between different products during developmental stages. In order to compare these results with clinical data, the same products should be given to patients. This poses a problem. Clinical testing will not involve multiple drug lead candidates. Also chances are that drugs given to patients in clinical testing will have different formulation, impurities and aggregation profiles than those during early development. A solution would be to include a reference drug during preclinical testing that has a known immunogenicity profile in patients. It is critical that such reference exhibits similar characteristics to the test product, such as target binding, size, and protein class, since these characteristics could all influence immunogenicity. For new drugs, having a reference with similar characteristics might be very challenging. These references, however, are very likely available for biosimilar development in the form of the original product against which the biosimilar should be tested.

Conclusion

The predictive value of the current *in silico*, *in vitro,* and *in vivo* models used to assess immunogenicity of therapeutic proteins is uncertain and in several cases only partial answers are obtained. In order to gain more knowledge about their predictive value and to potentially improve existing models, clinical validation and increased insights into the immune mechanism underlying immunogenicity should be aimed for. Predicted immunogenicity in these models may therefore not lead to a go/no go decision on individual drug leads, but instead could be used in the selection of one drug candidate over another for further (clinical) development. *In silico*, together with *in vitro* models would be most suitable to screen multiple drug leads for potential immunogenicity due to T- or B cell epitopes, activation of T- and B cells or due to a particular formulation. A selection of these leads, with assumed lowest immunogenicity potential, would then be tested for capability to form ADAs in animal models. These models could give an indication

of their relative potential immunogenicity by studying antibody incidences. Ideally for all predictive models, a reference product with known immunogenicity in patients would be tested in parallel. It is critical that such a product exhibits similar characteristics, such as target binding, size, and protein class. With the use of such a reference, better insight into immunogenicity potential of drug leads might be possible, however, it appears unlikely that for new drugs such reference products would be available. For now, clinical testing will stay critical for determining actual immunogenicity in patients.

References

1. Clarke JB. Mechanisms of adverse drug reactions to biologics. Handbook of Experimental Pharmacology. 2010. p. 453–74.

2. Bartelds GM, Wijbrandts C a, Nurmohamed MT, Stapel S, Lems WF, Aarden L, et al. Anti-infliximab and anti-adalimumab antibodies in relation to response to adalimumab in infliximab switchers and anti-tumour necrosis factor naive patients: a cohort study. Ann Rheum Dis. 2010;69(5):817–21.

3. Bertolotto A, Deisenhammer F, Gallo P, Sölberg Sørensen P. Immunogenicity of interferon beta: differences among products. J Neurol. 2004 Jun;251 Suppl(S2):II15–24.

4. De Vries MK, Wolbink GJ, Stapel SO, de Vrieze H, van Denderen JC, Dijkmans B a C, et al. Decreased clinical response to infliximab in ankylosing spondylitis is correlated with anti-infliximab formation. Ann Rheum Dis. 2007;66:1252–4.

5. Schellekens H. Immunologic mechanisms of EPO-associated pure red cell aplasia. Best Pract Res Clin. 2005;18(3):473–80.

6. McKoy JM, Stonecash RE, Cournoyer D, Rossert J, Nissenson AR, Raisch DW, et al. Epoetin-associated pure red cell aplasia: Past, present, and future considerations. Transfusion. 2008. p. 1754–62.

7. Yang J, Joo KW, Kim YS, Ahn C, Han JS, Kim S, et al. Two cases of pure red-cell aplasia due to anti-erythropoietin antibodies. J Nephrol. 2005;18(1):102–5.

8. Brinks V, Jiskoot W, Schellekens H. Immunogenicity of therapeutic proteins: The use of animal models. Pharmaceutical Research. 2011. p. 2379–85.

9. Bryson CJ, Jones TD, Baker MP. Prediction of immunogenicity of therapeutic proteins: validity of computational tools. BioDrugs. 2010 Jan;24(1):1–8.

10. Wullner D, Zhou L, Bramhall E, Kuck A, Goletz TJ, Swanson S, et al. Considerations for optimization and validation of an *in vitro* PBMC derived T cell assay for immunogenicity prediction of biotherapeutics. Clin Immunol. 2010;137(1):5–14.

11. Schellekens H. How to predict and prevent the immunogenicity of therapeutic proteins. Biotechnology Annual Review. 2008. p. 191–202.

12. Schellekens H. Factors influencing the immunogenicity of therapeutic proteins. Nephrol Dial Transplant. 2005;20(SUPPL. 6).

13. Stas P, Lasters I. Strategies for preclinical immunogenicity assessment of protein therapeutics. IDrugs. 2009;12(3):169–73.

14. De Groot a S, Knopp PM, Martin W. De-immunization of therapeutic proteins by T-cell epitope modification. Dev Biol (Basel). 2005;122:171–94.

15. Perry LCA, Jones TD, Baker MP. New approaches to prediction of immune responses to therapeutic proteins during preclinical development. Drugs in R and D. 2008. p. 385–96.

16. Baker MP, Reynolds HM, Lumicisi B, Bryson CJ. Immunogenicity of protein therapeutics: The key causes, consequences and challenges. Self Nonself. 2010;1(4):314–22.

17. Bian H, Hammer J. Discovery of promiscuous HLA-II-restricted T cell epitopes with TEPITOPE. Methods. 2004 Dec;34(4):468–75.

18. De Groot AS, Jesdale BM, Szu E, Schafer JR, Chicz RM, Deocampo G. An interactive Web site providing major histocompatibility ligand predictions: application to HIV research. AIDS Res Hum Retroviruses. 1997 May;13(7):529–31.

19. Dönnes P, Kohlbacher O. SVMHC: a server for prediction of MHC-binding peptides. Nucleic Acids Res. 2006 Jul;34(Web Server issue):W194–7.

20. Van Walle I, Gansemans Y, Parren PWHI, Stas P, Lasters I. Immunogenicity screening in protein drug development. Expert Opin Biol Ther. 2007 Mar;7(3):405–18.

21. Lata S, Bhasin M, Raghava GPS. Application of machine learning techniques in predicting MHC binders. Methods Mol Biol. 2007 Jan;409:201–15.

22. Davies MN, Sansom CE, Beazley C, Moss DS. A novel predictive technique for the MHC class II peptide-binding interaction. Mol Med. 9(9-12):220–5.

23. Desmet J, Meersseman G, Boutonnet N, Pletinckx J, De Clercq K, Debulpaep M, et al. Anchor profiles of HLA-specific peptides: analysis by a novel affinity scoring method and experimental validation. Proteins. 2005 Jan;58(1):53–69.

24. Lundegaard C, Hoof I, Lund O, Nielsen M. State of the art and challenges in sequence based T-cell epitope prediction. Immunome Research. 2010.

25. Fonseca SG, Coutinho-Silva A, Fonseca LAM, Segurado AC, Moraes SL, Rodrigues H, et al. Identification of novel consensus CD4 T-cell epitopes from clade B HIV-1 whole genome that are frequently recognized by HIV-1 infected patients. AIDS. 2006;20(18):2263–73.

26. Fonseca CT, Cunha-Neto E, Goldberg AC, Kalil J, De Jesus AR, Carvalho EM, et al. Identification of paramyosin T cell epitopes associated with human resistance to Schistosoma mansoni reinfection. Clin Exp Immunol. 2005;142(3):539–47.

27. Veeraraghavan S, Renzoni E a, Jeal H, Jones M, Hammer J, Wells a U, et al. Mapping of the immunodominant T cell epitopes of the protein topoisomerase I. Ann Rheum Dis. 2004;63(8):982–7.

28. Iwai LK, Yoshida M, Sidney J, Shikanai-Yasuda MA, Goldberg AC, Juliano MA, et al. *In silico* Prediction of Peptides Binding to Multiple HLA-DR Molecules Accurately Identifies Immunodominant Epitopes from gp43 of Paracoccidioides brasiliensis Frequently Recognized in Primary Peripheral Blood Mononuclear Cell Responses from Sensitized Ind. Mol Med. 2003;9(9-12):209–19.

29. Kapoerchan V V., Wiesner M, Hillaert U, Drijfhout JW, Overhand M, Alard P, et al. Design, synthesis and evaluation of high-affinity binders for the celiac disease associated HLA-DQ2 molecule. Mol Immunol. 2010;47(5):1091–7.

30. Anthony D. T-cell epitope mapping using the ELISPOT approach. Methods. 2003 Mar;29(3):260–9.

31. Mashishi T, Gray CM. The ELISPOT assay: an easily transferable method for measuring cellular responses and identifying T cell epitopes. Clin Chem Lab Med. 2002 Sep;40(9):903–10.

32. James EA, LaFond R, Durinovic-Bello I, Kwok W. Visualizing antigen specific CD4+ T cells using MHC class II tetramers. J Vis Exp. 2009 Jan;(25).

33. Hu G-L, Okita DK, Conti-Fine BM. T cell recognition of the A2 domain of coagulation factor VIII in hemophilia patients and healthy subjects. J Thromb Haemost. 2004;2(11):1908–17.

34. Jones TD, Phillips WJ, Smith BJ, Bamford CA, Nayee PD, Baglin TP, et al. Identification and removal of a promiscuous CD4+ T cell epitope from the C1 domain of factor VIII. J Thromb Haemost. 2005;3(5):991–1000.

35. Mond JJ. T cell independent antigens. Current Opinion in Immunology. 1995. p. 349–54.

36. Perini P, Facchinetti A, Bulian P, Massaro AR, Pascalis DD, Bertolotto A, et al. Interferon-beta (INF-beta) antibodies in interferon-beta1a- and interferon-beta1b-treated multiple sclerosis patients. Prevalence, kinetics, cross-reactivity, and factors enhancing interferon-beta immunogenicity *in vivo*. Eur Cytokine Netw. 2001;12(1):56–61.

37. Ben-Horin S, Mazor Y, Yanai H, Ron Y, Kopylov U, Yavzori M, et al. The decline of anti-drug antibody titres after discontinuation of anti-TNFs: Implications for predicting re-induction outcome in IBD. Aliment Pharmacol Ther. 2012;35(6):714–22.

38. Sauerborn M, Brinks V, Jiskoot W, Schellekens H. Immunological mechanism underlying the immune response to recombinant human protein therapeutics. Trends Pharmacol Sci. 2010 Feb;31(2):53–9.

39. Blythe MJ, Flower DR. Benchmarking B cell epitope prediction: underperformance of existing methods. Protein Sci. 2005 Jan;14(1):246–8.

40. De Groot AS, McMurry J, Moise L. Prediction of immunogenicity: *in silico* paradigms, ex vivo and *in vivo* correlates. Curr Opin Pharmacol. 2008 Oct;8(5):620–6.

41. Greenbaum JA, Andersen PH, Blythe M, Bui H, Cachau RE, Crowe J, et al. Towards a consensus on datasets and evaluation metrics for developing B-cell epitope prediction tools. J Mol Recognit. 2007;20(2):75–82.

42. Giaco L, Amicosante M, Fraziano M, Gherardini PF, Ausiello G, Helmer-Citterich M, et al. B-Pred, a structure based B-cell epitopes prediction server. Adv Appl Bioinform Chem. 2012;5:11–21.

43. Giese C, Lubitz A, Demmler CD, Reuschel J, Bergner K, Marx U. Immunological substance testing on human lymphatic micro-organoids *in vitro*. J Biotechnol. 2010;148(1):38–45.

44. Koren E, Zuckerman L a, Mire-Sluis a R. Immune responses to therapeutic proteins in humans--clinical significance, assessment and prediction. Curr Pharm Biotechnol. 2002;3(4):349–60.

45. Katsutani N, Yoshitake S, Takeuchi H, Kelliher JC, Couch RC, Shionoya H. Immunogenic properties of structurally modified human tissue plasminogen activators in chimpanzees and mice. Fundam Appl Toxicol. 1992;19(4):555–62.

46. Bellomi F, Mute A, Palmieri G, Focaccetti C, Dianzani C, Mattel M, et al. Immunogenicity comparison of Interferon Beta-1a preparations using the BALB/c mouse model: Assessment of a new formulation for use in multiple sclerosis. New Microbiol. 2007;30(3):241–6.

47. Braun A, Kwee L, Labow MA, Alsenz J. Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-alpha) in normal and transgenic mice. Pharm Res. 1997;14(10):1472–8.

48. Hermeling S, Schellekens H, Maas C, Gebbink MFBG, Crommelin DJA, Jiskoot W. Antibody response to aggregated human interferon alpha2b in wild-type and transgenic immune tolerant mice depends on type and level of aggregation. J Pharm Sci. 2006 May;95(5):1084–96.

49. Van Beers MMC, Sauerborn M, Gilli F, Brinks V, Schellekens H, Jiskoot W. Oxidized and aggregated recombinant human interferon beta is immunogenic in human interferon beta transgenic mice. Pharm Res. 2011;28(10):2393–402.

50. Filipe V, Jiskoot W, Basmeleh AH, Halim A, Schellekens H, Brinks V. Immunogenicity of different stressed IgG monoclonal antibody formulations in immune tolerant transgenic mice. MAbs. 2012;4(6):740–52.

51. Van Beers MMC, Sauerborn M, Gilli F, Hermeling S, Brinks V, Schellekens H, et al. Hybrid transgenic immune tolerant mouse model for assessing the breaking of B cell tolerance by human interferon beta. J Immunol Methods. Elsevier B.V.; 2010 Jan;352(1-2):32–7.

52. Ottesen JL, Nilsson P, Jami J, Weilguny D, Dührkop M, Bucchini D, et al. The potential immunogenicity of human insulin and insulin analogues evaluated in a transgenic mouse model. Diabetologia. 1994 Dec;37(12):1178–85.

53. Stewart TA, Hollingshead PG, Pitts SL, Chang R, Martin LE, Oakley H. Transgenic mice as a model to test the immunogenicity of proteins altered by site-specific mutagenesis. Mol Biol Med. 1989 Aug;6(4):275–81.

54. Lee HJ, Riley G, Johnson O, Cleland JL, Kim N, Charnis M, et al. *In vivo* characterization of sustained-release formulations of human growth hormone. J Pharmacol Exp Ther. 1997;281(3):1431–9.

55. Zwickl CM, Cocke KS, Tamura RN, Holzhausen LM, Brophy GT, Bick PH, et al. Comparison of the immunogenicity of recombinant and pituitary human growth hormone in rhesus monkeys. Fundam Appl Toxicol. 1991 Feb;16(2):275–87.

56. Zwickl CM, Smith HW, Zimmermann JL, Wierda D. Immunogenicity of biosynthetic human LysPro insulin compared to native-sequence human and purified porcine insulins in rhesus monkeys immunized over a 6-week period. Arzneimittelforschung. 1995 Apr;45(4):524–8.

57. Food and Drug Administration. Center for Biologics Evaluation and Research. Meeting of the Biological Response Modifiers Advisory Committee. In: Bethesda MD. 1999.

58. Gunn H. Immunogenicity of recombinant human interleukin-3. Clin Immunol Immunopathol. 1997 Apr;83(1):5–7.

59. Black KE, Murray J a, David CS. HLA-DQ determines the response to exogenous wheat proteins: a model of gluten sensitivity in transgenic knockout mice. J Immunol. 2002;169(10):5595–600.

60. Rosloniec EF, Brand DD, Myers LK, Whittington KB, Gumanovskaya M, Zaller DM, et al. An HLA-DR1 transgene confers susceptibility to collagen-induced arthritis elicited with human type II collagen. J Exp Med. 1997;185(6):1113–22.

61. Madoiwa S, Yamauchi T, Hakamata Y, Kobayashi E, Arai M, Sugo T, et al. Induction of immune tolerance by neonatal intravenous injection of human factor VIII in murine hemophilia A. J Thromb Haemost. 2004;2(5):754–62.

62. Madoiwa S, Yamauchi T, Kobayashi E, Hakamata Y, Dokai M, Makino N, et al. Induction of factor VIII-specific unresponsiveness by intrathymic factor VIII injection in murine hemophilia A. J Thromb Haemost. 2009;7(5):811–24.

63. Fishwild DM, O'Donnell SL, Bengoechea T, Hudson D V, Harding F, Bernhard SL, et al. High-avidity human IgG kappa monoclonal antibodies from a novel strain of minilocus transgenic mice. Nat Biotechnol. 1996;14(7):845–51.

64. Geluk a, Taneja V, van Meijgaarden KE, Zanelli E, Abou-Zeid C, Thole JE, et al. Identification of HLA class II-restricted determinants of Mycobacterium tuberculosis-derived proteins by using HLA-transgenic, class II-deficient mice. Proc Natl Acad Sci U S A. 1998;95(18):10797–802.

65. Mendez MJ, Green LL, Corvalan JR, Jia XC, Maynard-Currie CE, Yang XD, et al. Functional transplant of megabase human immunoglobulin loci recapitulates human antibody response in mice. Nat Genet. 1997;15(2):146–56.

66. Neeno T, Krco CJ, Harders J, Baisch J, Cheng S, David CS. HLA-DQ8 transgenic mice lacking endogenous class II molecules respond to house dust allergens: identification of antigenic epitopes. J Immunol. 1996;156(9):3191–5.

67. Ishikawa F, Yasukawa M, Lyons B, Yoshida S, Miyamoto T, Yoshimoto G, et al. Development of functional human blood and immune systems in NOD/SCID/IL2 receptor {gamma} chain(null) mice. Blood. 2005 Sep;106(5):1565–73.

68. Traggiai E, Chicha L, Mazzucchelli L, Bronz L, Piffaretti J-C, Lanzavecchia A, et al. Development of a human adaptive immune system in cord blood cell-transplanted mice. Science. 2004;304(5667):104–7.

69. Chan AC, Carter PJ. Therapeutic antibodies for autoimmunity and inflammation. Nat Rev Immunol. 2010;10(5):301–16.

70. Getts DR, Getts MT, McCarthy DP, Chastain EML, Miller SD. Have we overestimated the benefit of human(ized) antibodies? mAbs. 2010. p. 682–94.

71. Scott DW, De Groot AS. Can we prevent immunogenicity of human protein drugs? Ann Rheum Dis. 2010 Jan;69 Suppl 1:i72–6.

72. Watanabe Y, Takahashi T, Okajima A, Shiokawa M, Ishii N, Katano I, et al. The analysis of the functions of human B and T cells in humanized NOD/shi-scid/gammac(null) (NOG) mice (hu-HSC NOG mice). Int Immunol. 2009;21(7):843–58.

73. Shultz LD, Brehm M a, Garcia-Martinez JV, Greiner DL. Humanized mice for immune system investigation: progress, promise and challenges. Nat Rev Immunol. 2012;12(11):786–98.