Chapter 10
ICH S6 Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals

Joy Cavagnaro and Jennifer Sims

Abstract Since its publication in 1997, the ICH guideline on the Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals ("ICH S6") has fostered consistency while maintaining the necessary flexibility for testing within and across a variety of product classes. Successful implementation of a product-specific science-based "case-by-case" approach however has required individuals with a broad knowledge of toxicological processes and the ability to integrate data from molecular biology, pharmacology, physiology, pharmacokinetics, and pathology. Importantly, the "case-by-case" approach only works if there is an understanding of the science and an acceptance by both regulators and industry that the interpretation of the data has to reflect best scientific practice and that no study in experimental animals can predict with certainty the outcome in humans. As such, a greater dialogue between industry and regulatory authorities has been needed early and in some cases throughout development to ensure that the decision on how a product should be tested not only meets the stringencies of the regulatory authorities but is also designed to improve the predictive value for extrapolating to humans. This dialogue between industry and regulatory authorities continued to the ICH Expert Working Group charged with formulating the addendum to ICH S6(ICH S6R(1) finalized at step 4 in June 2011), guidance based on the accumulated and collective experience of the safety assessment of biotechnology-derived pharmaceuticals in the 14 years since the finalization of ICH S6.

J. Cavagnaro (✉)
Access BIO, Boyce, VA, USA
e-mail: jcavagnaro@accessbio.com

J. Sims
Integrated Biologix GmbH, Basel, Switzerland

10.1   Introduction

10.1.1   Historical Perspective

In the early 1980s, neither industry toxicologists nor regulatory scientists were sure of what constituted an appropriate toxicological assessment program for biopharmaceuticals. There were even some who believed that natural proteins were inherently safe thus the toxicity should be minimal or not relevant. However, in 1986, the biotechnology working party was established in Europe to focus on specific issues related to the development of biotechnology-derived pharmaceuticals. In July of that same year, a satellite symposium to the IV International Congress of Toxicology was held at the Keio Plaza Hotel, Tokyo, Japan. Attendees included government regulatory scientists, university scientists, and industrial scientists and research managers, all with an interest in the development of new biotechnology-derived products (Giss 1987; Dayan 1987; Galbraith 1987; Finkle 1987; Zbinden 1987).

10.1.2   Proposal for a Specific Guidance for Preclinical Safety Evaluation of Biotechnology-Derived Pharmaceuticals

Five years later, at the first ICH meeting in Brussels, Belgium, in 1991, it was questioned whether differing attitudes among the various regions towards development of biotechnology-derived pharmaceuticals were considered significant enough to actually justify a session. A session was held, and a "rational science-based approach" was acknowledged as critical to the successful and expeditious development of new and novel products (Hayakawa 1992; Cavagnaro 1992a; Hohbach 1992). One of the issues addressed in the workshop was whether common standards and attitudes that were evolving could be maintained without the issuance of formal guidance. The recommendations from the workshop were that in the short term, regulatory authorities should maintain a flexible approach to requirements for preclinical testing on a case-by-case basis, and the work should be initiated to prepare internationally accepted principles for the safety evaluation of drugs produced using biotechnology (Kikuchi 1992). Importantly, even in the early 1990s, it was recognized that the value of case-by-case for globalizing markets depended fully on a common understanding of all partners involved. If this was not achieved, there would be a continuous risk for inequality of advice on the requirements and standards from one country to another.

Supporting publications were also emerging questioning the relevance of the traditional pharmaceutical paradigm for the preclinical safety evaluation of biopharmaceuticals (Zbinden 1990, 1991; Bass et al. 1992; Hayes and Cavagnaro 1992; Cavagnaro 1992b; Claude 1992; Terrell and Green 1994; Dayan 1995; Thomas 1995; Henck et al. 1996). During this time period, there were both increases in the number of biopharmaceuticals under development and a rapidly increasing number of small companies coming into the field. At the second ICH

Table 10.1 Members of the ICH S6 Expert Working Group

Europe	Japan	United States
Dr. Jennifer Sims (EU)	Dr. Tohru Inoue (MHW)	Dr. James Green (PhRMA)
Prof. Giuseppe Vicari (EU)	Dr. Mashiro Nakadate (MHW)	Dr. Joy Cavagnaro (FDA) (Rapporteur)
Dr. Jorgen Carstensen (EFPIA)	Dr. Eliji Makai (JPMA)	
Dr. Wolfgang Neumann (EFPIA)	Dr. Mutsufumi Kawai (JPMA)	

meeting in Orlando, Florida, in 1993, biotechnology issues mainly focused on product quality issues although interest was increasing with rumblings for a more formal guidance for preclinical assessment of biotechnology-derived pharmaceuticals. Soon after this meeting, an ICH Expert Working Group (EWG) was established, and a concept paper was proposed by the FDA. A pre-step 2 document was released at the third ICH meeting in Yokohama, Japan, in 1995. A few years later in February of 1997, the 13th CMR International Workshop provided an opportunity for international experts to discuss experiences and difficulties encountered in designing scientifically based preclinical safety evaluation programs for biopharmaceuticals. This 2-day meeting brought together toxicologists and clinicians, from 32 pharmaceutical and biotechnology companies and regulators and regulatory advisors from the European Agency for the Evaluation of Medicines (EMEA, now European Medicines Agency, EMA) and 9 countries: Denmark, France, Germany, Italy, Japan, the Netherlands, Sweden, the UK, and the USA (Griffith and Lumley 1998). Recommendations arising from the CMR Workshop were taken into consideration by the expert working group for the final drafting of ICH S6 guideline, and agreement was reached at ICH 4 in Brussels in July 1997 (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997) (Table 10.1).

10.1.3   Implementation of ICH S6

Over the ensuing decade, the numbers, types, complexities, and indications for "biotech products" grew. Many of these novel products were successfully approved for market. Publications provided insight into experiences with the case-by-case approach strategies (Serabian and Pilaro 1999; Sims 2001; Ryan and Terrell 2002; Cavagnaro 2002; Brennan et al. 2004; Buckley et al. 2008). However, the explosion in new constructs and novel formats was also complicated with the arrival of second-generation products in the form of "biosimilars" and "biobetters" of the first-generation products approved for use in the 1990s. In parallel to the industry evolution, some key regulatory agencies also underwent reorganization, and there were also changes in industry access to regulatory authorities for informal and formal dialogue. This industry-regulatory evolution resulted in a combined industry-regulatory "creep" in terms of preclinical development programs to support biopharmaceuticals. A trend started to

emerge for an increasing number of questionable studies and the application of ICH guidance documents to biopharmaceuticals where biopharmaceuticals were specifically excluded in the scope of such guidance. There was also a concern for potential increases in regional guidance to aid in interpretation of ICH S6 (Nakazawa et al. 2004).

10.1.4   Rationale for Updating ICH S6

While there were reservations by some that updating ICH S6 could result in formalizing the emerging increase in studies, the perception of a considerable drift in the interpretation and application of the original intent of the ICH S6 guidance led to a series of regional industry-regulatory scientific meetings in June of 2007 to discuss specific topics identified as issues when applying the S6 guidance. The conclusions of these meetings were the need to evaluate the state of the art of safety testing of biopharmaceuticals. During this time under the auspices of BioSafe, a series of white papers were published on a series of topics (e.g., tissue cross-reactivity, species selection, immunogenicity, reproductive toxicity, carcinogenicity) and a review of scientific state-of-the-art best practice was published in Preclinical Safety Evaluation of Biopharmaceuticals: a science-based approach to facilitating clinical trials (ICH S6R). These publications would provide the necessary background for deliberations of the new ICH S6 EWG (Table 10.2).

10.1.5   Addendum to ICH S6: ICH S6(R1)

In June 2008, the ICH Steering Committee endorsed a concept paper on the proposal to establish an EWG to write an addendum to the ICH S6 guidance—the ICH S6R(1) addendum. The concept paper stated that there was a need for a clarification (and sometime amplification) of ICH S6 since substantial experience and new information has been gained since *step 4* (1997). The preclinical safety experts involved in ICH in S2/S9/M3 agreed that the flexible and case-by-case approach described in the original guidance is still valid and must be preserved. Based on the outcome of these discussions, it was agreed that the following topics would be addressed to facilitate the understanding and harmonized application of the guidance provided in S6:

- Species selection

  - How to justify the choice of a species
  - Clarify the role of tissue cross-reactivity
  - When to use a second species
  - Use of alternative models such as transgenics and homologous products

Table 10.2 Key papers outlining experiences and proposed best practices for preclinical assessment of biopharmaceuticals

---

Duration of chronic toxicity studies for biotechnology-derived pharmaceuticals: is 6 months still appropriate? (Clarke et al. 2008)[a]

Scientific review and recommendations on preclinical cardiovascular safety evaluation of biologics (Vargas et al. 2008)

Alternative strategies for toxicity testing of species-specific biopharmaceuticals (Bussiere et al. 2009)[a]

Consideration in assessing the developmental and reproductive toxicity potential of biopharmaceuticals (Martin et al. 2009)[a]

Practical approaches to dose selection for first-in-human clinical trials with novel biopharmaceuticals (Tibbitts et al. 2010)[a]

Use of tissue cross-reactivity studies in the development of antibody-based biopharmaceuticals: history, experience, methodology, and future directions (Leach et al. 2010)[a]

Carcinogenicity assessments of biotechnology-derived pharmaceuticals: a review of approved molecules and best practice recommendations (Vahle et al. 2010)[a]

Developmental toxicity testing of biopharmaceuticals in nonhuman primates: previous experience and future directions (Martin and Weinbauer 2010)

Preclinical safety evaluations supporting pediatric drug development with biopharmaceuticals: strategy, challenges, current practices (Cavagnaro 2008a)[a]

---

[a]Developed by ad hoc committees of the Biotechnology Industry Organization's (BIO) Preclinical Safety (BioSafe) Committee. BioSafe is a committee within the BIO Health Section. BioSafe is composed of BIO members working to serve as a resource for BIO members and BIO staff by identifying key scientific and regulatory issues and developments related to the preclinical safety evaluation of biopharmaceutical products and recommending appropriate science-based responses

- Study design

  – Scientific justification of duration of chronic toxicity study
  – High dose selection
  – Utility and length of recovery

- Reproductive/developmental toxicity

  – Justification of species selection including the use of rodents versus non-rodents and use of alternative models such as transgenics and homologous products
  – Considerations when using primates: use of combined study designs and timing of these studies, how to get data on fertility, impact of placental transfer, and how to get data from the F1 generation

- Carcinogenicity

  – Justification for the approach to address carcinogenic risk
  – Application of in vivo models: length of studies, use of proliferation indices, and use of homologous products

- Immunogenicity

  – Extent of characterization
  – Impact of neutralizing versus non-neutralizing

Table 10.3 Members of the ICH S6(R1) Expert Working Group

Europe	Japan	United States
Dr. Jan-Willem van der Laan (EU) (rapporteur from step 2)	Dr. Yoko Hirabayashi (NIHS)	Dr. Ann Pilaro (FDA)
	Dr. Kazushige Makai (PMDA)	Ms. Mercedes Serabian (FDA)
	Dr. Matusmoto Mineo (PMDA)	Dr. Abigail Jacobs (FDA)
Dr. Beatriz Silva Lima (EU)	Dr. Takahiro Nakazawa (JPMA)	Dr. David Jacobson-Kram (FDA)
Dr. Jennifer Sims (EFPIA) (rapporteur to step 2)	Dr. Atsushi Sanbuissho (JPMA)	
	Dr. Kazuichi Nakamura (JPMA)	Dr. Ruth Lightfoot-Dunn (PhRMA)
Dr. Maggie Dempster (EFPIA)		Dr. Helen Haggerty (PhRMA)

– Role of PD markers
– Assessment of recovery groups

ICH S6R(1) was finalized under *step 4* in June 2011. The harmonized addendum provides further complementary guidance to the S6 guidance and helps to define current recommendations and, hopefully, should reduce the likelihood that substantial differences will exist among regions. The addendum ICH S6R(1) is integrated as part II in the core S6 guideline (ICH S6R) (Table 10.3).

10.2   Definition of Biotechnology-Derived Pharmaceutical

The initial ICH S6 guidance was intended to recommend a basic framework for preclinical safety evaluation of biopharmaceuticals. Biotechnology-derived pharmaceuticals were defined as products derived from characterized cells including bacteria, yeast, insect, plant, and mammalian cells. The active substances include cytokines, growth factors, fusion proteins, toxin conjugates, enzymes, clotting factors, thrombolytics, soluble receptors, hormones, and monoclonal antibodies (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997). Importantly it was recognized that with each product class, there may also be variations. For example, over the years, monoclonal antibody products would evolve to include murine, chimeric, humanized, and fully humanized as well as "antibody-like" molecules and antibody derivatives. Products would span monospecific, bispecific, or trispecific variants; naked or conjugated; antagonist, agonist, or catalytic; targeting an endogenous epitope or a foreign epitope; with unique species specificity or with broad specificity; with no target or off-target binding on any "normal" animal species; or with specific binding to an epitope which is only upregulated in the disease state.

It was acknowledged that the principles outlined in the guidance may also be applicable to recombinant DNA protein vaccines, chemically synthesized peptides, plasma-derived products, endogenous proteins extracted from human tissue, and oligonucleotide-based drugs (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997).

10.3   Key Differences Between Biopharmaceuticals and Pharmaceuticals

Biopharmaceuticals and pharmaceuticals can be viewed as a product continuum based on size and complexity in molecular structure. However as products have evolved, there has been a blurring of product attributes. Small molecules have become larger as the result of alternative scaffolding technologies, e.g., protein conjugates and fusion proteins in order to improve exposure characteristics and dosing regimens. Large molecules have become smaller, e.g., antibody fragments and protein mimetics in order to improve distribution and decrease potential immunogenicity (Cavagnaro 2010). Novel delivery technologies are also enabling alternative routes of delivery for biopharmaceuticals, e.g., by the oral and inhalation routes. Some products such as oligonucleotide-based drugs (ONs) may have combined product attributes. For example, ONs are synthetically derived but have complex chemical profiles and are catabolized in ways similar to those followed by certain biopharmaceuticals. Although toxicity assessments are designed to address hybridization-independent effects, some ONs can also exhibit species specificity where analogous sequences may be needed to assess hybridization-dependent effects, i.e., toxicity related to exaggerated pharmacology. Thus, specific considerations are based upon product class and product attributes that influence program design. Table 10.4 provides a general comparison of product attributes across product classes. While there will be exceptions, the general distinctions provide the rationale for the different approaches to preclinical safety evaluation.

10.4   Key Considerations of ICH S6

A seminal principle of ICH S6 is that safety evaluation programs should include relevant species demonstrating pharmacological activity. Thus, a key challenge in the preclinical evaluation of biopharmaceuticals is species specificity. Unlike pharmaceuticals, one cannot assume that a molecule will be active in two species, e.g., rodent (rat or mouse) and non-rodent (rabbit, dog, nonhuman primate) traditionally used for toxicity testing. An even greater challenge is when a product is uniquely species specific, i.e., it is only pharmacologically active in humans. Determining biological activity is based on an understanding of in vitro receptor occupancy, affinity, and distribution and in vitro and in vivo pharmacological effects. Importantly, toxicity studies in nonrelevant species were discouraged (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997).

In general, 6-month duration for chronic dose studies was considered sufficient. However, it was acknowledged that specific considerations may require a longer duration study in some cases and shorter duration may also be acceptable in some cases. For example, formation of neutralizing antibodies could limit utility of longer-term dosing if there is significant impact on exposure.

Table 10.4 Comparative product attributes across product classes

Product attribute	Pharmaceutical	Oligonucleotide	Biopharmaceutical
Manufacture	Chemical synthesis	Chemical synthesis	Biological synthesis (cell culture, transgenic plants, transgenic animals)
Composition	Commonly organic chemical	Modified nucleic acids	Protein, carbohydrate
Size	Generally <1 kDa	Generally between 1 and 10 kDa	Generally >30 kDa up to 800 kDa
Purity	Homogeneous, single entity high chemical purity (except racemic mixtures)	Mostly single entity with product-related impurities	Heterogeneous mixture(microheterogeneity, aggregates)
Potency	Not determined	Not determined	Required. Generally in vivo. For some product, in vitro is acceptable
Route of administration	Oral, topical, inhalation	Parenteral, inhalation	Parenteral or targeted (e.g., intracardiac, intrathecal, and intra-articular)
Pharmacokinetics	Half-life usually minutes to hours	Half-life usually hours to days	Half-life usually days to weeks. PK can be impacted by target leading to target-mediated drug disposition (TMDD)
ADME	Rapid entry through blood capillaries, distribution to many organs/tissues, metabolized to active and non-active metabolites	Catabolized to nucleotides and other metabolites depending on modifications, selected tissue distribution	Distribution limited based on size to plasma and/or extracellular fluid, degraded (catabolism) to endogenous amino acids
Drug–drug interactions	Influenced by metabolism enzyme induction or inhibition can be significant, transporters	Not metabolized by P450	Interactions related to additive or synergistic pharmacological activity not P450-related metabolism
Dose response	Linear establishment of maximum tolerated dose (MTD)	Linear establishment of maximum tolerated dose (MTD)	Can be bell-shaped establishment of optimum biological dose (OBD), MFD
Targets	Intracellular or ligand–receptor	Intracellular and extracellular	Cell–matrix, cell–cell, or ligand–receptor
Species specificity	Species independent, preclinical assessments for general toxicity generally performed in one rodent (generally rat) and one non-rodent (generally dog)	Hybridization specific and nonspecific assessments Use of analogous/homologous molecules	Species specific. Nonhuman primates often the only relevant species Use of homologous/surrogate molecules

Toxicological effects	Unpredictable, related to MOA, chemistry, metabolites	Mostly related to chemistry	Often predictable based on known mechanism of action "exaggerated pharmacology." Animal models of disease are often used to asses safety in addition to activity
Immunogenicity (not including allergic or hypersensitivity reactions)	Never	Rare	Common. May effect exposure and exposure to bioactive drug (neutralizing antibodies)
Dosing interval for chronic use	Daily	Intermittent	Intermittent
Dose formulation	Complex	Complex, simple	Simple

Adapted Cavagnaro (1992a, b, 2002, 2008a, b, 2010)

During the implementation of ICH S6, there was a misconception that only one species was expected for assessing general toxicity of biopharmaceuticals. However, the language in ICH S6 explicitly stated "that safety evaluation programs should normally include two relevant species but, in certain justifiable cases, one relevant species may suffice (e.g., when only one relevant species can be identified or where the biological activity of the biopharmaceutical is well understood)." Importantly, the guidance intentionally did not specify use of the "most relevant" in order to avoid the routine consideration of use of higher primate species (e.g., greatest homology of a protein or a receptor with chimpanzees or baboons). There was also a growing confusion on how to define a relevant species.

10.5 Key Developments in Study Design Since ICH S6

The scientific discussions and guidance in the ICH S6 addendum ICH S6R(1) drafted by the ICH Expert Working Group was based on the accumulated experience of industry and regulators over the 14 years since ICH was finalized in 1997. A number of literature reviews on various aspects of the preclinical safety evaluation of biotechnology-derived products (see Table 10.2) were considered as well as anonymized case studies from the regulatory databases and the impact of the 2006 "Tegenero" incident in the United Kingdom.

10.5.1 Number of Species

The number of species required for safety assessment became a growing industry concern. In large part because there were requests by regulatory authorities for rodent studies with homologous products or rodent toxicology studies where the species was not a pharmacologically relevant species to satisfy the requirement for two species as standard for pharmaceuticals. The addendum therefore clarified that if there are two pharmacologically relevant species for the clinical candidate (one rodent and one non-rodent), then both species should be used for short-term general toxicology studies. The use of one species for all general toxicity studies is justified when the clinical candidate is pharmacologically active in only one species, generally the nonhuman primate. However, in such cases, where the only relevant species is the nonhuman primate, studies in a second species with a homologous product are not considered to add further value for risk assessment and are not recommended (ICH S6R).

If two relevant species exist, then short-term repeat dose toxicity studies are recommended. However, if the target organ profile is similar across species and/or similar class, effects are observed, and the dose selected in the clinical trials appears acceptable, then chronic toxicity studies in a single species may be justifiable.

10.5.2   *Selection of Relevant Species*

Clarification is provided in the addendum on the scientific data required to support the selection of a relevant species for safety assessment. This includes an evaluation of cross-species sequence homology, in vitro target binding and functional activity data, and in vivo pharmacodynamics markers such as evidence for target engagement, modulation of a known biological response, and/or pharmacological outcome. The aim of these in vitro assays and in vivo markers is to support species selection but also to provide data to make qualitative and quantitative cross-species comparison to provide confidence that a model is capable of demonstrating potentially adverse consequences of target modulation and to support translational PK–PD strategies (ICH S6R).

By 2007, the tissue cross-reactivity assay (TCR) inadvertently was becoming, either from industry or regulatory creep (or both), the primary means to select species for safety assessment of monoclonal antibodies. The history, experience, methodology, and future directions of TCR studies in the development of antibody-based biopharmaceuticals are reviewed in Leach et al. (2010). The authors state that TCR studies are screening assays recommended for antibody and antibody-like molecules that contain a complementary determining region (CDR), primarily to identify off-target binding and secondarily to identify sties of on-target binding that were not previously identified. This was also the intent of both *step 4* of ICH S6 and the FDA Points to Consider document in the manufacture and testing of monoclonal antibody products for human use (FDA 1997). This intent is now reconfirmed in note 1 of the addendum: "TCR studies are in vitro tissue-binding assays employing immunohistochemical (IHC) techniques conducted to characterize binding of monoclonal antibodies and related antibody-like products to antigenic determinants in tissues. Other technologies can be employed in place of IHC techniques to demonstrate target/binding site distribution." The addendum also clarifies the value of TCR for species selection: "assessment of TCR in animal tissues is of limited value for species selection" (ICH S6R).

The technical difficulties regarding the conduct of TCR studies are recognized, and there is an acknowledgement that a clinical candidate may not be a good immunohistochemical (IHC) reagent and a TCR study might not always be technically feasible. Issues relating to the technical conduct and interpretation of TCR studies are reviewed in detail in Leach et al. (2010) and in publication based on an industry survey on the use of the TCR IHC assay (Bussiere et al. 2011).

The addendum purposely provides very little additional guidance on the use of alternative models such as transgenic models and homologous products over the ICH S6 guidance, except to state that such models can be considered when no relevant species can be identified. The use of animal models of disease to aid safety assessment is recommended when such models are used to evaluate proof of principle for monoclonal antibodies directed at foreign targets (i.e., bacterial, viral targets, etc.). Alternative approaches for toxicity testing of species-specific biopharmaceuticals still include animal models of disease, genetically modified mice, or use of homologous product (Bussiere et al. 2009; Bussiere 2008; Bornstein et al. 2009).

10.5.3 Duration of Studies

The addendum confirmed that the duration of repeat dose toxicity studies for chronic use products and 6-month duration in rodents and non-rodents are considered sufficient. The EWG reviewed published data and anonymized case studies provided by regulatory agencies and reached the view that toxicity studies of longer duration have not generally provided useful information that changed the clinical course of development in terms of altering clinical study design or patient information (Clarke et al. 2008).

10.5.4 "Tegenero"

Another key development in the field of preclinical safety assessment of biopharmaceuticals between 1997 and 2007 was the 2006 Tegenero incident with TGN1412, a superagonistic CD28-specific monoclonal antibody, in which six healthy human volunteers had to be admitted to a critical care unit during a first-in-human (FIH) study (Suntharalingham et al. 2006). Much has been published relating to this incident including commentary on best practice in nonclinical safety assessment, setting safe starting doses for first-in-human studies, the introduction of MABEL to reemphasize the importance of taking account of the pharmacologically active dose (PAD) as well as the NOAEL and HED, and the design of FIH studies (Schneider et al. 2006; Liedert et al. 2007; Horvath and Milton 2009; Milton and Horvath 2009; Lowe et al. 2009; Hansel et al. 2010). The incident also had an impact on industry/regulatory practice and regulatory guidance such as the publication in 2007 of the CHMP guideline on strategies to identify and mitigate risks for first-in-human clinical trials with investigations on medicinal products (EMEA/CHMP/SWP/28367/07). The implications of the incident were relevant for the ICH S6R(1) discussions in relation to the use of pharmacologically relevant species for safety assessment. The data made available to the public in the IMPD did not provide evidence that the cynomolgus monkey was a pharmacologically relevant species for the safety assessment of TGN1412: data on CD28 binding affinity for cynomolgus monkey was provided in the IMPD but apparently no data on in vitro functionality (e.g., T cell proliferation) nor was there evidence for in vivo pharmacological effects even at doses resulting in full target saturation. Furthermore, other relevant data with parental and surrogate TGN1412 molecules in humanized mouse models and rodents, and in vitro human data showing T cell proliferative activity, were not used in the overall safety assessment and safe starting dose selection (Horvath and Milton 2009).

Subsequent to the incident, new data demonstrated that white blood cells from cynomolgus monkeys do not respond to TGN1412 in the same way as human white blood cells, whether the cells are stimulated in vitro or in vivo. Essentially, TGN1412 is superagonistic in humans, but not in cynomolgus monkeys (Stebbings et al. 2007,

2009). Further work by the same group at NIBSC, UK, showed that activation of CD4+ effector memory T cells by TGN1412 was likely to be responsible for the cytokine storm experienced by the healthy volunteers. Furthermore, lack of CD28 expression on the CD4+ effector memory T cells of species used for preclinical safety testing of TGN1412 offers an explanation for the failure to predict a cytokine storm in humans (Eastwood et al. 2010). This illustrates the importance of understanding the target biology and mechanism of action of the biopharmaceutical product, the selection of pharmacologically relevant species for safety assessment, and also for an understanding of the limitations of the selected animal species for predicting safety for humans and where necessary supplementing these limitations by appropriate in vitro human systems to aid optimal selection of safe starting doses for FIH studies.

Experience with many monoclonal antibodies suggests that nonhuman primates appear not to predict cytokine release well for humans (Horvath and Milton 2009), and for this reason, the Tegenero incident triggered multiple workshops and publications relating to the development of in vitro human systems to predict cytokine release with the aim of addressing this limitation of nonhuman primates (Bugelski et al. 2009; Vidal et al. 2010; Findlay et al. 2011).

10.5.5 *Dose Selection and Application of PK–PD Principles*

An example of the industry-regulatory creep that was apparent by 2007, 10 years after S6 was finalized, was high dose selection for general toxicology studies. The intent of the S6 guidance was to allow sponsors to provide a scientific justification for dose selection, tailored to the specific product attributes, to achieve the aim of understanding pharmacological/physiological and toxicological dose response relationships in a pharmacologically relevant species. The guidance acknowledged the need for a case-by-case approach such that for some classes of products with little or no toxicity, it may not be possible to define a specific maximum dose, but for products with a lower affinity to or potency in the cells of the selected species than in humans cells, testing of higher doses may be important. By 2007, requests for sponsors to use the maximum tolerated dose (MTD) or maximum feasible dose (MFD) approaches were becoming more frequent suggesting a drift towards the small molecule approach where the use of such limit doses are common.

Over the last 10 years, many sponsors began applying pharmacokinetic–pharmacodynamic (PK–PD) modeling as an integral part of the preclinical and clinical development of protein drugs (Tabrizi and Roskos 2007; Tabrizi et al. 2009; Roskos et al. 2011). Greater emphasis was placed on translational strategies using bioanalytical data from appropriately selected and well-characterized PK and PD biomarker assays to allow a quantitative relationship between protein drug exposure, target modulation and biochemical, and physiological and pathophysiological effects to be established (Roskos et al. 2011). The selection of PD biomarkers that assess target engagement and modulation and downstream cellular effects can

provide proof of mechanism and also define the magnitude and duration of target modulation following drug administration. This PK–PD data can guide the selection of doses and dosing schedules for preclinical studies and clinical trials.

The S6R(1) addendum recognized the development of these translational PK–PD approaches and recommends the use of such approaches for high dose selection in general toxicity studies by identifying (1) a dose which provides the maximum intended pharmacological effect in the preclinical species and (2) a dose which provides an approximately tenfold exposure multiple over the maximum exposure to be achieved in the clinic. Following step 2 of S6R(1), the EWG received many comments requesting further clarification of the term "exposure," e.g., AUC, C_{max}, and $C_{average}$. However, the EWG decided to refrain from detailed guidance on this to allow sponsors to provide a scientific justification for the approach taken. The addendum also recognizes that appropriate PD endpoints are not always available, and in these cases, high dose selection can be based on PK data and available in vitro functional data.

10.5.6   Reproductive/Developmental Toxicity

The need for reproductive/developmental toxicity studies is dependent on the product, clinical indication, and intended patient population. The specific study design may be modified based on issues related to species specificity, immunogenicity, biological activity, and/or a long elimination half-life (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997).

Both ICH S5A detection of toxicity to reproduction for medicinal products (ICHS5A Detection of Toxicity to Reproduction for Medicinal Products) and ICH S6 (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997) allow flexible design strategies based upon scientific justification. The principles for assessing reproductive and developmental toxicity are guided by ICH S5A; the practices for biopharmaceuticals are guided by ICH S6. Selection of relevant species is critical to generating relevant risk information. Traditional species (rodents and rabbits) if relevant are preferred. A variety of animal models are acceptable for assessing reproductive/development effects of biopharmaceuticals homologous products that have also been used. Strategies vary based upon product attributes and intended use. Different strategies are also acceptable across similar product classes and indications (Cavagnaro 2010).

Nonhuman primates (NHP) are best used when the objective of the study is to characterize a relatively certain reproductive toxicant, rather than detect a hazard. According to ICH S5A, if it can be shown by means of kinetic, pharmacological, and toxicological data that the species selected is a relevant model for the human, a single species can be sufficient (ICHS5A Detection of Toxicity to Reproduction for Medicinal Products). Relevant measures of male fertility performance can be included in repeat dose toxicity studies if animals are sexually mature although assessing fertility is limited when using nonhuman primates.

The means by which biopharmaceuticals cross the placenta if at all may be species dependent considering the notable differences between rodent and primate placenta. For biopharmaceuticals that do not cross the placenta, embryo–fetal development (EF) studies in both rodents and NHP are likely to be restricted to maternal effects rather than direct teratogenic effects; thus, a study in rodents with a homologous product could probably model these effects as effectively as a study in primates (Martin and Weinbauer 2010).

10.5.6.1   Key Developments in Assessment of Reproductive/Developmental Toxicity Since ICH S6

The considerations in assessing the developmental and reproductive toxicity potential of biopharmaceuticals in traditional and nontraditional animal species are well summarized in an extensive review by Martin et al. (2009, 2010). This review provides a framework for developing DART testing strategies for biopharmaceuticals. In addition, it provides an overview of the state of DART testing by highlighting various strategies that have been implemented over the past two decades for approved biopharmaceuticals, the lessons learned, and the current challenges in the evaluation of novel biopharmaceuticals.

The guidance on DART testing was very abbreviated in ICH S6 and related mainly to study design issues and adaption of study designs which may be needed for biopharmaceuticals, rather than issues relating to species selection. When S6 was finalized in 1997, there were a few approved products for non-oncology indications which also showed species-restricted pharmacological activity such that the nonhuman primate was the only relevant species. The experience was limited to the interferons, some cytokines, and a few monoclonal antibodies. Since the finalization of S6, there has been an explosion in the development of products for which assessment of toxicity to reproduction is needed but for which the nonhuman primate is the only relevant species. As a result, the number of nonhuman primates used for reproductive toxicity testing was increasing dramatically (Martin et al. 2009; Chapman et al. 2009; Chellman et al. 2009).

This situation led to many questions and divergent regulatory scientific advice about the relative merits of the use of rodents versus non-rodent species such as nonhuman primates and the use of alternative models such as transgenics and homologous products in rodent reproductive toxicity studies. In addition, there were many questions about the optimal design of nonhuman primate studies to address questions relating to assessment of developmental and reproductive toxicity. These two areas were the main focus of the EWG discussions for the addendum.

Firstly, the EWG reconfirmed that the principles of developmental and reproductive toxicity (DART) testing for biopharmaceuticals are similar to those for small molecule pharmaceuticals and in general follow the regulatory guidance outlined in ICH S5(R2) (ICHS5A Detection of Toxicity to Reproduction for Medicinal Products). This includes the use of rodents and rabbits for embryo–fetal development studies with biopharmaceuticals if the clinical candidate is pharmacologically

active in both species, unless clear developmental toxicity has been identified in one species. Several regulatory regions stated during the preparation of the S6 addendum that this requirement for two species for embryo–fetal development (EFD) studies was based on a review of internal databases and product labels and a lack of justification for the use of only a rodent or a rabbit.

An aspect which was considered by the EWG was the placental transfer of biopharmaceuticals. Small molecules (<1,000 Da) and their metabolites can diffuse across plasma membranes and the placenta by simple diffusion. In contrast, large molecule biopharmaceuticals do not appreciably diffuse across plasma membranes, including the placenta, and, therefore, have limited access to the conceptus. However, certain types of large molecules, such as monoclonal antibodies, can cross the placenta in mid- and late gestation by Fc receptor-mediated endocytosis via FcRn receptors (Martin et al. 2009; Simister 2003).

There are also species differences in placental transfer of antibodies between rodents and primates (Martin et al. 2009; Pentsuk and van der Laan 2009). In humans and nonhuman primates, transfer of antibodies across the placenta occurs primarily during the latter part of pregnancy, i.e., after organogenesis. This also appears to be the case for rabbits. In contrast, in rodents, transfer across the visceral yolk sac begins earlier in pregnancy, permitting exposure during organogenesis. Consequently, rodents may overestimate the human risk. However, the available data for some species is rather old and relates mainly to endogenous immunoglobulins induced by immunization to various antigens. The BioSafe group is in the process of gathering available data on placental transfer on a wide range of antibody and antibody-related products in development and plan to identify and fill data gaps to enable a better understanding of species differences in placental transfer.

One conclusion from the available information on the pattern of placental transfer in humans is that study designs that allow the detection of both indirect effects in early gestation plus the effects of direct fetal exposure in mid- and late gestation are recommended for developmental toxicity of monoclonal antibodies and related products.

There are increasing numbers of reports, many so far unpublished, of treatment-related fetal anomalies with monoclonal antibodies when administered to nonhuman primates only during the period of major organogenesis. One such published report related to figitumumab, an anti-insulin-like growth factor-1 receptor (IGF-1R) (Bowman et al. 2010). Thus, even low-level placental transfer and embryo–fetal exposure to potent monoclonal antibodies in early gestation may be sufficient to result in developmental toxicity.

Over several EWG meetings on the addendum, the regulators expressed a preference for DART testing with the clinical candidate, even if the only relevant species is a nonhuman primate. The EWG recognized the difficult balance between the limitations of a study in nonhuman primates with the clinical candidate versus the greater power of rodent developmental and reproductive toxicity studies but using a homologous product. However, although a preference is stated in the addendum, this does not mean that use of the nonhuman primate is the only acceptable option and a sponsor may still be able to provide a scientific justification for the use of an alternative DART testing strategy such as the use of alternative models including the use of homologous

products in rodent studies. This justification is likely to be based on the value of such alternative approaches to the communication and management of risk to humans.

There is now widespread industry and regulatory acceptance of the enhanced pre- and postnatal study (ePPND) design option when using nonhuman primates. This ePPND study combines the traditional "segmented" EFD study with the pre- and postnatal development (PPND) study into a single "enhanced" PPND study design where a single cohort of nonhuman primates is exposed throughout gestation and allowed to give birth naturally (Stewart 2009). The proposed "enhanced" PPND study design evaluates all the stages of the traditional two-study design using fewer animals. It also assesses the functional consequences of mid- to late gestational exposure (Martin and Weinbauer 2010; Chellman et al. 2009). This is of particular relevance to the risk assessment of monoclonal antibodies where fetal exposure to maternal IgG increases as pregnancy progresses and where morphologic examination of a preterm fetus may not be adequate to reveal the presence of adverse effects on functional development of key target organs. Another topic of hot debate in the EWG was the numbers of animals to be used in nonhuman primate ePPND studies. ICH S5(R2) note 13 states that for all but the rarest events (such as malformations, abortions, total litter loss), evaluation of between 16 and 20 litters for rodents and rabbits tends to provide a degree of consistency between studies. However, the same note also acknowledges that there is very little scientific basis underlying specified group sizes in past and existing guidelines nor in S5(R2). The numbers specified are educated guesses governed by the maximum study size that can be managed without undue loss of overall study control. The use of nonhuman primates carries additional ethical concerns, but number of animals per group should still be sufficient to allow meaningful interpretation of the data. An evaluation of pregnancy and infant loss in 1,069 vehicle-treated cynomolgus monkeys from 78 EFD studies and 14 PPND studies accrued during 1981–2007 was reported by Jarvis et al. (2010) to review the variability of pregnancy losses and impact on statistical power estimates and group size considerations. This evaluation indicated that based on the variability of pregnancy losses in this database and that in a PPND study with initial vehicle-control group sizes of 16 or 20, there is an 80 % likelihood of having 9 or 11 infants at day 7 postpartum, respectively.

After long debates on this topic, the addendum now states that "developmental toxicity studies in NHPs can only provide hazard identification. The number of animals per group should be sufficient to allow meaningful interpretation of the data (see Note 5)" (ICH S6R). Note 5 b refers to Jarvis et al. (2010) and recommends that group sizes in ePPND studies should yield a sufficient number of infants (6–8 per group at postnatal day 7) in order to assess postnatal development (Jarvis et al. 2010).

The addendum also outlines possibilities to reduce nonhuman primate use still further, e.g., by the use of fewer treatment groups (Chapman et al. 2012), reuse of vehicle-control maternal animals, early termination of animal accrual into the study if a treatment-related effect is noted during the course of the study, and use of a limited number of animals to confirm a likely hazard based on cause for concern based on mechanism of action (note: a study in rodents with a homologous product may also be justifiable in this case).

The evaluation of fertility is also problematic in nonhuman primates, and the addendum recognizes that mating studies are not practical for NHPs. Nonhuman primates are similar to humans with respect to the physiology and endocrinology of testicular and ovarian function (Chellman et al. 2009; Weinbauer et al. 2008), and potential effects on male and female fertility can be assessed by evaluation of the reproductive tract (organ weights and histopathological evaluation) in studies of at least 3-month duration using sexually mature nonhuman primates. The intent of the addendum was that the evaluation of potential effects on fertility in sexually mature nonhuman primates would be combined with the evaluation of general toxicity, usually the evaluation of chronic toxicity. Additional endpoints such as menstrual cyclicity, sperm counts, sperm morphology/motility, and male and female reproductive hormone levels are recommended if there is a specific cause for concern based on pharmacological activity or previous findings. Menstrual cyclicity is a fairly easy endpoint to monitor in cynomolgus monkeys by daily vaginal smears, and many sponsors chose to include this endpoint routinely in such studies rather than "for cause." However, the practical and logistical issues need to be recognized in order to get meaningful menstrual cyclicity data. When using social-housed female cynomolgus monkeys, it is essential to consider the housing history and familiarity between the animals prior to pair or group formations since this can lead to irregular cyclicity (Weinbauer et al. 2008). There is a need for several months of pre-study acclimation to the facility and cage mates.

Because mating studies are not practical for NHPs, there is a "data gap" in relation to a lack of information on the effects on conception and implantation. The addendum recommends that this data gap is addressed in several ways (1) experimentally using a homologous product in rodent studies or (2) risk mitigation through clinical trial management procedures, informed consent, and appropriate product labeling. It is not recommended to produce a homologous product or alternative animal model solely to conduct mating studies in rodents and to fill this data gap on effects on conception and implantation.

The timing of assessment of developmental and reproductive toxicity during clinical development was also a main topic for discussion in the EWG, in parallel to discussions ongoing in the ICH M3(R2) EWG. Both S6R(1) and M3(R2) recognize the difficulty of conducting developmental toxicity studies in nonhuman primates when this species is the only relevant species and allow for the conduct of such studies during phase III, providing there are sufficient precautions to prevent pregnancy, and the lack of animal reproductive toxicity data is communicated in the informed consent (ICH M3(R2), 2010).

Overall, while the addendum does express a preference for developmental and reproductive toxicity testing of the clinical candidate, various possibilities are suggested for reducing the overall number of monkeys in the reproductive toxicity testing strategy if use of the nonhuman primate is the only option for such testing. The use of homologous products in rodent studies rather than testing of the clinical candidate may also be appropriate where there is adequate scientific justification provided by the developer for the DART strategy proposed.

10.5.7   Genotoxicity

Genotoxicity testing is routinely conducted for pharmaceuticals to detect mutagenic and clastogenic compounds that may be carcinogens. Assays are designed to detect mutagenicity and clastogenicity, but not cellular proliferation. While uptake of low-molecular-weight compounds occurs through passive diffusion or nonspecific pinocytosis, large-molecular-weight compounds require active transport. Specific transporter mechanisms are typically not present in current assay systems thus "not relevant models" for assessing biopharmaceuticals (Cavagnaro 2010). False positives have been observed in the standard Ames test due to the presence of growth-promoting constituents in the test samples such as histidine or its precursors. Positive results have also been shown for lipase, glucagon, erythropoietin, and DNAse presumably based upon pharmacological activity hence considered predictable as exaggerated pharmacology.

   While studies may be applicable for protein conjugates with a chemical organic linker, consideration is warranted particularly when a residual organic linker is found in the product because of the instability of the conjugate during storage or upon dilution in the serum. Additionally, unlike pharmaceuticals where there may be a cause for concern for testing impurities for potential genotoxic potential, impurities associated with biopharmaceuticals are generally referred to as process related and include residual host cell proteins, fermentation components, column leachables, and detergents rather than organic chemicals and as such not considered to pose mutagenic risks.

   Biopharmaceuticals do not have the same distribution properties as small molecules and are therefore not expected to pass through cell and nuclear membranes to interact with DNA. Experience has confirmed that the standard battery of genotoxicity assays is not relevant for products that do not directly interfere with DNA or mitosis to induce gene mutations, chromosome aberrations, or DNA damage. While studies may be applicable for protein conjugates with a chemical organic linker, consideration is warranted if there is precedence of use with the linker or if there is no evidence of degradation of the protein conjugate. Additionally, unlike small molecules where there may be a cause for concern for testing for genotoxic impurities, process-related impurities associated with biopharmaceuticals include residual host cell proteins, fermentation components, column leachables, and detergents rather than organic chemicals.

10.5.7.1   Key Developments in Assessment of Genotoxicity Since ICH S6

Experience confirmed that the standard battery of genotoxicity assays is not relevant for products that do not directly interfere with DNA or mitosis to induce gene mutations, chromosome aberrations, or DNA damage. In a retrospective review of 78 compounds, mostly recombinant peptides and proteins, Gocke et al. (1999) concluded that genotoxicity testing of biological drugs was generally inappropriate and unnecessary.

10.5.8 Carcinogenicity

Carcinogenicity studies in two species are generally required for pharmaceuticals administered chronically. The need for carcinogenicity assessment of a biopharmaceutical is determined by a number of factors and is similar to those for pharmaceuticals. However, most of the early biotechnology molecules developed were for severe clinical indications and/or addressed unmet medical needs.

In cases where a biopharmaceutical is active and relatively non-immunogenic in rodents, and studies have not provided sufficient information to allow an assessment of carcinogenic potential, then a single bioassay has been considered per ICH S6 (e.g., a 2-year bioassay was performed for DNAse due to the mechanism of action and intended patient population). However, the standard bioassay was generally considered irrelevant for biopharmaceuticals (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997). One reason is that molecular structure excludes biopharmaceuticals from being intrinsically carcinogenic and as mentioned above, there would not be a concern for potential "carcinogenic metabolites." In addition, the rodent bioassay may otherwise not be relevant based on a high degree of antibody formation following repeat dosing of the clinical candidate, the lack of availability of an alternate product (e.g., homologous protein, surrogate molecule), or the lack of sufficient comparability.

ICH S6 guidance recommended incorporation of sensitive indices of cellular proliferation in chronic dose toxicity studies. However, it is recognized that while qualitative or quantitative increases in proliferation of target tissue and increases in organ weight signaling preneoplastic changes may represent early signals of epigenetic mechanisms, not all hyperplasia will result in neoplasia.

10.5.8.1 Key Developments in Assessment of Carcinogenicity Since ICH S6

The past and current practice over the last two decades regarding carcinogenicity assessments of biopharmaceuticals was reviewed by a collaborative effort of industry toxicologists involved in the preclinical development of biopharmaceuticals (Vahle et al. 2010). This review includes publicly available information on 80 approved biopharmaceuticals. No assessments related to carcinogenicity or tumor growth promotion were identified for 51 of the 80 molecules. For the 29 biopharmaceuticals in which assessments related to carcinogenicity were identified, various experimental approaches were employed. The review concluded that the traditional 2-year carcinogenicity assays should not be considered the default method for biopharmaceuticals and that if experimentation is considered warranted, it should be hypothesis driven and may include a variety of experimental models. Ultimately, it is important that preclinical data provide useful guidance in product labeling.

In parallel to the EWG discussion on assessment of carcinogenic potential, the value of the 2-year rodent bioassay for predicting carcinogenic hazard for humans of pharmaceutical products was also under review (Sistare et al. 2010; Friedrich and

Olejniczak 2010). Carcinogenicity data for pharmaceuticals and biopharmaceuticals approved via the European centralized procedure between 1995 and 2009 were evaluated; 65 % of compounds were deemed positive for carcinogenicity in at least one long-term carcinogenicity study or in repeat dose toxicity studies (Friedrich and Olejniczak 2010). These authors concluded that "due to the high number of rodent tumor findings with unlikely relevance for humans, the value of the currently used testing strategy for carcinogenicity appears questionable. A revision of the carcinogenicity testing paradigm is warranted." A pharmaceutical industry group made a proposal to refine regulatory criteria for conducting a 2-year rat study with pharmaceuticals to be based on assessment of histopathological findings from a rat 6-month study, evidence of hormonal perturbation, genetic toxicology results, and the findings of a 6-month transgenic mouse carcinogenicity study (Sistare et al. 2010).

Bugelski et al. (2010) reviewed the preclinical approaches to evaluate the potential of immunosuppressive drugs to influence human neoplasia. The authors concluded that the 2-year rodent bioassay performs poorly in identifying the mechanism of action-related hazard for developing certain tumor types, especially lymphomas and skin cancer. Classifying immunosuppressive drugs based on their mechanism of action and hazard identification from preclinical studies and a prospective pharmacovigilance program to monitor carcinogenic risk was proposed as a feasible way to manage patient safety during the clinical development program and post-marketing.

At the first EWG meeting for the addendum in 2008, there was a recognition that the issues encountered regarding the assessment of carcinogenic potential of biopharmaceuticals were likely related to the industry-regulatory creep and changing regulatory environment alluded to earlier. These issues were likely related more to implementation of the S6 guidance in some regulatory regions rather than lack of clarity of this guidance.

The S6 guidance started from the general philosophy that standard carcinogenicity bioassays are generally inappropriate for biopharmaceuticals but that a product-specific assessment of carcinogenicity may still be needed. By 2007, the general philosophy of some regulatory agencies was the same as for small molecules—"if you can do it, you should do it"—if such an assessment is needed according to the clinical population and treatment duration (ICH S1A).

The EWG reviewed the practice of carcinogenicity testing of biopharmaceuticals over the last two decades and also reviewed several case studies provided by some regulatory agencies. Overall, the general philosophy as outlined in the S6 guidance was upheld, and attempts were made to clarify certain aspects. When an assessment of carcinogenic potential is warranted, it is up to the sponsor to design a strategy to address the potential hazard, based on a weight of evidence approach and an understanding of target biology related to potential carcinogenic concern. Rodent bioassays (or short-term carcinogenicity studies) with homologous products were generally considered to be of limited value to assess carcinogenic potential of the clinical candidate. Ultimately, the product-specific assessment of carcinogenic potential is used to communicate risk and provide input to the risk management plan along with labeling proposals, clinical monitoring, post-marketing surveillance, or a combination of these approaches (Cavagnaro 2008b).

10.5.9 Immunogenicity

ICH S6 states, "Most biotechnology-derived pharmaceuticals intended for humans are immunogenic in animal" (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997). Traditional antigenicity studies or guinea pig anaphylaxis studies are not useful for predicting immunogenicity in humans and are now generally recognized as not being appropriate studies for biologics. When these studies were conducted with biopharmaceuticals, they were not surprisingly positive and led to adverse effects in animals. Since there is little to no predictive value in these studies, and they were not considered appropriate, such studies have not been conducted since publication of ICH S6.

Administration of human proteins in sufficient quantity into animals is expected to elicit an immunological response. Even homologous/surrogate molecules have induced immune responses in the respective species. Immunogenicity assessments are conducted to assist in the interpretation of the study results and design of subsequent studies rather than to predict potential immunogenicity in humans. The presence of neutralizing antibody can change the PK/PD profile and thus impact exposure margins and estimates of toxicity. In early studies with biopharmaceuticals, the development of antibodies in a toxicology study was considered a reason to stop studies; however, we now know that we can "dose through" in animals similar to dosing practices in humans. While the presence of antibodies in animals is generally not predictive for humans, the information has helped in defining relative immunogenicity and in identifying potential consequences of an immune response, e.g., neoantigenicity, autoantigenicity, immune complex deposition, complement activation, and the impact of antibodies crossing the placenta.

The two major areas of concern relating to the assessment of antigenic/immunogenic potential are (1) product/active ingredient and (2) process/excipient/final formulation. The formation of antibodies is monitored at various intervals throughout toxicity studies in order to be able to interpret the studies and determine if there is any impact on exposure. Information should be provided on the effect of antibody formation on the pharmacokinetic behavior of the product and whether antibodies interfere with the assay used to monitor the product in biological fluids. Clinically relevant antibodies include clearing antibodies, sustaining antibodies, neutralizing antibodies, and antibodies that cross-react with endogenous proteins. The presence of neutralizing antibodies and abrogation of subsequent pharmacological and/or toxicological effects can provide the justification for limiting the duration of repeated dose studies. However, the presence of antidrug antibodies in the absence of PK effects, neutralization of activity, or other toxicities is not sufficient to support study termination or shorter study durations.

10.5.9.1 Key Developments in Assessment of Immunogenicity Since ICH S6

By 2007, it had become apparent that immunogenicity testing was being largely driven by bioanalytical considerations with great emphasis being given to the S6 guidance

that "measurement of antibodies…should be performed when conducting repeated dose toxicity studies …" and "antibody responses should be characterized (e.g. titre, number of responding animals, neutralizing or non-neutralizing)" (ICH S6 Preclinical Safety Evaluation of Biotechnology-derived Pharmaceuticals 1997). The primary purpose of such immunogenicity testing in support of toxicity studies "in order to aid in the interpretation of these studies" seemed to be superseded by bioanalytical considerations. Because of assay sensitivity issues relating to drug interference, the perceived requirement to measure and characterize antibody responses in repeated dose toxicity studies in order to determine whether an animal was antidrug antibody (ADA) positive or negative was driving long treatment-free recovery periods, even in the absence of any toxicity findings needing an evaluation of reversibility.

A decision tree for conducting ADA analyses to support nonclinical study interpretation was provided by Ponce et al. (2009). This decision tree is intended to guide the investigator through a series of considerations to determine whether ADA analysis is necessary to aid in the interpretation of a study. The authors concluded that immunogenicity data should be integrated with available clinical and anatomic pathology, PK, and PD data to properly interpret nonclinical studies. PD markers of target engagement such as ligand capture (soluble ligand) or receptor occupancy (cell surface ligand), as well as downstream signaling markers or other in vivo mechanistic markers, also contain valuable information regarding the neutralizing potential of an ADA response evident as loss of target engagement or loss of functional or pharmacological activity. Where such PD markers are available, the need for specific neutralization assays may be obviated by the use of these alternative markers of functional activity (Buttel et al. 2011).

The S6R(1) addendum clarifies the purpose of immunogenicity testing in the first sentence: "immunogenicity assessments are conducted to assist in the interpretation of the study results and design of subsequent studies." The addendum provides clarification for when measurement of antidrug antibodies (ADA) in nonclinical studies should be evaluated and when characterization of neutralization potential is warranted. When no PD marker exists to demonstrate sustained activity in the in vivo toxicology studies, characterization of neutralizing potential is warranted, but the addendum provides clarification that this can be assessed indirectly with an ex vivo bioactivity assay or an appropriate combination of assay formats for PK–PD (Buttel et al. 2011) or directly in a specific neutralizing antibody assay.

10.6   Conclusions

Preclinical safety evaluation of biopharmaceuticals has evolved through the application of scientific insight, historical and anecdotal experiences, and common sense. The scientific community has relied on the exchange of ideas between academia, industry, and regulatory scientists. Many new challenges in biopharmaceutical clinical development lie ahead. New technologies and products not yet envisioned will continue to

challenge toxicologists. Additional challenges and advances will come from efforts devoted to site-directed delivery or site-specific expression. Open dialogue between scientists who are regulators, academic scientists, or industry scientists will be critical in ensuring that the new products that are safe and effective are made available without unnecessary delay. A regulatory environment that encourages innovation will make this possible.

Development practices for preclinical safety assessment of biopharmaceuticals have been and will continue to be a dynamic process that is strongly controlled by the expanding knowledge and the innovations in product design. However, the full investigation of the potential usefulness of biopharmaceuticals will require the development of reliable animal model systems that allow assessment of toxicity and provide pharmacokinetic data that can be successfully scaled to humans in order to reduce risk factors before clinical testing. There is also a need to develop and refine appropriate human in vitro systems to aid safety assessment in cases where reliable animal models do not exist but also to address specific limitations of animal studies, e.g., assessing the potential for cytokine release (Vidal et al. 2010). Once sufficient data have accrued, it is important to review experiences as was done in the case of the ICH S6 and recalibrate approaches if necessary.

The design of relevant preclinical safety evaluation programs is consistent with global initiatives to facilitate and to improve clinical development programs. In the coming years, stakeholders will be facing the issue of how to implement preclinical development programs for biopharmaceuticals and pharmaceuticals that better anticipate adverse effects including development of new test systems that produce reliable results faster and at lower cost. Hopefully, preclinical evaluation programs will evolve and mature concurrently with more novel products and will focus on improving the predictive value of preclinical safety testing, challenging toxicologists to provide information from the most appropriate studies.

Biotechnology has provided not only the hope of potential new therapies but also the necessary tools to evaluate new therapies. Toxicology as a science has benefited from this experience in many ways. The case-by-case approach to preclinical safety evaluation should continue to provide for scientific advancement in toxicology and the inducement of quality research into relevant safety assessment for the next generation of novel therapies.

References

Bass R, Kleeburg U, Schroder H et al (1992) Current guidelines for the preclinical safety assessment of therapeutic proteins. Toxicol Lett 64/65:339–347

Bornstein GG, Klakamp SL, Andrews L, Boyle WJ, Tabrizi M (2009) Surrogate approaches in development of monoclonal antibodies. Drug Discov Today 14(23/24):1159–1165

Bowman CJ, Chmielewski G, Oneda S, Finco D, Boucher MA, Todd M (2010) Embryo-fetal development toxicity of figitumumab, an anti-insulin-like growth factor-1 receptor (IGF-1R) monoclonal antibody, in cynomolgus monkeys. Birth Defects Res B Dev Reprod Toxicol 89(4):326–338

Brennan FR, Shaw L, Wing MG, Robinson C (2004) Preclinical safety testing of biopharmaceuticals. Mol Biotechnol 27:59–74

Buckley LA, Benson K, Davis-Bruno K, Dempster M, Finch GL, Harlow P, Haggerty HG, Hart T, Kinter L, Leighton JK, McNulty J, Roskos L, Saber H, Stauber A, Tabrizi M (2008) Nonclinical aspects of biopharmaceutical development: discussion of case studies at a PhRMA-FDA workshop. Int J Toxicol 27(4):303–312

Bugelski PJ, Achuthanandam R, Capocasale RJ, Treacy G, Bouman-Thio E (2009) Monoclonal antibody-induced cytokine release syndrome. Expert Rev Clin Immunol 5(5):499–521

Bugelski P, Volk A, Walker MR, Krayer JH, Martin P, Descotes J (2010) Critical review of preclinical approaches to evaluate the potential of immunosuppressive drugs to influence human neoplasia. Int J Toxicol 29(5):435–466

Bussiere J (2008) Species selection considerations for preclinical toxicology studies for biotherapeutics. Expert Opin Drug Metab Toxicol 4(7):871–877

Bussiere JL, Martin P, Horner M, Couch J, Flaherty M, Andrews L, Beyer J, Horvath C (2009) Alternative strategies for toxicity testing of species-specific biopharmaceuticals. Int J Toxicol 28:230–253

Bussiere JL, Leach MW, Price KD, Mounho BJ, Lighfoot-Dunn R (2011) Survey results on the use of the tissue cross-reactivity immunohistochemistry assay. Regul Toxicol Pharmacol 59:493–502

Buttel IC, Chamberlain P, Chowers Y et al (2011) Taking Immunogenicity assessment of therapeutic proteins to the next level. Biologicals 39:100–109

Cavagnaro JA (1992a) Misconceptions within biotechnology. In: D'Arcy PF, Harmon DWG (eds) Proceedings of the first conference on harmonisation, Brussels, 1991. The Queen's University of Belfast, Belfast, pp 301–307

Cavagnaro JA (1992b) Science-based approach to preclinical safety evaluation of biotechnology products. Pharmaceut Eng 12:32–33

Cavagnaro JA (2002) Preclinical safety evaluation of biotechnology-derived pharmaceuticals. Nat Rev Drug Discov 1:469–475

Cavagnaro JA (ed) (2008a) Preclinical safety evaluation of biopharmaceuticals: a science-based approach to facilitating clinical trials. Wiley, New York

Cavagnaro JA (2008b) Assessment of carcinogenic risk of biopharmaceuticals. In: Cavagnaro JA (ed) Preclinical safety evaluation of biopharmaceuticals: a science-based approach to facilitating clinical trials. Wiley, New York, pp 399–477

Cavagnaro J (2010) Considerations for the preclinical safety evaluation of biopharmaceuticals. In: Charlene A (ed) McQueen, comprehensive toxicology, vol 3. Academic, Oxford, pp 29–51

Chapman K, Pullen N, Coney L, Dempster M, Andrews L, Bajramovic J, Baldrick P, Buckley L, Jacobs A, Hale G, Green C, Ragan I, Robinson V (2009) Preclinical development of monoclonal antibodies: considerations for the use of non-human primates. mAbs 1(5):505–516

Chapman KL, Andrews L, Bajramovic JJ, Baldrick P, Black LE, Bowman CJ, Buckley LA, Coney LA, Couch J, Dempster AM, De Haan L, Jones K, Pullen N, de Boer AS, Sims J and Ragan CI (2012) The design of chronic toxicology studies of monoclonal antibodies: Implications for the reduction in the use of non-human primates. Regul Toxicol Pharmacol 62(2):347–354

Chellman GJ, Bussiere JL, Makori N, Martin PL, Ooshima Y, Weinbauer GF (2009) Developmental and reproductive toxicology studies in nonhuman primates. Birth Defects Res B 86:446–462

Clarke J, Hurst C, Martin P, Vahle J, Ponce R, Mounho B, Heidel S, Andrews L, Reynolds T, Cavagnaro J (2008) Duration of chronic toxicity studies for biotechnology-derived pharmaceuticals: is 6 months still appropriate? Regul Toxicol Pharmacol 50:2–22

Claude JR (1992) Difficulties in conceiving and applying guidelines for the safety evaluation of biotechnologically-produced drugs: some examples. Toxicol Lett 64(65):349–355

Dayan AD (1987) Rationality and regulatory requirements- a view from Britain. In: Grahm CE (ed) Preclinical safety of biotechnology products intended for human use: clinical and biological research, vol 235. Alan R. Liss, New York, pp 89–106

Dayan AD (1995) Safety evaluation of biological and biotechnology-derived medicines. Toxicology 105:59–68

Eastwood D, Findlay L, Poole S, Bird C, Wadhwa M, Moore M, Burns C, Thorpe R, Stebbings R (2010) Monoclonal antibody TGN1412 trial failure explained by species differences in CD28 expression on CD4[+] effector memory T-cells. Br J Pharmacol 161:512–526

EMEA/CHMP/SWP/28367/07

FDA (1997) Points to consider in the manufacture and testing of monoclonal antibody products for human use, 1997. http://www.fda.gov

Findlay L, Eastwood D, Ball C, Robinson CJ, Bird C, Wadhwa M, Thorpe R, Stebbings R, Poole S (2011) Comparison of novel methods for predicting the risks of pro-inflammatory clinical infusion reactions during monoclonal antibody therapy. J Immunol Methods 371(1–2):131–142

Finkle BS (1987) Genetically engineered drugs: toxicology with a difference. In: Grahm CE (ed) Preclinical safety of biotechnology products intended for human use: clinical and biological research, vol 235. Alan R. Liss, New York, pp 161–167

Friedrich A, Olejniczak K (2010) Evaluation of carcinogenicity studies for medicinal products for human use authorized via the European centralized procedure (1995–2009). Regul Toxicol Pharmacol 60(2):225–248

Galbraith WM (1987) Symposium discussion. In: CE Grahm (ed) Preclinical safety of biotechnology products intended for human use: clinical and biological research, vol 235. Alan R. Liss, New York, pp 189–206

Giss HE (1987) Foreword. In: Grahm CE (ed) Preclinical safety of biotechnology products intended for human use: clinical and biological research, vol 235. Alan R. Liss, New York, pp xiii–xv

Gocke E, Albertini S, Brendel-Schwaab S, Muller L, Suter W et al (1999) Genotoxicity testing of biotechnology-derived products. Report of GUM task force. Mutant Res 436:137–156

Griffith SA, Lumley CE (1998) Non-clinical safety studies for biotechnologically-derived pharmaceuticals: conclusions form an international workshop. Hum Exp Toxicol 17:63083

Hansel TT, Kropshofer H, Singer T, Mitchell JA, George AJT (2010) The safety and side effects of monoclonal antibodies. Nat Rev Drug Discov 9:325–338

Hayakawa T (1992) Current regulatory situation in Japan with respect to the preclinical safety testing for biotechnology products intended for human use. In: D'Arcy PF, Harmon DWG (eds) Proceedings of the first conference on harmonisation, Brussels 1991. The Queen's University of Belfast, Belfast, pp 296–301

Hayes TJ, Cavagnaro JA (1992) Progress and challenges in the preclinical assessment of cytokines. Toxicol Lett 64(65):291–297

Henck JW, Hilbish KG, Serabian MA, Cavagnaro JA, Hendrickx AG, Agnish ND, Kung ADH, Mordenti J (1996) Reproductive toxicity testing of therapeutic biotechnology agents. Teratology 53:185–195

Hohbach C (1992) Safety in biotechnology progress and problems, industrial point of view. In: D'Arcy PF, Harmon DWG (eds) Proceedings of the first conference on harmonisation, Brussels 1991. The Queen's University of Belfast, Belfast, pp 307–316

Horvath CJ, Milton MN (2009) The TeGenero incident and the Duff Report conclusions: a series of unfortunate events or an avoidable event? Toxicol Pathol 37:372–383

ICH M3(R2) (2010): Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals. http://www.ich.org

ICH S6 Preclinical safety evaluation of biotechnology-derived pharmaceuticals (1997) http://www.ich.org

ICH S6R(1): Preclinical safety evaluation of biotechnology-derived pharmaceuticals. http://www.ich.org

ICHS5A Detection of toxicity to reproduction for medicinal products

Jarvis P, Srivastav S, Vogelwedde E, Stewart J, Mitchard T, Weinbauer GF (2010) The cynomolgus monkey as a model for developmental toxicity studies: variability of pregnancy losses, statistical power estimates, and group size considerations. Birth Defects Res B 89:175–187

Kikuchi Y (1992) Panel presentation safety: biotechnology. In: D'Arcy PF, Harmon DWG (eds) Proceedings of the first conference on harmonisation, Brussels 1991. The Queen's University of Belfast, Belfast, pp 316–319

Leach MW, Halpern WG, Johnson CW, Rojko JL, MacLachlan TK, Chan CM, Galbreath EJ, Ndifor AM, Blanset DL, Polack E, Cavagnaro JA (2010) Use of tissue cross-reactivity studies in the development of antibody-based biopharmaceuticals; history, experience, methodology, and future directions. Toxicol Pathol 38:1138–1166

Liedert B, Bassus S, Schneider CK, Kalinke U, Lower J (2007) Safety of Phase I clinical trials with monoclonal antibodies in Germany – the regulatory requirements viewed in the aftermath of the TGN1412 disaster. Int J Clin Pharmacol Ther 45(1):1–9

Lowe P, Tannenbaum S, Wu K, Lloyd P, Sims J (2009) On setting the first dose in man: quantitating biotherapeutic drug-target binding through pharmacokinetic and pharmacodynamics models. Basic Clin Pharmacol Toxicol 106:195–209

Martin PL, Breslin W, Rocca M, Wright D, Cavagnaro J (2009) Consideration in assessing the developmental and reproductive toxicity potential of biopharmaceuticals. Birth Defects Res B 86:176–203

Martin PL, Weinbauer GF (2010) Developmental toxicity testing of biopharmaceuticals in nonhuman primates: previous experience and future directions. Int J Toxicol 29:552–568

Milton MN, Horvath CJ (2009) The EMEA guideline on first-in-human clinical trials and its impact on pharmaceutical development. Toxicol Pathol 37:363–371

Nakazawa T, Kai S, Kawai M, Maki E, Sagami F, Onodera H, Kitajima S, Inoue T (2004) "Points to Consider" regarding safety assessment of biotechnology-derived pharmaceuticals in nonclinical studies (English Translation). J Toxicol Sci 29:497–504

Pentsuk N, van der Laan JW (2009) An interspecies comparison of placental antibody transfer: new insights into developmental toxicity testing of monoclonal antibodies. Birth Defects Res B 86:328–344

Ponce R, Abad L, Amaravadi L, Gelzleichter T, Gore E, Green J, Gupta S, Herzyk D, Hurst C, Ivens I, Kawabata T, Mauer C, Mounho B, Rup B, Shankdar G, Smith H, Thomas P, Wierda D (2009) Immunogenicity of biotechnology-derived therapeutics: assessment and interpretation of nonclinical studies. Regul Toxicol Pharmacol 54:164–182

Roskos L, Schneider A, Vainshtein I, Schwickart M, Lee R, Lu H, Faggioni R, Liang M (2011) PK-PD modeling of protein drugs: implications in assay development. Bioanalysis 3(6):659–675

Ryan AM, Terrell TG (2002) Biotechnology and its products. In: Handbook of toxicologic pathology, 2nd edn. Academic, New York, pp 479–500

Schneider CK, Kalinke I, Lower J (2006) TGN1412 – a regulator's perspective. Nat Biotechnol 24(4):368

Serabian MA, Pilaro AM (1999) Safety assessment of biopharmaceuticals: ICH and beyond. Toxicol Pathol 27:27–31

Simister N (2003) Placental transport of immunoglobulin G. Vaccine 21:3365–3369

Sims J (2001) Assessment of biotechnology products for therapeutic use. Toxicol Lett 120:59–66

Sistare FD, Morton D, Alden C, Christensen J, Keller D et al (2010) An analysis of pharmaceutical experience with decades of rat carcinogenicity testing: support for a proposal to modify current regulatory guidelines. Toxicol Pathol 39(716–744):2011

Stebbings R, Findly L, Edwards C, Eastwood D, Bird C, North D, Mistry Y, Dilger P, Liefooghe E, Cludts I, Fox B, Tarrant G, Robinson J, Meager T, Dolman C, Thorpe SJ, Bristow A, Wadhwa M, Thorpe R, Poole S (2007) "Cytokine storm" in a Phase I trial of monoclonal antibody TGN1412: better understanding the causes to improve the preclinical testing of immunotherapeutics. J Immunol 179:3325–3331

Stebbings R, Pool S, Thorpe R (2009) Safety of biologics, lessons learnt from TGN1412. Curr Opin Biotechnol 20(6):673–677

Stewart J (2009) Developmental toxicity testing of monoclonal antibodies: an enhanced pre-and postnatal study design option. Reprod Toxicity 28:220–225

Suntharalingham G, Perry MR, Ward S, Brett SJ, Castello-Cortes A, Brunner MD, Panoskaltsis N (2006) Cytokine storm in a Phase I trial of the anti-CD28 monoclonal antibody TGN1412. N Engl J Med 355:1018–1028

Tabrizi MA, Roskos LK (2007) Preclinical and clinical safety of monoclonal antibodies. Drug Discov Today 12(13/14):540–547

Tabrizi MA, Bornstein GG, Klakamp SL, Drake A, Knight R, Roskos L (2009) Translational strategies for development of monoclonal antibodies from discovery to the clinic. Drug Discov Today 14(5–6):298–305

Terrell TG, Green JD (1994) Issues with biotechnology products in toxicologic pathology. Toxicol Pathol 22:187–193

Thomas JA (1995) Recent developments and perspectives of biotechnology-derived products. Toxicology 105:7–22

Tibbitts J, Cavagnaro JA, Haller CA, Marafino B, Andrews PA, Sullivan JT (2010) Practical approaches to dose selection for first-in-human clinical trials with novel biopharmaceuticals. Regul Toxicol Pharmacol 58:243–251

Vahle JL, Finch GL, Heidel SM, Hovland DN, Ivens I, Parker S, Ponce RA, Sachs C, Steigerwalt R, Short B, Todd MD (2010) Carcinogenicity assessments of biotechnology-derived pharmaceuticals: a review of approved molecules and best practice recommendations. Toxicol Pathol 38:522–553

Vargas HM, Bass AS, Breidenbach A, Feldman HS, Gintant GA, Harmer AR, Heath B, Hoffman P, Lagrutta A, Leishman D, McMahon N, Mittelstadt S, Polonchuk L, Pugsley MK, Satata JJ, Valentin J-P (2008) Scientific review and recommendations on preclinical cardiovascular safety evaluation of biologics. J Pharmacol Toxicol Methods 58:72–76

Vidal JM, Kawabata TT, Thorpe R, Silva-Lima B, Cederbrant K, Poole S, Mueller-Berghaus J, Pallardy M, Van der Laan JW (2010) In vitro cytokine release assays for predicting cytokine release syndrome: the current state-of-the-science. Report of a European Medicines Agency Workshop. Cytokine 51(2):213–215

Weinbauer GF, Niehoff M, Niehaus M, Srivastav S, Fuchs A, Van Esch E, Cline JM (2008) Physiology and endocrinology of the ovarian cycle in macaques. Toxicol Pathol 36(7S):7S–23S

Zbinden G (1987) Biotechnology products intended for human use, toxicological targets and research strategies. In: Grahm CE (ed) Preclinical safety of biotechnology products intended for human use: clinical and biological research, vol 235. Alan R. Liss, New York, pp 143–159

Zbinden G (1990) Safety evaluation of biotechnology products. Drug Saf 5(suppl 1):58–64

Zbinden G (1991) Predictive value of animal studies in toxicology. Regul Toxicol Pharmacol 14:167–177