

# Annex 4

## Guidelines on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant DNA technology

Replacement of Annex 3 of WHO Technical Report Series, No. 814

|  |     |
|--|-----|
| <b>Introduction</b>  | 177 |
| <b>Background</b>  | 177 |
| <b>Scope</b>   | 180 |
| <b>Terminology</b>   | 181 |
| <b>Part A. Manufacturing and quality control</b>                     | 184 |
| A.1 Definitions  | 184 |
| A.2 General manufacturing guidelines                                 | 184 |
| A.3 Control of starting/source materials                             | 186 |
| A.4 Control of the manufacturing process                             | 189 |
| A.5 Control of drug substance and drug product                       | 192 |
| A.6 Filling and containers   | 194 |
| A.7 Records, retained samples, labelling, distribution and transport | 194 |
| A.8 Stability, storage and expiry date                               | 194 |
| A.9 Manufacturing process changes                                    | 197 |
| <b>Part B. Nonclinical evaluation</b>                                | 198 |
| B.1 Introduction   | 198 |
| B.2 Pharmacodynamics   | 200 |
| B.3 Pharmacokinetics/toxicokinetics                                  | 201 |
| B.4 Toxicity studies   | 203 |
| <b>Part C. Clinical evaluation</b>                                   | 214 |
| C.1 Good clinical practice   | 214 |
| C.2 Clinical pharmacology (Phase I)                                  | 215 |
| C.3 Efficacy   | 222 |
| C.4 Statistical considerations                                       | 227 |
| C.5 Safety   | 229 |
| C.6 Immunogenicity   | 231 |
| C.7 Pharmacovigilance and risk management planning                   | 234 |
| C.8 Additional guidance  | 235 |
| <b>Authors and acknowledgements</b>                                  | 236 |
| <b>References</b>  | 239 |

|                   |   |     |
|-------------------|---|-----|
| <b>Appendix 1</b> | Manufacturing process validation  | 245 |
| <b>Appendix 2</b> | Characterization of rDNA-derived biotherapeutics                          | 247 |
| <b>Appendix 3</b> | Routine control of rDNA-derived biotherapeutics                           | 254 |
| <b>Appendix 4</b> | Product/indication-specific guidance in nonclinical evaluation (examples) | 257 |
| <b>Appendix 5</b> | Animal species/model selection  | 259 |
| <b>Appendix 6</b> | Explanatory notes   | 262 |

Guidelines published by WHO are intended to be scientific and advisory in nature. Each of the following sections constitutes guidance for national regulatory authorities (NRAs) and for manufacturers of biological products. If an NRA so desires, these Guidelines may be adopted as definitive national requirements, or modifications may be justified and made by the NRA. It is recommended that modifications to these Guidelines are made only on condition that such modifications ensure that the product is at least as safe and efficacious as that prepared in accordance with the guidance set out below.

## Introduction

These Guidelines are intended to provide guidance to national regulatory authorities (NRAs) and manufacturers on the quality, safety and efficacy of biotherapeutic protein products prepared by recombinant deoxyribonucleic acid (DNA) technology (rDNA-derived biotherapeutics) and intended for use in humans. The Guidelines are based on experience gained over three decades in this technically demanding field and replace Guidelines for assuring the quality of pharmaceutical and biological products prepared by recombinant DNA technology (1).

Part A of this annex sets out updated guidelines for the manufacture and quality control of rDNA-derived biotherapeutics, including consideration of the effects of manufacturing changes and of devices used in the delivery of the product and in its stability. Part B is new and provides guidelines on nonclinical evaluation, while Part C, also new, provides guidance on clinical evaluation. The nature and extent of characterization and testing (Part A) required for a product undergoing nonclinical and clinical studies will vary according to the nature of the product and its stage of development. The legal status of investigational products varies from country to country. The need for and extent of studies (e.g. on characterization) will depend on the product under consideration. Early communication between the manufacturer and the responsible NRA to agree on the requirements for, and the type of, studies is recommended. Some aspects of manufacturing and quality control in these Guidelines may apply to protein-based vaccine antigens made by rDNA technology. However, more detailed guidelines/recommendations on vaccine evaluation in terms of quality, safety and efficacy should be consulted (2, 3). Other product-specific vaccine-related WHO Recommendations and Guidelines are also available elsewhere.<sup>1</sup> Additional considerations for similar biotherapeutic products are addressed elsewhere (4).

## Background

Developments in molecular genetics and nucleic acid chemistry have enabled genes encoding natural biologically active proteins to be identified, modified and transferred from one organism to another in order to obtain highly efficient synthesis of their products. This has led to the production of new rDNA-derived biological medicines using a range of different expression systems such as bacteria, yeast, transformed cell lines of mammalian origin (including human

<sup>1</sup> See: <http://www.who.int/biologicals/vaccines/en/>

origin), insect and plant cells, as well as transgenic animals and plants. rDNA technology is also used to produce biologically active proteins that do not exist in nature, such as chimeric, humanized or fully human monoclonal antibodies, or antibody-related proteins or other engineered biological medicines such as fusion proteins.

There has been great progress in the ability to purify biologically active macromolecules. In addition, analytical technologies have improved tremendously since the early days of biotechnology, allowing the detailed characterization of many biological macromolecules, including their protein, lipid and oligosaccharide components.

Together these technologies have enabled the production of large quantities of medicinal products that are difficult to prepare from natural sources or were previously unavailable. Nevertheless, it is still not possible fully to predict the biological properties and clinical performance of these macromolecules on the basis of their physicochemical characteristics alone. In addition, the production processes are biological systems which are known to be inherently variable – a feature which has important consequences for the safety and efficacy of the resulting product. Therefore a prerequisite for introducing such biological substances into routine clinical use is to ensure consistency of quality from lot to lot, and for this purpose robust manufacturing processes are developed on the basis of process understanding and characterization, including appropriate in-process controls. Process understanding and consistency are critical since slight changes can occasionally lead to major adverse effects, such as immunogenicity, with potentially serious safety implications.

As with many other new technologies, a new set of safety issues for consideration by both industry and NRAs has been generated by these biotechnologies. Potential safety concerns arose from the novel processes used in manufacture, from product- and process-related impurities, and from the complex structural and biological properties of the products themselves. Factors that have received particular attention include the possible presence of contaminating oncogenic host-cell DNA in products derived from transformed mammalian cells (5), and the presence of adventitious viruses (5). Since the nature and production of these products are highly sophisticated, they require similarly sophisticated laboratory techniques to ensure their proper standardization and control. Although comprehensive analytical characterization of the drug substance and/or drug product is expected, considerable emphasis must also be given to the manufacturing process – i.e. process validation and in-process control. Adequate control measures relating to the starting materials and manufacturing process are, therefore, as important as analysis of the drug product. Thus, data on the host-cell quality, purity, freedom from adventitious agents, adequate in-process testing during production, and effectiveness of test methods are required for licensing.

At a very early stage in the development of rDNA-derived medicines, the European Medicines Agency (EMA) and the United States Food and Drug Administration produced guidelines and points to consider, respectively, for the development and evaluation of these new products (6, 7). Such guidelines, based as they were on long experience with traditional biological substances, set the scene for regulatory expectations both for clinical trials and for licensing. At the global level, WHO produced a series of guidance documents on the quality, safety and efficacy of rDNA-derived products, including specific guidance for products such as interferons and monoclonal antibodies (1, 8–10). These regulatory concepts have been instrumental in establishing expectations for the quality, safety and efficacy of rDNA-derived biotherapeutics which play a major role in today's medical practice.

As patents and data protection measures on biotechnology products have expired, or have neared expiry, considerable attention has turned to producing copies of the innovator products with a view to making more affordable products that may improve global access to these medicines. Since by definition it is not possible to produce identical biological substances, the normal method of licensing generic medicines – which relies primarily on bioequivalence data – is not appropriate for licensing such products. Consequently, the terms “similar biological product” and “biosimilar product” came into existence (4, 11). The concept of similar biological medicinal products was introduced first by the EMA (11) and subsequently by other NRAs (although the actual term used has varied slightly from agency to agency). The WHO Guidelines on evaluation of similar biotherapeutic products were produced in 2010 (4) and provided a set of globally acceptable principles regarding the regulatory evaluation of biosimilars, although it was recognized that these would not by themselves resolve all issues. During international consultations on the development of the biosimilar WHO Guidelines, and also during their implementation, it became clear that there was a need to update WHO guidance on the quality, safety and efficacy of rDNA-derived medicines and biotechnology products in general (12). In 2010, the International Conference of Drug Regulatory Authorities noted that WHO should supplement its guidance on the evaluation of similar biotherapeutic products by providing up-to-date Guidelines for the evaluation of biotherapeutic products in general.

The present Guidelines have been developed through international consultation and are intended as a replacement of those in Annex 3 of WHO Technical Report Series, No. 814 (1). They are considered to be a replacement and not a revision of the earlier Guidelines because they contain new sections on nonclinical and clinical evaluation of rDNA-derived biotherapeutics which were lacking in the original document. In addition, a section on issues related to manufacturing changes, both during development and once the product is

on the market, has also been introduced because considerable improvements to the production process and to the product itself can take place during all stages of development and post-licensure, especially in the immediate post-licensing years. These changes may unintentionally have an impact on the clinical performance of the product and therefore need to be handled carefully from a regulatory perspective.

Guidance on various aspects of rDNA-derived medicines is also available from several other bodies such as the International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), the EMA, and the United States Food and Drug Administration. These WHO Guidelines are not intended to conflict with, but rather to complement, these other documents.

## Scope

These WHO Guidelines provide guidance to NRAs and manufacturers on the quality, nonclinical and clinical aspects of rDNA-derived biotherapeutic protein products for the purpose of licensing. Relevant sections of the Guidelines may also be useful with regard to rDNA-derived biotherapeutic protein products intended for clinical trials; however, the amount and extent of data submitted for such products will be limited and will vary according to the nature of each product and its stage of development. In addition, the legal status of investigational products varies from country to country.

The Guidelines apply, in principle, to all biologically active protein products which are used in the treatment of human diseases and which are prepared by recombinant DNA technology using prokaryotic or eukaryotic cells. The Guidelines also apply to protein products used for *in vivo* diagnosis (e.g. monoclonal antibody products used for imaging), products used for *ex vivo* treatment, and those intentionally modified by, for example, PEGylation, conjugation with a cytotoxic drug, or modification of rDNA sequences. Some aspects of these Guidelines may apply to products produced in transgenic animals and plants. However, specific issues for such products can be found in the relevant documents published by WHO – for example, on products from transgenic plants (13) – and the responsible NRA should be consulted for specific advice on this.

Additional considerations for quality, safety, and efficacy of biosimilar products are available in the WHO Guidelines on evaluation of similar biotherapeutic products (4).

Some aspects of manufacturing and quality control in these Guidelines may apply to protein-based vaccine antigens made by rDNA technology. However, more-detailed guidelines/recommendations on vaccine evaluation in terms of quality, safety and efficacy should be consulted (2, 3). For example, WHO

guidance on vaccines such as yeast-derived hepatitis B vaccine or malaria vaccine produced by rDNA technology (14, 15). WHO guidance on DNA vaccines for therapeutic as well as prophylactic use, adopted by the WHO Expert Committee on Biological Standardization in 2005, is also available (16).

Protein products used for in vitro diagnosis are excluded.

## Terminology

The definitions given below apply to the terms used in this document. They may have different meaning in other contexts.

**Acceptance criteria:** numerical limits, ranges or other suitable measures for acceptance of the results of analytical procedures which the drug substance or drug product or materials at other stages of their manufacture should meet.

**Anti-drug antibody:** an antibody that binds to the active substance of a biotherapeutic product.

**Anti-product antibody:** an antibody that binds to the active substance, impurities or excipients of a biotherapeutic product.

**Biomarkers:** a laboratory measurement that reflects the activity of a disease process, correlates (either directly or inversely) with disease progression, and may also be an indicator of a therapeutic response. A genomic biomarker is a measurable DNA and/or RNA marker that measures the expression, function or regulation of a gene.

**Biotherapeutic:** a biological medicinal product with the indication of treating human diseases.

**Comparability exercise:** the activities – including study design, conduct of studies, and evaluation of data – that are designed to investigate whether a pre-change product and a post-change product are highly similar.

**Critical quality attribute:** a physical, chemical, biological or microbiological property or characteristic that is selected for its ability to help indicate the consistent quality of the product within an appropriate limit, range or distribution to ensure the desired product quality.

**Drug product:** a pharmaceutical product type in a defined container closure system that contains a drug substance, generally in association with excipients.

**Drug substance:** the active pharmaceutical ingredient and associated molecules that may be subsequently formulated, with excipients, to produce the drug product.

**Expiry date:** the date given on the individual container (usually on the label) of a product up to and including which the drug substance and drug product are expected to remain within specifications, if stored as recommended. The expiry date is established for each batch by adding the shelf-life period to the date of manufacture.

**Good clinical practice (GCP):** an international ethical and scientific quality standard for designing, conducting, recording and reporting trials that involve the participation of human subjects. Compliance with this standard provides public assurance that the rights, safety and well-being of trial subjects are protected, consistent with the principles that have their origin in the Declaration of Helsinki, and that the clinical trial data are credible.

**Good laboratory practice (GLP):** a quality system concerned with the organizational process and conditions under which nonclinical health and environmental safety studies are planned, performed, monitored, recorded, archived and reported.

**Good manufacturing practice (GMP):** that part of the pharmaceutical quality assurance process which ensures that products are consistently produced and meet the quality standards appropriate to their intended use as required by the marketing authorization. In these Guidelines, GMP refers to the current GMP guidelines published by WHO.

**Immunogenicity:** the ability of a substance to trigger an immune response or reaction (e.g. development of specific antibodies, T-cell response, or allergic or anaphylactic reaction).

**Impurity:** any component present in the drug substance or drug product that is not the desired product, a product-related substance, or excipient including buffer components. An impurity may be either process- or product-related.

**In-process control:** checks performed during production in order to monitor and, if necessary, to adjust the process to ensure that the intermediate or product conforms to its specifications. The control of the environment or equipment may also be regarded as a part of in-process control.

**In-silico modelling:** a computer-simulated model.

**Master cell bank (MCB):** an aliquot of a single pool of cells which generally has been prepared from the selected cell clone under defined conditions, dispensed into multiple containers and stored under defined conditions.

**Non-human primates (NHPs):** primates used as models for the study of the effects of drugs in humans prior to clinical studies.

**P450 (CYP) enzymes:** indicates the family of metabolizing enzymes which is the most common group.

**Pharmacodynamics (PD):** the study of the biochemical and physiological effects of drugs on the body and the mechanisms of drug action and the relationship between drug concentration and effect. One dominant example is drug–receptor interactions. PD is often summarized as the study of what a drug does to the body, as opposed to pharmacokinetics, which is the study of what the body does to a drug.

**Pharmacogenomics:** the study of the pharmacological correlation between drug response and variations in genetic elements has become of increasing importance for drug development. Such variations can have effects on

the risk of developing adverse drug reactions (ADRs) as well as on the response to treatment. Variations in drug pharmacokinetics and metabolic pathways can cause higher drug concentrations in some patients, resulting in increased drug toxicity, and/or lower drug concentrations in some patients, resulting in decreased drug effects.

**Pharmacokinetics (PK):** the study and characterization of the time course of drug absorption, distribution, metabolism and elimination. Pharmacokinetics is a quantitative analysis of how living systems handle foreign compounds.

**Pharmacovigilance:** the activities that are carried out after a medicinal product is marketed in order to observe and manage in a continuous manner the safety and the efficacy of the products.

**QT/QTc:** QT interval is a measure of the time between the start of the Q wave and the end of the T wave in the heart's electrical cycle on the electrocardiogram. It measures the conduction speed between the atria and the ventricles. A genetic predisposition to the prolongation of the QT interval can be triggered by several factors, including various medicinal products by themselves or as a result of their metabolic interaction. It is critical to understand whether a particular drug or a biological triggers the prolongation, because any prolongation of the QT interval outside of the normal limits determined for electrocardiograms indicates a potential for arrhythmia (disturbed heart rhythm), which is a serious adverse event during drug therapy. In extreme cases, this can lead to sudden death. Since the QT interval is affected by the heart rate, "corrected" QT (QTc) should also be used.

**rDNA-derived biotherapeutics:** biotherapeutics prepared by recombinant DNA technology, i.e. all biologically active protein products which are used in the treatment of human diseases and which are prepared by rDNA technology.

**Recombinant DNA technology:** technology that joins together (i.e. recombines) DNA segments from two or more different DNA molecules that are inserted into a host organism to produce new genetic combinations. It is also referred to as gene manipulation or genetic engineering because the original gene is artificially altered and changed. These new genes, when inserted into the expression system, form the basis for the production of rDNA-derived protein(s).

**Risk management plan:** a detailed description of the activities that continuously ensure patients' safety and their benefit from a medicinal ingredient. A risk management plan includes pharmacovigilance and many other elements.

**Shelf-life:** the period of time during which a drug substance or drug product, if stored correctly, is expected to comply with the specification, as determined by stability studies on a number of batches of the product. The shelf-life is used to establish the expiry date of each batch.

**Source material/starting material:** any substance of a defined quality used in the production of a biological medicinal product, but excluding packaging materials.

**Specification:** a list of tests, references to analytical procedures, and appropriate acceptance criteria which are numerical limits, ranges or other criteria for the tests described. Specifications are critical quality standards that are proposed and justified by the manufacturer and approved by regulatory authorities.

**Working cell bank (WCB):** the working cell bank is prepared from aliquots of a homogeneous suspension of cell obtained from culturing the master cell bank under defined culture conditions.

## Part A. Manufacturing and quality control

### A.1 Definitions

#### A.1.1 International name and proper name

Where an International Nonproprietary Name (INN) is available for an rDNA-derived biotherapeutic, the INN should be used (17). The proper name should be the equivalent of the INN in the language of the country of origin.

#### A.1.2 Descriptive definition

The description of an rDNA-derived biotherapeutic should indicate the biological system in which it is produced (e.g. bacterial, fungal or mammalian cells) as well as the presentation of the drug product.

#### A.1.3 International standards and reference materials

International standards and reference preparations have been established for a wide range of biological substances prepared by rDNA technology. These standards and materials are used either to calibrate assays directly or to calibrate secondary standards or manufacturers' working standards. A list of such materials is available on the WHO website.<sup>2</sup> Each standard or reference preparation is held by one of the WHO custodian laboratories (e.g. the National Institute for Biological Standards and Control, Potters Bar, England).

### A.2 General manufacturing guidelines

The present Guidelines cover the following three main areas:

- control of starting/source materials, including data both on the host cell and on the source, nature and sequence of the gene used in production;

<sup>2</sup> See: <http://www.who.int/bloodproducts/catalogue/AlphFeb2013.pdf>

- control of the manufacturing process;
- control of the drug substance and the drug product.

The quality, safety and efficacy of rDNA-derived products rely heavily on adequate control of the starting/source materials and on the manufacturing process, in addition to control tests on the drug substance and drug products themselves. These Guidelines therefore place considerable emphasis on the characterization and testing of host cell lines and other materials used during manufacturing and on validating the ability of the purification processes to remove or inactivate unwanted materials – especially possible viral contaminants and process-related impurities such as host-cell-derived proteins and DNA. The Guidelines also cover in-process controls in manufacturing and comprehensive characterization of the drug substance and the drug product.

Information should therefore be provided to describe adequately the starting/source materials, the manufacturing process and in-process controls. The description of the manufacturing process should be provided in the form of a flow diagram and sequential procedural narrative, and the in-process controls for each step or stage of the process should be indicated in this description. In addition, an explanation should be provided of how batches of the drug substance and drug product are defined (e.g. splitting and pooling of harvests or intermediates). Details of batch size or scale should also be included.

The manufacturing process should be validated before licensing. Process validation studies should include appropriate evaluation of the process and process steps (e.g. cell culture, harvest, purification, mixing, sterilization, filling) and the provision of evidence that they are capable of consistently delivering quality product and intermediates (i.e. meeting their predetermined specifications and quality attributes). The capacity of the purification procedures to remove product- and process-related impurities (e.g. unwanted variants, host-cell proteins, nucleic acids, resin leachates) should be investigated thoroughly (also see section A.4.2 and Appendix 1).

The general recommendations for manufacturing establishments contained in the WHO Good manufacturing practices: main principles for pharmaceutical products (18) and Good manufacturing practices for biological products (19), as well as those in the WHO Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks (5), should apply to establishments manufacturing rDNA-derived biotherapeutics. Not all of the requirements outlined in this Part are expected for each phase of clinical development (20).

### A.3 **Control of starting/source materials**

#### A.3.1 **Expression vector and host cell**

A description of the host cell, its source and history, and of the expression vector used in production, including source and history, should be provided in detail. The description should include details of the origin and identity of the gene being cloned as well as the construction, genetic elements contained and structure of the expression vector. An explanation of the source and function of the component parts of the vector, such as the origins of replication, promoters, or antibiotic markers, should be provided in addition to a restriction-enzyme map indicating at least those sites used in construction.

Methods used to amplify the expression constructs and to transform expression constructs into host cells, and the rationale used to select the cell clone for production, should be fully described. The vector within the cell, whether integrated or extrachromosomal, and the copy number, should be analysed. A host cell containing an expression vector should be cloned and used to establish a master cell bank (MCB) and the correct identity of the vector construct in the cell bank should be established. The genetic stability of the host-vector combination should be documented.

The nucleotide sequence of the cloned gene insert, including any codon optimization, and of the flanking control regions of the expression vector should be indicated and all relevant expressed sequences clearly delineated.

Any measures used to promote and control the expression of the cloned gene in the host cell during production should be described in detail.

#### A.3.2 **Cell bank system**

Typically, rDNA-derived biotherapeutics are produced using a cell bank system which involves a manufacturer's working cell bank (WCB) derived from an MCB. It is acknowledged that a WCB may not always be established in the early phases of development.

The type of cell bank system used, the size of the cell bank(s), the container (vials, ampoules, or other appropriate vessels) and closure system used, the methods for preparation of the cell bank(s) including the cryoprotectants and media used, and the conditions employed for cryopreservation or long-term storage should all be documented and described in detail.

Evidence should be provided for banked cell stability under defined storage conditions. Such evidence can be generated during the production of material from the banked cells and can be supported by a programme for stability monitoring that indicates attributes over time (e.g. data on cell viability upon thawing, stability of the host-vector expression system in the cell bank). Available data should be clearly documented and the proposed stability monitoring programme should be described in the marketing application. Evidence should

be provided of the stability of the host-vector expression in the cell bank both under storage and under recovery conditions.

For animal cells and animal-derived cell banks, reference should be made to the WHO Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks (5).

#### A.3.2.1 Control of cell banks

The characterization and testing of banked eukaryotic or prokaryotic cell substrates is a critical component of the control of rDNA-derived biotherapeutics. Cell banks should be tested to confirm the identity, purity and suitability of the cell substrate for the intended manufacturing use. The MCB should be characterized for relevant phenotypic and genotypic markers which should include the expression of the recombinant protein and/or presence of the expression construct. The testing programme chosen for a given cell substrate will vary according to the nature and biological properties of the cells (e.g. growth requirements) and its cultivation history (including use of human-derived or animal-derived biological reagents). The extent of characterization of a cell substrate may influence the type or level of routine testing needed at later stages of manufacturing. Molecular methods should be used to analyse the expression construct for copy number, insertions or deletions, and the number of integration sites. Requirements for bacterial systems expressing the protein from a plasmid or mammalian epigenetic expression should be distinguished from mammalian cell systems. The nucleic acid sequence should be shown to be identical to that determined for the expression construct and should correspond to that expected for the protein sequence. If any differences in nucleic acid sequences are identified, these should be clearly delineated and shown to be stable and capable of expressing the expected product consistently (see also section A.4.1.1).

Animal cell substrates are subject to contamination and have the capacity to propagate extraneous, adventitious organisms, such as mycoplasma and viruses. In addition, animal cells contain endogenous agents such as retroviruses that may raise safety concerns. Testing of cell substrates for both endogenous (e.g. retroviruses) and adventitious agents is critical. A strategy for testing cell banks for adventitious agents should be developed. This strategy should also involve an assessment of specific viruses and the families of viruses that may potentially contaminate the cell substrate. Such testing is described in detail in WHO's Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks (5) and the ICH guidelines Viral safety evaluation of biotechnology products derived from cell lines of human or animal origin (21).

Although cell substrates contaminated with microbial agents are generally not suitable for production, there are exceptions. For example, some murine cell lines that are widely used for the production of rDNA-derived biotherapeutics express endogenous retroviral particles. In such circumstances, risk mitigating strategies should be implemented, including the removal of such agents and/or their inactivation by physical, enzymatic and/or chemical treatment during processing of the rDNA-derived biotherapeutics.

In addition, tests of purity and identity should be performed once on each WCB. A specification that includes test methods and acceptance criteria should be established for the WCB.

#### A.3.2.2 Cell substrate genetic stability

The limit of in vitro cell age for production should be defined by the time of registration; it should be based on data derived from production cells expanded under pilot plant-scale or commercial-scale conditions to the proposed limit of in vitro cell age for production use or beyond. The production cells are generally obtained by expansion of cells from the WCB (22).

Specific traits of cells – which may include, for example, morphological or growth characteristics, biochemical or immunological markers, productivity of the desired product, or other relevant genotypic or phenotypic markers – may be useful for the assessment of cell substrate stability during the culture phase. The nucleotide sequence of the insert encoding the rDNA-derived biotherapeutic should be determined at least once after a full-scale culture for each MCB.

The molecular integrity of the gene being expressed and the phenotypic and genotypic characteristics of the host cell after long-term cultivation (i.e. end of production testing) should be established and defined by the time of registration.

#### A.3.3 Cell culture medium/other materials

Materials used in the manufacture of the drug substance (e.g. solvents, reagents, enzymes) should be listed, indicating where each material is used in the process. Information should be provided on the source, quality and control of these materials. There should also be information demonstrating that the materials (including biologically sourced materials, such as media components, monoclonal antibodies and enzymes) meet standards appropriate for their intended use (including the clearance or control of adventitious agents).

Media and other components should comply with the *WHO Guidelines on Transmissible Spongiform Encephalopathies in relation to Biological and Pharmaceutical Products* (23). The latest version of the *WHO Guidelines on Tissue Infectivity Distribution in Transmissible Spongiform Encephalopathies* (24) should also be consulted. Corresponding tables in this area are periodically updated as new data become available – see, for example, reference (25).

## A.4 **Control of the manufacturing process**

Adequate design of a process and knowledge of its capability are part of the strategy used to develop a manufacturing process which is controlled and reproducible, yielding a drug substance and drug product that consistently meet specifications. In this respect, limits are justified on the basis of information gathered from the entire process from early development through to commercial-scale production.

In-process controls are performed both at critical decision-making steps and at other steps where data serve not only to ensure the appropriate performance of the manufacturing process but also to demonstrate adequate quality during the production of both the drug substance and the drug product. Process parameters that are found to have an impact on the quality attributes of the drug substance or drug product should be controlled by suitable acceptance limits. Where appropriate, in-process controls may alleviate the need for routine testing of some quality attribute(s) at the level of the drug substance and/or drug product.

### A.4.1 **Cell culture**

#### A.4.1.1 **Production at finite passage**

Procedures and materials used both for cell growth and for the induction of the product should be described in detail. Acceptable limits for potential contamination should be set and the sensitivity of the methods used to detect it should be indicated. In case of contamination, the nature of the microbial contamination needs to be identified. Microbial and fungal contamination should be monitored according to Part A, section 5.2 of General requirements for the sterility of biological substances (26) or by methods approved by the NRA.

Data should be presented on the consistency of culture conditions and culture growth and on the maintenance of product yield. Criteria for the rejection of culture lots should be established. The maximum number of cell doublings or passage levels to be permitted during production should be specified taking into account the limit of in vitro cell age. For a process demonstrating consistent growth characteristics over the proposed cell age range for production, it may also be acceptable to define the cell age limit on the basis of the maximum number of permitted days in culture from thaw to the end of production.

Host-cell/vector characteristics at the end of production cycles should be monitored to establish consistency. For this purpose, information on the plasmid copy number or the degree of retention of the expression vector within the host cell may be of value, as may restriction enzyme mapping of the vector containing the gene insert. If the vector is present in multiple copies integrated into the host-cell genome, it may be difficult to confirm the rDNA sequence directly. In such cases, alternative approaches to confirming the sequence of insert-encoding

the rDNA-derived biotherapeutics should be considered and defined by the time of registration – e.g. restriction fragment length polymorphism (RFLP), fluorescence in situ hybridization (FISH), polymerase chain reaction single-strand conformation polymorphism (PCR-SSCP), Southern blot). For example, confirmation of protein sequence by peptide mapping might be an appropriate alternative to rDNA sequencing.

#### A.4.1.2 Continuous culture production

As recommended above, all procedures and materials used for cell culture and induction of the product should be described in detail and validated. In addition, particular consideration should be given to the procedures used in production control. Monitoring is necessary throughout the life of the culture, although the frequency and type of monitoring required depend on the nature of both the production system and the product.

Evidence should be produced to show that variations in yield or other culture parameters do not exceed specified limits. The acceptance of harvests for further processing should be clearly linked to the monitoring schedule being used, and a clear definition of “batch” of product should be established for further processing. Criteria for the rejection of harvests or termination of the culture should also be established. Tests for microbial contamination should be performed as appropriate to the harvesting strategy. In the case of continuous processing, multiple harvests from long fermentations could lead to a drift in some quality attributes, such as glycosylation, with the appearance of “new” variants with possible impacts on the quality, safety and efficacy of the product. Such drift should be appropriately addressed in process evaluation/validation studies.

The maximum period of continuous culture should be specified on the basis of information on the stability of the system and consistency of the product during and after this period. In long-term continuous culture, the cell line and product should be fully re-evaluated at intervals determined by information on the stability of the host-vector system and the characteristics of the product.

#### A.4.2 Purification

The methods used for the harvesting, extraction and purification of the product and related in-process controls, including acceptance criteria, should be described in detail and should be validated. Special attention should be given to the removal of viruses, nucleic acid, host-cell proteins and impurities considered to pose a risk of immunogenicity.

The ability of the purification procedure to remove unwanted product-related or process-related impurities (e.g. host-cell-derived proteins, nucleic acid, viruses and other impurities, including media-derived compounds and

undesirable chemicals introduced by the purification process itself) should be investigated thoroughly, as should the reproducibility of the process. Particular attention should be given to demonstrating the removal and/or inactivation of possible contaminating viruses and residual DNA from products manufactured using continuous cell lines.

#### A.4.2.1 Residual cellular DNA from continuous cell lines

The ability of the manufacturing process to reduce the amount of residual cellular DNA (rcDNA) to an acceptable level, to reduce the size of the rcDNA or to chemically inactivate the biological activity of this DNA should be demonstrated.

Acceptable limits on the amount of rcDNA, as well as points to be considered concerning the size of rcDNA in an rDNA-derived biotherapeutic, are discussed in WHO's Recommendations for the evaluation of animal cell substrates for the manufacture of biological medicinal products and for the characterization of cell banks (5). In setting these limits, there should be consideration of the characteristics of the cell substrate, the intended use and route of administration of the rDNA-derived biotherapeutics and, most importantly, the effect of the manufacturing process on the size, quantity and biological activity of the residual host-cell DNA fragments. In general, it has been possible to reduce rcDNA levels in rDNA-derived biotherapeutics to < 10 ng per dose. Alternatively, once validation studies (e.g. spiking studies using an adequate size distribution of DNA) have been performed, and once the reproducibility of the production process in reducing residual DNA to the level expected has been demonstrated, rcDNA testing may be omitted.

#### A.4.2.2 Virus clearance

For cell substrates of human or animal origin, virus removal or inactivation processes, individually and overall, should be shown to be able to remove/inactivate any contaminating viruses and to ensure viral safety in the drug substance.

Where appropriate, validation studies (see Appendix 1) should be undertaken using small-scale studies with carefully selected model viruses in order to evaluate the virus clearance/inactivation capability of selected process steps and of the overall process, aiming at a significant safety margins. The results will indicate the extent to which these contaminants can theoretically be inactivated and removed during purification.

The overall manufacturing process – including the testing and selection of the cells and source materials, as well as the validation of the ability of the purification process to adequately remove possible contaminants – should ensure the absence of infectious agents in the drug product. Nevertheless, to complement such approaches, routine testing of the fermentation process for the absence of contamination by infectious viruses is also recommended. A sample of the

unprocessed bulk following fermentation constitutes one of the most suitable levels at which adventitious virus contamination can be determined with a high probability of detection. A programme of ongoing assessment of adventitious viruses in fermentation should be undertaken. The scope, extent and frequency of virus testing on the unprocessed bulk should take into account the nature of the cell lines used, the results and extent of virus testing performed during the qualification of the MCB and WCB, the cultivation method, the source materials used, and the results of virus clearance studies. In vitro screening tests using one or more cell lines are generally used to test unprocessed bulk. If appropriate, a PCR test or other suitable methods may be used.

If contamination by adventitious viruses is detected in the unprocessed bulk, the manufacturing process should be carefully checked to determine the cause of the contamination and to decide on appropriate action. Typically, adventitious virus contamination leads to the batch being discarded.

Further considerations of the detection, elimination and inactivation of viruses in animal cell substrates used in the production of rDNA-derived biotherapeutics, as well as the problem of rcDNA, can be found in the WHO Recommendations for the evaluation of animal cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks (5) as well as in the ICH guidelines Viral safety evaluation of biotechnology products derived from cell lines of human or animal origin (21).

## A.5 Control of drug substance and drug product

### A.5.1 Characterization

Rigorous characterization of the rDNA-derived biotherapeutics by chemical, physicochemical and biological methods is essential. Characterization is typically performed in the development phase to determine the physicochemical properties, biological activity, immunochemical properties, purity and impurities of the product, and – following significant process changes and/or for periodic monitoring – to confirm the quality of the product. Characterization allows appropriate release specifications to be established.

Particular attention should be given to using a wide range of analytical techniques that exploit different physicochemical properties of the molecule (e.g. size, charge, isoelectric point, amino acid sequence, hydrophobicity). Post-translational modifications such as glycosylation should be identified and adequately characterized. It may also be necessary to include suitable tests to establish that the product has the desired conformation and higher order structure. In addition to evaluation of purity, there should also be investigation of impurities (e.g. aggregates including dimers and higher multiples of the desired product). The rationale for selecting the methods used for characterization should

be provided and their suitability should be justified since the characterization of the product is intended to identify attributes that may be important to the overall safety and efficacy of the product. Details of the expected characterization of an rDNA-derived biotherapeutic and techniques suitable for such purposes are set out in Appendix 2. The specific technical approach employed will vary from product to product; alternative approaches, other than those included in Appendix 2, will be appropriate in many cases. New analytical technologies and modifications to existing technologies are continually being developed and should be utilized when appropriate.

Where relevant and possible, characteristics of the properties of the product should be compared with those of its natural counterpart. For example, post-translational modifications such as glycosylation are likely to differ from those found in the natural counterpart and may influence the biological, pharmacological and immunological properties of the rDNA-derived biotherapeutics.

#### A.5.2 Routine control

Not all the characterization and testing described in section A.5.1 and in Appendix 2 needs to be carried out on each batch of drug substance and drug product prior to release for licensing or clinical use. Some tests may need to be performed only initially and/or periodically to establish or verify the validity or acceptability of a product and of its manufacturing process. Other tests may be required on a routine basis. A comprehensive analysis of the initial production batches is expected in order to establish consistency with regard to identity, purity and potency. A more limited series of tests is appropriate for routine control, as outlined below and in more detail in Appendix 3. Tests for use in routine control should be chosen to confirm quality. The rationale and justification for including and/or excluding testing for specific quality attributes should be provided.

An acceptable number of consecutive batches should be characterized to determine the consistency of analytical parameters at the time of licensing, unless otherwise justified. Any differences between one batch and another should be noted. Data obtained from such studies, as well as knowledge gained from clinical and nonclinical development and during stability studies, should be used as the basis for establishing product specifications.

The selection of tests to be included in the routine control programme will be product-specific and should take into account the quality attributes (e.g. potential influence on safety, efficacy or stability), the process performance (e.g. clearance capability, content), the controls in place through the manufacturing process (e.g. multiple testing points), and the material used in relevant nonclinical and clinical studies. These tests should include criteria such as potency, the nature and quantity of product-related substances, product-related impurities, process-related impurities, and absence of contaminants.

## A.6 Filling and containers

The general requirements for filling and containers given in WHO Good manufacturing practices for biological products (19) should apply.

A description of the container closure systems for the drug substance and the drug product should be provided, including a specification for their component materials.

Evidence shows that formulated proteins can interact chemically or physically with the formulation excipients and/or the container closure system, and could therefore influence the quality, safety or pharmacological properties of the product. Some products have been shown to form aggregates with excipients, and such aggregates may lead to the formation of potentially immunogenic complexes – see, for example references (27–29). The suitability of the container closure system should be evaluated and described for its intended use. This should cover evaluation of the compatibility of the container construction materials with the formulated product, including adsorption to the container, leaching, and other chemical or physical interactions between the product and the materials in contact with it. The integrity of the closure and its ability to protect the formulation from contamination and to maintain sterility needs to be ensured.

When a delivery device is presented as part of the drug product (e.g. prefilled syringe, single-use autoinjector), it is important to demonstrate the functionality of such a combination – such as the reproducibility and accuracy of the dispensed dose under testing conditions which should simulate the use of the drug product as closely as possible. For multi-use containers such as vials or cartridges for a pen injector, proper in-use stability studies should be performed to evaluate the impact of the in-use period of the vial or the assembled device on the formulation and the functionality of the pen injector. Dose accuracy should be demonstrated for the first and last dose delivered. In addition, the effect of multiple injections/withdrawals on the closure system should be evaluated.

## A.7 Records, retained samples, labelling, distribution and transport

The requirements given in the WHO Good manufacturing practices for biological products (19) should apply.

The conditions of shipping should be such as to ensure that the products are maintained in appropriate conditions.

## A.8 Stability, storage and expiry date

### A.8.1 Stability studies

While the expectations outlined in this section are primarily applicable to the marketing application stage, products in clinical development should be tested for stability concurrently with clinical trials. For proteins, maintenance of biological

activity is generally dependent on maintaining molecular conformation. Such products can be particularly sensitive to environmental factors such as temperature changes, oxidizing factors, and light exposure. In order to ensure the maintenance of biological activity and to avoid degradation, appropriate conditions for storage are necessary.

A detailed protocol for the assessment of the stability of both drug substance and drug product in support of the proposed storage conditions and expiration dating periods should be developed. This should include all information necessary to demonstrate the stability of the rDNA-derived biotherapeutics throughout the proposed shelf-life, including, for example, well-defined specifications and test intervals.

Each product should remain within its specification for stability-indicating attributes, including potency throughout its proposed shelf-life. Specifications should be derived from all available information using appropriate statistical methods at the time of licensing. There is no single stability-indicating assay or parameter that profiles the stability characteristics of an rDNA-derived biotherapeutic. Consequently, the manufacturer should develop a stability-indicating programme that provides assurance that changes in the quality and potency of the product will be detected.

Primary data to support a requested storage period for both drug substance and drug product should be based on long-term, real-time, and real-condition stability studies, and these should be further supported by accelerated- and stress-condition stability data, as available, to justify the claimed shelf-life. In cases where the stability of the product is influenced by the storage of intermediates (e.g. a significant degradation trend is observed during storage of an intermediate), a cumulative stability study should be considered. This study should include all intermediates stored at the longest storage time claimed, or a selection of the most storage-sensitive intermediates, as appropriate. In view of the time necessary to generate the data, inclusion of study results may not be feasible at the time of licensing. The absence of such a cumulative study could be justified on the basis of a proposed stability programme that will include such monitoring. In addition, stability studies should include an evaluation of the impact of the container closure system on the formulated rDNA-derived biotherapeutics throughout the shelf-life. In order to ensure that the formulated product is in contact with all material of the container closure system, stability studies should include samples maintained in the inverted or horizontal position (i.e. in contact with the closure). Data should be supplied for all different container closure combinations that will be marketed.

Stability information should be provided on at least three batches for which manufacture and storage are representative of the commercial process.

When shelf-lives of 1 year or less are proposed, real-time stability studies should be conducted monthly for the first 3 months and at 3-month intervals

thereafter. For products with proposed shelf-lives greater than 1 year, the studies should be conducted every 3 months during the first year of storage, every 6 months during the second year, and annually thereafter. A minimum of 6 months' data at the time of submission should be submitted in cases where storage periods greater than 6 months are requested, unless otherwise justified. For storage periods of less than 6 months, the minimum amount of stability data in the initial submission should be determined on a case-by-case basis.

It is recommended that stability studies under accelerated and stress conditions, including the impact of the container closure system (see section A.6), should also be conducted on the drug product. Studies under accelerated conditions may: (a) provide useful supportive data for establishing the expiry date; (b) provide product stability information for future product development (e.g. preliminary assessment of proposed manufacturing changes such as changes in formulation or scale-up); (c) assist in validation of analytical methods for the stability programme; or (d) generate information which may help elucidate the degradation profile of the rDNA-derived biotherapeutics. Studies under stress conditions may also be useful for determining whether accidental exposures to conditions other than those proposed (e.g. during transportation) are deleterious to the product and for evaluating which specific test parameters may be the best indicators of product stability.

Further guidance on both general and specific aspects of stability testing of an rDNA-derived biotherapeutic can be obtained by consulting the WHO guidelines on the stability testing of active pharmaceutical ingredients and finished pharmaceutical products (30), as well as the WHO Guidelines on stability evaluation of vaccines (31).

#### A.8.2 Drug product requirements

Stability information should be provided on at least three batches of drug product that are representative of that which will be used in commercial manufacture, and presented in the final container. Where possible, the drug product batches included in stability testing should be derived from different batches of drug substance.

Where one product is distributed in multiple presentations, the samples to be entered into the stability programme may be selected on the basis of a matrix system and/or by bracketing. Where the same strength and exact container/closure system is used for three or more fill contents, the manufacturer may elect to place only the smallest and largest container size into the stability programme (i.e. bracketing). The design of a protocol that incorporates bracketing assumes that the stability of the intermediate condition samples is represented by those at the extremes. In certain cases, data may be needed to demonstrate that all samples are properly represented by data collected for the extremes.

Matrixing (i.e. the statistical design of a stability study in which account is taken of factors such as the tests, process characteristics, presentation characteristics and different testing time-points) should be applied only when appropriate documentation is provided confirming that the stability of the samples tested represents the stability of all samples. The differences in the samples for the same drug product should be identified as, for example, covering different batches, different strengths, different sizes of the same closure and, possibly in some cases, different container/closure systems. Matrixing should not be applied to samples with differences that may affect stability, such as different strengths and different containers/closures, where it cannot be confirmed that the products respond similarly under storage conditions.

For preparations intended for use after reconstitution, dilution or mixing, in-use stability data should be obtained. The stability should be demonstrated up to and beyond the storage conditions and the maximum storage period claimed.

In addition to the standard data necessary for a conventional single-use vial, it should be shown that the closure used with a multiple-dose vial is capable of withstanding the conditions of repeated insertions and withdrawals so that the product retains its identity, strength, potency, purity and quality for the maximum period specified in the instructions for use on containers, packages and/or package inserts.

#### A.9 **Manufacturing process changes**

Changes to the manufacturing process of an rDNA-derived biotherapeutic often occur both during development and after approval. The reasons for such changes include, for example, improvement of the manufacturing process, increase in scale, a site change, improvement of product stability, or compliance with changes in regulatory requirements. When substantial changes are made to the manufacturing process, a comparability exercise to evaluate the impact of the change(s) on the quality, safety and efficacy of the rDNA-derived biotherapeutics should be considered. The extent of such an exercise depends on the potential impact of the process change(s) as well as on the manufacturer's experience in the process and knowledge of the product. The demonstration of comparability does not necessarily mean that the quality attributes of the pre-change and post-change product are identical, but rather that they are highly similar and that the existing knowledge is sufficiently predictive to ensure that any differences in quality attributes have no adverse impact on the safety or efficacy of the rDNA-derived biotherapeutics. The reason for each significant change should be explained, together with an assessment of its potential to impact on quality, safety and efficacy.

The extent of a comparability exercise depends on the potential impact of the process change(s) on the quality, safety and efficacy of the product.

A comparability exercise can range from analytical testing alone (e.g. where process changes lead to no changes in any quality attribute) to a comprehensive exercise requiring nonclinical and clinical bridging studies (e.g. the establishment of a new host cell line with altered properties resulting in more pronounced changes in quality attributes). If assurance of comparability can be shown through analytical studies alone, nonclinical or clinical studies with the post-change product may not be necessary. However, where the relationship between specific quality attributes and safety and efficacy has not been established, and differences between quality attributes of the pre-change and post-change product are observed, it may be appropriate to include a combination of quality, nonclinical and/or clinical studies in the comparability exercise.

Further considerations of manufacturing changes can be found in guidelines provided by the ICH (32), the EMA (33), the United States Food and Drug Administration (34) and other major NRAs.

## Part B. Nonclinical evaluation

### B.1 Introduction

The general aim of nonclinical evaluation is to determine whether new medicinal products possess the desired pharmacodynamic (PD) activity and whether they have the potential to cause unexpected and undesirable effects. However, classic PD, safety or toxicological testing, as recommended for chemical drugs, may be of only limited relevance for rDNA-derived biotherapeutics due to the latter's unique and diverse structural and biological properties, including species specificity, immunogenicity, and unpredicted pleiotropic activities. These properties pose particular problems in relation to nonclinical testing in animals, and their pharmacological and safety evaluation will have to take a large number of factors into account. Thus, a flexible approach is necessary for the nonclinical evaluation of rDNA-derived biotherapeutics. For example, certain proteins (e.g. interferons) are highly species-specific, so that the human protein is pharmacologically much more active in humans than in any animal species. Furthermore, human proteins frequently produce immunological responses in animal species which may ultimately modify their biological effects and may result in toxicity (e.g. due to immune complex formation). Such toxicity has little bearing on the safety of the product in the intended human host.

Although some safety testing will be required for most products, the range of tests that need to be carried out should be decided on a case-by-case basis (see Appendix 4) in consultation with the NRA or national control laboratory. A wide range of pharmacological, biochemical, immunological, toxicological and histopathological investigative techniques should be used, where appropriate, to assess a product's effect over an appropriate range of doses and, in accordance

with the desired clinical indication(s), during both acute and chronic exposure. However, the points made above concerning species specificity and antibody formation should always be taken into consideration.

Additional information on specific safety issues – such as, for example, carcinogenic potential, reproductive toxicity or safety pharmacology – is provided in respective ICH safety guidelines (35–37). Relevant sections of this part may be useful with regard to products intended for clinical trials; however, the amount and the extent of data submitted for a product will be limited and will need to take into account the nature of the product and its stage of development. Recommendations concerning timing and interplay of nonclinical and clinical studies in drug development are given in the ICH Guidance on nonclinical safety studies for the conduct for human clinical trials and marketing authorization for pharmaceuticals (38) and in the ICH guideline Preclinical safety evaluation of biotechnology-derived pharmaceuticals (39).

#### B.1.1 Objectives of the nonclinical evaluation

The objectives of nonclinical studies are to define pharmacological and toxicological effects throughout clinical development, not only prior to initiation of human studies.

The primary goals are:

- to identify an initial safe dose and subsequent dose escalation schemes in humans;
- to identify potential target organs for toxicity and for the study of whether such toxicity is reversible;
- to identify safety parameters for clinical monitoring.

Nonclinical evaluation should consider:

- selection of the pharmacologically or toxicologically relevant animal species;
- the age of the animals;
- the physiological state of the animals (e.g. whether healthy/diseased animals are used, whether treatment-naïve animals are used);
- the weight of the animals;
- the manner of delivery, including relevant dose or amount, route of administration, and treatment regimen;
- stability of the test material under the conditions of use;
- interpretation of results.

Both in vitro and in vivo studies can contribute to this characterization.

rDNA-derived biotherapeutics that belong structurally and pharmacologically to a product class for which there is wide experience in clinical practice may need less-extensive toxicity testing.

### B.1.2 Product development and characterization

In general, the product that is used in the definitive pharmacology and toxicology studies should be representative of the product proposed for the initial clinical studies. However, it is appreciated that during the course of development programs, changes normally occur in the manufacturing process in order to improve product quality and yields. The potential impact of such changes for extrapolation of the animal findings to humans should be considered, including the impact of post-translational modifications.

The comparability of the test material should be demonstrated when a new or modified manufacturing process or other significant changes in the product or formulation are made in an ongoing development program. Comparability can be evaluated on the basis of biochemical and biological characterization (i.e. identity, purity, stability and potency). In some cases, additional studies may be needed (i.e. PK, PD and/or safety). The scientific rationale for the approach taken should be provided.

### B.1.3 Good laboratory practice

Pivotal (toxicity) studies should be performed in compliance with good laboratory practice (GLP) (40, 41). However, it is recognized that some studies employing specialized test systems which are often needed for rDNA-derived biotherapeutics may not comply fully with GLP. Areas of non-compliance should be identified and their significance evaluated relative to the overall nonclinical assessment. In some cases, lack of full GLP compliance does not necessarily mean that the data from these studies cannot be used to support clinical trials and marketing authorization. However, justification which is supported with data, such as method validation should be provided for the data quality assurance.

## B.2 Pharmacodynamics

### B.2.1 Primary and secondary pharmacodynamics/biological activity

Biological activity may be evaluated by the use of in vitro assays to determine which effects of the product may be related to clinical activity. The use of cell lines and/or primary cell cultures can be useful to examine the direct effects on cellular phenotype and proliferation. Due to the species specificity of many rDNA-derived biotherapeutics, it is important to select relevant animal species for testing (see Appendix 5). Non-human primates (NHPs) are often the only pharmacologically or toxicologically relevant species; however, other species should also be evaluated for relevant biological activity. In vitro cell lines derived

from mammalian cells can be used to predict specific aspects of in vivo activity and to assess quantitatively the relative sensitivity of various species, including humans, to the biotherapeutics. Such studies may be designed to determine, for example, receptor occupancy, receptor affinity, and/or pharmacological effects, and to assist in the selection of an appropriate animal species for further in vivo pharmacology and toxicology studies. The combined results from in vitro and in vivo studies assist in the extrapolation of the findings to humans. In vivo studies to assess pharmacological activity, including defining mechanism(s) of action, are often used to support the rationale for the proposed use of a product in clinical studies. When feasible, PD end-points can be incorporated into general toxicity studies (e.g. haemoglobin blood concentration in repeated dose toxicity studies with erythropoietins).

### B.2.2 Safety pharmacology

According to the target or mechanism of action of the product, it is important to investigate the potential for undesirable pharmacological activity in appropriate animal models. The aim of the safety pharmacology studies is to reveal any functional effects on the major physiological systems (e.g. cardiovascular, respiratory, central nervous system). These functional indices may be investigated in separate studies or incorporated in the design of toxicity studies and/or clinical studies. Investigations may include the use of isolated organs or other test systems not involving intact animals. All of these studies may allow for a mechanistically based explanation of specific organ effects/toxicities, which should be considered carefully with respect to applicability for human use and indication(s).

## B.3 Pharmacokinetics/toxicokinetics

### B.3.1 General principles

It is difficult to establish uniform guidelines for PK studies for rDNA-derived biotherapeutics. Single-dose and multiple-dose PK, toxicokinetics (TK) and tissue distribution studies in relevant species are useful; however, routine studies that attempt to assess mass balance are not useful. Differences in PK between animal species may have a significant impact on the predictiveness of animal studies or on the assessment of dose–response relationships in toxicity studies. Scientific justification should be provided for the selection of the animal species used for PK/TK evaluation, taking into account that the PK profile in the chosen animal species should ideally reflect the PK profile in humans. Alterations in the PK profile due to immune-mediated clearance mechanisms may affect the kinetic profiles and the interpretation of the toxicity data (see also section B.4.8.1). For some products there may also be significant inherent delays in the expression of PD effects relative to the PK profile (e.g. cytokines) or there may be prolonged expression of PD effects relative to plasma levels.

PK studies should, whenever possible, utilize preparations that are representative of that intended for toxicity testing and clinical use, and should employ a route of administration that is relevant to the anticipated clinical studies. Patterns of absorption may be influenced by formulation, active substance concentration, application site, and/or application volume. Whenever possible, systemic exposure should be monitored during the toxicity studies. When feasible, PK/TK evaluations can be incorporated into general toxicity studies.

Some information on absorption, disposition and clearance in relevant animal models should be available prior to clinical studies in order to predict margins of safety based on exposure and dose. Understanding the behaviour of the biotherapeutic in the biological matrix (e.g. plasma, serum, cerebral spinal fluid) and the possible influence of binding proteins is important for understanding the PD effect.

### B.3.2 Assays

The use of one or more assay methods should be addressed on a case-by-case basis and the scientific rationale should be provided. One validated method is usually considered sufficient. For example, quantitation of trichloroacetic acid (TCA)-precipitable radioactivity following administration of a radiolabelled protein may provide adequate information, but a specific assay for the analyte is preferred. Ideally, the assay methods should be the same for animal and human studies. The possible influence of plasma-binding proteins and/or antibodies in plasma/serum on the performance of the assay should be determined.

### B.3.3 Distribution

Unlike small chemical drugs that readily diffuse, rDNA-derived biotherapeutics, due to their molecular weight, usually do not readily do so but, following intravenous application, are initially confined to the vascular system. However, with time they may distribute to the extravascular space as a result of various factors, including bulk flow and active transport.

As a supplement to standard tissue distribution studies, complementary information about the tissue distribution of molecular targets for rDNA-derived biotherapeutics may be obtained from tissue cross-reactivity (TCR) studies, if appropriate (see section B.4.8.3).

Tissue concentrations of radioactivity and/or autoradiography data using radiolabelled proteins may be difficult to interpret due to rapid protein metabolism in vivo or unstable radiolabelled linkage. Care should be taken in interpreting studies using radioactive tracers incorporated into specific amino acids because of the possibility of recycling of amino acids into non-drug-related proteins/peptides.

### B.3.4 Metabolism

The expected consequence of metabolism of rDNA-derived biotherapeutics is degradation to small peptides and individual amino acids. Therefore, the metabolic pathways are generally understood. Classic biotransformation studies, as performed for pharmaceuticals, are not needed.

## B.4 Toxicity studies

### B.4.1 General principles

#### B.4.1.1 Number/gender of animals

For ethical reasons, it is desirable to apply the 3Rs concept of “Replace Reduce Refine” to minimize the use of animals in research, and consideration should be given to the use of appropriate in vitro alternative methods for safety evaluation (42).

The number of animals used per dose has a direct bearing on the ability to detect toxicity. A small sample size may lead to failure to observe toxic events due to observed frequency alone, regardless of severity. The limitations that are imposed by sample size, as often is the case for NHP studies, may in part be compensated by increasing the frequency and duration of monitoring. Both genders should generally be used or justification given for specific omissions. As an example, the minimum sample size for a pivotal GLP toxicity study in NHPs is considered to be three animals per sex and, if a recovery group is included in the study, an additional minimum of two animals per sex would be included.

#### B.4.1.2 Administration/dose selection and application of PK/PD principles

The route and frequency of administration should be as close as possible to that proposed for clinical use. Consideration should be given to the pharmacokinetics and bioavailability of the product in the species being used, as well as the volume that can be safely and humanely administered to the test animals. For example, the frequency of administration in laboratory animals may be increased compared to the proposed schedule for the human clinical studies in order to compensate for faster clearance rates or low solubility of the active ingredient. In these cases, the level of exposure of the test animal should be defined relative to the clinical exposure. Consideration should also be given to the effects of application volume, active substance concentration, formulation, and site of administration. The use of routes of administration other than those used clinically may be acceptable if the route must be modified due to limited bioavailability, limitations due to the route of administration, or to size/physiology of the animal species used.

If feasible, dosage levels should be selected in order to provide information on a dose–response relationship, including a toxic dose and a “no observed adverse effect level” (NOAEL). These data may be used for estimating the maximum

recommended starting dose in initial clinical trials (43). In addition, for selection of a safe starting dose for first-in-human clinical trials (38), the identification of the “minimum anticipated biological effect level” (MABEL) (44) should be considered. For oncology products, see Appendix 4.

The toxicity of most rDNA-derived biotherapeutics is related to their targeted mechanism of action; therefore, relatively high doses can elicit adverse effects which are apparent as exaggerated pharmacology. For some classes of product which show little or no toxicity it may not be possible to define a specific maximum dose. In these cases, a scientific justification of the rationale for the dose selection and projected multiples of human exposure should be provided. To justify selection of a high dose, consideration should be given to the expected pharmacological/physiological effects and the intended clinical use. Where a product has a lower affinity for, or potency in, the cells of the selected species than for human cells, testing of higher doses may be important. The multiples of the human dose that are needed to determine adequate safety margins may vary with each class of rDNA-derived biotherapeutics and its clinical indication(s).

A rationale should be provided for dose selection, taking into account the characteristics of the dose–response relationship. PK-PD approaches (e.g. simple exposure–response relationships or more complex modelling and simulation approaches) can assist in high-dose selection by identifying: (a) a dose which provides the maximum intended pharmacological effect in the selected animal species; and (b) a dose which provides an approximately 10-fold exposure multiple over the maximum exposure to be achieved in the clinic. The higher of these two doses should be chosen for the high-dose group in nonclinical toxicity studies unless there is a justification for using a lower dose (e.g. maximum feasible dose).

Where *in vivo/ex vivo* PD end-points are not available, the high-dose selection can be based on PK data and on available *in vitro* binding and/or pharmacology data. Corrections for differences in target binding and *in vitro* pharmacological activity between the nonclinical species and humans should be taken into account to adjust the exposure margin over the highest anticipated clinical exposure. For example, a large relative difference in binding affinity and/or *in vitro* potency might suggest that testing higher doses in the nonclinical studies is appropriate. In the event that toxicity cannot be demonstrated at the doses selected using this approach, then additional toxicity studies at higher multiples of human dosing are unlikely to provide additional useful information.

#### B.4.1.3 Use of one or two species

With regard to the use of one or two species for toxicity studies, see Appendix 5.

#### B.4.1.4 Study duration

For chronic-use products, repeat dose toxicity studies of 6 months’ duration in rodents or non-rodents are usually considered sufficient so long as the high dose is

selected in accordance with the principles above. Studies of longer duration have not generally provided useful information that has changed the clinical course of development (see also section B.4.3). Performance of (6-month) chronic toxicity studies may not always be feasible (e.g. if an induction of anti-drug antibodies prevents a meaningful study interpretation). For chronic use of rDNA-derived biotherapeutics developed for patients with advanced cancer, see Appendix 4.

#### B.4.1.5 Evaluation of immunogenicity

Many rDNA-derived biotherapeutics intended for human use are immunogenic in animals. Therefore, plasma samples from animals subjected to repeated dose toxicity studies should be stored at an appropriate temperature and analysed for the presence of anti-drug antibody when considered necessary for study interpretation (see section B.4.8.1).

#### B.4.2 Single-dose toxicity studies

In general, single-dose toxicity studies should be pursued only in cases where significant toxicity is anticipated and the information is needed to select doses for repeated dose studies (38, 39). Single-dose studies may generate useful data to describe the relationship of dose to systemic and/or local toxicity. These data can be used to select doses for repeated dose toxicity studies. Information on dose–response relationships may be gathered through the conduct of a single-dose toxicity study as a component of pharmacology or animal-model efficacy studies. The incorporation of safety pharmacology parameters in the design of these studies should be considered.

#### B.4.3 Repeated dose toxicity studies

For consideration of the selection of animal species for repeated dose studies, see section B.4.1. The route and dosing regimen (e.g. daily versus intermittent dosing) should reflect the intended clinical use or exposure. When feasible, these studies should include TK measurements, but interpretation should consider the formation of possible anti-drug antibodies (see section B.4.8.1).

##### B.4.3.1 Study duration

The duration of repeated dose studies should be based on the intended duration of clinical exposure and disease indication. Duration of animal dosing has generally been 1–3 months for most rDNA-derived biotherapeutics. For rDNA-derived biotherapeutics intended for short-term use (e.g. <7 days) and for acute life-threatening diseases, repeated dose studies of up to 2 weeks' duration have been considered adequate to support clinical studies as well as marketing authorization. For those rDNA-derived biotherapeutics intended for chronic indications, studies of 6 months' duration have generally been appropriate,

although in some cases shorter or longer durations have supported marketing authorizations. For rDNA-derived biotherapeutics intended for chronic use, the duration of long-term toxicity studies should be scientifically justified.

#### B.4.3.2 Recovery period

Recovery from pharmacological and toxicological effects with potential adverse clinical impact should be understood when they occur at clinically relevant exposures. This information can be obtained by understanding that the particular effect observed is generally reversible/nonreversible, or by including a non-dosing period in at least one study, at least at one dose level, to be justified by the sponsor. The purpose of the non-dosing period is to examine reversibility of these effects and not to assess delayed toxicity. The demonstration of complete recovery is not considered essential. The addition of a recovery period for the sole purpose of assessing the potential for immunogenicity is not required.

#### B.4.4 Genotoxicity studies

The range and type of genotoxicity studies routinely conducted for pharmaceuticals are not applicable to rDNA-derived biotherapeutics and are therefore not needed. Moreover, the administration of large quantities of peptides/proteins may yield un-interpretable results. It is not expected that these substances will interact directly with DNA or other chromosomal material.

With some rDNA-derived biotherapeutics there is a potential concern about accumulation of spontaneously mutated cells (e.g. via facilitating a selective advantage of proliferation), leading to carcinogenicity. The standard battery of genotoxicity tests is not designed to detect these conditions. Alternative *in vitro* or *in vivo* models to address such concerns may have to be developed and evaluated (see section B.4.5).

Studies in available and relevant systems, including newly developed systems, should be performed in those cases where there is cause for concern about the product (e.g. because of the presence of an organic linker molecule in a conjugated protein product).

The use of standard genotoxicity studies for assessing the genotoxic potential of process contaminants is usually not considered appropriate. If performed for this purpose, however, the rationale should be provided.

#### B.4.5 Carcinogenicity studies

##### B.4.5.1 General principles

Carcinogenicity is, in the strict sense, increased probability of development of new tumours. However, activation of proliferation and progression of existing tumour cells/tumours should also be considered.

The need for a product-specific assessment of the carcinogenic potential of rDNA-derived biotherapeutics should be determined with regard to the intended clinical population and treatment duration – see, for example, reference (36). When an assessment is warranted, the sponsor should design a strategy to address the potential hazard. This strategy could be based on a review of relevant data from a variety of sources. The data sources can include published data (e.g. information from transgenic, knock-out or animal disease models, and human genetic diseases), information on class effects, detailed information on target biology and mechanism of action, in vitro data, and data from chronic toxicity studies and clinical data. In some cases, the available information can be sufficient to address carcinogenic potential and inform clinical risk without additional nonclinical studies.

The mechanism of action of some rDNA-derived biotherapeutics may raise concern regarding potential for carcinogenicity (e.g. immunosuppressives and growth factors). If the review of all available data (see above) supports this concern, rodent bioassays are not warranted. In this case, potential hazard can be best addressed by product labelling and risk management practices. When a review of all available data suggests that there is no carcinogenic concern, no additional testing is needed. However, if the potential for carcinogenicity remains unclear after a review of all available data, the sponsor can propose additional studies that could mitigate the mechanism-based concern – see, for example, reference (39). When a review of all available data, including the additional study data, supports concern regarding carcinogenic potential, this is best addressed by product labelling and risk management practices. Correspondingly, if the potential for carcinogenicity remains unclear after the extended data review, this should also be addressed by product labelling and risk management practices. In case the concern regarding carcinogenicity is mitigated by the additional study data, this should be reflected in the product information.

For products where there is insufficient knowledge about specific product characteristics and mode of action in relation to carcinogenic potential, a more extensive assessment might be appropriate (e.g. understanding of target biology related to potential carcinogenic concern, and inclusion of additional end-points in toxicity studies). If the review of all data from this more extensive assessment does not suggest a carcinogenic potential, no additional nonclinical testing is recommended. Alternatively, if the review of all data available suggests a concern about carcinogenic potential, then the sponsor can propose additional nonclinical studies that could mitigate the concern (see above), or the label should reflect the concern.

The selection of animal models for the assessment of tumour growth potential should take into account that rDNA-derived biotherapeutics may have secondary, unspecific effects on tumour growth which would be clinically irrelevant. Careful design and choice of controls should be used to avoid misinterpretations.

#### B.4.5.2 Use of homologous proteins

A homologous protein is defined as a protein of animal origin (e.g. from mouse, rat, dog, rabbit or non-human primates) that recognizes the appropriate target(s) in the respective species with similar potency as the clinical candidate recognizes the corresponding human target(s) (45). Rodent bioassays (or short-term carcinogenicity studies) with homologous products are generally of limited value for assessing the carcinogenic potential of the clinical candidate. Since the production process, range of impurities/contaminants, pharmacokinetics, and exact pharmacological mechanism(s) may differ between the homologous form and the product intended for clinical use, studies with homologous proteins are generally not useful for quantitative risk assessment (see Appendix 5).

#### B.4.5.3 Risk communication

The product-specific assessment of carcinogenic potential is used to communicate risk and provide input to the risk management plan along with labelling proposals, clinical monitoring, post-marketing surveillance, or a combination of these approaches.

### B.4.6 Reproductive performance and developmental toxicity studies

#### B.4.6.1 General principles

The need for reproductive/developmental toxicity studies is dependent upon the product, the clinical indication and the intended patient population. The specific study design and dosing schedule may be modified on the basis of issues related to species specificity, immunogenicity, biological activity and/or a long elimination half-life. For example, concerns regarding potential developmental immunotoxicity, which may apply particularly to certain monoclonal antibodies with prolonged immunological effects, could be addressed in a study design modified to assess immune function of the neonate.

##### B.4.6.1.1 *Products with expected/probable adverse effects on fertility/pregnancy outcome*

When the available data (e.g. mechanism of action, phenotypic data from genetically modified animals, class effects) clearly suggest that there will be an adverse effect on fertility or pregnancy outcome, these data can provide adequate information to communicate risk to reproduction and, under appropriate circumstances, additional nonclinical studies might not be warranted. There may be extensive public information available regarding the potential reproductive and/or developmental effects of a particular class of compounds (e.g. interferons) where the only relevant species is the non-human primate. In such cases, mechanistic studies indicating that similar effects are likely to be caused by a new but related molecule may obviate the need for formal reproductive/developmental

toxicity studies. In each case, the scientific basis for assessing the potential for possible effects on reproduction/development should be provided.

#### B.4.6.1.2 *Products with unclear potential for adverse effects on fertility/pregnancy outcome*

The specific study design and dosing schedule can be modified on the basis of an understanding of species specificity, the nature of the product and its mechanism of action, immunogenicity and/or PK behaviour, and embryo-fetal exposure.

**Species selection** – an assessment of reproductive toxicity of the clinical candidate should usually be conducted only in pharmacologically relevant species. When the clinical candidate is pharmacologically active in rodents and rabbits, both species should be used for embryo-fetal development (EFD) studies, unless embryo-fetal lethality or teratogenicity has been identified in one species. Developmental toxicity studies should be conducted in NHPs only when they are the only relevant species. When the clinical candidate is pharmacologically active only in NHPs, there is still a preference to test the clinical candidate. However, an alternative model can be used in place of NHPs if appropriate scientific justification is provided.

**Alternative evaluation in the absence of a relevant species** – when no relevant animal species exist(s) for testing the clinical candidate, the use of transgenic mice expressing the human target or homologous protein in a species expressing an orthologue of the human target can be considered, assuming that sufficient background knowledge (e.g. historical background data) exists for the model.

#### B.4.6.1.3 *Products for which adverse effects on fertility/pregnancy outcome are not expected*

For products that are directed at a foreign target such as bacteria and viruses, in general no reproductive toxicity studies would be expected.

### B.4.6.2 **Fertility**

For products where mice and rats are pharmacologically relevant species, an assessment of fertility can be made in one of these rodent species (35). Study designs can be adapted for other species provided they are pharmacologically relevant. In such cases the designs should be amended as appropriate – for example, to address the nature of the product and the potential for immunogenicity.

It is recognized that mating studies are not practical for NHPs. However, when the NHP is the only relevant species, the potential for effects on male and female fertility can be assessed by evaluation of the reproductive tract (organ weights and histopathological evaluation) in repeat-dose toxicity studies of at least 3 months' duration, using sexually mature NHPs. If there is a specific cause for concern based on pharmacological activity or previous findings, specialized

assessments such as menstrual cyclicity, sperm count, sperm morphology/motility, and male or female reproductive hormone levels can be evaluated in a repeat-dose toxicity study.

If the pharmacological activity leads to a specific concern about potential effects on conception/implantation and the NHP is the only relevant species, the concern should be addressed experimentally. A homologous product or transgenic model could be the only practical means to assess potential effects on conception or implantation when those are of specific concern. However, it is not recommended to produce a homologous product or transgenic model solely to conduct mating studies in rodents. In the absence of nonclinical information, the risk to patients should be mitigated through clinical trial management procedures, informed consent and appropriate product labelling.

#### B.4.6.3 EFD and pre/postnatal development

##### B.4.6.3.1 *Selection of study design*

Potential differences in placental transfer of rDNA-derived biotherapeutics should be considered in the design and interpretation of developmental toxicity studies (see Appendix 6, Note 1).

For products that are pharmacologically active only in NHPs, several study designs can be considered according to intended clinical use and expected pharmacology. Separate EFD and/or pre/postnatal development (PPND) studies, or other study designs (justified by the sponsor) can be appropriate, particularly when there is some concern that the mechanism of action may lead to an adverse effect on EFD or pregnancy loss. However, one well-designed study in NHPs which includes dosing from day 20 of gestation to birth – “enhanced PPND” (ePPND) – can be considered rather than separate EFD and/or PPND studies.

##### B.4.6.3.2 *ePPND studies*

For the single ePPND study design described above, no caesarean section group is warranted, but assessment of pregnancy outcome at natural delivery should be performed. This study should also evaluate offspring viability, external malformations, skeletal effects (e.g. by X-ray) and, ultimately, visceral morphology at necropsy. Ultrasound is useful for tracking the maintenance of pregnancy but is not appropriate for detecting malformations. These latter data are derived from postpartum observations. Because of potential adverse effects of treatment on maternal care of offspring, dosing of the mother postpartum is generally not recommended. Other end-points in the offspring can also be evaluated if relevant to the pharmacological activity. The duration of the postnatal phase will depend on which additional end-points are considered relevant in view of the mechanism of action (see Appendix 6, Note 2).

Developmental toxicity studies in NHPs can provide only hazard identification. The number of animals per group should be sufficient to allow meaningful interpretation of the data (see Appendix 6, Note 3).

The study design should be justified if species other than the cynomolgus monkey are used. The developmental toxicity studies in NHPs, as outlined above, are hazard identification studies; therefore, it may be possible to conduct these studies using a control group and one dose group, provided there is a scientific justification for the dose level selected (see Appendix 6, Note 4).

#### B.4.6.4 **Timing of studies**

If women of childbearing potential are included in clinical trials prior to acquiring information on the effects on EFD, suitable clinical risk management is appropriate – such as the use of highly effective methods of contraception (38). For rDNA-derived biotherapeutics pharmacologically active only in NHPs, where there are sufficient precautions to prevent pregnancy an EFD or ePPND study can be conducted during phase III and the report submitted at the time of marketing application. When a sponsor cannot take sufficient precautions to prevent pregnancy in clinical trials, either a complete report of an EFD study or an interim report of an ePPND study should be submitted before initiation of phase III (see Appendix 6, Note 5). Where the product is pharmacologically active only in NHPs and its mechanism of action raises serious concern about embryo-fetal development, the label should reflect the concern without warranting a developmental toxicity study in NHPs and the administration to women of childbearing potential should be avoided.

If the rodent or rabbit is a relevant species, timing of reproductive toxicity/fertility studies should follow the recommendations given – see, for example, reference (38).

For oncology products, see Appendix 4.

#### B.4.7 **Local tolerance studies**

Local tolerance should be evaluated. Ideally, the formulation intended for marketing should be tested; however, in certain justified cases, the testing of representative formulations may be acceptable. If feasible, the potential adverse effects of the product can be evaluated in single- or repeated-dose toxicity studies, thus obviating the need for separate local tolerance studies.

#### B.4.8 **Other toxicity studies**

##### B.4.8.1 **Antibody formation**

Immunogenicity assessments in animals should be conducted only to assist in the interpretation of the study results and to improve the design of subsequent studies. Such analyses in animal studies are usually not relevant in terms of

predicting potential immunogenicity of human or humanized proteins in humans. Since antibody formation to human proteins in animal studies is usually not predictive of the clinical situation, concerns regarding antibody formation to the endogenous hormones (as in the case of erythropoietin or somatropin) will have to be addressed on a clinical safety level.

Measurement of anti-drug antibodies in nonclinical studies should be evaluated when there is: (a) evidence of altered PD activity; (b) unexpected change in exposure in the absence of a PD marker; or (c) evidence of immune-mediated reactions (immune complex disease, vasculitis, anaphylaxis, etc.). Since it is difficult to predict prior to study completion whether such analysis will be necessary, it is often useful to obtain appropriate samples during the course of the study so that these can subsequently be analysed when warranted to aid in interpretation of the study results.

When anti-drug antibodies are detected, their impact on the interpretation of the study results should be assessed. Antibody responses should be characterized (e.g. titre, number of responding animals, neutralizing or non-neutralizing activity), and their appearance should be correlated with any pharmacological and/or toxicological changes. Specifically, the effects of antibody formation on PK/PD parameters, incidence and/or severity of adverse effects, complement activation, or the emergence of new toxic effects should be considered when interpreting the data. Attention should also be paid to the evaluation of possible pathological changes related to immune complex formation and deposition.

Characterization of neutralizing potential is warranted when anti-drug antibodies are detected and there is no PD marker to demonstrate sustained activity in the *in vivo* toxicology studies. Neutralizing antibody activity can be assessed indirectly with an *ex vivo* bioactivity assay or an appropriate combination of assay formats for PK-PD, or directly in a specific neutralizing antibody assay.

The detection of antibodies should not be the sole criterion for the early termination of a nonclinical safety study or modification in the duration of the study design, unless the immune response neutralizes the pharmacological and/or toxicological effects of the rDNA-derived biotherapeutics in a large proportion of the animals. In most cases, the immune response to rDNA-derived biotherapeutics is variable, similar to that observed in humans. If the interpretation of the data from the safety study is not compromised by these issues, then no special significance should be ascribed to the antibody response.

#### B.4.8.1.1 *Anaphylaxis tests*

The occurrence of severe anaphylactic responses to rDNA-derived biotherapeutics is uncommon in humans. In this regard, the results of guinea pig anaphylaxis tests, which are generally positive for protein products, are usually not predictive for reactions in humans and are usually not conducted.

#### B.4.8.2 Immunotoxicity studies

One aspect of immunotoxicological evaluation is the assessment of potential immunogenicity (see sections B.4.1 and B.4.8.1). Many rDNA-derived biotherapeutics are intended to stimulate or suppress the immune system and, therefore, may affect humoral as well as cell-mediated immunity. Inflammatory reactions at the injection site may be indicative of a stimulatory response. It is important to recognize, however, that simple injection trauma and/or specific toxic effects caused by the formulation vehicle may result in toxic changes at the injection site. The expression of surface antigens on target cells may be altered, with implications for autoimmune potential. Immunotoxicological testing strategies may require screening studies followed by mechanistic studies to clarify such issues. Routine tiered testing approaches or standard testing batteries, however, are not recommended for rDNA-derived biotherapeutics.

The following modes of action may require special attention (44):

- A mode of action that involves a target which is connected to multiple signalling pathways (a target with pleiotropic effects), e.g. leading to various physiological effects, or targets that are ubiquitously expressed, as often seen in the immune system.
- A biological cascade or cytokine release, including one leading to an amplification of an effect that might not be sufficiently controlled by a physiological feedback mechanism (as in the immune system or blood coagulation system). The so-called cytokine release syndrome (CRS) is characterized by the uncontrolled release of cytokines (such as interleukin-6, tumour necrosis factor or interferon gamma). CD3 or CD28 (super-)agonists may serve as an example. In severe cases, a “cytokine storm” (hypercytokinaemia) with potentially fatal consequences might be induced (46).

Currently available tests for prediction of the potential of an rDNA-derived biotherapeutic with immunomodulatory properties to induce a CRS could, for example, include on a case-by-case basis whole blood assays, peripheral blood mononuclear cell (PBMC)-based assays and biomimetic cell models (47).

#### B.4.8.3 Tissue cross-reactivity studies

Tissue cross-reactivity (TCR) studies are in vitro tissue-binding assays employing immunohistochemical (IHC) techniques that are conducted to characterize the 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 distribution to the target/binding site.

A TCR study with a panel of human tissues is a recommended component of the safety assessment package supporting initial clinical dosing

of these products – see, for example, references (48, 49). However, in some cases the clinical candidate is not a good IHC reagent and a TCR study may not be technically feasible.

TCR studies can provide useful information to supplement knowledge of target distribution and can provide information on potential unexpected binding. Tissue binding does not as such indicate biological activity *in vivo*. In addition, binding to areas not typically accessible to the active substance *in vivo* (i.e. cytoplasm) is generally not therapeutically relevant. Findings should be evaluated and interpreted in the context of the overall pharmacology and safety assessment data package. When there is unexpected binding (i.e. cross-reactivity) to human tissues, a TCR evaluation of selected tissues for the animal species chosen for the nonclinical toxicity studies can provide supplementary information on potential correlations or the lack thereof, with preclinical toxicity. TCR using a full panel of animal tissues is not recommended.

When a bi-specific antibody product is to be evaluated in a TCR study using a panel of human tissues, there is no need to study the individual binding components. Evaluating the tissue binding of homologous products does not provide additional value when TCR studies have been conducted with the clinical candidate in a human tissue panel, and is not recommended. TCR studies are not expected to detect subtle changes in critical quality attributes. Therefore TCR studies are not recommended for assessing the comparability of the test article as a result of process changes over the course of a development programme.

#### B.4.8.4 Impurities

Safety concerns may arise as a result of the presence of impurities or contaminants. There are potential risks associated with host-cell contaminants, whether derived from bacteria, yeast, insect, plant or mammalian cells. The presence of cellular host contaminants can result in allergic reactions and other immunopathological effects. The adverse effects associated with nucleic acid contaminants are theoretical but include potential integration into the host genome (5). For products derived from insect, plant and mammalian cells, or transgenic plants and animals, there may be an additional risk of viral infections. However, it is preferable to rely on manufacturing and quality control processes to deal with these issues (Part A) rather than to establish a preclinical testing programme for their qualification.

## Part C. Clinical evaluation

### C.1 Good clinical practice

All clinical trials should be conducted under the principles described in the WHO Guidelines for good clinical practice (GCP) for trials on pharmaceutical products (50).

## C.2 Clinical pharmacology (Phase I)

### C.2.1 Initial safety and tolerability studies

Initial safety and tolerability studies are the first-in-human studies of drugs after the completion of essential nonclinical studies (35–39, 51). The safety of clinical study participants is the paramount consideration when proceeding to first-in-human studies. Decisions on strategies for the development of a new medicine and the experimental approaches used to assemble information relevant to the safety of first-in-human studies must be science-based and ethically acceptable. Such studies should be closely monitored and are generally conducted with small numbers of subjects who may be healthy volunteers or patients. However, products that are designed to bind a target or receptor present only in patients are normally studied in the intended target population. Study protocols should define stopping rules for individual subjects, for cohorts and for the trial itself. Initial safety and tolerability studies are designed to detect common adverse reactions, the tolerated dose range and the potential drug effect. The ultimate goal of the studies is to obtain adequate safety and pharmacokinetic data to permit the design of sufficiently valid phase II studies.

Initial safety and tolerability studies should preferably be randomized, placebo-controlled studies but may also be single-arm studies with no comparator; they may range from single-dose studies to studies involving multiple doses and lasting for an extended period of time. Drug doses usually start at low levels, and study participants are monitored very carefully as the dose is escalated. In some settings, and depending on the study protocol, individual participants receive only one dose (see sections C.2.3 and C.2.4).

From a clinical perspective, rDNA-derived biotherapeutics present particular challenges compared with chemically derived small molecule drugs, and special safety issues should be addressed in the initial safety and tolerability studies, as follows:

- Currently, the nonclinical data are not completely predictive of safety in humans. In particular, since rDNA-derived biotherapeutics typically contain non-host proteins and polysaccharides, nonclinical studies are usually not predictive for immunogenicity (i.e. a test species may not react to an rDNA-derived biotherapeutic, which could cause serious adverse reactions in humans, or a test species may react when humans do not).
- Data from healthy volunteers may also not be fully predictive of safety/efficacy in patients, especially in the case of monoclonal antibodies which exhibit a target-mediated effect.

Predicting the potential for severe ADRs for first-in-human use of an investigational medicinal product, involves the identification of risk factors,

which may be related to: (a) the mode of action; (b) the nature of the target; and/or (c) the relevance of animal models. High-risk biological substances (e.g. TGN1412, an anti-CD28 superagonist which caused an acute cytokine storm in humans that was not predicted from animal studies) require extended safety measures, which may include strict sequential inclusion of trial participants with clear stopping rules and extremely careful calculation of the first dose in man (51).

The toxicity of most rDNA-derived biotherapeutics is related to their targeted mechanism of action; therefore, relatively high doses can elicit adverse effects which are apparent as exaggerated pharmacology. A rationale should be provided for dose selection, taking into account the characteristics of the dose–response relationship in non-human (in vitro and/or in vivo) PK/PD studies in a relevant animal model. PK-PD approaches (e.g. simple exposure–response relationships or more complex modelling and simulation approaches) can assist in high-dose selection. Where in vivo/ex vivo PD end-points are not available, the high-dose selection can be based on PK data and on available in vitro binding and/or pharmacology data.

### C.2.2 Pharmacogenomics

Pharmacogenomic studies performed early during drug development can provide useful information for the design of robust phase III trials – such as identifying receptor, genetic or phenotypic characteristics and drug response in populations; using biomarkers to identify dose response in individuals; and identifying patients with genetic polymorphisms whose drug dosages should be adjusted for improved safety and/or efficacy or for whom a particular treatment should not be used (52, 53). However, pharmacogenomic effects are not commonly seen with rDNA-derived biotherapeutics. The most recent guidance documents on this topic from appropriate regulatory agencies should be consulted.

### C.2.3 Pharmacokinetics

The PK profile is an essential part of the basic description of a medicinal product and should always be investigated. PK studies should be performed for the intended dose range and routes of administration (4). In general, the PKs (absorption, distribution and elimination) of rDNA-derived biotherapeutics should be characterized during single-dose and steady-state conditions in relevant populations. However, historically, the PK evaluation of peptide or protein products has suffered from limitations in the assay methodology, thus limiting the usefulness of such studies. Immunoassays and bioassays are most frequently used for assaying therapeutic proteins in biological matrices. Special emphasis should, therefore, be given to the analytical method selected and its capability to detect and follow the time course of the protein (the parent

molecule and/or degradation and/or metabolic products) in a complex biological matrix that contains many other proteins. The method should be optimized for satisfactory specificity, sensitivity and a range of quantification with adequate accuracy and precision (4).

The choice of the study population as well as the choice of single-dose and/or multiple-dose studies should be justified (4). If part of the PK information is gathered in healthy volunteers, the validity of extrapolation of that information to the target population needs to be addressed (54). A prospective plan for defining the dosing schedule on the basis of observed/calculated PK parameters should be developed and should be included in the PK study protocol (55). It should be kept in mind that changes in the manufacturing process may alter the quality attributes, thereby potentially altering the PK profiles of rDNA-derived biotherapeutics. In such cases a comparison of the pre-change and post-change products is indicated and it may be necessary to repeat PK studies with the post-change product.

#### C.2.3.1 Absorption

Most biological products are administered parenterally through intravenous, subcutaneous or intramuscular administration. Alternative routes proposed for delivery of proteins may be considered (e.g. nasal and pulmonary administration) which bypass the interstitial subcutaneous or intramuscular environment. Oral delivery of proteins for systemic effects is still rare due to low bioavailability (54).

Unless the intravenous route is used exclusively, appropriate *in vivo* studies should be conducted in healthy volunteers or patients to describe the absorption characteristics of the rDNA-derived biotherapeutics – i.e. the rate and extent of absorption. Single-dose studies are generally sufficient to characterize absorption and to compare different administration routes (56). It should be noted that the rate of absorption following intramuscular or subcutaneous administration may vary according to the site and depth of the injection, and the concentration and volume of the solution injected, and may also be influenced by patient-specific factors (54, 56). These factors which have an influence on the PK/PD parameters should be identified, described and controlled for through established methodologies as far as is possible in order to allow for a better interpretation of the observed outcomes.

Protein therapeutics administered by the subcutaneous route exhibit limited transport into blood capillaries and enter the systemic circulation indirectly through the lymphatics. Passage through the lymphatic system usually results in presystemic elimination, and consequently a bioavailability of less than 100% is obtained. In addition, small proteins may undergo proteolytic degradation in tissues as a first-pass mechanism (54). Since proteases can be affected by disease states and are reported to be upregulated with disease progression, consideration should be given to patient-specific circumstances (56).

### C.2.3.2 Distribution

Tissue distribution studies should be undertaken unless otherwise justified. The volume of distribution of a drug is determined largely by its physicochemical properties (e.g. charge, lipophilicity) and its dependency on active transport processes. Because most rDNA-derived biotherapeutics are large in size, their volume of distribution is usually small and is limited to the volume of the extracellular space due to their limited mobility resulting from impaired passage through biomembranes. Site-specific and target-oriented receptor-mediated tissue uptake and binding to intravascular and extravascular proteins, however, can substantially increase the volume of distribution of rDNA-derived biotherapeutics (57).

The binding capacity to plasma proteins (albumin,  $\alpha$ -acid glycoprotein) should be studied when considered relevant (57).

PK calculations of steady-state volume of distribution may be problematic for some rDNA-derived biotherapeutics. Noncompartmental determination using statistical moment theory assumes first-order disposition processes with elimination occurring from the rapidly equilibrating or central compartment. This basic assumption, however, is not fulfilled for numerous recombinant peptide and protein products, as proteolysis in peripheral tissues may constitute a substantial fraction of the overall elimination process for such rDNA-derived biotherapeutics (57). There is an inverse correlation between the steady-state volume of distribution and molecular weights, and a similar relationship is also seen between permeability and molecular weight. Unlike in the case of small-molecule chemical drugs, distribution to tissues (i.e. cellular uptake) is often part of the elimination process and not part of the distribution process as such, thus contributing to the small distribution volumes. Thus, a small steady-state volume of distribution should not necessarily be interpreted as indicating low tissue penetration, and adequate concentrations may be reached in a single target organ due to receptor-mediated uptake (54).

### C.2.3.3 Elimination

The main elimination pathway, including the major organs of elimination, should be identified. Radiolabelled proteins can be used for this purpose (57). However, for therapeutic proteins, the main elimination pathway *in vivo* can be predicted to a large extent by the molecular size; consequently, specific studies may not be necessary.

Breakdown products may have different PK profiles when compared with the parent rDNA-derived biotherapeutics. However, in cases where measurement of separate active peptide fragments is not technically feasible, the PKs of the active moiety could be determined (54).

Catabolism of small proteins and peptides (molecular weight (MW) < 50 000 Da) appears to occur mainly in the kidneys. The liver may also play a major role in the metabolism of peptides and proteins, mediated by substance-specific enzymes such as for insulin, glucagon, epidermal growth factor, antibodies, and tissue plasminogen activators (57). If biliary excretion of peptides and proteins occurs, it generally results in subsequent breakdown and metabolism of these compounds in the gastrointestinal tract (57).

Catabolism of proteins usually occurs by proteolysis via the same catabolic pathways as for endogenous or dietary proteins. Proteolytic enzymes such as proteases and peptidases are ubiquitously available throughout the body. Thus, locations of intensive peptide and protein metabolism also include blood and various body tissues (57).

If elimination of the protein is largely dependent on target receptor uptake, differences in receptor density between healthy volunteers and target populations, such as over expression of receptors in tumours or inflamed tissues, can create important pharmacokinetic differences in half-life. These differences should be considered when using healthy volunteer data for predictions to the target population (54). After subcutaneous administration of proteins with relatively rapid elimination, the rate of absorption can be slower than the rate of elimination, leading to longer apparent half-lives (flip-flop kinetics) and prolonged exposure when compared to intravenous administration. As a consequence, dosing frequency may have to be reduced (58).

#### C.2.3.4 Subpopulations

The clinical development programme should involve studies to support the approval in subpopulations such as patients with organ dysfunction. Whether such studies are necessary depends on the elimination characteristics of the compound. If no study is conducted, this should be justified by the applicant. An understanding of the influence of intrinsic factors, such as age and body weight, should be provided. Such information might arise from dedicated studies in the respective population or from population PK analyses of phase II/III data (54).

##### C.2.3.4.1 Renal impairment

For proteins with MW lower than 50 000 Da, renal excretion is important for elimination (increasing in importance with lower MW) and consequently for the half-life of the protein. Thus, for these products, PK studies in patients with renal impairment are recommended. It is also conceivable that renal impairment itself may affect functioning of other organs and tissues (e.g. by up- or down-regulation of enzymes or receptors), thereby influencing the PKs and/or PDs of the experimental compound. This should be taken into account in the planning of the clinical pharmacology programme (54).

#### C.2.3.4.2 *Hepatic impairment*

Reduced hepatic function may decrease the elimination of a protein for which hepatic degradation is an important elimination pathway. Where relevant, PK studies in patients with different degrees of hepatic impairment are recommended (54).

#### C.2.3.5 **Interaction studies**

Therapeutic proteins may influence the pharmacokinetics of conventional drugs metabolized by cytochrome P450 enzymes (CYPs) even if the proteins are not metabolized by CYPs (59). Therefore it is important that drug interaction studies are also conducted with therapeutic proteins, unless sufficient evidence is provided from published data or sufficient scientific rationale is provided on the basis of biological plausibility. Additionally, since elimination of proteins may involve capacity-limited steps such as receptor-binding, the inhibition or induction of receptors may have an impact on pharmacokinetics. However, there is currently a lack of knowledge about suitable tools to explore such interactions.

##### C.2.3.5.1 *Dose-dependency and time-dependency*

The dose–concentration relationship may be nonproportional, depending on the relative impact of capacity-limited barriers on distribution and elimination of the product. The dose proportionality should be evaluated in single-dose or multiple-dose studies and the clinical consequences should be discussed. Time-dependent changes in PK parameters may occur during multiple-dose treatment (e.g. due to down- or up-regulation of receptors responsible for (part of) the elimination of the rDNA-derived biotherapeutics or due to formation of anti-drug/product antibodies). Using appropriate methods, soluble receptors may be measured before and during treatment, differentiating between free and bound receptors. The effect on the PKs should be evaluated and the clinical relevance discussed (39).

It is recommended that PKs should be determined at several dose levels on several occasions during long-term studies. Population PK analysis of data from long-term trials could be considered (54).

##### C.2.3.6 **Pharmacokinetic data analysis**

As in the case of small-molecule products, the pharmacokinetics may be analysed using compartment or noncompartment methods. The choice of the PK model used to derive PK parameters should be justified. Mean (or median) and individual results should always be included in a licensure submission. The inter-subject variability should be estimated and, if possible, the important sources of the variability (e.g. demographic factors such as weight and age) should be identified. Potential additional sources of inter-subject variability specific to

therapeutic proteins are the formation of antibodies, absorption variability (e.g. differences in site of injection), variable levels of binding components in blood, variability in target burden (e.g. tumour load), and variability in degradation rate (e.g. of de-PEGylation) or in degradation pattern. Based on the results, individualized dosing should be considered if necessary from safety and/or efficacy perspectives. For products intended for multiple-dose administration, the variability within an individual should also be quantified, since knowledge about the variability between occasions is especially valuable for products for which titration is recommended. Population PK analysis of phase II/III data using a sparse sample approach is recommended for characterizing the PKs, the variability of the PK parameters and possible covariate relationships (54). Population analyses may thus support the individualization of doses.

#### C.2.4 Pharmacodynamics

In many cases, PD parameters are investigated in the context of combined PK/PD studies. Such studies may provide useful information on the relationship between dose/exposure and effect, particularly if performed at different dose levels. PD markers should be selected according to their clinical relevance.

Studies in relevant animal models, if available, provide important information on the PD properties of a biological medicinal product and may guide PD studies in humans. If no animal model is available, a suitable human population must be chosen. In any case, relevant PD effects should always be confirmed in human subjects, either in patients with the disease that is being targeted by the biological medicinal product or in healthy volunteers when the mechanism of action/receptor(s) is the same as in patients. Human PD studies are usually carried out during phase I or phase II studies. Phase II studies can also be called proof-of-concept clinical studies and are important for the subsequent development of the product by helping to determine the dose to be used in further confirmatory trials, and by providing some level of confidence that the biotherapeutic is pharmacologically active and can do what it is intended to do.

#### C.2.5 Pharmacokinetic/pharmacodynamic relationship

The relationship between drug concentration and PD response (PK/PD relationship) should be evaluated as part of drug development. If feasible, markers for both efficacy and safety should be measured, preferably in the same study. It should be noted that PK and PD for a biological medicinal product may not necessarily be entirely and fully correlated (e.g. ceiling effect due to saturation of target receptors) and both may be altered by modifications to the molecule, binding to blood components, or formation of anti-drug/product antibodies. Early preclinical and clinical data can be evaluated using appropriate models for a mechanistic understanding of the disease and the PK/PD relationship. PK/PD models may be developed to account for the time delay between plasma

concentrations and measured effect. Models may also need to take into account the presence or absence of the therapeutic target (e.g. presence of antigen in the case of anticancer monoclonal antibodies). PK/PD models may allow extrapolation from volunteers to the target population if suitable assumptions have been made (e.g. regarding the influence of disease-related factors). These models may provide guidance for dose selection and are helpful when interpreting changes in the PKs in important subpopulations or when evaluating comparability in the context of a change in the manufacturing process. Efforts to explore relevant biomarkers and their link (surrogacy) to safety and efficacy end-points are encouraged (54).

#### C.2.6 **Modifications of pharmacokinetic and pharmacodynamic profiles of therapeutic proteins**

Many protein drugs display suboptimal therapeutic efficacies due to their inherent poor molecular stability, low systemic bioavailability and, as a consequence of their innate susceptibility to various clearance mechanisms, short circulatory lifetimes. Higher protein concentrations and increased dosing frequencies are therefore often employed to achieve favourable therapeutic responses. Approaches to improve these factors, and thus in vivo efficacy, include targeted mutations, the generation of fusion proteins and conjugates, glycosylation engineering, and PEGylation (60).

Glycosylation may influence a variety of physiological processes at both the cellular level (e.g. intracellular targeting) and the protein level (e.g. protein–protein binding, protein molecular stability, plasma persistence lifetimes). Since the glycosylation pattern of a biological medicinal product may be influenced even by subtle changes in the manufacturing process, the potential effects on PK and PD profiles need to be considered when evaluating comparability of pre-change and post-change product in the context of a change in the manufacturing process. PEGylation increases the size of a protein, which prolongs its half-life by reducing renal clearance. PEGylation can also provide water solubility to hydrophobic drugs and proteins.

### C.3 **Efficacy**

#### C.3.1 **Phase II**

Phase II studies provide the first test of efficacy in patients with the disease targeted by the rDNA-derived biotherapeutics. The studies aim to determine the correct dosage, identify common short-term side-effects and determine the best regimen to be used in pivotal clinical trials.

Conventionally, the first step (frequently called phase IIa) is focused on an initial proof of concept. This step aims to demonstrate that the rDNA-derived biotherapeutics interacts correctly with its molecular target and, in turn, alters the disease or its symptoms. Subsequent trials (frequently called phase IIb trials)

are larger and may use placebo, and/or active comparator agents and a broader dosage range to obtain a much more robust proof of concept and additional guidance on dose selection.

For initial proof of concept, single-arm trials may be used with their results interpreted relative to historical control subjects. However, this design could introduce bias since, for example, current study participants may be different from historical control subjects in ways that affect the outcome of interest or because changes in supportive care may limit the validity of the conclusions. Therefore, comparative randomized phase II trials are generally preferred.

Phase II trials usually explore a variety of possible end-points (e.g. time-to-event end-points, change in a continuous end-point of tumour size) and provide opportunities for biomarker discovery. A variety of study designs can be used, including the randomized parallel-group design, randomized discontinuation design, single-stage and two-stage designs, delayed-start design and adaptive (Bayesian) designs. In all cases, clear decision rules should be in place.

Standard study designs for assessing dose–response have been described (61), such as randomized parallel dose–response studies. However, the approaches to selecting the optimal dose may differ for rDNA-derived biotherapeutics compared to small chemical molecules. For example, biological agents developed in oncology are usually cytostatic and their maximal activities may occur at doses lower than their maximum tolerated doses.

Combination therapy is an important treatment modality in many disease settings such as cancer. Increased understanding of the pathophysiological processes that underlie complex diseases has provided further impetus for therapeutic approaches using combinations of (new) products directed at multiple therapeutic targets to improve treatment response, minimize development of resistance or improve tolerability. This requires the use of flexible designs and new modelling approaches for the design of clinical trials.

As observed for small-molecule chemical drugs, rDNA-derived biotherapeutics may affect cardiac electrical activity either directly or indirectly. The amount and type of electrocardiogram data considered appropriate should be individualized according to the type of product and the nonclinical findings regarding its cardiotoxic potential. A thorough QT/QTc (TQT) study (62), or a study that incorporates many of the key components of a TQT study, should be considered (62). However, this may not be necessary if electrocardiogram data are collected in at least a subset of patients during clinical development and reviewed by respective experts, preferably in a blinded manner.

### C.3.2 Confirmatory phase III

Phase III clinical trials are designed to evaluate the benefit of the rDNA-derived biotherapeutics in a carefully selected patient population with the disease.

These trials are carried out to confirm efficacy at the chosen dose(s) and dosing regimen(s), to further evaluate safety and monitor side-effects, and sometimes to compare the candidate product to commonly used treatments. Confirmatory phase III clinical trials should be adequately sized and powered to meet the primary objectives.

Confirmatory trials should be prospective randomized trials comparing the test agent against placebo (in addition to the best supportive care) or an active comparator, usually the best available evidence-based current standard. If no such comparator is available (e.g. in patients who have failed several lines of therapies), the comparator may be the investigator's best choice. Ideally, trials should be double-blinded, where neither the patient nor the investigator knows the nature of the product received by the patient. Blinding or masking is intended to limit the occurrence of conscious or unconscious bias in the conduct and interpretation of a clinical trial (63).

The design of the trials depends on the hypothesis to be tested – superiority to placebo or active comparator, or equivalence or non-inferiority to an active comparator (64).

The choice of end-points depends on the therapeutic indication; there should be sufficient evidence that the primary end-point can provide a valid and reliable measure of clinically relevant and important treatment benefit in the targeted patient population. If a single primary variable cannot be selected, a composite end-point integrating or combining multiple measurements into a single variable, using a predefined algorithm, can also be used. Such validated end points are commonly used in inflammatory diseases (e.g. ACR20 in rheumatoid arthritis, ASAS20 in ankylosing spondylitis, CDAI in Crohn disease, PASI in psoriasis) or in oncology (disease progression, disease-free survival, or overall survival). Patient-reported outcomes and quality-of-life scales are also important end-points and may already be included in some of these composite end-points.

When direct assessment of the clinical benefit to the patient is not practical, a surrogate end-point can be considered. The strength of the evidence for surrogacy depends on: (a) the biological plausibility of the relationship; (b) the demonstration of the prognostic value of the surrogate for the clinical outcome in epidemiological studies; and (c) evidence from clinical trials that treatment effects on the surrogate correspond to effects on the clinical outcome. Most surrogate end-points are not formally validated, but such end-points can be used if they are reasonably likely to predict the desired clinical benefit (e.g. the effect on tumour size, as assessed by imaging, in patients refractory to available treatments). In some cases, particularly for rare diseases, a biomarker could be considered acceptable as the primary study end-point on the basis of biological plausibility and mechanism of action of the product.

Specific decisions about the size of the study group will depend on factors such as the magnitude of the effect of interest, characteristics of the study population, and study design (see section C.4).

Preferably, two confirmatory trials should be performed in order to show that the results can be replicated. However, one controlled study with statistically compelling and clinically relevant results may be sufficient, especially with regard to life-threatening conditions or rare disorders. If the biological medicinal product shows promising efficacy for a serious or life-threatening condition where no other treatment option exists, licensing based on a limited amount of data may be possible with further confirmatory efficacy data being provided post-marketing. Because most rare diseases have a more homogeneous genetic pattern than common diseases and because they are often characterized by similar or identical genetic or epigenetic defects, patients with these diseases could be expected to have a more uniform therapeutic response. This should reduce the size of phase III studies required to demonstrate efficacy. The use of historical controls (or possibly no controls) may also be justified if the rare disease has a defined course in the absence of treatment that will permit comparisons with the results for the investigational rDNA-derived biotherapeutics.

### C.3.3 Biomarkers for patient selection

Biomarkers have the potential to enhance the benefit–risk profile of rDNA-derived biotherapeutics by enabling the selection of patients who are more likely to respond, especially with molecules that target serum or cell markers. In such a case, the treatment may benefit only a subset of patients defined by the biomarker (e.g. those with tumours overexpressing HER-2 or negative for KRAS mutations). The biomarker evaluation process should consist of the following three steps: (a) analytical validation; (b) qualification (i.e. assessment of available evidence on associations between the biomarker and disease states, including data showing effects of interventions on both the biomarker and clinical outcomes); and (c) utilization (i.e. contextual analysis based on the specific use proposed and the applicability of available evidence to this use) (64–66). In principle, biomarker qualification should occur prior to its use as the inclusion and exclusion criteria for patient selections in confirmatory phase III trials. However, these trials can also be used for qualification or identification of other (new) biomarkers.

### C.3.4 Manufacturing and formulation changes

While manufacturing and formulation changes may be expected during product development, the phase III trials should be conducted with the test rDNA-derived biotherapeutics manufactured according to the final manufacturing (commercial) process. If this is not the case, a comparability exercise between the clinical and commercial products is necessary to ensure that the change would

not have an adverse impact on the clinical performance of the product (32, 33). This comparability exercise should normally follow a stepwise approach, starting with a comparison of quality attributes of the active substance and relevant intermediates. A comparability exercise should not be limited to release testing but should include more extensive characterization, using a range of suitable analytical methods as appropriate to the product and process changes in question (see section A.9). If differences are detected that may have an impact on the clinical properties of the product, nonclinical and/or clinical bridging studies, such as PK/PD studies and possibly immunogenicity studies, will generally be needed.

### C.3.5 Special populations

As in any clinical development programme, studies in special populations would be expected where relevant to the indications (e.g. in the elderly and in paediatric patients). The elderly population is arbitrarily defined as those patients aged 65 years or older. However, patients 75 years and above should also be considered to the extent possible (67). Recommended age categories for the paediatric population include preterm and term newborn infants, infants to toddlers, children, and adolescents (68).

Some rDNA-derived biotherapeutics that may be of particular importance to elderly patients are those developed for cancer, Parkinson's disease, Alzheimer's disease, coronary heart disease and diabetes mellitus. It is important to determine whether the PK profile of an rDNA-derived biotherapeutic is different in elderly as compared to younger subjects since impairment of organ function such as renal or hepatic function is more frequent in an aged population. The elderly subpopulation should also be represented sufficiently in the clinical trials to permit the comparison of treatment effects, dose response and safety between older and younger patients. Where the disease to be treated is characteristically associated with ageing, it is expected that elderly patients will constitute the major portion of the clinical database (67).

The extent of the studies needed in children depends on the possibility of extrapolation from adults and children of other age groups. Some rDNA-derived biotherapeutics may be used in children from the early stages of drug development, especially those targeting genetic diseases where manifestations occur early in life. Evaluation should be made in the appropriate age group and it is usually recommended to begin with older children before extending the trial to younger children and then infants (68). Where justified, extrapolation of efficacy data from adult to paediatric patients may be based on PK and/or PD data (e.g. when a similar effect can be expected with similar exposure). However, safety data usually cannot be extrapolated and need to be generated in children (see section C.5).

### C.3.6 **Post-marketing: Phase IV**

Phase IV trials may be required to evaluate further an approved rDNA-derived biotherapeutic and to obtain more information about safety or effectiveness, or both, especially if the biotherapeutic has been approved on the basis of a surrogate end-point.

## C.4 **Statistical considerations**

### C.4.1 **General considerations**

The application of sound statistical principles to the design, conduct, analysis and interpretation of clinical trials should be considered an important and integral component of the overall development of an rDNA-derived biotherapeutic. The success of a trial depends on the appropriateness of the study design, the conduct of the trial and the analysis of trial results. Statistical principles are relevant to all three aspects of the clinical trial. In general, details on these aspects should be specified in the trial protocol which should be written and finalized prior to the start of the trial. Any subsequent amendments to the protocol should be clearly justified, should be documented in a formal amendment to the protocol, and should include the statistical consequences of the proposed changes.

The scientific integrity of the trial and the credibility of the data from the trial depend substantially on the trial design (69). The study protocol should include a clear description of the specific design selected for a particular trial. Additional details regarding the primary end-point, which is directly related to the primary objective of the trial, should also be included. If multiple primary end-points are defined, the criteria for achieving study success should be clearly laid out in order to avoid potential problems with the interpretation of the trial results. The protocol should also clearly define secondary end-points, and their role in the interpretation of the trial results should be stated. Details on measures that have been put in place to avoid or minimize bias in the trial (e.g. randomization and blinding) should also be provided.

With regard to the type of hypothesis to be tested in a specific trial, it should be clear in the protocol whether the trial is designed to show superiority, non-inferiority, or equivalence. The statistical issues involved in the design, conduct, analysis and interpretation of equivalence and non-inferiority trials are complex and subtle, and they require that all aspects of these trials are carefully evaluated. Sample size and power are important for the success of a clinical trial and should be given careful consideration at the trial design stage. In determining sample size, the specific hypothesis being tested should be taken into consideration.

It is important to ensure that the protocol will provide good quality data that permit an adequate evaluation of the efficacy (and safety) of the product

under development. In addition, if formal interim analyses are planned, then the details governing such analyses should be pre-specified in the protocol.

In an era when it is recognized that improvements in the drug development process are needed in order to increase the likelihood of trial success, decrease costs and increase the efficiency with which efficacious and safe medicines are brought to market, adaptive clinical trial designs are increasingly considered as one tool through which these improvements can be achieved. Adaptive design refers to a clinical study design that uses the accumulation of data as a basis for modifying aspects of the study as it continues, without undermining the validity and integrity of the trial (70, 71). A key statistical issue for adaptive designs is the preservation of the Type I error rate. The methods used to control the Type I error rate properly should be described in the study protocol, with additional details provided in the Statistical Analysis Plan (SAP).

Details regarding the statistical methodology to be applied to the clinical trial should be provided in the protocol, with the more technical details being captured in the SAP. The SAP should be prepared and finalized prior to unblinding the clinical study. Any amendments to the SAP must also be finalized prior to unblinding.

#### C.4.2 **Specific considerations for rDNA-derived biotherapeutics**

Since rDNA-derived biotherapeutics are often indicated to treat severe and/or life-threatening diseases and chronic diseases, trials for rDNA-derived biotherapeutics present unique statistical challenges.

##### C.4.2.1 **Trials in small populations and single-arm studies**

Some rDNA-derived biotherapeutics are intended for the treatment of rare diseases for which the target population is very small. Consequently, trials that are considered confirmatory for rare disease indications are often based on a limited number of subjects. While such studies must still be designed with the rigour of traditional trials, and should be conducted with high quality in order to provide reliable and valid data for assessing efficacy and safety, some flexibility is needed with regard to the statistical methods that will be utilized in these trials. Single-arm studies with comparisons made to an external control can sometimes be justified.

##### C.4.2.2 **Tumour-based end-points in oncology trials and composite end-points**

In confirmatory oncology trials for rDNA-derived biotherapeutics, the use of tumour-based end-points such as disease-free survival and progression-free survival as the primary end-point is not uncommon (72). The use of a tumour-based end-point as the primary end-point creates several statistical challenges, and considerations for the collection and analysis of such end-points have been

discussed – see, for example, reference (73). Clinical trials may involve the use of a composite primary end-point arising from the combination of multiple clinical measurements or outcomes (e.g. major adverse cardiac events (MACE), which is the most commonly used composite end-point in cardiovascular studies). For such a composite end-point, it is important that the individual components are analysed separately (usually as secondary end-points) in order to ensure that the treatment effect is shown across all components and is of similar magnitude.

#### C.4.2.3 Missing data

Missing data is a common problem in long-term trials of rDNA-derived biotherapeutics targeting chronic diseases such as diabetes and rheumatoid arthritis, although it is usually not a problem in short-term trials. The impact of missing data on the validity of trial results should be carefully assessed using sensitivity analyses with appropriate underlying assumptions.

### C.5 Safety

Pre-licensing safety data should be obtained in a sufficient number of patients in order to characterize and quantify the safety profile – including type, frequency and severity of ADRs – of the rDNA-derived biotherapeutics. The safety evaluation should cover a reasonable duration of time, taking into account the intended duration of use of the drug, so as to assess potential changes in the ADR profile over time and to capture delayed ADRs.

For drugs intended for long-term treatment of non-life-threatening conditions, a 12-month exposure of at least 100 patients to the investigational medicinal product at the intended clinical dosage should be considered (74). When no serious ADR is observed in a 1-year exposure period, this number of patients can provide reasonable assurance that the true cumulative 1-year incidence is no greater than 3%. This estimate is based on the statistical “rule of three” which states that if no major ADR occurred in a group of  $n$  people, there can be 95% confidence that the chance of a major ADR is less than one in  $n/3$  (or equivalently, less than 3 in  $n$ ). This estimate is considered a good approximation for  $n > 30$ .

The safety database may need to be larger or may require longer patient observation if a safety signal is identified, if the drug is expected to cause late-developing ADRs, or if ADRs increase in severity or frequency over time. Concerns requiring a larger safety database may arise from nonclinical or early clinical data, or from experience with other products of the same or related pharmacological class. A smaller safety database may be acceptable if the intended treatment population is small. Safety data should be obtained from prospective, and preferably controlled, studies including a placebo or active comparator arm since comparison with an external control group (e.g. with

published data) is usually hampered by differences in the investigated patient population, concomitant therapy, observation period and/or reporting. Causality assessment – i.e. whether the observed adverse event is causally related to the investigational drug – is usually easiest in placebo-controlled studies. Generally accepted definitions and terminology, as well as procedures, are important for harmonizing the way to gather and, if necessary, to take action on important clinical safety information arising during clinical development (75). The term “adverse event” describes any untoward medical occurrence developing with administration of a pharmaceutical product irrespective of a causal relationship. The term “adverse drug reaction”, on the other hand, should be used only for adverse events that have at least a reasonably possible causal relationship to the pharmaceutical agent.

Standardized reporting is important for the transmission of pre- or post-marketing safety information – for example, between the reporting source or pharmaceutical industry and regulatory authorities, or between regulatory authorities and the WHO Collaborating Centre for International Drug Monitoring (76). Data elements to be included in individual case safety reports should comprise all important information on the primary source, date, sender and receiver of the information, the type, seriousness, duration and outcome of the adverse event or ADR, detailed patient characteristics and drug information, actions taken with the drug (e.g. dose reduction, discontinuation), and an assessment of the degree of suspected relatedness of the drug to the adverse event (76).

To facilitate international sharing of regulatory safety information for medical products used by humans, specific MedDRA terminology has been developed. This is a rich and highly specific standardized medical terminology for accurate and consistent safety information that allows for the aggregation of reported terms in medically meaningful groupings (77). Products covered by the scope of MedDRA include pharmaceuticals, vaccines and drug device combination products.

Since safety data obtained from pre-marketing clinical trials can be expected to detect mainly common and shorter-term ADRs, further monitoring of clinical safety of the biological product to detect rare but sometimes serious adverse effects and an ongoing benefit–risk evaluation are necessary in the post-marketing phase (see section C.7).

### C.5.1 Special populations

#### C.5.1.1 Elderly population

The safety of rDNA-derived biotherapeutics should be investigated in elderly patients during clinical drug development (67, 78), except where there is no intention to use these biotherapeutics in this age group. Elderly patients are

more prone to adverse effects since they often have comorbidities and are taking concomitant medication that could interact with the investigational drug. The adverse effects can be more severe, or less tolerated, and may have more serious consequences than in the non-elderly population. Depending on the mechanism of action of the drug and/or the characteristics of the disease, specific effects on cognitive function, balance and falls, urinary incontinence or retention, weight loss and sarcopenia should be investigated.

Elderly patients may be included in the main phase III or phase II/III studies, or in separate studies. Inclusion of younger and elderly patients in the same studies has the advantage of allowing direct comparisons using data collected in similar ways. Certain assessments, however, such as studies of cognitive function, require special planning and can be best accomplished in separate studies.

Where enrolment of elderly patients has been insufficient despite the efforts of the applicant, a specific plan to collect post-marketing data should be presented in the marketing application.

#### C.5.1.2 Paediatric population

Data on the safety of medicinal products in the pediatric population should be generated unless their use is clearly inappropriate (68). During clinical development, the timing of paediatric studies will depend on the medicinal product, the type of disease being treated, safety considerations, and the efficacy and safety of alternative treatments. Justification for the timing and approach to the clinical programme needs to be clearly addressed with the regulatory authorities.

Medicinal products may affect physical and cognitive growth and development, and the adverse event profile may differ in paediatric compared to adult patients. In addition, adverse effects may not be seen immediately but may become apparent only at a later stage of development. Long-term studies, or surveillance data while patients are on chronic therapy and/or during the post-therapy period, may be needed to determine possible effects on skeletal, behavioural, cognitive, sexual and immune maturation and development.

### C.6 Immunogenicity

rDNA-derived biotherapeutics may induce unwanted humoral and/or cellular immune responses in recipients. Immunogenicity of rDNA-derived biotherapeutics should therefore always be investigated prior to authorization (79). Since animal data are usually not predictive of the immune response in humans, immunogenicity needs to be investigated in the target population. Although in-silico modelling may help in identifying T-cell epitopes related to immunogenicity (i.e. T-helper epitopes), it does not predict whether immunogenicity will occur. The frequency and type of product antibodies induced against the active substance, impurity

or excipient, as well as possible clinical consequences of the immune response, should be thoroughly assessed.

The immune response against a biotherapeutic is influenced by many factors – such as the nature of the drug substance, product- and process-related impurities (e.g. host-cell proteins, aggregates), excipients and stability of the product, the route of administration (subcutaneous administration is usually more immunogenic than intravenous administration), the dosing regimen (intermittent use is usually more immunogenic than continuous use), and patient-related, disease-related and/or therapy-related factors (e.g. antibody development is more likely in an immune-competent than in an immunosuppressed state and is potentially enhanced in the presence of autoimmune disease). The consequences of unwanted immunogenicity on safety may vary considerably, ranging from clinically irrelevant to serious and life-threatening (e.g. serious infusion/anaphylactic) reactions. Neutralizing antibodies may directly alter the PD effect of a product (i.e. by blocking the active site of the protein), leading to reduction or loss of efficacy. Binding antibodies often affect pharmacokinetics and may indirectly influence pharmacodynamics. Thus, an altered effect of the product over time due to anti-drug antibody formation might be a composite of pharmacokinetic, PD and safety effects.

The proposed antibody testing strategy – including the selection, assessment, and characterization of assays, the identification of appropriate sampling time points (including baseline samples), sample storage and processing, and selection of statistical methods for analysis of data – should be appropriately justified (79). The studies to be considered for immunogenicity testing (e.g. short-term and/or long-term clinical trials or even single-dose studies) and the sampling time points depend on the expected appearance of antibodies and the clinical consequences of such antibodies. For example, some rDNA-derived biotherapeutics are highly immunogenic and may elicit an immune response after the first dose, others may require prolonged or intermittent exposure to mount an immune response, and some may have a very low immunogenic potential. Anti-product antibody screening and subsequent characterization for confirmation, titre, neutralizing activity, isotype, subclass, etc. should be determined early as the sponsor performs an immunogenicity risk assessment, mitigation and management strategy. The assessment should consider the immunogenic factors listed above and the potential clinical consequences if antibodies develop.

Antibody assays (screening, confirmation, and neutralizing) should be validated for their intended purpose. Validation studies need to establish appropriately linear responses to relevant analytes as well as appropriate accuracy, precision, sensitivity, specificity and robustness of the assay(s) (79–81). Possible interference of the circulating antigen with the antibody assay(s) should be taken

into account. A highly sensitive screening assay should be used for antibody detection and a confirmatory assay should be used to confirm the presence of antibodies and eliminate false-positive results. To achieve confirmation of specificity, it is necessary to include an assay which evaluates specificity. A neutralization assay should be available for further characterization of antibodies. The determination of the phase of clinical testing at which the need for characterization (e.g. neutralizing, isotype, subclass, etc.) of detected anti-drug antibodies is warranted, is commensurate with the potential safety risk to patients, and may be based on knowledge and experience with the substance class.

If the rDNA-derived biotherapeutic is a monoclonal antibody (mAb), the development of assays to detect antibodies against this mAb can be technically challenging (80). Many standard assay formats involve the use of anti-immunoglobulin reagents such as antibodies against immunoglobulins, protein A or protein G, but these are inappropriate for use in detecting antibodies against mAbs as they often bind to the product itself. Different assay approaches have been developed to overcome this problem, such as the bridging enzyme-linked immunosorbent assay (ELISA) format or surface plasmon resonance (SPR) procedures which do not require anti-immunoglobulin reagents but may be less sensitive than other immunoassay methods (2).

Detected antibodies should be further characterized with regard to antibody content (concentration/titre) and possibly, depending on case-by-case considerations, other criteria such as antibody class and subclass (isotype), affinity and specificity. For example, the isotype of the antibodies could be determined if this may be predictive of safety (such as the development of IgE antibodies causing allergic and anaphylactic responses). Potential clinical implications of detected antibodies regarding safety, efficacy and pharmacokinetics should always be evaluated. Special attention should be paid to the possibility that the immune response seriously affects the endogenous protein and its unique biological function (e.g. neutralizing anti-erythropoietin antibodies cross-reacting with endogenous erythropoietin and causing pure red cell aplasia).

The required observation/monitoring period for immunogenicity testing will depend on the intended duration of therapy and the expected time of antibody development, if known, and should be justified. In the case of chronic administration, 1-year data will usually be appropriate prior to licensing to assess antibody incidence and possible clinical implications. If considered clinically relevant, development of antibody titres, their persistence over time, potential changes in the character of the antibody response and the possible clinical implications should be assessed pre- and post-marketing.

Since pre-licensing immunogenicity data are often limited, further characterization of the immunogenicity profile may be necessary post-marketing,

particularly if rare but clinically meaningful, or even serious, antibody-related ADRs have been encountered with biological agents of the same or related substance class that are not likely to be detected in the pre-marketing phase.

### C.7 Pharmacovigilance and risk management planning

NRAs should be vigilant to ensure that the health of the public is protected. The aim is to ensure that the risks associated with rDNA-derived biotherapeutics are actively minimized. Patient safety is a key concern for all medicinal products that are on the market, and rDNA-derived biotherapeutics are no exception. Due to the specific characteristics of rDNA-derived biotherapeutics already discussed in these Guidelines, pharmacovigilance activities required for rDNA-derived biotherapeutics may differ in some respects from those required for small-molecule drugs. For example, biotherapeutic use may lead to antibody formation with consequences for clinical efficacy and/or safety.

A risk management plan should be submitted and agreed to by the NRA. The key components of a risk management plan may include:

- safety specifications, which summarize the known and potential safety issues and missing information about the rDNA-derived biotherapeutic;
- a pharmacovigilance plan to further evaluate important known or potential safety concerns and to provide post-marketing data where relevant information is missing;
- a risk minimization plan, which provides proposals on how to minimize any identified or potential safety risk.

In the risk management plan, the known or potential risks may be described with pharmacovigilance, and risk minimization activities may be proposed to identify, characterize, prevent or minimize risks related to the use of the rDNA-derived biotherapeutics, to assess the effectiveness of those interventions, and to communicate those risks to both patients and health-care providers.

Pharmacovigilance and risk minimization activities that might be included in a risk management plan usually fall into two categories: (a) routine activities, which would generally be conducted for any medicine where no special safety concerns have arisen; and (b) additional activities designed to address identified and potential safety concerns that could have an impact on the benefit–risk balance of a product. Routine pharmacovigilance activities would include the monitoring and reporting of spontaneous adverse events post-approval and any safety evaluations incorporated in clinical trials that may be initiated by the marketing authorization holder following marketing authorization for a wide variety of reasons. In case there are relevant safety issues,

NRAs may request additional pharmacovigilance activities in the form of active surveillance (e.g. registries), epidemiology studies, further clinical studies, and drug utilization studies. Routine risk minimization activities would ensure that suitable contraindications and warnings are included in the product information and that this information is updated on an ongoing basis. A risk minimization plan can further specify other risk minimization activities, as appropriate, which could include: (a) specific educational material about the product and its use; (b) patient-oriented or physician-oriented training programmes; (c) restricted use of the product; and (d) registration programmes for patients, physicians and/or pharmacists.

Once on the market, manufacturers should monitor the effectiveness of their risk minimization plans and revise them if new safety and effectiveness concerns are identified. Changes in the manufacturing processes introduced post-marketing could also influence the safety profile (e.g. by enhancing immunogenicity) of rDNA-derived biotherapeutics and may necessitate enhanced safety monitoring.

In case a relevant or even serious potentially drug-related adverse event occurs, it is important to be able to identify the specific biological causing this event. Therefore, all ADR reports should carry information unique to the product, including the proprietary (brand) name, the INN, the identification code (if there is one), and the lot information of the respective biological to help trace an ADR to a specific product and ascertain any relation to causality.

A risk management plan will not reduce the scientific and clinical standards or the data requirements for the market authorization of rDNA-derived biotherapeutics, nor will it replace the precautionary approach that is taken to managing the risks associated with those products. On the contrary, implementation of a risk management plan will further strengthen the rigour of post-marketing surveillance, allowing for earlier identification of risks associated with rDNA-derived biotherapeutics and earlier interventions to minimize those risks.

## C.8 Additional guidance

Further guidance on various aspects of clinical trials is available from several other bodies such as the ICH, the EMA and the United States Food and Drug Administration, as well as from several other NRAs. These WHO Guidelines are not intended to conflict with, but rather to complement, these other documents with respect to medicinal products prepared by rDNA technology. Relevant sections of this part may be useful with regard to products intended for clinical trials; however, the amount and extent of data submitted for a product will be limited and should take into account the nature of the product and its stage of development.

## Authors and acknowledgements

The preliminary draft of these Guidelines was prepared by the following WHO Drafting Group members after a meeting held at the National Institute for Biological Standards and Control, England on 19–20 March 2012: Dr M-C. Bielsky, Medicines and Healthcare Products Regulatory Agency, England; Dr E. Griffiths, Consultant, England; Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr H-N. Kang, World Health Organization, Switzerland; Dr R. Thorpe, National Institute for Biological Standards and Control, England; Dr M. Wadhwa, National Institute for Biological Standards and Control, England; Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany.

The first draft was prepared by the following authors for the parts indicated: Dr E. Griffiths, Consultant, England (Introduction); Dr K. Ho, Agence nationale de sécurité du médicament et des produits de santé, France, Dr R. Thorpe, National Institute for Biological Standards and Control, England and Dr M. Wadhwa, National Institute for Biological Standards and Control, England (Part A); Dr L. Gomes Castanheira, National Health Surveillance Agency, Brazil and Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany (Part B); Dr M-C. Bielsky, Medicines and Healthcare Products Regulatory Agency, England, Dr A. Klein, Health Canada, Canada, Dr C. Njue, Health Canada, Canada, Dr J. Wang, Health Canada, Canada and Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany (Part C); with support from Dr H-N. Kang and Dr I. Knezevic, World Health Organization, Switzerland, and taking into consideration discussions held at the WHO Informal Consultation on the Revision of the Guidelines on the quality, safety and efficacy of biological medicinal products prepared by recombinant DNA technology held in Xiamen, China on 31 May–1 June 2012 attended by: Mrs A. Abas, Ministry of Health Malaysia, Malaysia; Dr W.S. Alhaqaish, Jordan Food and Drug Administration, Jordan; Ms J. Archer (*International Generic Pharmaceutical Alliance representative*), Hospira, Australia; Dr B. Boonyapiwat, Ministry of Public Health, Thailand; Dr L. Gomes Castanheira, National Health Surveillance Agency, Brazil; Dr W. Chang, State Food and Drug Administration, China; Dr R. Chakrabarti (*United States Pharmacopoeial Convention representative*), United States Pharmacopoeia–India, India; Mr D. Cheng, Beijing Four Rings Bio-Pharmaceutical Co., Ltd., China; Dr L. Chenggang, National Institutes for Food and Drug Control, China; Dr Y. Choi, Korea Food and Drug Administration, Republic of Korea; Ms J. Dahlan, National Agency of Drug and Food Control, Indonesia; Mr G. Eich (*IFPMA representative*), Amgen Inc., USA; Dr K. Gao, National Institutes for Food and Drug Control, China; Mr T. Go, Health Sciences Authority, Singapore; Dr E. Griffiths, England; Dr L. Gu, Shenyang Sunshine Pharmaceutical Co. Ltd., China; Dr Z. Guo (*Chinese Pharmacopoeia representative*), Chinese Pharmacopoeia

Commission, China; Dr N. Hassannia, Biological Office Food and Drug Organization, Islamic Republic of Iran; Dr K. Ho, Agence nationale de sécurité du médicament et des produits de santé, France; Dr S. Hufton, National Institute for Biological Standards and Control, England; Mrs W. Jariyapan, Ministry of Public Health, Thailand; Mr R. Jian, Health Sciences Authority, Singapore; Dr J. Joung, Korea Food and Drug Administration, Republic of Korea; Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr H-N. Kang, World Health Organization, Switzerland; Dr Y. Kishioka (*Japanese Pharmacopoeia representative*), Pharmaceutical and Medical Devices Agency, Japan; Dr I. Knezevic, World Health Organization, Switzerland; Mr J. Leong, Health Sciences Authority, Singapore; Dr J. Li, Shanghai CP-Guojian Pharmaceutical Co. Ltd., China; Dr J. Luo, State Food and Drug Administration, China; Mrs V. Madrigal, Recepta Biopharma, Brazil; Dr C. Njue, Health Canada, Canada; Mrs Y. Hechavarria Nunez, Centro para el Control Estatal de la Calidad de los Medicamentos, Cuba; Dr P.H. Pan (*DCVMN representative*), Innovax Biotech Co. Ltd., China; Dr S. Pluschkell (*IFPMA representative*), Pfizer Inc., USA; Professor C. Rao, National Institutes for Food and Drug Control, China; Dr M. Schiestl (*International Generic Pharmaceutical Alliance representative*), Sandoz GmbH, Austria; Dr T. Schreitmueller (*IFPMA representative*), F. Hoffmann-La Roche Ltd., Switzerland; Dr S. Shani, Ministry of Health and Social Welfare, India; Dr X. Shen, China Bio-Tech Group, China; Dr L. Sun, Xiamen Amoytop Biotech Co. Ltd., China; Dr R. Thorpe, National Institute for Biological Standards and Control, England; Mrs C. Ulm (*European Generic Medicines Association representative*), Mylan GmbH, Switzerland; Dr A. Vallin, Centre of Molecular Immunology, Cuba; Dr J. Wang, Health Canada, Canada; Dr J. Wang, National Institutes for Food and Drug Control, China; Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany; Dr M. Xu, National Institutes for Food and Drug Control, China.

The second draft of the Guidelines was prepared by the following authors for the parts indicated: Dr E. Griffiths, Consultant, England (Introduction); Dr E. Griffiths, Consultant, England, Dr K. Ho, Agence nationale de sécurité du médicament et des produits de santé, France, Dr J. Joung, Korea Food and Drug Administration, Republic of Korea, Dr R. Thorpe, National Institute for Biological Standards and Control, England, Dr M. Wadhwa, National Institute for Biological Standards and Control, England and Dr J. Wang, National Institutes for Food and Drug Control, China (Part A); Dr L. Gomes Castanheira, National Health Surveillance Agency, Brazil and Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany (Part B); Dr A. Klein, Health Canada, Canada, Dr F. Lentz, Federal Institute for Drugs and Medical Devices, Germany, Dr C. Njue, Health Canada, Canada, Dr J. Wang, Health Canada, Canada and Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany (Part C); with support from Dr H-N. Kang, Dr J-W. Kim and Dr I. Knezevic, World Health

Organization, Switzerland, and taking into account comments received from: Dr J. Archer, Hospira, Australia; Dr J. Bernat, IFPMA, Switzerland; Dr B. Brake, Federal Institute for Drugs and Medical Devices, Germany; Dr T. Go, Health Sciences Authority, Singapore; Mrs W. Jariyapan, Ministry of Public Health, Thailand; Mrs Y. Hechavarria Nunez, Centro para el Control Estatal de la Calidad de los Medicamentos, Cuba; Dr M. Schiestl, Sandoz, Austria; Dr Y. Sohn, Korea Food and Drug Administration, Republic of Korea; Dr G.R. Soni, Ministry of Health and Family Welfare, India; Dr T. Yamaguchi, National Institute of Health Sciences, Japan.

The draft Guidelines were then posted on the WHO website for the first round of public consultation from 20 March to 19 April 2013.

The document WHO/BS/2013.2213 was prepared by Dr M-C. Bielsky, Medicines and Healthcare Products Regulatory Agency, England; Dr E. Griffiths, Consultant, England; Dr H-K. Heim, Federal Institute for Drugs and Medical Devices, Germany; Dr K. Ho, Agence nationale de sécurité du médicament et des produits de santé, France; Dr J. Joung, Ministry of Food and Drug Safety (formerly Korea Food and Drug Administration), Republic of Korea; Dr H-N. Kang, World Health Organization, Switzerland; Dr A. Klein, Health Canada, Canada; Dr I. Knezevic, World Health Organization, Switzerland; Dr F. Lentz, Federal Institute for Drugs and Medical Devices, Germany; Dr R. Thorpe, National Institute for Biological Standards and Control, England; Dr M. Wadhwa, National Institute for Biological Standards and Control, England; Dr M. Weise, Federal Institute for Drugs and Medical Devices, Germany, taking into account comments received from the following reviewers: Ms P. Agsiri, Ministry of Public Health, Thailand; Mrs J. Bernat on behalf of the IFPMA, Switzerland (reviewed by Dr C. Phillips, Eli Lilly and Company, and Dr S. Ramanan, Amgen); Dr B. Brake, Federal Institute for Drugs and Medical Devices, Germany; Mrs J. Dahlan, National Agency of Drug and Food Control, Indonesia; Ms C. Dubeaux on behalf of GlaxoSmithKline, Belgium (reviewed by Dr C. Lecomte, Dr F. Mortiaux and Dr C. Saille); Dr T. Go, Health Sciences Authority, Singapore; Dr S. Jadhav (*DCVMN representative*), Serum Institute of India, India; Dr Y. Kishioka, Pharmaceutical and Medical Devices Agency, Japan; Dr B. Lan (*DCVMN representative*), China National Biotec Group, China; Dr T. Morris, United States Pharmacopoeial Convention, USA; Dr C. Njue, Health Canada, Canada; Dr G. Raychaudhuri on behalf of the United States Food and Drug Administration Center for Biologics Evaluation and Research (reviewed by Dr I. Mahmood, Dr C. Kimchi-Sarfaty and Dr M. Serabian); Dr Y. Ren, Chinese Pharmacopoeia, China; Dr I. Shin on behalf of the Ministry of Food and Drug Safety, Republic of Korea (reviewed by Mr D. Baek, Dr J. Joung, Mrs Y. Kim and Mr O. Kwon); Dr S. Sontakke, Health Canada, Canada; Dr J. Southern (*Developing Country Vaccine Regulators' Network representative*), South Africa; Dr R. Thorpe, National Institute for Biological Standards and Control,

England; Dr J. Wang, Health Canada, Canada; Dr J. Wang, National Institutes for Food and Drug Control, China; Dr A. Womack, Biotechnology Industry Organization, USA.

The document WHO/BS/2013.2213 was then posted on the WHO website for the second round of public consultation from 1 July to 20 September 2013, and comments were received from the following reviewers: Mrs J. Bernat and Dr A. Womack on behalf of the IFPMA, Switzerland; Mrs M. Bevzyuk on behalf of Apotex Inc., Canada; Dr G. Castellero, Dirección Nacional de Farmacia y Drogas, Panama; Dr C. Dubeaux on behalf of GlaxoSmithKline, Belgium (reviewed by Dr C. Lecomte, Dr F. Mortiaux, Dr C. Saille and Dr M. Duchêne); Dr P. Duffy, National Institutes of Health, USA; Dr K. Heidenreich, Novartis, Switzerland; Dr Y. Huang, China Food and Drug Administration, China; Dr S. Jadhav, Serum Institute of India, India; Dr Y. Kishioka, Pharmaceutical and Medical Devices Agency, Japan; Mrs D. Kusmiaty, National Agency of Drug and Food Control, Indonesia; Mrs D. Lucía Mesa Lautero on behalf of the Instituto Nacional de Vigilancia de Medicamentos y Alimentos, Colombia (reviewed by Mr A. Alarcón, Mr D. Duarte, Mrs A. Fula, Mrs J. García, Mrs F. Gil, Mrs L. Hidalgo and Mrs C. Niño); Dr R. Mody, Lupin Ltd., India; Dr T. Morris, United States Pharmacopeial Convention, USA; Dr J. Petricciani, Consultant, USA; Dr G. Raychaudhuri on behalf of the United States Food and Drug Administration Center for Biologics Evaluation and Research (reviewed by Dr I. Mahmood and Dr C. Kimchi-Sarfaty); Dr J. Schletter on behalf of the European Biosimilars Group (a sector group of the European Generic Medicines Association), Belgium; Dr G. Unger, Paul-Ehrlich-Institut, Germany; Dr K. Zoon, National Institutes of Health, USA.

Acknowledgement is also given to Dr M. Schiestl, Sandoz GmbH, Austria and to Dr T. Schreitmüller, F. Hoffmann-La Roche Ltd., Switzerland for providing expertise on the section on drug product container closure system and delivery devices.

## References

1. Guidelines for assuring the quality of pharmaceutical and biological products prepared by recombinant DNA technology. In: *WHO Expert Committee on Biological Standardization. Forty-first report*. Geneva, World Health Organization, 1991 (WHO Technical Report Series, No. 814), Annex 3.
2. WHO guidelines on nonclinical evaluation of vaccines. In: *WHO Expert Committee on Biological Standardization. Fifty-fourth report*. Geneva, World Health Organization, 2005 (WHO Technical Report Series, No. 927), Annex 1.
3. Guidelines on clinical evaluation of vaccines: regulatory expectations. In: *WHO Expert Committee on Biological Standardization. Fifty-second report*. Geneva, World Health Organization, 2004 (WHO Technical Report Series, No. 924), Annex 1.

4. Guidelines on evaluation of similar biotherapeutic products (SBPs). In: *WHO Expert Committee on Biological Standardization. Sixtieth report*. Geneva, World Health Organization, 2013 (WHO Technical Report Series, No. 977), Annex 2.
5. Recommendations for the evaluation of animal-cell cultures as substrates for the manufacture of biological medicinal products and for the characterization of cell banks. In: *WHO Expert Committee on Biological Standardization. Sixty-first report*. Geneva, World Health Organization, 2013 (WHO Technical Report Series, No. 978), Annex 3.
6. Ad hoc Working Party on Biotechnology/Pharmacy. Guidelines on the production and quality control of medicinal products derived by recombinant DNA technology. *Trends in Biotechnology*, 1987, 5:G1–G4.
7. *Points to consider in the production and testing of new drugs and biologicals produced by recombinant DNA technology*. Bethesda, MD, Food and Drug Administration, 1985.
8. Requirements for human interferons made by recombinant DNA techniques. In: *WHO Expert Committee on Biological Standardization. Thirty-eighth report*. Geneva, World Health Organization, 1987 (WHO Technical Report Series, No. 771), Annex 7.
9. Requirements for human interferons prepared from lymphoblastoid cells. In: *WHO Expert Committee on Biological Standardization. Thirty-ninth report*. Geneva, World Health Organization, 1988 (WHO Technical Report Series, No. 786), Annex 3.
10. Guidelines for assuring the quality of monoclonal antibodies for use in humans. In: *WHO Expert Committee on Biological Standardization. Forty-second report*. Geneva, World Health Organization, 1991 (WHO Technical Report Series, No. 822), Annex 3.
11. Committee for Medicinal Products for Human Use. *Guidelines on similar biological medicinal products*. London, European Medicines Agency, 2005 (CHMP/437/04).
12. WHO/KFDA joint workshop on implementing WHO guidelines on evaluating similar biotherapeutic products, Seoul, Republic of Korea, 24–26 August 2010. *Biologicals*, 2011, 39:349–357.
13. *WHO Informal Consultation on Scientific basis for regulatory evaluation of candidate human vaccines from plants*. Geneva, Switzerland, 24–25 January 2005. Meeting report. Geneva, World Health Organization, 2005 (<http://www.who.int/biologicals/Plant%20Vaccine%20Final%20Mtg%20Report%20Jan.2005.pdf?ua=1>, accessed 28 February 2014).
14. Recommendations to assure the quality, safety and efficacy of recombinant hepatitis B vaccines. In: *WHO Expert Committee on Biological Standardization. Sixty-first report*. Geneva, World Health Organization, 2013 (WHO Technical Report Series, No. 978), Annex 4.
15. Guidelines on the quality, safety and efficacy of recombinant malaria vaccines targeting the pre-erythrocytic and blood stages of *Plasmodium falciparum*. In: *WHO Expert Committee on Biological Standardization. Sixty-third report*. Geneva, World Health Organization, 2014 (WHO Technical Report Series, No. 980), Annex 3.
16. Guidelines for assuring the quality and non-clinical safety evaluation of DNA vaccines. In: *WHO Expert Committee on Biological Standardization. Fifty-sixth report*. Geneva, World Health Organization, 2007 (WHO Technical Report Series, No. 941), Annex 1.
17. *Guidelines on the use of international nonproprietary names (INNs) for pharmaceutical substances. Annex 4: Specific groups of biological compounds*. Geneva, World Health Organization, 1997 ([http://whqlibdoc.who.int/hq/1997/WHO\\_PHARM\\_S\\_NOM\\_1570.pdf](http://whqlibdoc.who.int/hq/1997/WHO_PHARM_S_NOM_1570.pdf), accessed 7 November 2013).
18. Good manufacturing practices: main principles for pharmaceutical products. In: *WHO Expert Committee on Specifications for Pharmaceutical Preparations. Forty-fifth report*. Geneva, World Health Organization, 2011 (WHO Technical Report Series, No. 961), Annex 3.

19. Good manufacturing practices for biological products. In: *WHO Expert Committee on Biological Standardization. Forty-second report*. Geneva, World Health Organization, 1992 (WHO Technical Report Series, No. 822), Annex 1.
20. Good manufacturing practices: supplementary guidelines for the manufacture of investigational pharmaceutical products for clinical trials in humans. In: *WHO Expert Committee on Specifications for Pharmaceutical Preparations. Thirty-fourth report*. Geneva, World Health Organization, 1996 (WHO Technical Report Series, No. 863), Annex 7.
21. *Viral safety evaluation of biotechnology products derived from cell lines of human or animal origin*. ICH Guideline Q5A. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1999.
22. *Analysis of the expression construct in cells used for production of r-DNA derived protein products*. ICH Guideline Q5B. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1995.
23. *WHO Guidelines on Transmissible Spongiform Encephalopathies in relation to Biological and Pharmaceutical Products*. Geneva, World Health Organization, 2003 (WHO/BCT/QSD/03.01) (<http://www.who.int/biologicals/publications/en/whotse2003.pdf>, accessed 11 January 2013).
24. *WHO Guidelines on Tissue Infectivity Distribution in Transmissible Spongiform Encephalopathies*. Geneva, World Health Organization, 2006 (<http://www.who.int/bloodproducts/tse/WHO%20TSE%20Guidelines%20FINAL-22%20JuneupdatedNL.pdf>, accessed 11 January 2013).
25. *WHO Tables on Tissue Infectivity Distribution in Transmissible Spongiform Encephalopathies*. Geneva, World Health Organization, 2010 (WHO/EMP/QSM/2010.1) (<http://www.who.int/bloodproducts/tablestissueinfectivity.pdf>, accessed 11 January 2013).
26. General requirements for the sterility of biological substances (Revised 1973). In: *WHO Expert Committee on Biological Standardization. Twenty-fifth report*. Geneva, World Health Organization, 1973 (WHO Technical Report Series, No. 530), Annex 4.
27. Hochuli E. Interferon immunogenicity: technical evaluation of interferon-alpha 2a. *Journal of interferon & cytokine research*, 1997, 17(Suppl. 1):S15–S21.
28. Seidl A et al. Tungsten-induced denaturation and aggregation of epoetin alfa during primary packaging as a cause of immunogenicity. *Pharmaceutical Research*, 2012, 29(6):1454–1467.
29. Braun A et al. Protein aggregates seem to play a key role among the parameters influencing the antigenicity of interferon alpha (IFN-alpha) in normal and transgenic mice. *Pharmaceutical Research*, 1997, 14(10):1472–1478.
30. Stability testing of active pharmaceutical ingredients and finished pharmaceutical products. In: *WHO Expert Committee on Specifications for Pharmaceutical Preparations. Forty-third report*. Geneva, World Health Organization, 2009 (WHO Technical Report Series, No. 953), Annex 2.
31. Guidelines on stability evaluation of vaccines. In: *WHO Expert Committee on Biological Standardization. Fifty-seventh report*. Geneva, World Health Organization, 2011 (WHO Technical Report Series, No. 962), Annex 3.
32. *Comparability of biotechnological/biological products subject to changes in their manufacturing process*. ICH Guideline Q5E. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2004.
33. Committee for Medicinal Products for Human Use. *Guideline on comparability of biotechnology-derived medicinal products after a change in the manufacturing process – non-clinical and clinical issues*. London, European Medicines Agency, 2007 (EMA/CHMP/BMWP/101695/2006).
34. *Guidance concerning demonstration of comparability of human biological products, including therapeutic biotechnology-derived products*. Bethesda, MD, Food and Drug Administration, 1996.

35. *Detection of toxicity to reproduction for medicinal products & toxicity to male fertility*. ICH Guideline S5(R2). Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2000.
36. *Guideline on the need for carcinogenicity studies*. ICH Guideline S1A. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1995.
37. *Safety pharmacology studies for human pharmaceuticals*. ICH Guideline S7A. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2000.
38. *Guidance on nonclinical safety studies for the conduct of human clinical trials and marketing authorization for pharmaceuticals*. Tripartite Harmonised Guideline M3(R2). Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2009.
39. *Preclinical safety evaluation of biotechnology-derived pharmaceuticals*. ICH Guideline S6(R1). Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2011.
40. *Handbook: good laboratory practice (GLP): quality practices for regulated non-clinical research and development*, 2nd ed. Geneva, UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, 2009 (<http://www.who.int/tdr/publications/documents/glp-handbook.pdf>, accessed 26 June 2013).
41. *OECD Principles on Good Laboratory Practice*. Paris, Organisation for Economic Cooperation and Development, 1998 ([http://www.oecd-ilibrary.org/environment/oecd-principles-on-good-laboratory-practice\\_9789264078536-en](http://www.oecd-ilibrary.org/environment/oecd-principles-on-good-laboratory-practice_9789264078536-en), accessed 2 March 2014).
42. Shin J et al. International regulatory requirements for vaccine safety and potency testing: a WHO perspective. *Procedia in Vaccinology*, 2011, 5:164–170.
43. *Guidance for industry. Estimating the maximum safe starting dose in initial clinical trials for therapeutics in adult healthy volunteers*. Rockville, MD, Center for Drug Evaluation and Research, 2005.
44. Committee for Medicinal Products for Human Use. *Guideline on strategies to identify and mitigate risks for first-in-human clinical trials with investigational medicinal products*. London, European Medicines Agency, 2007 (EMA/CHMP/SWP/28367/07).
45. Chapman K et al. Preclinical development of monoclonal antibodies. Considerations for the use of non-human primates. *mAbs*, 2009, 1(5):505–516.
46. Hansel TT et al. The safety and side effects of monoclonal antibodies. *Nature Reviews Drug Discovery*, 2010, 9(4):325–338.
47. Sathish JG et al. Challenges and approaches for the development of safer immunomodulatory biologics. *Nature Reviews Drug Discovery*, 2013, 12:306–324.
48. *Guideline on development, production, characterization and specifications for monoclonal antibodies and related products*. London, European Medicines Agency, 2008.
49. *Points to consider in the manufacture and testing of monoclonal antibody products for human use*. Silver Spring, MD, Food and Drug Administration, 1997.
50. Guidelines for Good Clinical Practice (GCP) for Trials on Pharmaceutical Products. In: *WHO Expert Committee on Selection and Use of Essential Medicines. Sixth Report*. Geneva, World Health Organization, 1995 (WHO Technical Report Series, No. 850), Annex 3 (<http://apps.who.int/medicinedocs/en/d/Jwhozip13e/>, accessed 1 March 2014).

51. *Guideline on requirements for first-in-man clinical trials for potential high-risk medicinal products.* London, European Medicines Agency, 2007 (EMA/CHMP/SWP/28367/2007 Corr.).
52. *Draft guidance for industry. Clinical pharmacogenomics: premarketing evaluation in early phase clinical studies.* Silver Spring, MD, Food and Drug Administration, 2011.
53. Hudson KL. Genomics, health care, and society. *New England Journal of Medicine*, 2011, 365:1033–1041.
54. Committee for Medicinal Products for Human Use. *Guideline on the clinical investigation of the pharmacokinetics of therapeutic proteins.* London, European Medicines Agency, 2004.
55. *Guidance for industry. Safety, efficacy, and pharmacokinetic studies to support marketing of immune globulin intravenous (human) as replacement therapy for primary humoral immunodeficiency.* Silver Spring, MD, Food and Drug Administration, 2008.
56. Richter et al. Mechanistic determinants of biotherapeutics absorption following SC administration. *The AAPS Journal*, 2012, 14(3):559–570.
57. Tang L et al. Pharmacokinetic aspects of biotechnology products. *Journal of Pharmaceutical Sciences*, 2004, 93(9):2184–2204.
58. Porter et al. Lymphatic transport of proteins after subcutaneous administration. *Journal of Pharmaceutical Sciences*, 2000, 89(3):297–310.
59. Mahmood I et al. Drug interaction studies of therapeutic proteins or monoclonal antibodies. *Journal of Clinical Pharmacology*, 2007, 47:1540–1554.
60. Solá RJ, Griebenow K. Glycosylation of therapeutic proteins. *BioDrugs*, 2010, 24(1):9–21.
61. *Dose response studies to support drug registration.* ICH Guideline E4. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1994.
62. *The clinical evaluation of QT/QTc interval prolongation and proarrhythmic potential for non-antiarrhythmic drugs.* ICH Guideline E14. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2005.
63. *Choice of control group and related issues in clinical trials.* ICH Guideline E10. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2000.
64. *Statistical principles for clinical trials.* ICH Guideline E9. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1998.
65. *Definitions for genomic biomarkers, pharmacogenomics, pharmacogenetics, genomic data and sample coding categories.* ICH Guideline E15. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2007.
66. *Biomarkers related to drug or biotechnology product development: context, structure and format of qualification submissions.* ICH Guideline E16. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2010.
67. *Studies in support of special populations. Geriatrics, questions & answers.* ICH Guideline E7. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2010.
68. *Clinical investigation of medicinal products in the paediatric population.* ICH Guideline E11. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2000.

69. *Guideline for good clinical practice*. ICH Guideline E6. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1996.
70. Adaptive Designs in Clinical Drug Development – An Executive Summary of the PhRMA Working Group. *Journal of Biopharmaceutical Statistics*, 2006, 16:275–283.
71. *Guidance for industry. Adaptive design clinical trials for drugs and biologics*. Silver Spring, MD, Food and Drug Administration, 2010.
72. Committee for Medicinal Products for Human Use. *Guideline on missing data in confirmatory clinical trials*. London, European Medicines Agency, 2011.
73. *Guidance for industry. Clinical trial endpoints for the approval of cancer drugs and biologics*. Silver Spring, MD, Food and Drug Administration, 2007.
74. *The extent of population exposure to assess clinical safety for drugs intended for long-term treatment of non-life-threatening conditions*. ICH Guideline E1. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1994.
75. *Clinical safety data management: definitions and standards for expedited reporting*. ICH Guideline E2A. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1994.
76. *Maintenance of the ICH guideline on clinical safety data management: data elements for transmission of individual case safety reports*. ICH Guideline E2B. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2001.
77. MedDRA terminology M1. Medical Dictionary for Regulatory Activities. Points to consider ICH-endorsed. Guide for MedDRA Users on Data Output, release 3.3 based on MedDRA Version 15.0, 1 April 2012.
78. Studies in support of special populations. Geriatrics. ICH Guideline E7. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 1993.
79. Committee for Medicinal Products for Human Use. *Guideline on immunogenicity assessment of biotechnology-derived therapeutic proteins*. London, European Medicines Agency, 2007 (CHMP/BMWP/14327).
80. Committee for Medicinal Products for Human Use. *Guideline on immunogenicity assessment of monoclonal antibodies intended for in vivo clinical use*. London, European Medicines Agency, 2012 (EMA/CHMP/BMWP/86289/2010).
81. *Guidance for industry. Immunogenicity assessment for therapeutic protein products*. Rockville, MD, Center for Biologics Evaluation and Research, 2013 (<http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM338856.pdf>, accessed 7 November 2013).

# Appendix 1

## Manufacturing process validation

- Process validation is the documented evidence that the process, operated within established parameters, can perform effectively and reproducibly to produce a drug product, drug substance or intermediate that meets its predetermined specifications and quality attributes.
- Process validation should include the collection and evaluation of data throughout production in order to establish scientific evidence that a process is capable of consistently delivering a quality drug substance. It generally includes collection of data on an appropriate number of production batches. The number of batches can depend on several factors that include, but are not limited to: (a) the complexity of the process being validated; (b) the level of process variability; and (c) the amount of experimental data and/or process knowledge available on the specific process.
- Process conditions (e.g. column loading capacity, column regeneration and sanitization, height) should be appropriately evaluated. Columns should also be evaluated throughout their expected lifespan with regard to their purification ability (e.g. impurity clearance, collection of intended variants), leaching of ligands (e.g. dye, affinity ligand) and/or chromatographic material (e.g. resin). Process validation activities should normally include the evaluation of resin lifetime, including maximum cycles and/or maximum time duration, using small-scale studies to ensure proper performance and integrity of the columns. In addition, the results should normally be verified at full scale through the life-cycle of the product. These studies should also confirm the suitability of the column cleaning, storage and regeneration procedures.
- Where hold times are applied to intermediates (e.g. harvest, column eluate), the impact of hold times and hold conditions on the product quality (e.g. degradation) should be appropriately evaluated.
- Evaluation of selected steps (e.g. steps for which high impurity or viral clearance are claimed) operating in worst-case and/or challenging conditions (e.g. maximum hold times, spiking challenge) could be performed to demonstrate the robustness of the process. Depending on the relevance of the experimental model

with regard to the final process (e.g. scale, materials, equipment, operating conditions), these studies could be leveraged in support of process validation and/or quality control data requirements.

- The information provided in the dossier in support of process validation usually contains both commercial-scale process validation studies and small-scale studies. Process validation batches should be representative of the commercial process, taking into account the batch definition as detailed in the process description.
- Process changes at the level of fermentation and/or purification during progression to full-scale commercial production may have considerable consequences for the quality of the product, the yield and/or quantitative and qualitative differences in impurities. Consequently the contribution of data from small-scale studies to the overall validation package will depend on demonstration that the small-scale model is an appropriate representation of the proposed commercial scale. Data demonstrating that the model is scalable and representative of the proposed commercial process should be provided. Successful demonstration of the suitability of the small-scale model can enable manufacturers to propose process validation with reduced dependence on testing of commercial-scale batches. Data derived from commercial-scale batches should confirm results obtained from small-scale studies used to generate data in support of process validation. Scientific rationale or reference to guidelines can be an appropriate justification to conduct certain studies (e.g. viral removal) only at small scale.
- In order to demonstrate viral safety of purification processes used to manufacture drug substance for clinical trials, in-house data from previous validation studies may be used. If in-house experience with highly robust and well understood process steps is available, it may be justified to reduce the product-specific validation effort (1).

## Reference

1. Committee for Medicinal Products for Human Use. *Guideline on virus safety evaluation of biotechnological investigational medicinal products*. London, European Medicines Agency, 2006 (EMA/CHMP/BWP/398498/2005).

## Appendix 2

### Characterization of rDNA-derived biotherapeutics

This appendix provides details of suggested approaches that can be applied to the characterization of an rDNA-derived biotherapeutic. It also provides examples of technical approaches which may be considered for structural characterization and confirmation, and for evaluation of physicochemical and biological properties of the desired product, drug substance and/or drug product. The methods should provide an understanding of the product with a sufficient level of detail (e.g. complete primary structure, properties for the higher order structure, qualitative and quantitative analysis of product-related substances and product- and process-related impurities, assessment of biological functions).

A subset of the methods described in this appendix can be used for routine batch release testing. Others are subject to extended characterization of the desired product during product and process development and are also often used to support process evaluation/validation and/or comparability studies (e.g. after making significant process changes). The selection of release testing methods depends on the overall design of quality control for which release testing is only one element among others. For example, if a certain quality attribute can be controlled by in-process tests, parametric controls and/or demonstrated manufacturing process capability (e.g. high impurity clearance), that attribute may not need to be tested routinely on every batch.

#### 1. Physicochemical characterization

##### 1.1 Primary structure

The primary structure – i.e. amino acid sequence, including the disulfide linkages – of the desired product can be determined as far as possible using combined approaches such as those described in items (a) and (b) below and then compared with the sequence of the amino acids deduced from the gene sequence of the desired product. Attention should be paid to the possible presence of N-terminal methionine (e.g. in *Escherichia coli*-derived products), signal or leader sequences, other possible N-terminal and C-terminal modifications (such as acetylation, amidation or partial degradation by exopeptidases), and any heterogeneity (e.g. C-terminal processing, N-terminal pyroglutamation, deamidation, oxidation, isomerization, fragmentation, disulfide bond mismatch, N-linked and O-linked oligosaccharide, glycation, aggregation). The variability of N-terminal and C-terminal amino acid sequences should be analysed (e.g. C-terminal lysine(s)).

Free sulphhydryl groups and disulfide bridges should be determined. Disulfide bridge integrity and mismatch should be analysed. Experimentally

determined disulfide bonding patterns should be compared to the predicted structure based on the class of the molecule.

*(a) Peptide map* – selective fragmentation of the product into discrete peptides is performed by using suitable enzymes or chemicals. The resulting peptide fragments are analysed by high-performance liquid chromatography (HPLC) or other appropriate analytical procedures. The peptide fragments should be identified as far as possible using appropriate techniques such as mass spectrometry (MS) methods (e.g. electrospray ionization MS, matrix-assisted laser-desorption ionization time-of-flight MS). The use of MS/MS coupling should also be considered as it could reveal more detailed sequence information about the analysed peptide fragment. If one fragmentation method does not deliver the complete amino acid sequence, the use of an orthogonal enzyme or chemical cleavage method can increase the sequence coverage. The correct formation of the disulfide bridges may be characterized by the use of peptide mapping under reducing and non-reducing conditions.

*(b) Molecular weight determination by mass spectrometry* – the molecular weight of the intact molecule, as determined by MS, serves as an additional confirmation of the primary structure. For smaller peptides, MS/MS sequencing can provide the complete amino acid sequence. MS can be performed under reduced and non-reduced conditions and under deglycosylated and intact conditions for multi-subunit and glycosylated protein molecules such as monoclonal antibodies.

## 1.2 Glycan structure

Glycosylation should be identified and adequately characterized. The glycan content (neutral sugars, amino sugars and sialic acids) should be determined if it is linked to clearance or activity. In addition, the structure of the glycan chains, the glycan pattern (antennary profile native glycan profile and site-specific glycan analysis), and the glycosylation site(s) of the polypeptide chain are analysed as far as possible. This task can be achieved by the combination of enzymatic or chemical hydrolytic cleavage with a variety of separation methods (HPLC, electrophoresis) and detection/identification methods (MS including MS/MS, ultraviolet, fluorescence detection, electrochemical detection). The quantitative oligosaccharide analysis (chemical or enzymatic cleavage followed by HPLC) provides additional useful qualitative and quantitative information on the glycan structure.

Measurement of the quantitative charge patterns of the intact glycoprotein, such as by measuring the charge-based isoforms using an appropriate method (e.g. capillary electrophoresis, isoelectric focusing), may be useful as an overall

measure of the degree of sialylation and antennary profile. Particular attention should be paid to glycan structures that may be associated with adverse effects, such as non-human structures or residues. Further tests to be conducted include analysis of charge heterogeneity.

### 1.3 Higher-order structure

Higher-order structure should be characterized by appropriate physicochemical methodologies and confirmed by biological function. The analysis of PEGylated proteins should include, though should not be limited to, the average rate of modification, the location of modification and the analysis of site occupancy.

The complete assessment of the three-dimensional chemical structure in the context of product characterization is rarely achieved because absolute methods such as X-ray crystallography or nuclear magnetic resonance (NMR) with isotope-labelled amino acids deliver only an approximation to the structure of the product of interest. They measure the product either in a nonrelevant state or require a separate production of the isotope-labelled sample. However, the use of applicable but relative orthogonal methods as described below enables the determination and characterization of discrete folding and the assessment of changes in the higher-order structure (e.g. in the case of comparability studies).

The higher-order structure of the product should be examined using appropriate procedures such as circular dichroism, Fourier transform infrared spectroscopy (FT-IR), fluorescence, differential scanning calorimetry, proton nuclear magnetic resonance (<sup>1</sup>H-NMR) and/or other suitable techniques such as hydrogen-deuterium exchange MS. FT-IR and CD in the far ultraviolet range deliver information on the secondary structure, whereas CD in the near ultraviolet reflects to some extent the tertiary and quaternary structure. When using these methods, their capabilities and limitations need to be considered (e.g. impact of protein concentration).

In vitro or in vivo assays that illustrate the functional activity of the therapeutic may also serve as additional confirmation of the higher-order structure in addition to demonstrating biological function.

## 2. Biological activity

Assessment of the biological properties of a product constitutes an essential step in establishing a complete characterization profile. The biological activity describes the specific ability or capacity of a product to achieve a defined biological effect. Description of a relevant biological assay to measure the biological activity should be provided by the manufacturer.

The biological activity should be assessed by in vitro, in vivo, biochemical (including immunochemical assays) and/or physicochemical assays as appropriate.

For antibody products, where effector function may play a role in the mechanism of action and/or have an impact on the product safety and efficacy, a detailed analysis of biological activity demonstrating the mechanism of action (e.g. antibody-dependent cellular cytotoxicity, complement-dependent cytotoxicity, apoptosis), ability for complement-binding and activation, and other effector functions – including Fc gamma receptor-binding activity and neonatal Fc receptor-binding activity – should be provided as appropriate.

The mechanism of action should be discussed and, where relevant, the importance (or consequences) of other functions (e.g. effector functions) with regard to the safety and efficacy of the product should be included.

Potency (expressed, for example, in units or international units (IU)) is the quantitative measure of biological activity based on the attribute of the product which is linked to the relevant biological properties, whereas quantity (expressed in mass) is a physicochemical measure of product content. For assessing potency, use of bioassays that reflect the biological activity in the clinical situation is preferable but is not always possible or necessary for lot release. For example, bioassays which assess some functional aspect of the protein or mechanism of action (rather than the intended clinical effect) can also be used as the basis for a potency assay.

Examples of procedures used to measure biological activity include:

- animal-based biological assays, which measure an organism's biological response to the product;
- cell-based biological assays, which measure biochemical or physiological response at the cellular level;
- biochemical assays, which measure biological activities such as receptor- or ligand-binding, enzymatic reaction rates or biological responses induced by immunological interactions.

### 3. Immunochemical properties

Where relevant (e.g. for monoclonal antibody products), the immunochemical properties should be extensively characterized. Binding assays using purified antigens and defined regions of antigens should be performed, where feasible, to determine affinity, avidity and immunoreactivity (including cross-reactivity with other structurally homologous proteins).

The part of the target molecule bearing the relevant epitope should be characterized to the extent that this is possible. This should include biochemical identification of these structures (e.g. protein, oligosaccharide, glycoprotein, glycolipid) and relevant characterization studies (amino acid sequence, carbohydrate structure) as appropriate.

Since glycosylation and PEGylation may have an impact on the pharmacological properties of the product and may modulate its immunogenic properties, appropriate characterization studies should be conducted.

Unless otherwise justified, the ability for complement-binding and activation, and/or other effector functions, should be evaluated even if the intended biological activity does not require such functions.

#### 4. Purity, impurity and contaminant

Biotechnological products commonly display several sources of heterogeneity (e.g. C-terminal processing, N-terminal pyroglutamation, deamidation, oxidation, isomerization, fragmentation, disulfide bond mismatch, N-linked and O-linked oligosaccharide, glycation, aggregation), which lead to a complex purity/impurity profile comprising several molecular entities or variants. This purity/impurity profile should be assessed by a combination of methods, and individual and/or collective acceptance criteria should be established for relevant product-related substances and impurities. These methods generally include the determination of physicochemical properties such as molecular weight or size, isoform pattern, determination of hydrophobicity, electrophoretic profiles, chromatographic data including peptide mapping and spectroscopic profiles including mass spectroscopy. Multimers and aggregates should also be appropriately characterized using a combination of methods. Unless otherwise justified, the formation of aggregates and subvisible and visible particulates in the drug product is important and should be investigated and closely monitored at the time of release and during stability studies.

Impurities may be either process-related or product-related. These materials should be characterized as far as is possible and their impact on biological activity should be evaluated if appropriate.

Potential process-related impurities (e.g. host-cell protein, host-cell DNA, cell culture residues, downstream processing residues) should be identified and evaluated qualitatively and/or quantitatively, as appropriate.

Contaminants, which include all adventitiously introduced materials not intended to be part of the manufacturing process (e.g. microbial species, endotoxins) should be strictly avoided and/or suitably controlled. Where non-endotoxin pro-inflammatory contaminants, such as peptidoglycan, are suspected, the use of additional testing should be considered.

##### 4.1 Process-related impurities and contaminants

Process-related impurities are derived from the manufacturing process itself and can be classified in three major categories: (a) cell substrate-derived; (b) cell culture-derived; and (c) downstream-derived. Contaminants, on the other hand,

are unwanted materials, such as adventitious viruses, that are introduced by unintentional means into the manufacturing process.

*(a) Cell substrate-derived impurities* – include, but are not limited to, proteins derived from the host organism, and nucleic acid (host-cell genomic, vector, or total DNA). For host-cell proteins, a sensitive assay (e.g. immunoassay) capable of detecting a wide range of protein impurities is generally utilized. In the case of an immunoassay, polyclonal antibodies used in the test are typically generated by the immunization of animals with an appropriate preparation derived from the production cell minus the product-coding gene, which have been cultured in conditions representative of the intended culture and appropriately collected (e.g. filtered harvest, partial purification).

The level of DNA from the host cells can be detected by direct analysis on the product (e.g. qPCR, immunoenzymatic techniques). Clearance studies, which could include spiking experiments conducted at small scale, to demonstrate the removal of cell substrate-derived impurities such as nucleic acids and host-cell proteins, may sometimes be used to eliminate the need for establishing acceptance criteria for these impurities.

*(b) Cell culture-derived impurities* – include, but are not limited to, inducers, antibiotics, serum and other media components. These impurities need to be tested and evaluated on a case-by-case basis using a risk-assessment and risk-management approach. In the case of a potential impact on the safety of the product, the removal of such impurities to acceptably low levels during downstream purification may need to be validated or end-product testing and specification limits established.

*(c) Downstream-derived impurities* – include, but are not limited to, enzymes, chemical and biochemical processing reagents (e.g. guanidine, dyes, oxidizing and reducing agents), inorganic salts (e.g. heavy metals, non-metallic ions), solvents, carriers, ligands (e.g. protein A) and other leachables. As for cell culture-derived impurities, these impurities should be evaluated on a case-by-case basis using a risk-assessment and risk-management approach. Where appropriate, development of analytical methods for these impurities and validation of their removal could be considered.

#### 4.2 **Product-related substances and impurities, including degradation products**

Molecular variants of the desired product may need considerable effort in isolation and characterization in order to identify the type of modification(s). When the activity of these variants is comparable to the desired product, the variants should be included in the product purity profile. Degradation products arising in significant amounts during manufacture and/or storage

should be appropriately considered. The most frequently encountered molecular variants of the desired product, and relevant technology for their assessment, are listed below.

*(a) Truncated forms* – hydrolytic enzymes or chemicals may catalyse the cleavage of peptide bonds. This may lead to terminal heterogeneity (e.g. for C-terminal Lys in monoclonal antibodies). These may be detected by HPLC and/or electrophoretic methods and verified by mass spectrometry. Peptide mapping may also be useful, depending on the property of the variant.

*(b) Amino acid modifications* – individual amino acid modification may include deamidation (Asp/Gln to Asp, Glu), oxidation (e.g. Met to Met-sulfoxide), spontaneous formation of pyroglutamate out of N-terminal Glu or Gln residues, glycation of Lys residues and others. These forms may be detected and characterized by relevant analytical methods (e.g. HPLC, capillary electrophoresis, mass spectrometry). In some cases peptide mapping is important to clearly identify and localize the site and nature of the amino acid modification.

*(c) High molecular weight species and particles* – high molecular weight species (HMWS) includes dimers and higher oligomers of the desired product. Particles include intrinsic visible particles of the desired product. HMWS are generally resolved from the desired product and product-related substances, and are quantitated by appropriate separation procedures (e.g. size exclusion chromatography, field flow fractionation, analytical ultracentrifugation) coupled with sensitive detection methods (e.g. ultraviolet, fluorescence, light scattering). Using orthogonal methods and/or procedures with overlapping analytical windows (e.g. light obscuration testing, micro-flow imaging for testing of subvisible particles) can greatly enhance the characterization of aggregates and particles. Foreign particles are not intended to be part of the product and should be minimized.

## 5. Quantity

Quantity should be determined by use of an appropriate physicochemical and/or immunochemical assay. The protein content (expressed in mass units) can be determined by measuring the sample against an appropriate reference standard using a suitable method (e.g. HPLC). The protein content can also be measured in an absolute way – such as by ultraviolet photometry using an extinction coefficient (e.g. at 280 nm). If the deviation is too large, redetermination by another method can be considered.

## Appendix 3

### Routine control of rDNA-derived biotherapeutics

This appendix discusses approaches to routine control of an rDNA-derived biotherapeutic.

#### 1. Specification

A specification is defined as a list of tests, references to analytical procedures, and appropriate acceptance criteria which are numerical limits, ranges, or other criteria for the tests described. A specification establishes the set of criteria to which a drug substance and drug product – or materials at other stages of the manufacture – should conform in order to be considered acceptable for its intended use. “Conformance to specification” means that the drug substance and drug product, when tested according to the listed analytical procedures, will meet the acceptance criteria. The justification of specification should take into account relevant development data and data from nonclinical, clinical and stability studies. The setting of acceptance ranges should also take into account the sensitivity of the analytical method used.

The selection of tests to be included in the specifications is product-specific and should take into account the quality attributes (e.g. potential influence on safety, efficacy or stability), the process performance (e.g. clearance capability, content), the controls in place through the manufacturing process (e.g. multiple testing points), and the material used in relevant nonclinical and clinical studies. These tests could include criteria such as potency, the nature and quantity of product-related substances, product-related impurities, process-related impurities, and absence of contaminants. Such attributes can be assessed by multiple analytical procedures, each yielding different results. Since specifications are chosen to confirm quality rather than to characterize the product, the rationale and justification for including and/or excluding testing for specific quality attributes should be provided.

The rationale used to establish the acceptable range of acceptance criteria should be described. Acceptance criteria should be established and justified on the basis of data obtained from lots used in nonclinical and/or clinical studies. Nevertheless, where appropriately justified, data from lots used for stability studies, or relevant development data, could support limits beyond ranges used in clinical studies.

## 2. Identity

The identity test(s) should be highly specific and should be based on unique aspects of the product's molecular structure and/or other specific properties (e.g. peptide map, anti-idiotypic immunoassay, or other appropriate method). Depending on the product, more than one test (physicochemical, biological and/or immunochemical) may be necessary to establish identity, and such test(s) should possess sufficient specificity to discriminate other products that may be manufactured in the same facility.

## 3. Purity and impurities

As noted in the characterization section, recombinant proteins may display a complex purity/impurity profile that should be assessed by a combination of orthogonal methods, and for which individual and/or collective acceptance criteria should be established for relevant product-related variants. Chromatographic and/or electrophoretic methods capable of detecting product truncation, dissociation and aggregation should be included, and quantitative limits should be proposed for these, as appropriate. In addition, as appropriate, such control could further confirm the consistency of the product.

The control of relevant process-related impurities should be included in the plan for quality control. Control of process-related impurities (e.g. protein A, host-cell protein, DNA, and other potential culture or purification residues) is typically part of the drug substance specification, as appropriate. In some situations, and where appropriately demonstrated, their control may be performed on an intermediate product at an appropriate process step. Routine testing may not be necessary for some impurities for which the process has been demonstrated to achieve high reduction levels.

## 4. Potency

Potency is the quantitative measure of biological activity based on the attribute of the product which is linked to the relevant biological properties. A relevant potency assay should be part of the specifications for drug substance and/or drug product, and should reflect the presumed mechanism of action whenever possible. Specific activity (units of biological activity per mg of product) is of considerable value in demonstrating consistency of production.

The potency of each batch of the drug substance and the final dosage form should be established using, wherever possible, an appropriate national or international reference material – see, for example, section A.1.3 – which is normally calibrated in units of biological activity such as IU. In the absence of

such preparations, an approved in-house reference preparation may be used for assay standardization.

For biological substances with antagonist activity, it may be appropriate to calibrate the potency assay using the standard/reference preparation for the agonist and to express activity of the antagonist in terms of inhibition of biological activity – i.e. units of the agonist. For example, for tumour necrosis factor (TNF) antagonists, bioassays can be calibrated using the international standard for TNF- $\alpha$  and activity expressed as the number of IUs of TNF neutralized by the amount of the antagonist.

## 5. Quantity

The quantity of the drug substance and drug product, usually based on protein content, should be determined using an appropriate assay.

## 6. General tests

General tests should be performed in accordance with relevant monographs, which could include appearance (e.g. form, colour), solubility, pH, osmolality, extractable volume, sterility, bacterial endotoxins, stabilizer and water, and visible and subvisible particulate, as appropriate.

## Appendix 4

### Product-/indication-specific guidance in nonclinical evaluation (examples)

#### 1. Anticancer rDNA-derived biotherapeutics

For anticancer rDNA-derived biotherapeutics, nonclinical evaluations are intended to identify the pharmacological properties, establish a safe initial dose level for the first human exposure and understand the toxicological profile (e.g. identification of the target organ, estimation of the safety margin and reversibility). In the development of anticancer drugs, most clinical studies involve cancer patients whose disease condition is often progressive and fatal. In addition, the clinical dose levels are often close to or at the adverse effect dose levels. For these reasons, the type and timing and flexibility called for in designing of nonclinical studies of anticancer pharmaceuticals can have a different pattern from those for other pharmaceuticals (1, 2).

##### 1.1 Starting dose for clinical studies

Nonclinical evaluations should identify a pharmacologically active and safe dose. For selection of the starting dose for first-in-human clinical trials, a MABEL approach should be considered (3). Toxicology studies to determine a NOAEL/NOEL (no observed effect level) are not considered essential to support clinical use of an anticancer medicinal product.

##### 1.2 Study duration

For medicinal products intended for the treatment of patients with advanced cancer, nonclinical studies of 3 months' duration are usually considered sufficient to support phase III clinical studies and, in most cases, licensing.

##### 1.3 Reproductive toxicity

With regard to reproduction toxicology, an embryo-fetal toxicity study should be available for licensing but is not considered essential to support clinical trials in patients with advanced cancer. Fertility and prenatal and postnatal toxicological studies are in general not warranted to support clinical trials or licensing for rDNA-derived biotherapeutics intended for the treatment of patients with advanced cancer (2).

## 2. Monoclonal antibodies

For monoclonal antibodies, the immunological properties of the antibody should be described in detail, including its antigenic specificity, complement-binding, and any unintentional reactivity and/or cytotoxicity towards human tissues distinct from the intended target. For monoclonal antibodies and other related antibody products directed at foreign targets (i.e. bacterial, viral targets, etc.), a short-term (i.e. 2 weeks' duration) safety study in one species (with the choice of species justified by the sponsor) can be considered; no additional toxicity studies, including reproductive toxicity studies, are needed. When animal models of disease are used to obtain proof of principle, a safety assessment can be included to provide information on potential target-associated safety aspects. Where this is not feasible, appropriate risk mitigation strategies should be adopted for clinical trials.

### 2.1 Antibody-drug/toxin conjugates

Species selection for an antibody-drug/toxin conjugate (ADC) incorporating a novel toxin/toxicant should follow the same general principles as an unconjugated antibody. If two species have been used to assess the safety of the ADC, an additional short-term study or an arm in a short-term study should be conducted in at least one species with the unconjugated toxin. In these cases a rodent is preferred unless the toxin is not active in the rodent. If only one pharmacologically relevant species is available, then the ADC should be tested in this species. A novel toxicant calls for an approach to species selection similar to that used for a new chemical entity on a case-by-case approach – see, for example, reference (2) below. For toxins or toxicants which are not novel and for which a sufficient body of scientific information is available, separate evaluation of the unconjugated toxin is not warranted. Data should be provided to compare the metabolic stability of the ADC in animals with humans.

## References

1. *Preclinical safety evaluation of biotechnology-derived pharmaceuticals*. ICH Guideline S6(R1). Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2011.
2. *Nonclinical evaluation for anticancer pharmaceuticals*. ICH Guideline S9. Geneva, International Conference on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use, 2009.
3. *Guideline on strategies to identify and mitigate risks for first-in-human clinical trials with investigational medicinal products*. London, European Medicines Agency, 2007 (EMA/CHMP/SWP/ 28367/07).

# Appendix 5

## Animal species/model selection

### 1. Species selection

The biological activity, together with species- and/or tissue-specificity, of many rDNA-derived biotherapeutics often precludes standard pharmacological/toxicity testing designs in commonly used species (e.g. rats and dogs). Pharmacological and safety evaluation programmes should include the use of relevant species. A relevant species is one in which the test material is pharmacologically active due to the expression of the receptor or an epitope (in the case of monoclonal antibodies). In addition to receptor expression, the cellular/tissue distribution of receptors is an important consideration in the selection of appropriate species.

A number of factors should be taken into account when determining species relevancy. Comparisons of target sequence homology between species can be an appropriate starting point, followed by in vitro assays to make qualitative and quantitative cross-species comparisons of relative target-binding affinities and receptor/ligand occupancy and kinetics. Assessments of functional activity are also recommended. Functional activity can be demonstrated in species-specific cell-based systems and/or in vivo pharmacology or toxicology studies. Modulation of a known biological response or of a PD marker can provide evidence for functional activity to support species relevance.

Consideration of species differences in target-binding and functional activity in the context of the intended dosing regimens should provide confidence that a model is capable of demonstrating potentially adverse consequences of target modulation. When the target is expressed at very low levels in typical healthy preclinical species (e.g. inflammatory cytokines or tumour antigens), binding affinity and activity in cell-based systems can be sufficient to guide species selection.

Tissue cross-reactivity in animal tissues is of limited value for species selection. However, in specific cases (i.e. where the approaches described above cannot be used to demonstrate a pharmacologically relevant species) TCR studies can be used to guide the selection of species to be used in toxicology studies by comparison of tissue-binding profiles in human and those animal tissues where target-binding is expected (see also section B.3.3). An animal species which does not express the desired epitope may still be of some relevance for assessing toxicity if comparable unintentional tissue cross-reactivity to humans is demonstrated.

When no relevant species exists, the use of relevant transgenic animals expressing the human receptor or the use of homologous proteins should be considered.

## 2. Number of species

Safety evaluation programmes should normally include two relevant species. However, in certain justified cases one relevant species may suffice (e.g. when only one relevant species can be identified or when the biological activity of the biotherapeutic is well understood).

In addition, even where two species may be necessary to characterize toxicity in short-term studies, it may be possible to justify the use of only one species for subsequent long-term toxicity studies. If there are two pharmacologically relevant species for the clinical candidate (one rodent and one non-rodent), both species should be used for short-term (up to 1 month's duration) general toxicology studies. If the toxicological findings of these studies are similar, or the findings are understood from the mechanism of action of the product, then longer-term general toxicity studies in one species are usually considered sufficient. The rodent species should be considered unless there is a scientific rationale for using non-rodents. Studies in two non-rodent species are not appropriate.

The use of one species for all general toxicity studies is justified when the clinical candidate is pharmacologically active in only one species. Studies in a second species with a homologous product (see below) are not considered to add further value for risk assessment and are not recommended.

### 2.1 Transgenic animals

When no relevant animal species exists for testing the clinical candidate, the use of a transgenic animal expressing the human target can be considered, assuming that data exist on comparable expression and distribution of the target orthologue, and on the biology of the target in the model, and that sufficient background knowledge on the strain/model (e.g. historical background data) exist.

### 2.2 Homologous proteins

While useful information may also be gained from the use of homologous proteins, it should be noted that the production process, range of impurities/contaminants, PK and exact pharmacological mechanism(s) may differ between the homologous form and the product intended for clinical use. Studies with homologous proteins can be used for hazard detection and for understanding the potential for adverse effects due to exaggerated pharmacology, but are generally not useful for quantitative risk assessment. Therefore, for the purposes of hazard identification it can be possible to conduct safety evaluation studies using a control group and one treatment group, provided there is a scientific justification for the study design and the dose(s) selected (e.g. maximum pharmacological dose).

### 2.3 Nonclinical testing in a nonrelevant species

Pharmacological/toxicity studies in nonrelevant species may be misleading and are generally discouraged. However, where it is not possible to identify a relevant species or to use transgenic animal models, or if it is not possible to use a homologous protein for testing purposes, it may still be prudent to assess some aspects of potential toxicity in a limited toxicity evaluation in a single species (e.g. a repeated dose toxicity study of < 14 days' duration that includes an evaluation of important functional end-points such as cardiovascular and respiratory end-points).

## 3. Animal models of disease

In recent years there has been much progress in the development of animal models that are thought to be similar to the human disease. These animal models include induced and spontaneous models of disease, gene knock-out(s) or knock-in(s), and transgenic animals. These models may provide further insight in determining the pharmacological action of the product, PK and dosimetry, and may also be useful in the determination of safety (e.g. evaluation of undesirable promotion of disease progression). In certain cases, studies performed in animal models of disease may be used as an acceptable alternative to toxicity studies in normal animals.

Animal models of disease may be useful in the definition of toxicity end-points; selection of clinical indications; and determination of appropriate formulations, route of administration and treatment regimen. It should be noted that with these models of disease there is often a paucity of historical data for use as a reference when evaluating study results. Therefore, the collection of concurrent control and baseline data is critical for optimizing study design.

The scientific justification should be provided for the use of these animal models of disease to support safety.

## Appendix 6

### Explanatory notes

**Note 1:** The species-specific profile of embryo-fetal exposure during gestation should be considered in interpreting studies. High molecular weight proteins (> 5000 D) do not cross the placenta by simple diffusion. For monoclonal antibodies with molecular weight as high as 150 000 D, there exists a specific transport mechanism – the neonatal Fc receptor – which determines fetal exposure and varies across species.

In the NHPs and humans, IgG placental transfer is low in the period of organogenesis and begins to increase in the early second trimester, reaching the highest levels late in the third trimester. Therefore, standard embryo-fetal studies in NHPs, which are dosed from early pregnancy up to gestation day 50, may not be of value in assessing direct embryo-fetal effects in the period of organogenesis, although effects on embryo-fetal development as an indirect result of maternal effects can be evaluated. Furthermore, maternal dosing in NHPs after delivery is generally without relevance since IgG is excreted in the milk only initially (i.e. in the colostrum), and not later during the lactation and nursing phase.

Rodents differ from the NHPs and humans, as IgG crosses the yolk sac in rodents by neonatal Fc receptor transport mechanisms and exposure can occur relatively earlier in gestation than with NHPs and humans. In addition, delivery of rodents occurs at a stage of development when the pups are not as mature as those of the NHP or the human neonate. Therefore, rat/mouse dams should be dosed during lactation in order to expose pups via the milk up to at least day 9 of lactation when the offspring are at an equivalent stage of development as human neonates.

**Note 2:** The minimum duration of postnatal follow-up should be 1 month to cover early functional testing (e.g. growth and behaviour). In general, if there is evidence for adverse effects on the immune system (or immune function) in the general toxicology studies, immune function testing in the offspring during the postpartum phase of the ePPND study is warranted. When appropriate, immunophenotyping can be obtained as early as postnatal day 28. The duration of postnatal follow-up for the assessment of immune function can be 3–6 months depending on the functional tests used.

Neurobehavioural assessment can be limited to clinical behavioural observations. Instrumental learning calls for a training period, which would result in a postnatal duration of at least 9 months and is not recommended.

**Note 3:** A detailed discussion of the approach to determining group sizes in cynomolgus monkey ePPND studies is available (1). 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 and provide the opportunity for specialist evaluation if necessary (e.g. immune system).

Most ePPND studies accrue pregnant animals over weeks and months. Consideration should be given to terminating further accrual of pregnant animals into the study and adapting the study design (e.g. by caesarean section) when prenatal losses in a test item group indicate a treatment-related effect. Reuse of vehicle-control treated maternal animals is encouraged. If there is some cause for concern that the mechanism of action may lead to an effect on EFD or pregnancy loss, studies can be conducted in a limited number of animals in order to confirm the hazard.

**Note 4:** An example of an appropriate scientific justification would be a monoclonal antibody which binds a soluble target with a clinical dosing regimen intended to saturate target-binding. If such a saturation of target-binding can be demonstrated in the animal species selected and there is an exposure multiple of up to 10-fold the therapeutic drug levels, a single-dose level and control group would provide adequate evidence of hazard to embryo-fetal development.

**Note 5:** End-points to be included in an interim report of an ePPND study in NHPs are:

- Dam data – survival, clinical observations, bodyweight, gestational exposure data (if available), any specific PD end-points.
- Pregnancy data – number of pregnant animals started on study, pregnancy status at both the end of organogenesis (gestation day 50) and at gestation day 100, occurrence of abortions and timing of abortions. There is no need for ultrasound determinations of fetal size in the interim report; these are not considered essential since actual birth weight will be available.
- Pregnancy outcome data – number of live births/still births, infant birth weight, infant survival and bodyweight at day 7 postpartum, qualitative external morphological assessment (i.e. confirming appearance is within normal limits), infant exposure data (if available), any specific PD end-points in the infant if appropriate.

## Reference

1. Jarvis P et al. The cynomolgus monkey as a model for developmental toxicity studies: variability of pregnancy losses, statistical power estimates, and group size considerations. *Birth Defects Research (Part B)*, 2010, 89:175–187.