How can the biosimilar concept be applied to more complex proteins such as monoclonals?
The European Union has led the way in establishing an approval process for what is termed “similar biological medicinal products” and popularly referred to as “biosimilars”.
Over the past quarter of a century, recombinant DNA technology has spawned an array of blockbuster medicines. These have advanced the frontiers of medical care in the treatment of diseases as diverse as anemia, cancer and rheumatoid arthritis to name but a few.

Now that many of these products are coming to the end of their patent protection period, the race is on to produce replica versions, which could prove to be crucial in managing the escalating cost of health care. But this is not as simple as it sounds, since it is not considered possible to produce identical biological products using manufacturing processes that will inevitably differ between comparator and originator process. Therefore, follow-on proteins cannot be approved as simple generics but will need additional nonclinical and clinical data to demonstrate they have equivalent safety and efficacy profiles to the originator product.

Experience with the approval of follow-on proteins varies across the world, but the European Union has led the way in establishing an approval process for what is termed “similar biological medicinal products” and popularly referred to as “biosimilars”. Within the EU the first Scientific Advice for Biosimilar development was given by CHMP as far back as 2003 and the number of scientific advice procedures has risen year on year since then.

The EU biosimilar marketing authorization (MA) approval process allows similarity to be demonstrated to the innovator product based on a combination of physico-chemical and biological data supplemented by limited relevant nonclinical and clinical data which is far less than would be required for a new stand-alone MA application.

The EU to date has approved a number of biosimilars based on this concept, including formulations of somatropin and epoetin and, most recently, filgrastim.

Right now a new question is looming: can the “biosimilar” concept be applied to follow-on monoclonals and other complex proteins? Certainly, there is already extensive knowledge of monoclonal antibodies, in terms of their therapeutic value, manufacture and characterization. In the EU, as of late 2009, there had been 25 novel monoclonal antibodies authorized by CHMP.

— Continued
While the development of more-complex biosimilar proteins may be more challenging, there is no compelling reason why most recombinant proteins should not qualify for marketing approval via a "biosimilar" route once the period of data exclusivity has elapsed. Modern production methods and analytical techniques have evolved to a level that now allow the production of highly similar monoclonals and other complex proteins.

The monoclonal market is certainly of great interest to the biosimilars industry. Global mAb sales reached $26 billion in 2007, and are forecast to almost double to $49 billion by 2013. Sales of Avastin, Herceptin, Humira, Remicade and Rituxan account for over half of total global mAb revenues. This market is expected to achieve good annual growth rates both by expansion of indications e.g. Herceptin has recently been approved for gastric cancer and by introduction of new products; there are about 25 products in Phase 3 development.

In the USA, a legal pathway for complex biosimilars has only just been established, but as yet there is no indication as to how this will be regulated and what supporting data might be required. Elsewhere, in Japan and throughout the globe, biosimilar regulatory pathways are emerging or are already established. In Europe, the biosimilar revolution is beginning to embrace monoclonals with the publication by the EMA of a Concept Paper titled Development of a Guideline on Similar Biological Medicinal Products Containing Monoclonal Antibodies. This concept paper acknowledges the increased complexity regarding the design of a biosimilar development program for monoclonals, which require careful consideration and exploration of further science-based approaches.

In the USA, a legal pathway for complex biosimilars has only just been established, but as yet there is no indication as to how this will be regulated and what supporting data might be required.
Monoclonals are structurally more complex

Undoubtedly monoclonals are far more complex than the first biosimilars to be approved in Europe, which were the human growth hormones – Omnitrope, approved in April 2006 and Valtropin, approved a month later. In fact human growth hormone is a relatively small (22,00 daltons) non-glycosylated protein comprising 191 amino acids and is hardly challenging to copy. However, since then more complex proteins have been approved, notably the epoetins towards the end of 2007. The interesting point here is, that while epoetin is not a large protein, it is heavily glycosylated with one third of its mass being contributed by sugar residues and it, consequently, displays a high degree of heterogeneity, making it not only more difficult to produce as a biosimilar but also more difficult to demonstrate similarity. Yet biosimilar epoetins have been shown to be comparable to their reference products to the satisfaction of the European Commission. Of note, is the fact that while the biosimilar manufacturers were unable to match exactly the glyco-profile of the innovator product, they were able to scientifically justify these differences and this has relevance to monoclonals, which are also glycosylated but to a lesser extent. The structure of a typical monoclonal and associated heterogeneity is illustrated in Figure 1.

Figure 1.
According to the current European regulatory view, any difference in amino acid sequence or primary structure means that a protein must be treated as a different entity and cannot be considered a Biosimilar.

Monoclonals, at 150,000 Daltons, are significantly larger than other biosimilars approved to date, so to conclusively demonstrate a primary structure identical to the reference product may prove to be challenging but it certainly is not insurmountable. Firstly monoclonals can easily be broken down into their four constituent peptides – the two heavy and two light chains; this reduces the sequencing burden. Secondly, new methods such as MS-MS provide powerful tools to characterize complex proteins down to a single atomic mass unit. These technologies combined with multiple peptide mapping, C and N terminal sequencing, amino acid analysis and sequencing of the recombinant gene and mRNA can provide solid proof for matching primary structures.

Techniques such as circular dichroism (CD), Fourier transform infra-red spectroscopy (FTIR) and micro-calorimetry, make it possible to differentiate between slight differences in folding.
As mentioned, monoclonals are glycosylated but while epoetin possesses four glycosylation sites, the number of glycan structures found in most IgG’s is limited to one asparagine (Asn) on each Fc region on the heavy chains. Specifically these are biantenary glyco-structures, designated G0, G1, and G2 with increasing presence of terminal galactose. This structure can strongly influence the binding capacity of the Fc region to specific receptors of the immune system e.g. Fc gamma IIIa found on natural killer cells and some macrophages and is involved in cytotoxicity, which contributes to the killing of tumor cells. A plethora of analytical techniques such as MS, weak anion exchange and glycoanalysis enable the glycoprofile to be well characterized. Using these techniques it is often possible to identify more than a dozen glycoforms. Additional heterogeneity comes from deamidation, oxidation and C terminal lysine (Lys) processing, which generate numerous charge variants detected by ion exchange or electrophoretic methods. Variants such as split forms can be detected by techniques that separate molecules on the basis of size such as SDS PAGE (sodium dodecyl sulphate polyacrylamaide gel electrophoresis), SE-HPLC (size-exclusion high performance liquid chromatography) and analytical ultracentrifugation; these techniques are also useful for detecting aggregates. Aggregates which may form during manufacture, on storage or during reconstitution are undesirable and often a concern with monoclonals because of the high concentrations needed to obtain a therapeutic effect and the natural tendency for immunoglobulins to aggregate. Appropriate formulation studies will need to be performed to find an optimal formulation that is stable with respect to formation of particulates at release and during storage; the sponsor should not rely just on copying the innovator formulation and in fact sometimes copying the innovator formulations is precluded by patent limitations.

In view of this heterogeneity of monoclonals, it is not simply a question of being able to demonstrate that one molecule is similar to another but that populations of molecules are similar. This is further confounded by the fact that each batch of reference product will differ from the last and, therefore, there is a need for the biosimilar manufacturer to understand the acceptable range of variability of the reference product, which cannot be done by studying just one batch. The key challenge is not so much the ability to detect differences but to understand the significance of the inevitable differences that will be detected and indeed many structural variations such as C-terminal lysine heterogeneity may have no impact on activity, while slight differences in glycosylation might have a profound effect.

According to the current European regulatory view, any difference in amino acid sequence or primary structure means that a protein must be treated as a different entity and cannot be considered a Biosimilar.
Detecting differences that have structure-activity significance

Biological characterization can play an important role in distinguishing changes which are inconsequential, from those that are not. Schneider and Kalinke write “in most cases, in vitro potency assays are available that allow estimation of biological function, enabling a comparison in vitro of reference and biosimilar mAbs at a functional level.”² With monoclonals, complexities exist in that multiple biological activities may be involved in contributing to the therapeutic effect.

This stems from the fact that antibodies not only bind to the target epitope but also interact with the cells of the immune system and bind complement to precipitate a range of effects, which may culminate in inflammatory responses and ultimately in cell destruction or cytotoxicity. This is illustrated in Figure 2, which outlines the actions of the anti-Her 2 monoclonal, trastuzumab.

The CHMP advises applicants to employ a wide range of in vitro and in vivo studies, including receptor-binding studies and cell-based assays.

Figure 2
In the case of monoclonals, these tests will need to include binding to the target epitope, the Fc receptors from cells of the immune system and to complement, using techniques such as ELISA and surface plasmon resonance and application of functional assays such as antibody dependent cytotoxicity.

Structural changes such as differences in glyco-profile can impact the activity of a monoclonal profoundly e.g. the loss of the fucose sugar near the base of the glyco-structure on the heavy chain, significantly enhances antibody dependent cellular cytotoxicity (ADCC). In fact as illustrated in Figure 3, it has been demonstrated that a rise of only a few percent in defucosylated form, leads to a many fold increase of ADCC activity. As another example, the presence of mannose may increase complement dependent cytotoxicity (CDC) activity (see Figure 4).

In order to more fully understand the interrelationships between target receptor binding and ADCC and CDC activity, cellular and in vivo assays reflective of the therapeutic action of monoclonals may have additional value. However, there can remain many hurdles to the full utilization of biological testing methods including the availability of binding proteins, receptors, suitable cell lines and appropriate models. It is also important to understand the variability associated with the different assays in order to ensure that interpretable results can be generated.
The value of nonclinical data

The value of the physico-chemical and biological testing program in providing a high level of assurance of similarity to the originator product cannot be underestimated. Nevertheless, no matter how compelling this evidence, patient safety for use both in the clinical trial and post marketing phases will need to be assured. The next step in demonstrating biosimilarity requires leveraging current knowledge and filling the gaps by conducting appropriate nonclinical and subsequently clinical studies. Any observed toxicity needs to be judged in relation to immunogenicity, pharmacokinetics and pharmacodynamic effects and therefore these need to be studied in parallel and preferably in the same study.

Monoclonals present special challenges in the design of appropriate and relevant nonclinical studies. The CHMP nonclinical and clinical comparability guideline stresses the need for in vivo nonclinical studies. However, such studies may prove of little value or even not feasible for some monoclonals because of immunogenicity and/or species specificity. In reality, for many biologicals, a comparative study of longer than four to twelve weeks will not be practical, primarily due to immunogenicity and this is recognized by the published product specific guidelines, which generally specify the need for just four weeks data in one relevant species, which could be the rat. For compounds with a prolonged pharmacological effect such as monoclonals, there could be an argument for conducting studies of longer duration e.g. three months but for reasons discussed below, the value of such studies can be questioned.

The above mentioned CHMP guideline requires that the nonclinical study needs to demonstrate similarity and will need to be appropriately sized for that purpose. However generally, monoclonals are only active in higher species, meaning that the monkey is often the only suitable species if indeed a suitable species exists. This will necessitate small group sizes and thus generally the nonclinical studies cannot be designed or powered to demonstrate “similarity.” The best that can be achieved is a qualitative analysis and probably the most meaningful comparison will be with respect to the pharmacokinetic and toxicokinetic data, which can perhaps be done in lower species anyway. For example, Kanda et. al. have demonstrated that the impact of mannose 5 glycoform on pharmacokinetics is apparent from studies in mice (Figure 5), nevertheless it is likely that studies in mice and even primates will not detect all possible effects on human PK since, for example, differences in Fc receptors will exist and Fc receptor binding, particularly FcRn can play a role in clearance.
Generally the primary value of the nonclinical data will be to provide initial assurance that the biosimilar under development is safe to enter clinical trials. Even this will be difficult to demonstrate where no suitable species for toxicity testing exists at all. One point for debate is whether there is value in studying off-target effects in a non-relevant species and of the value of transgenic models, which are often not validated and their relevance questionable. Conventionally the expectation is that clinical trials of new compounds must be supported by nonclinical data but biosimilars are not new entities and in this respect the value of undertaking nonclinical studies can be questioned. Here, the Concept Paper reflects current thinking and states that “experience gained with numerous mAbs over the past decade(s) will have to be considered, together with a differential discussion on the toxicity that should be in the focus (specific toxicity, based on the mechanism of action; or unspecific toxicity, based on impurities etc) and to what extent nonclinical pharmacodynamic studies can be done or are even needed in view of the clinical data to be gathered.”

What is clear, is that there is the need to maximize whatever data can be gleaned from using in vitro based techniques and a minimum of animals. However regardless of the above debate, local tolerance studies are likely to be required, as these may be influenced by changes in the impurity profile and formulation.

Clearly as for other Biosimilars, studies such as mutagenicity studies, animal reproduction and carcinogenicity studies will not be required, the latter two categories are, in any case, often impractical and have generally not been generated for the originator products.
Changes in protein structure may impact their clinical effect in a number of ways. For example, changed glycosylation patterns may affect bioactivity such as ADCC (defucosylation), CDC (high mannose), pharmacokinetics (PK) (high mannose) (see Figures 3 - 5), and immunogenicity (1,3-á-gal sugars and sialation). Changes in the conformation of the protein may impact on receptor affinity and expose novel epitopes altering the immunogenicity. Despite the expansion in understanding of the structural activity relationships, there remains the concern that structural differences that might impact safety and efficacy may not be detected by physico-chemical and biological studies alone, so that some clinical trial data are considered mandatory.

Since clinical programs are associated with significant costs and, moreover, considerable ethical constraints, there is a compelling need to optimize the clinical program so as to limit trial sizes. Thus an efficient clinical program needs to be formulated and justified to the regulatory authorities. It is also critical that these studies are carried out on the
manufacturing scale batches and formulation intended for marketing. Furthermore, as things stand, in order to meet EU requirements, the reference product will need to be sourced from the European Economic Area. The use of EU sourced reference product should allow approval in most other regions although the future stance of the United States Food and Drug Agency (FDA) on this point is not clear at this stage.

As for any clinical development program, monoclonal development will need to commence with a Phase 1 study, which the CHMP guidelines require to be completed before advancing to confirmatory efficacy and safety studies. Generally these Phase 1 studies are performed in healthy volunteers, but with monoclonals this may not always be ethical or feasible because of the intense pharmacodynamic effect. Furthermore, single dose studies in patients might not be feasible either, as often monoclonals provide critical and sometimes life-saving therapy and a course of therapy cannot ethically be interrupted. Since monoclonals generally have half-lives of several weeks, single dose PK comparison of treatments would require a separation of dosing of many months, which in most cases would simply not be in the interest of the patient. This means that it may only be possible to study comparative PK in patients at steady state as part of the Phase 3 program.

Pharmacodynamic (PD) studies with monoclonals are also often challenging as there may be no appropriate pharmacodynamic marker and where a marker exists this may not be well correlated with therapeutic effect or dose response e.g. the fall in CD20 associated with rituximab therapy has not been correlated with therapeutic effect or dose. Thus the value of the standard Phase 1 studies in the development of biosimilar monoclonals can be limited. Certainly the standard Phase 1 approach advocated by the EU CHMP Guideline on Similar Biological Medicinal Products Containing Biotechnology-Derived Proteins as Active Substance: Nonclinical and Clinical Issues will often not be appropriate or feasible and further consideration as to the nature of such early phase studies is required. Nevertheless the expectation will be that some data are available before commencement of Phase 3 in order to mitigate against any risk to trial subjects.

Changes in the conformation of the protein may impact on receptor affinity and expose novel epitopes altering the immunogenicity.
If a monoclonal product demonstrates similar physico-chemical and biological properties and displays similar *in vitro* and *in vivo* effects to the reference product, there already exists a strong indication for a similar efficacy profile. Furthermore for a biosimilar, the relationship between PD, dose and therapeutic response is better understood than for a novel biological entity. The clinical comparability program is, therefore, confirmatory not exploratory and the magnitude and complexity of the plan should be reduced by an increasing level of understanding of the structure, impurity profile, and biological, nonclinical and clinical properties, together with the interaction of these factors and with an understanding of their relevance to the safety and efficacy of the biosimilar. This is a very important concept because it is not a question of demonstrating efficacy all over again but understanding the science and providing just sufficient evidence to demonstrate expected therapeutic equivalence. Thus a limited program ought to be acceptable although in effect the test population needs be representative of the complete target population and sensitive to differences that might impact the efficacy, safety and immunogenicity profile.

Monoclonals are often approved for multiple indications. The current CHMP nonclinical and clinical biosimilar guideline, requires that “efficacy and safety has to be justified, or if necessary, demonstrated separately for each of the claimed indications.”7 Under the current product specific guidelines for somatropin, epoetin and G-CSF all the indications approved for the reference product can be gained based on studies in a single or just some indications.8, 9, 10, 11 For monoclonals, there are no guidelines, but extrapolation from one indication to others will need to be based on the totality of the data and an understanding of the mechanism of action. As already discussed monoclonals are multifunctional and their therapeutic effect may depend on both blocking the target receptor, and/or interfering with downstream signaling and on eliciting an immune assault on pathogenic tissue. The balance of these effects could theoretically differ from one indication to another as might the ability to penetrate different target tissues. For example the anti-tumor necrosis factor (TNF) fusion protein etanercept is effective in joint diseases such as rheumatoid arthritis (RA), but not gastrointestinal conditions such as Crohn’s Disease, whereas the anti-TNF monoclonals have been shown to be effective in all these indications. However etanercept and monoclonals

**Confirmatory efficacy trials are challenging**
are fundamentally different molecules and if two molecules are demonstrated to be similar using a range of orthogonal and state of the art physico-chemical and biological tests and furthermore display similar nonclinical, PK and PD properties, then it seems reasonable to assume that if equivalent efficacy and safety is demonstrated in one adequately sensitive population, a similar result would be expected in other populations. In fact this approach is fundamental to facilitating the development of biosimilar monoclonals. In this respect Schneider and Kalinke write “one imagines that other indications could be extrapolated by adequate justification, or bridged by short-term efficacy studies using surrogate clinical endpoints, especially when for other indications endpoints exist only where differences are not detected with sufficient sensitivity. (Schneider and Kalinke, Nature Biotech 2008). Such scenarios are, however, without precedent and would thus require discussion with regulators.”

One also needs to bear in mind that the clinical trials on the original biological entity were generally designed to demonstrate efficacy against placebo. However, placebo trials will not be ethical, at least for serious and life threatening diseases, once an effective therapy has been established. This necessitates a non-inferiority or equivalence trial, which are more challenging and will require trial sizes far larger than the original placebo controlled trials.
The fact is, that it is statistically impossible to show that two products are 100% equivalent; there has to be a margin of acceptable difference which needs to be defined prospectively. Defining this margin is not easy – ICH E9\(^3\) states that “this margin is the largest difference that can be judged as being clinically acceptable and should be smaller than differences observed in superiority trials of the active comparator.” Clinical acceptability is very much an individual judgment. However, it is clear that the margin should exclude products which are evidently not significantly better than the placebo. What this means is that if only a small effect over placebo was determined for the reference product, which indeed is the case for some monoclonals, even a study of thousands of patients might not be adequate to satisfactorily demonstrate equivalence. Thus the whole approach to developing a biosimilar by necessity must differ from the originator program.

Choice of end point represents another area for controversy, particularly in the oncology arena where overall survival has been considered the gold standard. However overall survival is often not a suitable endpoint for equivalence trials as a high number of events would be required in order to fall within the bounds of an acceptable equivalence margin and that may take many, many years to reach. Secondly it may not even be possible to demonstrate survival as new next line therapies may have been developed to which patients will be transferred at the time of progression, this is even a problem in the development of new entities. Therefore, progression free survival (PFS) is often the best end-point that can be achieved and that is recognized by the CHMP guidelines, which state: “in situations where there is a large effect on PFS, a long expected survival after progression, or a clearly favourable safety profile, precise estimates of OS may not be needed for approval.”\(^1\)
Unfortunately even using PFS, it can take many years to demonstrate equivalence. Another more practicable option would be to use overall response rate as the primary end point and to include progression free survival and, where possible, overall survival as secondary endpoints. Granted this does not accord with the CHMP Guideline. However, the guideline is intended for novel oncology products. In the case of biosimilars, the goal of this clinical program is not to prove efficacy, but to show comparability between two products. This approach is supported by literature and regulatory precedent exists for the use of overall response rate as a primary endpoint in equivalence trials of this nature. At the EMEA/CHMP Biomarkers Workshop held on 16 December 2005, Dr Francesco Pignatti, MD of the European Medicines Agency (EMEA) London presented slides reporting that “response rate” had been used as a primary endpoint in Phase 3 (presumably confirmatory studies) in 4 out of a total of 27 studies examined. All four of these studies related to investigating variants of existing products. The fact is that demonstrating therapeutic equivalence for monoclonals will be challenging and companies will need to consider their options carefully. There is the need for novel approaches and novel thinking by both the biosimilar industry and regulators.

Safety
The CHMP guideline states that “even if the efficacy is shown to be comparable, the similar biological medicinal product may exhibit a difference in the safety profile (in terms of nature, seriousness, or incidence of adverse reactions). Prelicensing safety data should be obtained in a number of patients sufficient to address the adverse effect profiles of the test and the reference medicinal product.”

Safety is best demonstrated in the most sensitive populations, for example patients most prone to develop an immune response and/or patient groups receiving the highest doses and this population may prove to be different to the most suitable efficacy population. The nature and extent of safety data will differ depending on the profile for the originator product. However, generally six month’s data from a minimum of 300 subjects would meet current ICH requirements for chronic therapies and at least this number of patients will in any case likely be required to demonstrate equivalence with respect to efficacy.

Nevertheless, in theory, serious adverse reactions may only emerge after extensive exposure and usage such as PML reported to potentially be associated with anti-inflammatory monoclonals. Thus there is a requirement to follow safety post-marketing; and post-marketing commitments will need to be made, which might include the need for pharmacovigilance registries.
There is the need to apply state of the art methodology, novel approaches and novel thinking by both the biosimilar industry and the regulators for biosimilar monoclonals to indeed become a reality.

**Immunogenicity**

The potential for immunogenicity is highly influenced by the patient population and is generally less of an issue in immunocompromised patients. The route of administration will also affect immunogenicity with the subcutaneous route being reported to be associated with the greatest immunogenicity.

The immunogenicity monitoring for biosimilars is no different than for novel biological entities, which is addressed in a draft CHMP immunogenicity guideline. In general safety studies should include comparative testing against reference product, which should be performed at screening and at appropriate intervals thereafter until the end of the study. It is generally not possible to compare results with historical or literature data as results are highly dependent on the assay used.

Neutralizing antibodies are of particular concern for replacement therapies as these can be associated with reduced clinical efficacy or in the case of epoetin and factor VIII for instance, autoimmunity, although since monoclonals are in effect foreign rather than copies of endogenous proteins, autoimmunity has not proven to be a problem to date. Nevertheless, since monoclonals have long half lives binding antibodies may also reduce efficacy by rapidly clearing the monoclonal from the circulation.

**Conclusion**

Producing biosimilar monoclonals is certainly feasible. However the fact is that demonstrating physico-chemical, biological and therapeutic equivalence for monoclonals will be challenging and companies will need to consider their options carefully. There is the need to apply state of the art methodology, novel approaches and novel thinking by both the biosimilar industry and the regulators for biosimilar monoclonals to indeed become a reality.
References


3) ICH E9, Statistical Principles for Clinical Trials


7) CHMP Guideline on Similar Biological Medicinal Products containing Biotechnology-Derived Proteins as Active Substance: Nonclinical and Clinical Issues (CHMP/42832/05) effective June 2006


11) Annex to Guideline on Similar Biological Medicinal Products Containing Biotechnology-derived Proteins as Drug Substance – Nonclinical and Clinical Issues containing Recombinant Human Insulin, February 2006, CPMP/32775/05)

12) Guideline On The Evaluation Of Anticancer Medicinal Products In Man (CPMP/EWP/205/95/rev.3/corr.2)


To Learn More About PAREXEL Consulting
Please Contact:

Paul Bridges
Vice President
Global Product Development
PAREXEL Consulting
paul.bridges@parexel.com

Ron Kraus
Vice President
Global Customer Strategy
PAREXEL Consulting
ron.kraus@parexel.com