Look out: here come the proteins

Glenn W. Dale, Morphochem AG, Wro-1055, Schwarzwaldale 215, 4058 Basel, Switzerland. tel: +41 61 697 5064, fax: +41 61 697 5066, e-mail: glenn.dale@morphochem.ch

Since the regulatory approval of recombinant insulin in 1982, an increasing number of engineered proteins have entered the market and have been successful in both clinical and financial terms. A recent survey of the pharmaceutical/biotechnology sector estimated that over the past ten years, more than 100 drugs from the biotechnology industry have been approved. Newcomers to the world of protein pharmaceuticals, attending IIR's recent *Engineering Protein Pharmaceuticals* conference in London, UK, were brought up-to-date with the rapid advancements in this fast-growing, billion-dollar industry. The speakers provided exciting insights into the various aspects of the process of identifying drug candidates as well as engineering the protein to improve its efficacy, stability and pharmacokinetic properties.

Many early attempts at introducing protein drugs derived from biotechnology failed for various reasons. In most cases, the primary molecules had either sub-optimal affinity or a poor half-life *in vivo* leading to poor efficacy. In other cases, the molecules were derived from non-human sources and therefore stimulated immune responses to the drug. This is particularly true for monoclonal antibodies (mABs) raised in rodents, as a major limitation in their use clinically is the anti-globulin response they invoke. Moreover, as the vast majority of proteins entering the clinic are recombinant, post-translational modifications (particularly glycosylation) often result in a structurally heterogeneous product that can effect the affinity, stability, immunogenicity and clearance rate *in vivo*. A large part of the conference was dedicated to different methods employed to overcome many of these barriers.

As the discovery of new genes has increased dramatically, high-throughput methods for the validation of potential drug candidates has become a key area of interest whereby protein expression and purification remain the major focus in this process. Charlie Birse (Human Genome Sciences; HGS, Rockville, MD, USA) initiated this conference with a discussion on how HGS uses functional genomics and proteomics in the preclinical development process. There was a clear consensus among the delegates that protein engineering is presently the major bottleneck.

Breakthrough in antibody technology

The design and development of Herceptin, a humanized version of the anti-HER2 antibody for the treatment of breast cancer was outlined by Paul Carter (Immunex Corporation, Seattle, WA, USA). In this case, a reduction in immunogenicity was accomplished by grafting the complementary-determining regions (CDR), which are the hypervariable loops involved in antigen binding, from mouse antibodies onto human V domains. This approach has already been proven to be successful with at least two similar therapeutic products already on the market (Zenapax developed by Protein Design Labs and Synagis developed by MedImmune) and more than 30 others currently in clinical trials. Carter also discussed two promising strategies for mAB therapy: the antibody–small-molecule toxin conjugates (mAB–maytansinoid conjugates from Immunogen and mAB–calicheamicin conjugates from Wyeth Ayerst/CellTech) and the bispecific mAB (anti-CD3 × anti-HER2). The identification of suitable targets is now the rate-limiting step in mAB therapy.

Kevin Johnson (Cambridge Antibody Technology; CAT, Cambridge, UK) described the evolution of the therapeutic mABs from mouse (1970s) to chimeric (1980s) to grafted CDR (1990s) to the future as fully human mABs (2000s). As a result of the large size of the CAT phage antibody library (1×10^{11}), their methodology regularly provides mABs with affinities for the target in the sub-nanomolar range, a necessity for the target validation process. Antibodies screened in functional assays using technology originating in HTS laboratories readily identify suitable clones. Using the ProAB® technology (CAT), it is foreseen that mABs will be used not only to validate targets and identify indications but will themselves become potential leads.

Roy Jeffries (University of Birmingham, Birmingham, UK) addressed the importance of post-translational modifications and described the glycosylation of human IgG and its relevance to therapeutic applications. His talk elegantly summarized the experiments undertaken by his laboratory to provide evidence that the heterogeneity in the glycosylation of IgG can effect both the stability and biological activity of an antibody. These experiments would indicate that it is possible to selectively generate monoclonal IgG antibodies that are optimally glycosylated for specific therapeutic applications. John Lund (University of Birmingham) also described his work using protein engineering to influence the glycosylation of IgG. The importance of the various residues in the C_H2–C_H3 domain interface was also presented, as was convincing evidence that the oligosaccharide profile can be manipulated by directed mutagenesis.

1359-6446/01/$ – see front matter ©2001 Elsevier Science Ltd. All rights reserved. PII: S1359-6446(01)01758-5

Optimizing stability, bioactivity and bioavailability

David James (University of Kent, Canterbury, UK) began his seminar by describing the various types of glycosylation and the importance of the in herent heterogeneity for the biophar maceutical industry. As the regulatory authorities are now demanding an increasingly sophisticated carbohydrate analysis as part of process validation, there is a requirement for advanced analytical techniques. The various ways of controlling the glycosylation of recombinant proteins were discussed stressing that the most realistic goal would be to achieve a consistent, reproducible, 'human-like' glycosylation pattern. The most crucial factor for the achievement of this goal is the choice of expression system, a view re-emphasized by Nigel Jenkins (Eli Lilly & Co, Indianapolis, IN, USA). Additional strategies include controlling the bioprocess environment, engineering the host cell, mutating the protein, or even *in vitro* remodeling of the carbohydrates.

The limitations of mammalian cell culture systems to secrete proteins that are optimally glycosylated was discussed by David Zopf (Neose Technologies, Horsham, PA, USA). GlycoAdvance™, an enzymatic carbohydrate remodeling technology that uses isolated glycosyltransferases to complete carbohydrate chains *in vitro*, was presented. The effectiveness of this technology was supported by examples of its successful application. Zopf also stressed the opportunities for protein production engineers to concentrate on optimizing secretion rather than the degree and heterogeneity of glycosylation.

The issue of protein stability and *in vivo* half-life was addressed by Homa Sadeghi (Principia Pharmaceuticals Corporation; PPC, Norristown, PA, USA; a subsidiary of HGS). Sadeghi pointed out that most therapeutic proteins and peptides have relatively short *in vivo* half-lives in the circulation, which is generally dependent on the size of the protein: small peptides of <3 kDa, seconds to minutes; small proteins of 16–35 kDa, minutes to hours; and plasma proteins of >50 kDa, weeks. To increase these half-lives, PPC genetically fuse proteins or peptides to human serum albumin (HAS) and produce the recombinant proteins in yeast. The data for HAS–IFN and HAS–hGH indicated *in vitro* activity, but these results were contradicted by a much higher activity *in vivo*, which was attributed to the increased stability. It was suggested that the HAS fusion system has many advantages such as compliance (fewer injections), improved efficacy and an increased stability, enabling a convenient liquid formulation and packaging. PPC, it appears, are in the process of creating and testing several other HAS fusion proteins.

The use of three-dimensional structural information of target proteins in the lead identification and drug discovery process has seen a renaissance in recent years. A major stumbling block in the past has been the production of suitable protein crystals for structure determination by X-ray crystallography. Glenn Dale (Morphochem AG, Basel, Switzerland) described how protein engineering has been used to solve some of the major problems through the generation of mutant proteins to improve the crystallization properties. Dale described a new method in which the detailed structural knowledge of the active site of the enzyme was used to develop a pharmacophore hypothesis. Together with mutagenesis and biophysical methods (surface plasmon resonance, analytical ultra-centrifugation), this was used to rapidly identify leads from a small directed library.

Insulin delivery

John Patton (Inhale Therapeutic Systems, San Carlos, CA, USA) presented Inhale's approach to providing a less invasive delivery of insulin. The inhalation of insulin for meal-time glucose control was found to be safe, efficacious and reliable in Type I and Type II diabetics. The relative efficiency of insulin delivery by aerosol (estimated from the dose measured at the exit point of the aerosol device) compared with subcutaneous injection was found to range between 8 and 25% of levels reached by the subcutaneous route. To date, they have tested over 1000 patients with only one incidence of an adverse side effect reported. For those of us with family or friends who are diabetics, this technology certainly has the potential to provide them with a significant increase in quality of life.

Improved expression systems

Martin Fussenegger (ETH, Zurich, Switzerland) began the session with a discussion on metabolically engineering cells to increase protein production. Optimizing the vector for the system of choice was emphasized. Dee Athwal (Celltech Chiroscience, Slough, Berkshire, UK) described the production of therapeutic antibody fragments in *Escherichia coli*. The microbial expression system has the advantages of reproducibility, scalability and the low cost of goods. Finally, he described how pegylation of proteins can be used to increase the *in vivo* half-life.

The Semliki Forest Virus (SFV) expression system for recombinant protein production in mammalian cells was introduced by Kenneth Lundstrom (F. Hoffmann-La-Roche, Basel, Switzerland). The advantages of the SFV system include rapid generation of high-titer SFV particles, high expression levels, scale-up to produce hundreds of milligrams of recombinant protein, but perhaps most importantly is its broad host range. The system is ideally suited to rapidly screening various cell lines to define the optimal host system. Lundstrom also provided various examples where the SFV system was employed to produce the desired protein. Ian Cottingham (PPL Therapeutics, Edinburgh, UK) discussed the expression of recombinant proteins in the milk of transgenic animals, and

suggested that the advantages of the system are high expression levels, linear scale-up and the low cost of goods. Additionally, the protein is produced in a purification-friendly environment and the post-translational modifications are human-like. PPL has used the technology to produce a number of proteins, one of which (α1 antitrypsin) is in Phase III clinical trials.

Concluding remarks

The take-home message for the industry outsider was that the lessons learned from the mistakes made during the initial attempts to bring protein therapeutics to the market are now being addressed at a very early stage in the discovery process. Parallels can be drawn with the experience gained by the small-molecule therapeutic discovery process in that it is now recognized that multidimensional optimization of the compounds throughout the discovery process will lead to decreased costs and higher success rates in the clinic. The industry as a whole has laid the foundations for financial success but, most importantly, it is poised to provide many new medicines to treat the various medical needs.

Biocomputing: impact of the genomic revolution

Rebecca Lawrence, News & Features Editor

The annual meeting of the *Pacific Symposium on Biocomputing*, held in Mauna Lani (HI, USA) was greatly oversubscribed this year due to the growing importance of computational biology and bioinformatics in the post-genomic era. The meeting covered a wide range of aspects, from DNA and protein structure, protein–DNA interactions and expression, protein evolution, and human genome variation to phylogenetics, high-performance computing, natural language processing and bioethics. This report will provide an overview of the main talks relating to drug discovery. The full papers presented at the meeting are available in the symposium book[1].

The keynote speaker, David Haussler, provided an overview of the effort to produce the working draft of the human genome sequence. He admitted there will be some cross-contamination present in the sequence as all the data that passes the NCBI filters is used in the final assembly. He also pointed out that the bioinformatics tools used are not perfect and miss-alignments have created some artefactual duplication. He reiterated the plan to complete the sequence by 2003, and that the next steps will be to identify the complete set of human genes, explore gene regulation, research human, mammalian and vertebrate diversity, and connect genomic data to clinical data.

Comparing sequences

A range of different computational approaches for sequence comparison was discussed. William Martins (Department of Electrical and Computer Engineering, University of Delaware, Newark, DE, USA) and colleagues have developed a dynamic programming algorithm that uses a rigorous mathematical approach based on a multi-threading system. This system can be scaled up and actually performs better with longer sequences, providing the capability to compare whole genomes. A further advantage of this method is that it enables the visual display of the matching sections of the genomes.

A new algorithm based on Monte Carlo multiple sequence alignment techniques was used by C. Guda (San Diego Supercomputer Center, University of California San Diego, La Jolla, CA, USA) and colleagues. Four different types of moves were designed to generate random changes in the alignment – shifting (shifts residues in one randomly chosen structure); expanding (expands the alignment block by acquiring one residue); shrinking (shrinks the alignment block by removing one residue); and splitting-and-shrinking (splits longer blocks in two and then shrinks one of the new blocks). Changes in distance-based scores were examined for each trial move and changes were only accepted if the alignment score increased. This process was repeated until the alignment scores converged. Each step only examines changes of one residue at a time, but many were concerned that more global changes might be missed that could have significantly improved the score.

The use of Kestrel single-instruction multiple-data (SIMD) parallel processor methods for sequence analysis was

1359-6446/01/$ – see front matter ©2001 Elsevier Science Ltd. All rights reserved. PII: S1359-6446(01)01737-8