Agency for Technology (TEKES), The Instrumentarium Foundation and The Small Business Center Foundation of Finland (to Outi Nieminen) are gratefully acknowledged.

References

1 Chignon, T. (2003) Regulation of human tissue/cell products in Europe. *Keynote Presentation, Regulatory Affairs Professional Services Discussion Forum: development and regulation of cell-based and related therapies* (20–21 March 2003, London, UK) (unpublished)
2 Warner, J. (2003) The status of regulation of cell therapy products by the US FDA. *Keynote Presentation, Regulatory Affairs Professional Services Discussion Forum: development and regulation of cell-based and related therapies* (20–21 March 2003, London, UK) (unpublished)
3 Gee, A.P. (2002) The impact of regulatory policy on the development of somatic cell therapies. *Transpl. Immunol.* 9, 295–300
4 Hirsjarvi, S. and Hurme, H. (2000) *Tutkimushaastattelu – Teemahaastattelun teoria ja käytäntö* (1st edn), Yliopistopaino
5 Persidis, A. (1999) Xenotransplantation. *Nat. Biotechnol.* 17, 205–206
6 Dove, A. (2002) Cell-based therapies go live. *Nat. Biotechnol.* 20, 339–343

The *Discussion Forum* provides a medium for airing your views on any issues related to the pharmaceutical industry and obtaining feedback and discussion on these views from others in the field. You can discuss issues that get you hot under the collar, practical problems at the bench, recently published literature, or just something bizarre or humorous that you wish to share. Publication of letters in this section is subject to editorial discretion and company-promotional letters will be rejected immediately. Furthermore, the views provided are those of the authors and are not intended to represent the views of the companies they work for. Moreover, these views do not reflect those of Elsevier, *Drug Discovery Today* or its editorial team. Please submit all letters to Steve Carney, Editor, *Drug Discovery Today*, e-mail: S.Carney@elsevier.com

The analysis of tandem mass spectrometric datasets: high-throughput investigations require high-quality validation

As the various technologies involved in proteomic analyses have matured to become capable of rapidly generating high-quality data, recent studies have focussed on high-throughput, automated applications of these technologies. However, the quality of the output data must be considered. Given this, the timely review by Nesvizhskii and Aebersold [1] in a recent issue of *Drug Discovery Today* summarizes many of the key requirements for the statistical analysis and validation of the results of large-scale proteomic experiments [1].

One of the primary factors that requires consideration is the amount of information that is needed from the generated dataset. Studies carried out within our laboratory have indicated that extremely limited sequence data are sufficient for the identification of the gene from which peptide ion-fragmentation data derives, with a high degree of confidence in many cases (as in the so-called 'just-enough diagnostic information' approach [2]). Indeed, this approach has the advantage that the ion current of tandem mass spectrometry (MS–MS) spectra can be concentrated on a smaller number of fragment ions, thus increasing their abundance and, in cases where high-quality sequence databases are available, significantly increasing the confidence of the matched sequences identified by database searching.

Although decreasing the rate of false-positive peptide identification in database searching is clearly desirable, there is also a strong case to be made for allowing more minor sequence variations. This could enable peptides with high-interspecies identity to be matched to orthologues and errors in sequence databases and post-translational modifications to be more readily incorporated into searches.

Another requirement that many high-throughput proteomic strategies overlook is the biological significance of the data. In a typical MS–MS-based approach, peptides are selected for MS–MS based on their relative abundance. Clearly, the nature of many of the peptides selected using these criteria will be the products of housekeeping genes (i.e. proteins that are essential for cell function, but the levels of which fluctuate insignificantly). However, in the analysis of biological processes such genes are rarely informative. Hence, the preselection of ions having relative intensities that vary significantly between cells under different conditions would clearly give a more biologically relevant dataset. This type of strategy exploits the developed stable isotope methods, although inter- and intra-experimental variability require further investigation [3].

For proteomic data to be shared and compared between different laboratories, standardized criteria for the

  PII: S1359-6446(04)03030-2

validation and evaluation of proteomic data must be established. PEDRo is one such schema that proposes to standardize the format of archiving and management of data. PEDRo is currently being implemented into the initial high-throughput projects of the Human Proteome Organisation (http://www.hupo.org) [4].

A final point is a more philosophical one; the concept of 'protein identification' is somewhat ill defined. In a typical 2D-gel approach, multiple distinct spots are frequently identified as being derived from the same gene, but they clearly differ at some level. In multidimensional protein identification technology (MudPIT)-based approaches, such information is lost and protein identification is frequently based on MS–MS of single peptides [3]. Such peptide-based identification does not indicate the number or heterogeneity of isoforms present and, therefore, this data must be interpreted with caution. Clearly, high-throughput strategies must always be supported by thorough bioinformatic and biochemical validation.

References

1 Nesvizhskii, A.I. and Aebersold, R. (2004) Analysis, statistical validation and dissemination of large-scale proteomics datasets generated by tandem mass spectrometry. *Drug Discov. Today* 9, 173–181

2 Riba-Garcia, I. and Gaskell, S.J. (2003) Peptide derivatisation for routine improvement in identification of proteins from 1D and 2D gels. *American Society for Mass Spectrometry Annual Conference* (8–12 June 2003; Montreal, Canada) (http://www.inmerge.com/asms03pdf/A031678.pdf)

3 Washburn, M.P. *et al.* (2003) Reproducibility of quantitative proteomic analyses of complex biological mixtures by multidimensional protein identification technology. *Anal. Chem.* 75, 5054–5061

4 Taylor, C.F. *et al.* (2003) A systematic approach to modeling, capturing, and disseminating proteomics experimental data. *Nat. Biotechnol.* 21, 247–254

Sarah R. Hart *and* ***Isabel Riba-Garcia****

Michael Barber Centre
UMIST
Manchester, UK M60 1DQ
**e-mail: isabel.riba-garcia@umist.ac.uk*

Log(BB), PS products and *in silico* models of drug brain penetration

In silico models of drug brain penetration attempt to predict blood–brain barrier (BBB) permeability on the basis of parameters such as hydrogen bonding, lipid solubility and molecular weight. The majority of *in silico* models use log(BB) as the index of BBB permeability, where BB is equal to the brain:blood drug concentration ratio at some defined time point, such as 60 minutes after drug administration. In a recent issue of *Drug Discovery Today*, Martin [1] questioned the validity of the log(BB) parameter and suggested that it is time that log(BB) is no longer used to predict BBB penetration. The BB is a volume of distribution that is determined largely by cytoplasmic binding of drugs in brain and much less by BBB permeability. Two different drugs could have comparable log(BB) values yet differ in BBB permeability by tenfold. Any pharmaceutical company that relies heavily on log(BB) values is at risk of following poor drug leads and missing promising drug leads. If log(BB) values are not to be used, then what parameter should be used? Alternative parameters of BBB permeability include intra-cerebral micro-dialysis fibers, drug concentrations in cerebrospinal fluid (CSF) and *in vitro* BBB models that measure the BBB permeability–surface area (PS) product. However, it can be shown that none of these parameters offers any advantages over log(BB). The one methodology that does give reliable measures of BBB permeability is the *in vivo* quantification of the BBB PS product.

Drug action in the brain is a function of drug receptor occupancy. This receptor saturation by drug is a function of the free drug in brain cells and not the total drug concentration, which is reflected by log(BB). Total drug is the sum of free drug and the drug that is bound to cytoplasmic proteins in brain. Most drugs are avidly bound in the brain, such that the total drug concentration in brain is many-fold greater than the free drug concentration in brain. Unfortunately, it is not so easy to measure free drug in brain experimentally. Intra-cerebral dialysis fibers have been used, but these measure free drug under conditions of brain injury in which BBB permeability is artifactually increased [2]. An alternative to the measurement of free drug in brain is the determination of drug levels in CSF [1]. However, measurement of drug distribution into CSF is not a measure of BBB permeability. Drug distribution into CSF is a function of drug transport across the choroid plexus, which forms the blood–CSF barrier. By contrast, free drug in brain is a function of drug transport across the brain capillary endothelium, which forms the BBB. These two membrane systems in brain, the BBB and the blood–CSF barrier, can have very different permeability profiles [3]. For example, azidothymidine (AZT) readily distributes into CSF, but does not cross the BBB because of the selective expression of an AZT active efflux transporter at the BBB, which is not expressed at the choroid plexus. As a result of the rapid rate of CSF bulk flow through the brain and the slow rate of drug diffusion within the brain, the CSF and brain interstitial fluid (ISF) occupy functionally separate compartments [3]. It should not be assumed that the drug level in CSF is equal to the free drug in brain ISF, as illustrated by the case of AZT.

The best index of BBB permeability is the BBB PS product, which has units of $\mu l\ min^{-1}\ g^{-1}$ and is a measure of unidirectional clearance from blood to brain across the BBB. The BBB PS product predicts the level of free drug in brain, because the level of free drug is determined by: (i) the total drug concentration in plasma; (ii) the PS

1359-6446/04/$ – see front matter PII: S1359-6446(04)03065-X