

Although the work of Giarratana *et al.* is a significant milestone in cell engineering, its practical implications should be put into context. Last year, 80 million units of blood were collected worldwide for transfusions. With the new protocol, one unit of cord blood containing 2–5 million CD34⁺ cells can generate 4–10 trillion mature red blood cells. Since each unit of blood used for transfusion contains some 2.5 trillion red blood cells, the protocol can, under optimal conditions, yield between 2 and 4 units of transfusable red blood cells from one unit of cord blood, the richest source of CD34⁺ cells currently available.

However, the complex engineering needed to implement the three-step protocol, including the use of multiple growth factors and stromal cells, would make each unit of blood produced prohibitively expensive. Thus, the practicality of the method for producing red blood cells on a scale that would have an impact on the world blood supply in the next decade or two may be viewed with some skepticism.

Nevertheless, I am very enthusiastic about the potential of this method to provide new

insights into mechanisms of terminal erythroid differentiation. Of all the approaches now available to us, it is the only one that yields large numbers of red blood cells with normal physiological functions, and it should therefore find wide application in this important and demanding area of research.

It is my hope that further technical advances in cell engineering in the next decade, along with improved understanding of the biology of embryonic stem cells, will lead to processes for the production of cultured red blood cells at reasonable cost. Such a development would have no small impact on the quality of health-care worldwide.

1. Giarratana, M.-C. *et al.* *Nat. Biotechnol.* **23**, 69–74 (2005).
2. Winslo, R.M. *J. Intern. Med.* **253**, 508–517 (2003).
3. Gregory, C.J. & Eaves, A.C. *Blood* **51**, 527–537 (1978).
4. Fibach, E. *et al.* *Blood* **73**, 100–103 (1989).
5. Malik, P. *et al.* *Blood* **91**, 2664–2671 (1998).
6. Migliaccio, G. *et al.* *Blood Cells Mol. Dis.* **28**, 169–180 (2002).
7. Lichtman, M.A. *Exp. Hematol.* **9**, 391–410 (1981).
8. Mohandas, N. & Prenant, M. *Blood* **51**, 633–643 (1978).
9. Hanspal, M. & Hanspal, J.S. *Blood* **84**, 3494–3504 (1994).

post-translationally, are often critical to protein function and may influence cellular homeostasis.

Phosphorylation of seryl, threonyl and tyrosyl residues is arguably the most important of the post-translational modifications, being responsible for activating and repressing the activity of proteins involved in nearly all signal transduction cascades. The importance of phosphorylation to cellular response and homeostasis has naturally spurred efforts to develop efficient methods for characterizing large numbers of these sites across the entire proteome, with the aim of identifying novel or aberrant signaling events that may be associated with disease.

The relative abundance of phosphorylated peptides compared to the entire proteome has required the development of enrichment techniques for these specific peptides. The limitations of existing techniques for enriching phosphotyrosine peptides is illustrated by two recent studies on the global identification of phosphorylation sites, in which less than 1% of the phosphorylation sites identified were phosphotyrosine residues^{4,5}. Although these results correspond to the relative cellular abundance of phosphotyrosine compared with phosphoserine and phosphothreonine residues, they also illustrate the need for methods specific to phosphotyrosine residue characterization.

In the analysis of complex mixtures such as cell lysates, tandem mass spectrometry is limited in the number of peptides it can sample per experiment and therefore typically identifies those that are present in high abundance. Low abundance peptides, such as those that are phosphorylated, are infrequently identified unless some measure is taken to exclude high abundance peptides. Isolating a mixture enriched for pTyr peptides makes this class of peptide the most abundant in the sample and forces the mass spectrometer to sample these peptides, resulting in a much greater number being identified.

The elegance of the study by Rush *et al.* is that it uses readily available reagents and does not require any sophisticated chemistry or chromatography to enrich for phosphotyrosine peptides (Fig. 1). Proteome samples extracted from cells are digested into peptides that are passed through an immunoaffinity column containing a phosphotyrosine monoclonal antibody coupled to protein G agarose. The captured peptides are fractionated using reversed-phase liquid chromatography and identified by tandem mass spectrometry.

As Rush *et al.* point out, other investigators have used pTyr-specific antibodies to immunoprecipitate proteins from cell

An enriched look at tyrosine phosphorylation

Thomas P Conrads & Timothy D Veenstra

Immunoaffinity isolation enables the identification of phosphotyrosine peptides on a global level.

Although phosphorylated tyrosyl residues represent only 0.05% of the phosphoamino acids within a cell, they are indispensable in many signal transduction cascades¹. Abnormalities in tyrosine phosphorylation are directly responsible for the pathogenesis of numerous inherited and acquired human diseases, ranging from cancer to immune deficiencies². Unfortunately, existing proteomics techniques for detecting proteins containing phosphotyrosine continue to leave these proteins under-represented. In this issue, Rush *et al.*³ address this deficiency with a sample-preparation

technique specific for the enrichment of phosphotyrosine-containing peptides.

Current global proteomic investigations are discovery-driven, enabling information on large numbers of proteins to be gathered relatively quickly. The ultimate goal is to use such information to construct detailed and testable hypotheses about changes in the proteome that reflect cellular function or dysfunction, instead of having to rely on the unfeasible or laborious analyses of individual proteins.

Because of the complexity and heterogeneity of cellular proteins, proteome characterization represents a much greater analytical challenge than genomic sequencing or quantitative comparison of gene transcripts. Not only are proteins made up of a large number of amino acids with divergent physicochemical properties, they are often chemically modified in various ways. These chemical modifications, which are usually introduced

Thomas P. Conrads and Timothy D. Veenstra are at the Laboratory of Proteomics and Analytical Technologies, SAIC-Frederick, Inc., National Cancer Institute at Frederick, P.O. Box B, Frederick, MD 21702-1201, USA.
e-mail: conrads@ncifcrf.gov, veenstra@ncifcrf.gov

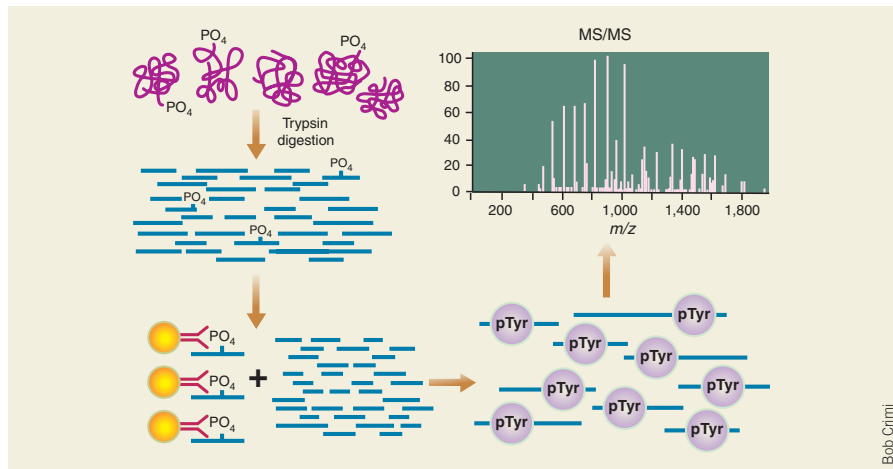


Figure 1 In the study by Rush *et al.*, the proteome is digested into peptides that are then passed through an anti-phosphotyrosine monoclonal antibody column to extract phosphotyrosine (pTyr)-containing peptides. This enriched mixture is analyzed by liquid chromatography coupled on-line with tandem mass spectrometry (MS/MS) to identify the exact sequence of the pTyr peptides.

extracts. Unfortunately, after the targeted protein was proteolytically digested and the collection of resultant peptides analyzed by mass spectrometry, the Tyr-phosphorylated peptides within the targeted proteins were not identified. What makes the method of Rush *et al.* unique is that the protein extract is proteolytically digested into peptides before immunoprecipitation. By directly immunoprecipitating the pTyr peptides, this method ensures that only peptides with phosphorylated Tyr residues will be introduced into the mass spectrometer for identification.

In two validation analyses of pervanadate-treated Jurkat cells and NIH-3T3 cells constitutively expressing active c-Src (that is, 3T3-Src), Rush *et al.* identified 194 and 185 phosphotyrosine sites, respectively. In both experiments, a sizable number of known phosphotyrosine sites were identified, as well as a large proportion (~70%) of novel sites.

The phosphotyrosine sites identified in these cell lines were compared with those identified in a similar analysis of two cell lines derived from anaplastic large cell lymphomas, Karpas 299 and SUD-DHL-1, which express a constitutively active oncogenic fusion tyrosine kinase (NPM-ALK). Whereas the overlap of the phosphotyrosine sites identified in both Karpas 299 and SUD-DHL-1 cells was high (72%), the overlap between these cells and Jurkat and 3T3-Src cells was very low. The results from the anaplastic large cell lymphoma cell lines showed phosphorylation of metabolic enzymes and ribosomal proteins, proteins that were found to be unmodified in Jurkat or 3T3-Src cell lines. Obvious differ-

ences were observed in a comparison of the flanking sequences surrounding the identified phosphotyrosine sites, consistent with the differences in the dominating kinase activity present in each of the cell types.

The results from this type of analysis demonstrate the utility of identifying the primary types of kinases that are functional within a deranged cancer cell and provide provoca-

The elegance of the study by Rush *et al.* is that it uses readily available reagents and does not require any sophisticated chemistry or chromatography to enrich for phosphotyrosine peptides.

tive targets for therapeutic development. The value of such information has already been shown in the development of the anti-chronic myeloid leukemia drug, imatinib mesylate⁶ (Novartis, Basel, Switzerland). The onset of chronic myeloid leukemia is attributed to a genetic abnormality that results in the expression of the fusion protein BCR-ABL, a constitutively activated tyrosine kinase⁶. By specifically inhibiting BCR-ABL, imatinib mesylate significantly reduces the leukemic activity in chronic myeloid leukemia patients without affecting normal cells. The development of proteomic tools capable of recognizing abnormal or differential kinase activity

provides a first step to understanding disease mechanisms for identifying possible drug-gable targets.

The study of Rush *et al.* highlights one of the major differences among genomic, transcriptomic and proteomic analyses. While DNA and mRNA can be extracted en masse from a system for genome sequencing or total mRNA analysis, such inclusive strategies are not very effective for proteome characterization for two reasons. First, no single existing method is capable of extracting with equal efficiency every expressed protein in a proteome owing to the vast differences in their physiochemical properties. Second, even if such a gross extraction of cellular proteins were possible, the proteome analysis would pale in comparison to genomic and transcriptomic analyses because of the finite dynamic range of detection possible with current separation methods and mass spectrometry instrumentation. These limitations effectively hamper proteome analyses to the extent that typically only 1% of the total number of reliably identified peptides is post-translationally modified.

Arguably, the greatest need in proteomics today is for continued development of more effective methods and tools to separate the proteome into various definable components (that is, phosphopeptides, glycopeptides, sub-cellular components, complexes, etc.) to allow more complete characterizations of these sub-proteomes. The future challenge will be to systematically collate these subproteomic data to assemble a comprehensive view of cellular networks and events and how they contribute to cellular homeostasis.

The magnitude of this undertaking leaves the field of proteomics at a crossroads. Should it focus on studies that target a subproteome (like phosphotyrosine sites in Rush *et al.*), or should it aim to characterize cells at a more global level so that predictions concerning protein localization, modification status, quantitation, interacting partners and so forth can be accurately made based on a particular cell's environment? Given the available technology and the number of investigators in the field, undoubtedly both avenues will continue to be pursued, and the direction will depend on the success of the discoveries and the testable hypothesis that are generated.

1. Pawson, T. & Scott, J.D. *Science* **278**, 2075–2080 (1997).
2. Alonso, A. *et al.* *Cell* **117**, 699–711 (2004).
3. Rush, J. *et al.* *Nat. Biotechnol.* **23**, 94–101 (2005).
4. Ficarro, S.B. *et al.* *Nat. Biotechnol.* **20**, 301–305 (2002).
5. Beausoleil, S.A. *et al.* *Proc. Natl. Acad. Sci. USA* **101**, 12130–12135 (2004).
6. Druker, B.J. *et al.* *N. Engl. J. Med.* **344**, 1031–1037 (2001).