Article Watch: September 2014

*Clive A. Slaughter**GRU-UGA Medical Partnership, Athens, GA*

This column highlights recently published articles that are of interest to the readership of this publication. We encourage ABRF members to forward information on articles they feel are important and useful to Clive Slaughter, Georgia Regents University-University of Georgia Medical Partnership, 1425 Prince Ave., Athens, GA 30606, USA. Phone: 706-713-2216; Fax: 706-713-2221; E-mail: cslaught@uga.edu; or to any member of the Editorial Board. Article summaries reflect the reviewer's opinions and not necessarily those of the association.

MACROMOLECULAR SYNTHESIS AND
SYNTHETIC BIOLOGY

Malyshev DA, Dhami K, Laverne T, Chen T, Dai N, Foster JM, Correa IR, Romesberg FE. A semi-synthetic organism with an expanded genetic alphabet. *Nature* 509;2014: 385–388.

Extension of the genetic code beyond the two canonical base pairs of DNA (adenine-thymine and cytosine-guanine) has long been an aspiration of molecular biology. Malyshev et al. here demonstrate that a third, unnatural base pair can be propagated in *Escherichia coli*. They construct a plasmid containing a pair of unnatural bases that rely on hydrophobic interaction rather than hydrogen bonding for their complementarity. The base pair has been shown previously to be capable of efficient transcription and replication in vitro. Propagation of such a base pair in vivo requires first that the unnatural nucleotides be available inside of the cell. This is achieved by expression of an algal nucleotide triphosphate transporter that efficiently transports the triphosphates into *E. coli* cells. The next requirement is that the unnatural triphosphates be faithfully incorporated into DNA to support replication. The unnatural base pair is here placed in a region of the plasmid that is predicted to rely on DNA polymerase I for its replication. This enzyme normally fills in gaps in DNA or connects Okazaki fragments. The enzyme was shown previously to incorporate the unnatural base pair efficiently in vitro. Recovery and analysis of the plasmid after 15 h of cell growth indicate retention of the unnatural base pair to be 99.4%/doubling in the present study. The third requirement is that the base pair is not excised by DNA repair pathways. In the present work, it appears to be stable in exponential and stationary-phase

growth. The results therefore qualify as the first demonstration of a cell whose DNA stably harbors three base pairs. These results encourage the hope that cells may eventually be re-engineered to enable site-specific labeling of nucleic acids, construction of promoter sequences subject to orthogonal regulation, and expression of proteins with multiple, different unnatural amino acids.

SMALL MOLECULE ANALYSIS AND METABOLOMICS

Zhou Y, Park H, Kim P, Jiang Y, Costello CE. Surface oxidation under ambient air—not only a fast and economical method to identify double bond positions in unsaturated lipids but also a reminder of proper lipid processing. *Analytical Chemistry* 86;2014:5697–5705.

The authors of this paper describe a new method for assigning the positions of double bonds in the fatty acyl chains of unsaturated lipids. The method requires no chemical derivatization or special mass spectrometry (MS) instrumentation: it is conducted here with an electrospray orthogonal quadrupole time-of-flight (Q-TOF) instrument. For convenience, the mass spectrometer is coupled with a liquid-extraction, surface-analysis device to sample analytes after they have been separated by thin-layer chromatography (TLC). The technique involves oxidation on dried TLC plates by exposure to air and light, usually for 30 min, although the reaction rate is slower for some double bonds. Aldehydes, the most common oxidation products formed at double bonds, are localized by collision-induced dissociation. Lipid oxidation is found to be slower if the concentration of ozone is increased. The authors propose that singlet oxygen is the oxidizing agent responsible for generating the products observed. In addition to providing an inexpensive way to locate double bonds, the authors point out that biological samples extracted from TLC plates should be protected from artifactual oxidation

doi: 10.7171/jbt.14-2503-002

brought about by exposure to ambient air when native lipids are being studied.

Sarafian M H, Gaudin M, Lewis M R, Martin F-P, Holmes E, Nicholson J K, Dumas M-E. Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid chromatography–mass spectrometry. *Analytical Chemistry* 86;2014:5766–5774.

Sarafian et al. compare eight different methods of preparing lipid samples from blood plasma for high-throughput, untargeted ultra-high-performance liquid chromatography–MS analysis. They test four protein precipitation methods (methanol, acetonitrile, isopropanol, and isopropanol–acetonitrile) and four liquid–liquid extraction methods (methanol combined with chloroform, dichloromethane, and methyl-*t*-butyl ether or isopropanol with hexane). Precipitation methods are found to be slightly more efficient than extraction methods, and isopropanol precipitation is determined to be the most straightforward and robust method. It provides the best protein removal, repeatability, lipid recovery, and coverage.

CARBOHYDRATES AND GLYCOCONJUGATES

Chen W, Smeekens JM, Wu R. A universal chemical enrichment method for mapping the yeast N-glycoproteome by mass spectrometry (MS). *Molecular & Cellular Proteomics* 13;2014:1563–1572.

Low abundance and great diversity of glycoforms complicate the isolation of glycopeptides in studies aimed at documenting protein glycosylation sites. The present work exploits the well-known interaction between boronic acid and 1,2- and 1,3-*cis*-diols as the basis for a glycopeptide purification method. The method is used to survey *N*-glycosylation sites on yeast proteins. Glycopeptides from yeast cell lysates are extracted by covalent binding to boronic acid-conjugated magnetic beads at pH 10.0. The bound glycopeptides are eluted from the beads by washing in acid solution, which dissociates the chemical bonds between boronic acid and diols. The glycopeptides are then digested with PNGase F to remove the glycan moieties. The resulting mass shift of the encoded asparagine at the linkage site is augmented by performing the digestion in H₂¹⁸O. This precaution minimizes false-positive glycopeptide identification during mass spectrometric analysis. Application of the methodology here results in the identification of 816 *N*-glycosylation sites on 332 yeast proteins, demonstrating its suitability for large-scale analysis of protein glycosylation.

Meuris L, Santens F, Elson G, Festjens N, Boone M, Dos Santos A, Devos S, Rousseau F, Plets E, Houthuys E, Malinge P, Magistrelli G, Cons L, Chatel L, Devreese B,

Callewaert N. GlycoDelete engineering of mammalian cells simplifies N-glycosylation of recombinant proteins. *Nature Biotechnology* 32;2014:485–489.

Many protein drug molecules bear glycan side-chains. The glycan moiety is frequently required for correct protein folding. Glycoprotein pharmaceuticals are produced mainly in mammalian expression systems, such as Chinese hamster ovary cells, so that immunogenicity of the glycan moieties in patients is minimized. However, the glycans added in mammalian expression systems are heterogeneous. The heterogeneity causes difficulties in quality control and contributes to variation in pharmacokinetics and biological activity. Meuris et al. here describe an engineered mammalian expression system that yields shortened, uniform glycan moieties of human-like structure. In mammalian cells with an intact glycosylation machinery, oligomannose glycans entering the Golgi apparatus are trimmed to a moiety containing two *N*-acetyl-D-glucosamine (GlcNAc) groups and five mannose (Man) groups (Man₅GlcNAc₂). The Golgi enzyme *N*-acetylglucosaminyltransferase 1 (GnT1) variably modifies this structure by addition of GlcNAc residues to one or more of the terminal Man groups. This modification initiates assembly of hybrid- or complex-type glycans. The cells engineered by Meuris et al. lack GnT1. The cells also express an endoglycosidase, endo-β-*N*-acetylglucosaminidase (Endo T) that is targeted to the Golgi, where it cleaves the Man₅GlcNAc₂ core between the first two GlcNAc residues, leaving a single GlcNAc residue *N*-linked to the protein. This single residue is then extended by galactosyltransferases and sialyltransferases in the Golgi. The result is a small, sialylated trisaccharide. Protein molecules bearing this moiety are selected by lectin affinity chromatography with concanavalin A. The engineered expression system is tested by producing two therapeutic glycoproteins, GM-CSF, and an anti-CD20 mAb. Both proteins prove to have similar function, stability, and antigenicity to their normally glycosylated counterparts. It is hoped that this methodology will enable production of therapeutic glycoproteins with greater consistency and potentially with enhanced biopharmaceutical function.

MASS SPECTROMETRY

Robb DB, Brown JM, Morris M, Blades MW. Tandem mass spectrometry using the atmospheric pressure electron capture dissociation ion source. *Analytical Chemistry* 86; 2014:4439–4446.

Electron transfer dissociation (ETD) has become a popular fragmentation method for sequencing peptides and particularly, for localizing phosphorylation and glycosylation sites. However, the instrumentation required to implement the method is complex and expensive, involv-

ing a dedicated second ion source to generate the reagent anions for the electron transfer reactions and provision of special ion-ion storage and manipulation schemes to accomplish the fragmentation. Conventional electron capture dissociation (ECD) produces product ion spectra similar to those of ETD but is hardly an advantage, as it generally requires a Fourier transform ion cyclotron resonance instrument. An implementation of ECD has been described that can readily be retrofit to instruments not originally designed for ECD or ETD experiments. The process is atmospheric pressure-ECD. It requires only a photoionization detector lamp to produce electrons from a photoionizable dopant in the source region of the instrument. Unfortunately, the location of the electron capture reaction to the ion source prevents isolation of precursor ions for fragmentation and therefore, seems incapable of supporting tandem MS (MS/MS) experiments. The present paper, however, shows how true MS/MS spectra can be acquired using the method. During electron capture, charge-reduced products are formed in which a covalent bond has been cleaved without immediate dissociation of the products, which remain held together by noncovalent forces. These nondissociated products may be isolated in a quadrupole mass filter. They can then dissociate by gentle collision in a collision cell. The product ions are then available for separation in a high-resolution mass analyzer to produce a product ion spectrum. The authors demonstrate this principle using a Xevo G2-S Q-TOF instrument from Waters Micromass (Milford, MA, USA). They describe MS/MS results for model peptides and demonstrate high sequence coverage and low femtomole sensitivity. They find some limitations in the method: cleavage *N*-terminal to proline is not observed, and some cleavages are missed at the ends of the precursors. However, despite these limitations, it is hoped that the method will provide a low-cost way for laboratories to deploy electron-based fragmentation.

PROTEOMICS

Dickhut C, Feldmann I, Lambert J, Zahedi RP. Impact of digestion conditions on phosphoproteomics. *Journal of Proteome Research* 13;2014:2761–2770.

Phosphorylation of amino acid residues in proximity to trypsin cleavage sites is known to affect the efficiency of cleavage by trypsin. This compromises the detection of phosphorylation sites in mass spectrometric surveys of tryptic peptides and complicates calculation of phosphorylation stoichiometry. Dickhut et al. here report a systematic study of the effect of phosphorylation on tryptic cleavage using a panel of model peptides. They find strong suppression of cleavage in the vicinity of phosphoserine

(pSer) and phosphothreonine (pThr) residues, but little effect is noted for phosphotyrosine. The authors hypothesize that suppression of cleavage is a result of the formation of salt bridges between Arg/Lys at the site of cleavage and the pSer or pThr residues nearby. Digestion efficiency of phosphopeptides that are resistant to cleavage can be improved substantially by adding 10% acetonitrile or 5% trifluoroethanol to the digestion solvent or by increasing the ratio of trypsin to peptide substrate from 1:50 to 1:10. Nevertheless, some peptides remain sufficiently resistant even under such optimized conditions to hamper calculation of stoichiometry.

Kim M-S, Pinto SM, Getnet D, Nirujogi RS, Manda SS, Chaerkady R, Madugundu AK, Kelkar DS, Isserlin R, Jain S, Thomas JK, Muthusamy B, Leal-Rojas P, Kumar P, Sahasrabudhe NA, Balakrishnan L, Advani J, George B, Renuse S, Selvan LDN, Patil AH, Nanjappa V, Radhakrishnan A, Prasad S, Subbannayya T, Raju R, Kumar M, Sreenivasamurthy SK, Marimuthu A, Sathe GJ, Chavan S, Datta KK, Subbannayya Y, Sahu A, Yelamanchi SD, Jayaram S, Rajagopalan P, Sharma J, Murthy KR, Syed N, Goel R, Khan AA, Ahmad S, Dey G, Mudgal K, Chatterjee A, Huang T-C, Zhong J, Wu X, Shaw PG, Freed D, Zahari MS, Mukherjee KK, Shankar S, Mahadevan A, Lam H, Mitchell CJ, Shankar SK, Satishchandra P, Schroeder JT, Sirdeshmukh R, Maitra A, Leach SD, Drake CG, Halushka MK, Prasad TSK, Hruban RH, Kerr CL, Bader GD, Iacobuzio-Donahue CA, Gowda H, Pandey A. A draft map of the human proteome. *Nature* 509;2014:575–581.

Wilhelm M, Schlegl J, Hahne H, Gholami AM, Lieberenz M, Savitski MM, Ziegler E, Butzmann L, Gessulat S, Marx H, Mathieson T, Lemeier S, Schnatbaum K, Reimer U, Wenschuh H, Mollenhauer M, Slotta-Huspenina J, Boese J-H, Bantscheff M, Gerstmair A, Faerber F, Kuster B. Mass-spectrometry-based draft of the human proteome. *Nature* 509;2014:582–587.

These two papers describe compilations of data from in-depth proteomic analyses of human cells or tissues. Kim et al. identify the protein products of 17,294 genes, and Wilhelm et al. identify the products of 18,097 genes. Of particular interest in both papers is the evidence for translation of pseudogenes, novel open-reading frames (including proteins mapping to transcripts annotated as noncoding RNAs), novel coding regions within previously annotated genes, and products of extension events at the gene, protein, or exon levels. Novel *N*-termini and signal peptide cleavage sites are also reported. Special search strategies and heightened stringency of validation criteria (up to and including confirmation by peptide synthesis) are required to enable and support such observations, and the methods used in these studies will be especially helpful to investigators in the field.

FUNCTIONAL GENOMICS AND PROTEOMICS

Kim S, Kim D, Cho SW, Kim J, Kim J-S. Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research* 24; 2014:1012–1019.

Ramakrishna S, Kwaku Dad A-B, Beloor J, Gopalappa R, Lee S-K, Kim H. Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. *Genome Research* 24;2014:1020–1027.

RNA-guided endonucleases have provided exceptional flexibility and convenience in directed genome editing. They work by introducing a double-stranded break in chromosomal DNA at the site specified by the sequence of an RNA guide. The repair of such breaks by endogenous DNA repair processes gives rise to targeted mutagenesis and chromosomal rearrangements. DNA cleavage is mediated by the clustered, regularly interspaced, short palindromic repeat-associated 9 (Cas9) protein. Recently, off-target effects of RNA-guided endonucleases have been documented. One source of such effects is the delivery of Cas9 and the RNA guide by transfection with plasmids encoding the protein and RNA. Such plasmids may undergo random integration into the host genome, either at the target site or elsewhere. Off-target insertion can be difficult to detect and may provoke host-immune responses. Prolonged expression of RNA-guided endonucleases can exacerbate off-target effects. DNA-transfected cells are also classified as “genetically modified” and therefore, require compliance with the relevant regulatory procedures. Two groups circumvent these difficulties by introducing Cas9 into cells as a recombinant protein along with in vitro-transcribed guide RNA. Kim et al. accomplish this by electroporation. Ramakrishna et al. do it by conjugation of Cas9 to a cell-penetrating peptide and complexation of the RNA guide to the cell-penetrating peptide. Chromosomal DNA is cleaved quickly following delivery, and the ribonucleoprotein is degraded rapidly. The results show that delivery by these methods supports efficient genome editing, even in cells hitherto not amenable to RNA-guided endonuclease engineering because of their resistance to DNA transfection or because they are vulnerable to the stress produced by transfection. Together with appropriate management of other sources of off-target effects, the approaches described encourage continued use of RNA-guided endonuclease genome engineering.

CELL BIOLOGY AND TISSUE ENGINEERING

Genovese P, Schirotti G, Escobar G, Di Tomaso T, Firrito C, Calabria A, Moi D, Mazziere R, Bonini C, Holmes MC, Gregory PD, van der Burg M, Gentner B, Montini E,

Lombardo A, Naldini L. Targeted genome editing in human repopulating haematopoietic stem cells. *Nature* 510; 2014:235–240.

Gene therapy for inherited diseases of hematopoiesis has long been approached by insertion of a functional copy of the defective gene into hematopoietic stem cells. Unfortunately, this approach risks insertion of the gene into inopportune genomic locations and consequent oncogenesis, toxicity, or elimination of the genetically modified cells. Artificial endonucleases that produce targeted double-stranded breaks (zinc finger nucleases, transcription activator-like effector nucleases, and RNA-guided nucleases) now offer ways to circumvent these problems. They accomplish repair of defective genes rather than insertion of new copies. Homologous recombination in the recipient cell can edit a defective gene sequence when provided with an exogenous DNA template flanked by sequences homologous to the nuclease target site. A major hurdle in making this approach work is overcome by Genovese et al. The problem is that homologous recombination is active, mainly during the S/G2 phase of the cell cycle. However, hematopoietic stem cells rarely enter mitosis, as they cycle slowly. Genovese et al. pretreat their stem cells with cytokines to promote cell division and apply dimethyl PGE2 and the aryl hydrocarbon receptor inhibitor protein SR1 to prevent premature differentiation and maintain the stem-cell state. After 2 days, the authors introduce by electroporation a zinc-finger nuclease and a DNA template for repair of the targeted gene. The therapeutic potential of this methodology is demonstrated by reconstituting the *IL2RG* gene in hematopoietic stem cells from a patient with X-linked severe combined immunodeficiency disease, a condition that results from a defect in this gene. The edited stem cells display a growth advantage over endogenous, defective cells and are shown to support normal hematopoiesis in the patient.

Chong JJH, Yang X, Don CW, Minami E, Liu Y-W, Weyers JJ, Mahoney WM, Van Biber B, Cook SM, Palpant NJ, Gantz JA, Fugate JA, Muskheli V, Gough G M, Vogel KW, Astley CA, Hotchkiss CE, Baldessari A, Pabon L, Reinecke H, Gill EA, Nelson V, Kiem H-P, Laflamme MA, Murry CE. Human embryonic-stem-cell-derived cardiomyocytes regenerate non-human primate hearts. *Nature* 510;2014: 273–277.

Chong et al. bring one step closer to reality the therapeutic regeneration of injured heart muscle by cardiomyocytes derived from human embryonic stem cells. The investigators use a nonhuman primate model of myocardial infarction. The very large number of cells (1×10^9) estimated to be required for heart-muscle repopulation in this model necessitates cryopreservation to accumulate the requisite batch size of cells. This is shown to be feasible with

good retention of viability. The cells are delivered directly into the infarct region, and their fate is determined upon necropsy by observing their expression of a GFP marker. One animal is tested on Day 14 after implantation, two on Day 28, and one on Day 84. Graft size averaged 2.1% of the left ventricle (range: 0.7–5.3%), and accounted for 40% of the infarct volume. More than 98% of the cells expressed α -actinin, indicating cardiomyocyte identity. The cells showed progressive, although incomplete, maturation during the duration of the experiment, indicated by increasing myofibril alignment, sarcomere registration, and cardiomyocyte diameter. The implanted cells were perfused by host vasculature. Electromechanical junctions between graft and host cardiomyocytes were observed by Day 14 and showed calcium transients synchronized to the host electrocardiogram. Nonfatal ventricular arrhythmias, however, were observed. This complication must be addressed in future studies. Nevertheless, the results encourage further development of the technology for therapeutic purposes.

TOXICITY TESTING

Kleinstreuer NC, Yang J, Berg EL, Knudsen TB, Richard AM, Martin MT, Reif DM, Judson RS, Polokoff M, Dix DJ, Kavlock RJ, Houck KA. Phenotypic screening of the ToxCast chemical library to classify toxic and therapeutic mechanisms. *Nature Biotechnology* 32;2014:583–591.

Most people agree on the importance of ascertaining the potential toxicity of pharmaceutical and environmental chemicals to which we may be exposed. However, there is increasingly widespread ethical concern about the use of animals to acquire this information. Lack of toxicity data for thousands of environmental chemicals and the high failure rate of investigational drugs as a result of adverse reactions are among the circumstances indicating that animal testing is also limited in its effectiveness. The present

study is aimed at evaluating the use of human cells in culture to screen chemicals for their ability to disrupt physiologically important pathways. Low-passage cells that retain in vivo signaling patterns are chosen for the work. They include primary endothelial cells, peripheral blood mononuclear cells, bronchial epithelial cells, dermal fibroblasts, keratinocytes, and coronary artery smooth muscle cells. Responses to chemicals that are measured are not restricted to cell viability but include expression of cell adhesion proteins, cytokines, matrix metalloproteases, and cell-surface receptors. Such responses are measured over a range of exposure doses. Environmental, industrial chemicals, and failed drugs are tested. Many of these chemicals have unknown toxicity or unknown mechanisms of toxicity, but statistical clustering methods reveal some interesting associations with the effects of substances with known mechanisms of toxicity, thus allowing discrete bioactivity profiles to be recognized. For example, a donated pharmaceutical compound, DPharm86, is found to stimulate production of the inflammatory molecules VCAM-1 and IFN- γ -induced protein 10 in dermal fibroblasts. These markers are also stimulated by several molecules associated with skin rash in clinical trials, an adverse outcome also noted with DPharm86. In another example, estrogen receptor pathway modulators, aryl hydrocarbon receptor activators, and mammalian target-of-rapamycin inhibitors are found to share the ability to up-regulate tissue-factor expression in venous endothelial cells stimulated with the cytokines IL-1 β , TNF- α , and IFN- γ . This indicates an association with increased risk of thrombosis. It is hoped that the use of cultured cells to assemble large data sets in this way will help elucidate toxicity mechanisms, define pathways leading to adverse outcomes, enable construction of predictive models, and ultimately, provide a first-tier alternative to animal testing.