Journal of Biomolecular Techniques • Volume 33(1); 2022 Apr

Article Watch: April, 2022

Published on: Apr 12, 2022

DOI: 10.7171/3fc1f5fe.f0cb1a82

License: Copyright © 2022 Association of Biomolecular Resource Facilities. All rights reserved.

ABSTRACT

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, MCG-UGA Medical Partnership, 1425 Prince Avenue, Athens GA 30606. Tel; (706) 713-2216: Fax; (706) 713-2221: Email; 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.

NUCLEIC ACID SEQUENCING

Santos I C, Lanzillotti M, Shilov I, Basanta-Sanchez M, Roushan A, Lawler R, Tang W, Bern M, Brodbelt J S. Ultraviolet photodissociation and activated electron photodetachment mass spectrometry for top-down sequencing of modified oligoribonucleotides. *Journal of the American Society for Mass Spectrometry* 33;2022:510-520.

Interest in the use of RNAs for therapeutic purposes is rapidly expanding. RNAs are employed to encode immunogenic proteins, disrupt the expression of endogenous genes, and modulate splicing. Synthetic RNAs are commonly deployed as small-interfering RNAs (siRNAs), antisense RNAs, microRNAs (miRNAs) or aptamers. Nucleotides covalently modified in various ways are frequently incorporated to stabilize these molecules, or increase their biological lifetime, or minimize immune responses they provoke. Modifications include changes in bases; incorporation of non-phosphate linkages, *e.g.* phosphorothioate linkages; and changes to the ribose, *e.g.* locked nucleic acids (LNAs). Various mass spectrometric methods are available for characterization of such species. Santos *et al.* evaluate various ion activation methods for the generation of informative fragmentation patterns in MS/MS analysis of modified oligonucleotides. Collisional methods can result in loss of bases and produce other uninformative processes. Furthermore, LNAs suppress fragmentation by collisional methods. In the present study, higher-energy collisional activation (HCD) is compared with ultraviolet photodissociation (UVPD) at 193 and 213 nm wavelengths, and activated electron photodetachment dissociation (a-EPD), a technique in which EPD is coupled with UVPD. The authors' group has previously shown that UVPD alone provides favorable retention of bases, including modified ones, although with modest overall backbone fragmentation; and that a-EPD increases backbone dissociation to provide good yield of diagnostic fragment ions belonging to the *w*, *z*, *a*, and *d* series. The present paper documents the performance of these techniques with RNA

containing methylated bases, phosphorothioate linkages and LNAs. Both UVPD and a-EPD produce high yields of *a*/*w*, *b*/*x*, *c*/*y* and *d*/*z* fragment ions. Phosphorothioate bonds don't suppress backbone fragmentation in either method. The methylene bridges in LNAs suppress fragmentation in UVPD, but less so in a-EPD. Both 213 and 193 nm irradiation provide extensive backbone fragmentation in UVPD. The authors demonstrate excellent results with extensively modified 21-mer oligonucleotides, and anticipate useful results with even longer modified RNAs.

GLYCANS

Pupo E, Van Der Ley P, Meiring H D. Nanoflow LC–MS method allowing in-depth characterization of natural heterogeneity of complex bacterial lipopolysaccharides. *Analytical Chemistry* 93;2021:15832-15839.

The cell-surface lipopolysaccharides (LPSs) of gram-negative bacteria are amphipathic components of high-abundance in bacterial cell surfaces and are major determinants of bacterial pathogenicity and immunogenicity. LPS structures are variable both within and among bacterial strains, and structures may even vary at successive stages in the evolution of an infection. This variation occurs within a common, architectural, covalently bonded plan. This framework consists of a lipid A moiety that inserts into the bacterial cell membrane, a core oligosaccharide bonded to it *via* a 3-deoxy-*D*-manno-oct-2-ulosonic acid (Kdo) monosaccharide moiety, and, linked to the core oligosaccharide, an O-antigen polysaccharide containing a repeating oligosaccharide structure. Pupo *et al.* provide methodology for the separation and structural analysis of LPS molecules by combined liquid chromatography and mass spectrometry (LC-MS) to assist in the characterization of LPS variation. They fractionate mixtures of intact LPS molecules by nanoflow LC using water/triethylamine/acetic acid mobile phases and 2-propanol elution. On-line electrospray ionization (ESI) and Fourier transform mass spectrometry in an Orbitrap analyzer provide MS analysis and MS/MS analyses with collisional dissociation. Structural heterogeneity of the lipid A domain is characterized in detail, and *O*-polysaccharide chains varying from 3 to 15 kDa are analyzed.

Chao H-C, McLuckey S A. Manipulation of ion types via gas-phase ion/ion chemistry for the structural characterization of the glycan moiety on gangliosides. *Analytical Chemistry* 93;2021:15752-15760.

Gangliosides are amphipathic components of eukaryotic cell membranes consisting of a hydrophilic glycan head group linked to a hydrophobic ceramide moiety. Gangliosides are associated especially with lipid rafts, and are particularly abundant in

the central nervous system where they constitute the predominant glycan source. They mediate cellular interactions and are associated with neurodegenerative and neurodevelopmental diseases. Their structures are diverse. Classes of gangliosides differ in the number of sialic acids and subclasses differ in the linkages between sialic acids. Chao and McLuckey here contribute methods utilizing gas-phase ion/ion chemistry suitable for use in shotgun studies to analyze the glycan linkages present in biological samples. Deprotonated ganglioside ions are generated by ESI - $[M\text{-}H]^-$ ions for GM gangliosides; $[M\text{-}2H]^{2-}$ ions for GD; and $[M\text{-}3H]^{3-}$ ions for GT. Unfortunately, these products are contaminated with ions of different charge states. The authors isolate the desired ions with the help of gas phase chemical reactions. A terpyridine (Terpy) magnesium complex, $[Mg(Terpy)_2]^{2+}$, is used to yield charge-inverted (positive) product ions complexed with Mg. A 'proton sponge,' *N,N,N',N'*-tetramethyl-1,8-naphthalenediamine (PrS), is used to produce charge-reduced anions of some ganglioside series prior to reaction with Terpy. Because the contaminants do not react the same way as the desired ions, the reactions effect purification of species for subsequent structural study, which is performed by collision-induced dissociation (CID). The authors use this methodology for shotgun analysis of gangliosides from a porcine brain extract. They characterize 34 different gangliosides from precursor ions with just 20 m/z values through resolution of isomers.

Vainauskas S, Guntz H, McLeod E, McClung C, Ruse C, Shi X, Taron C H. A broad-specificity O-glycoprotease that enables improved analysis of glycoproteins and glycopeptides containing intact complex O-glycans. *Analytical Chemistry* 94;2022:1060-1069.

A number of new *O*-glycosidases have recently been described for release of *O*-glycan chains from proteins in glycan structural characterization. However, these enzymes are restricted in their specificity either by a requirement for the absence of sialic acid residues or by the particular amino acid sequences that need to be present at the site of glycan attachment amenable for them to cleave. So the hunt for more useful *O*-glycosidase reagent enzymes continues. Vainauskas *et al.* here characterize the immunomodulating metalloprotease (IMPa) from *Pseudomonas aeruginosa* for this purpose. The enzyme cleaves on the *N*-terminal side of *O*-glycosylated serine and threonine residues and is not affected by sialylation. Using an array of synthetic glycopeptides, the authors show that it is broadly tolerant of the amino acids adjacent to the cleavage site. It doesn't cleave between directly adjacent *O*-glycosites, but it does cleave if the sites are separated by as few as 2 amino acid residues. Application of the reagent enzyme is demonstrated in studies of granulocyte colony-stimulating factor

and receptor-type tyrosine-protein phosphatase C. The enzyme is commercially available from New England Biolabs, Ipswich. MA.

Burt R A, Dejanovic B, Peckham H J, Lee K A, Li X, Ounadjela J R, Rao A, Malaker S A, Carr S A, Myers S A. Novel antibodies for the simple and efficient enrichment of native O-GlcNAc modified peptides. *Molecular & Cellular Proteomics* 20;2021:100167.

Addition of the monosaccharide *N*-acetylglucosamine to serine or threonine residues (*O*-GlcNAc) is a post-translational modification associated with the control of an extensive list of biological processes. Mapping of *O*-GlcNAc sites has therefore become the focus of widespread interest. Burt *et al.* seek a simple, highly specific strategy for enrichment of native *O*-GlcNAc-modified peptides that would help in the identification of sires bearing this modification. Here they characterize a set of 4 new monoclonal anti-*O*-GlcNAc antibodies to add to the list of anti-*O*-GlcNAc immunologic reagents already described. Their antibodies are commercially available from Cell Signaling Technology, Inc., Danvers, MA. These new antibodies are used as a mixture coupled to Protein A-coated agarose beads. The mixed immunoaffinity reagent is shown to capture O-HexNAc-modified peptides but not peptides modified with complex glycans that contain HexNAc. Evidence indicating that the reagent binds *O*-GlcNAc peptides in preference to peptides *O*-linked to the isomeric monosaccharide *O*-GalNAc (*i.e.* Tn antigen) is also provided. The authors deploy the reagent in a study of mouse brain synaptosomes in which over 1300 *O*-GlcNAc-modified peptides are identified. It is hoped that the methodology for isolation of *O*-GlcNAc-conjugated peptides will stimulate investigation of this common modification.

METABOLOMICS

Ghorasaini M, Mohammed Y, Adamski J, Bettcher L, Bowden J A, Cabruja M, Contrepois K, Ellenberger M, Gajera B, Haid M, Hornburg D, Hunter C, Jones C M, Klein T, Mayboroda O, Mirzaian M, Moaddel R, Ferrucci L, Lovett J, Nazir K, Pearson M, Ubhi B K, Raftery D, Riols F, Sayers R, Sijbrands E J G, Snyder M P, Su B, Velagapudi V, Williams K J, De Rijke Y B, Giera M. Cross-laboratory standardization of preclinical lipidomics using differential mobility spectrometry and multiple reaction monitoring. *Analytical Chemistry* 93;2021:16369-16378.

Application of analytical technology in the clinical arena requires extensive, documented standardization. The present, interlaboratory study evaluates the

Lipidyzer platform for quantification of plasma lipids. The platform employs differential mobility spectrometry and multiple reaction monitoring on a QTRAP mass spectrometer to resolve isomeric and isobaric lipid species. It can resolve more than 1000 lipid species from 13 lipid classes: cholesteryl ester, ceramide, diacylglyerol, dihydroceramide, free fatty acid, hexosylceramide, lactosylceramide, lysophosphatidylcholine, lysophosphatidylethanolamine, phosphatidylcholine, phosphatidylethanolamine, sphingomyelin, and triacylglycerol. (Phosphatidylinositol and phosphatidylserine are not presently included.) In the present work, 54 deuterated internal standards are employed for quantification. Methyl-*t*-butyl ether is used for lipid extraction in preference to the classic chloroform/methanol method of Bligh & Dyer. All steps in the work-flow are standardized, including sample extraction, standard preparation, data acquisition, data processing and reporting of results. Such standardization contributes to the high reproducibility among 9 laboratories which the study demonstrates. The results reported in this work may also be used as reference values for other laboratories in the field.

MACROMOLECULAR SYNTHESIS & SYNTHETIC BIOLOGY

Sahtoe D D, Praetorius F, Courbet A, Hsia Y, Wicky B I M, Edman N I, Miller L M, Timmermans B J R, Decarreau J, Morris H M, Kang A, Bera A K, Baker D. Reconfigurable asymmetric protein assemblies through implicit negative design. *Science* 375;2022:eabj7662.

Many key biological processes are performed by multiprotein complexes that reversibly add or remove subunits as part of their mechanism of action. Examples include DNA replication and transcription. The design of proteins that can form reconfigurable assemblies is not straight forward. It must fulfill three requirements. Firstly, the subunits that undergo heteromeric interaction must be stable in solution on their own as monomers so that they can be added or removed at will. Furnishing them with a hydrophobic face might allow them to interact but will also result in the formation of unwanted homo-oligomers. Secondly, binding between interacting protein pairs must be specific, although dissociation rates must be high enough to support subunit exchange upon demand. Thirdly, the binding partners should have defined folded structures as monomers. It is more difficult to design structurally defined interactions when one of the proteins is an unfolded monomer that acquires structure only upon binding. The authors design 12 different heterodimer interfaces that fulfill these requirements. They provide the individual protomers with a substantial hydrophobic core for rigidity. The protomers are constructed with mixed α-β scaffolds

in which an exposed β-strand of limited length at the edge of a β-sheet in one protomer interacts with a complementary exposed β-strand at the edge of a β-sheet in the other protomer to form a continuous β-sheet in the heterodimer. Non-target interactions such as homodimeric interactions are rendered unfavorable by the energetic cost of burying polar groups without satisfying their hydrogen bonding propensity. The authors employ the protomers as connectors between helical repeat proteins to which they're fused. In this way they assemble linear heterodimers, heterotrimers, heterotetramers, heteropentamers and a heterohexamer with distinct shapes. They also assemble branched heterotetramers by using trivalent connectors. Many applications are envisioned for this methodology, including design of protein logic gates, assembly of processive multienzyme complexes, programmable intracellular control for synthetic biology, and reprogramming cells by manipulation of receptor modules.

Anishchenko I, Pellock S J, Chidyausiku T M, Ramelot T A, Ovchinnikov S, Hao J, Bafna K, Norn C, Kang A, Bera A K, Dimaio F, Carter L, Chow C M, Montelione G T, Baker D. De novo protein design by deep network hallucination. *Nature* 600;2021:547-552.

Last year saw remarkable successes in the use of deep neural networks for the prediction of protein 3-D structure from amino acid sequence based upon biophysical constraints alone, without invoking patterns gleaned from genetic homology. Following these developments, Anishchenko *et al.* now ask whether such neural networks could also be used to design novel polypeptide sequences, unrelated to naturally occurring proteins, that can fold up to produce definable 3-D structures. The authors use a process analogous to Google's DeepDream network trained to recognize patterns in images corresponding, for example, to faces. If the trained network is presented with an arbitrary input image, it will select among serial variants of that image ones that look to the network successively more like a face. The image resulting from such a process is referred to as a 'hallucination' because it doesn't represent any particular face but rather something the network recognizes as possessing features generically similar to faces. For protein work, the authors employ the trRosetta protein structure network. They input completely random sequences of 100 amino acids, which have a vanishingly small chance of folding into a definable structure. They then execute a Monte Carlo simulated healing trajectory (a probabilistic technique for approximating the global optimum of a given function) in sequence space. Amino acids are changed successively and the network chooses changes that optimize the sharpness of maps specifying distances and orientations against the background random noise produced

by random sequences. After about 20,000 Monte Carlo steps, the maps become as sharply featured as those predicted for naturally occurring protein sequences. In this way, the authors generate 2000 new proteins with sequences predicted to fold into well-defined structures, and show that these hallucinated proteins are different from any proteins in the structure database on which the network was initially trained. A sample of 129 of the novel proteins is expressed in *E. coli* and the proteins purified. Twenty-seven yielded monodispersed species with physical characteristics concordant with the hallucinated structures. These results indicate that deep neural networks are capable of contributing to molecular design for potentially diverse applications.

MASS SPECTROMETRY

Peters-Clarke T M, Riley N M, Westphall M S, Coon J J. Practical effects of intramolecular hydrogen rearrangement in electron transfer dissociation-based proteomics. *Journal of the American Society for Mass Spectrometry* 33;2022:100-110.

Electron-based peptide fragmentation methods – electron capture dissociation (ECD) and electron transfer dissociation (ETD) – cleave mainly at N–C_α bonds to form c- and $z^\bullet$-type ions. Using the nomenclature of Chu *et al.* (*International Journal of Mass Spectrometry* 390;2015:24-27), these ion types may be represented as $[c + 2H]^{1+}$ and $[z + H]^{\bullet 1+}$ respectively. It is well known that when ECD or ETD fragment ions of low charge density there is a tendency for the product ions to remain held together by non-covalent intramolecular interactions. Various strategies have proven effective for disruption of these undissociated product ions produced by ETD. They include following ETD with beam-type higher-energy collisional dissociation (EThcD), or resonant energy activation (ETcaD), or infrared photoirradiation concurrently with ETD in the same reaction cell (AI-ETD). Unfortunately, under circumstances in which these post-activation methods are utilized with most benefit, a side reaction may also occur in which a hydrogen atom migrates from the c-type ion to the $z^\bullet$-type in to form $[c + H]^{1+}$ and $[z + 2H]^{1+}$ products. Such hydrogen rearrangements shift the monoisotopic mass of the c-type ion 1 Da lower and that of the $z^\bullet$-type ion 1 Da higher, thereby diminishing the monoisotopic peak intensity and distorting isotopic distributions. These effects interfere with correct peptide identification during database searching. Hydrogen atom rearrangement has been studied in ETD. Here, the authors study the phenomenon in post-activation methods. They find that the propensity for hydrogen migration depends upon the amino acids surrounding the cleavage site, and the time even- and odd-electron product ions are held in close

proximity. More hydrogen migration is observed with ETcaD and EThcD than with either AI-ETD or ETD alone. The presence of D, G, N, S, or T directly *C*-terminal to the cleavage site exacerbates the effect. Thus, while supplemental activation yields more product ions, a proportional benefit in peptide identifications is only seen with AI-ETD. Manual spectral interpretation and database searching is required for the remaining products.

Jordan J S, Xia Z, Williams E R. Tips on making tiny tips: secrets to submicron nanoelectrospray emitters. *Journal of the American Society for Mass Spectrometry* 33;2022:607-611.

Emitters (tips) for nanoelectrospray ionization (nESI) typically have orifice diameters of >1 µm, but sub-µm emitters are advantageous for various purposes, including ionization of macromolecular analytes and desalting that occurs while generating the spray. Jordan *et al.* summarize advice about making one's own sub-µm emitters using a P-87 micropipette puller from Sutter Instrument, Novato, CA. They draw upon information from systematic experiments, and collect advice from various quarters including their own experience over a period of years. Of the 5 parameters of the P-87 duty cycle under user control, the "Pull" parameter, which controls the force applied to the pulley arms holding the borosilicate or quartz capillary from which the emitter is to be made, has the greatest effect on orifice diameter. The "Velocity," "Pressure," and "Heat" parameters have smaller effects, and the "Time" parameter makes little difference provided an appropriate threshold value is exceeded. Some parameters change the tip shape. The authors also offer advice on how to use emitters to analyze protein samples in solutions containing NaCl and non-volatile buffers such has Tris. It is hoped that the practical information they provide will encourage more extensive use of nESI for analysis of native proteins.

PROTEOMICS

Boekweg H, Van Der Watt D, Truong T, Johnston S M, Guise A J, Plowey E D, Kelly R T, Payne S H. Features of peptide fragmentation spectra in single-cell proteomics. *Journal of Proteome Research* 21;2022:182-188.

Aspirations to investigate cell-to-cell variation in proteomes has led to advances in capabilities for identification and quantification of proteins from single cells on the basis of fragment ion mass spectra derived from peptides. As in the proteomic analysis of bulk samples, the likelihood that a match between features in an observed spectrum and expected features for a database sequence correctly identifies a peptide provides

the basis for peptide/protein identification. Obviously, when available sample quantity is severely limiting, as it is in single-cell work, diminishing spectral quality is expected to impair peptide identification. Boekweg *et al.* investigate how. Current matching algorithms calculate quality of a match based on 3 criteria: (1) the number of observed fragment ion peaks that match the theoretical spectrum, (2) relative peak intensity of fragment ions and elevation above noise level, and (3) the degree of similarity among repeated spectra of the same peptide. The authors examine how single cell data differ from bulk data in each of these 3 respects. They demonstrate that peptides from single cells yield fewer annotated (and total) fragment peaks than bulk data. The peptides show intensity compression, thereby blurring the distinction between signal and noise. And although high quality single-cell spectra are internally consistent, they are distinct from bulk spectra. These differences affect the distribution of matching scores. The study highlights a need to adjust search algorithms to compensate for the differences between bulk and single-cell data for use in work at the limits of sensitivity.

Ctortecka C, Stejskal K, Krššáková G, Mendjan S, Mechtler K. Quantitative accuracy and precision in multiplexed single-cell proteomics. *Analytical Chemistry* 94;2022:2434-2443.

Ctortecka *et al.* assess quantification accuracy of single-cell proteomic analyses performed with tandem mass tags (TMT) in the presence of a carrier. Carriers reduce absorptive losses prior to MS analysis, augment MS1 signals during precursor scanning, and provide fragment ions for peptide identification. In multiplexed experiments carriers were originally deployed at very high carrier/sample ratio per single-cell sample (25x – 500x), but, more recently, 20x has been proposed to be optimal. In the present study, a single-cell sample is emulated by diluting a bulk HeLa cell digest to provide a standard input of 150 pg. A *maximum* carrier of 20x is employed and various TMT protocols evaluated. A dual carrier is also evaluated in which the carrier is distributed equally between 2 TMT channels to ameliorate extreme ratios but still boost MS1 signals. As expected, the authors observe increasing numbers of peptide identifications with increasing carrier ratio across the range tested, although identification rates are lower in split-carrier experiments. Compression in the reporter ion ratio occurs, but its severity is modest when carrier is present. Especially concerning, however, is that filtering of scans based on signal/noise ratio results in rejection of a large majority of spectra even when conducted at low stringency, indicating that careful attention to stochastic variation in quantitative results in single-cell experiments is essential.

FUNCTIONAL GENOMICS/PROTEOMICS

Batut P J, Bing X Y, Sisco Z, Raimundo J, Levo M, Levine M S. Genome organization controls transcriptional dynamics during development. *Science* 375;2022:566-570.

A major surge in research on 3-D genome organization is presently under way. At the basis of much of this work is the premise that chromatin contacts, and more particularly chromosome looping, control gene expression. The role of interactions between enhancer elements and distant promoters is well established. Topologically associating domains (TADs) – loops delimited by boundary elements that interact *via* insulator proteins – are believed to be involved in facilitating looping between enhancers and promoters. This notion is congruent with the placement of enhancers and their cognate promoters within the same TAD. But disruption of boundaries between TADs often fails to affect gene expression, and TAD boundaries are often shared between cell types differing in the expression of genes within them. Insight into this dilemma is provided by the study of Batut *et al.* The authors investigate regulation of *Drosophila* homeobox (*Hox*) genes in the *Antennapedia* gene complex, which specifies segment identity during a 60-min period at the beginning of gastrulation by precise temporal patterning of gene expression. Using the Micro-C technique to define chromatin interactions at single-nucleosome resolution, they define TAD boundaries. But the authors also identify distal tethering elements (DTEs) that mediate specific focal contacts *within* TADs occurring between promoters and their distal regulatory regions. TAD boundaries are distinguished from DTEs by H3 histone modifications, and by the proteins that they bind – insulators, *e.g.* CTCF, at the TAD boundaries, and pioneer factors, *e.g.* Trithorax-like (Trl), grainyhead (grh), and Zelda (zld) at the DTEs. The authors systematically disrupt DTEs and TAD boundaries throughout the gene complex and observe the resulting changes in *Hox* gene expression by quantitative live-imaging methods. The data support a model in which the tethering elements foster enhancer-promoter interactions and mediate the fast transcriptional activation kinetics essential for normal development, whereas the TAD boundaries enforce regulatory specificity by preventing enhancer-promoter interactions across neighboring TADs. Transcriptional control is thereby exerted at two distinct levels of organization. This insight will be of interest to all investigators studying 3-D genome architecture in efforts to understand its influence on gene expression.

MACROMOLECULAR CHARACTERIZATION

Humphreys I R, Pei J, Baek M, Krishnakumar A, Anishchenko I, Ovchinnikov S, Zhang J, Ness T J, Banjade S, Bagde S R, Stancheva V G, Li X-H, Liu K, Zheng Z, Barrero D J, Roy U, Kuper J, Fernández I S, Szakal B, Branzei D, Rizo J, Kisker C, Greene E C, Biggins S, Keeney S, Miller E A, Fromme J C, Hendrickson T L, Cong Q, Baker D. Computed structures of core eukaryotic protein complexes. *Science* 374;2021:eabm4805.

Following recent successes in the use of the neural networks AlphaFold and RoseTTAFold for prediction of protein 3-D structure, Humphreys *et al.* extend these methods to elucidate protein-protein interactions in multisubunit protein assemblies. Computational methods built upon patterns of coevolution between interacting proteins have long been used for this purpose in prokaryotes, although they work less well in eukaryotes because coevolutionary signals are weaker in more complex proteomes. The authors' approach is to combine a rapidly computable version of RoseTTAFold with the slower, more accurate AlphaFold to investigate interactions that are initially suggested by coevolutionary patterns. In a study of yeast proteins, interactions among which have been extensively studied, the authors first ascertain the accuracy of predictions based on 699 protein pairs for which experimental structures are available. They then construct models for 106 protein assemblies that had not previously been identified, and a further 806 assemblies that were known to exist but had not yet been structurally characterized. The biological information that can be derived from this methodology is illustrated by new insights into the mechanism of DNA repair by the Rad55-Rad57-Rad51 complex, and the hitherto uncharacterized structure and mechanism of the glycosylphosphatidylinositol transaminase (GPI-T) complex, which attaches GPI anchors to protein *C*-termini. Among the many areas of research expected to be stimulated by the new methodology are assignment of functions to orphan receptors and characterization of weaker, more transient interactions.

IMAGING

Wang L, Xing X, Zeng X, Jackson S R, Teslaa T, Al-Dalahmah O, Samarah L Z, Goodwin K, Yang L, McReynolds M R, Li X, Wolff J J, Rabinowitz J D, Davidson S M. Spatially resolved isotope tracing reveals tissue metabolic activity. *Nature Methods* 19;2022:223-230.

Wang *et al.* demonstrate the use of imaging mass spectrometry to assess spatially defined metabolic flux in organs and tissues by visualizing isotope labeling patterns in tissue sections from mice into which selected isotopically-labeled metabolites have been infused. Imaging is performed by matrix-assisted laser desorption-ionization (MALDI). *N*-(1-naphthyl) ethylenediamine dihydrochloride (NEDC) is chosen as matrix because of its low interference with metabolites and broad metabolite coverage. Spatial resolution in this methodology is limited to ~50 µm. ^{13}C- or ^{15}N-labeled metabolites are infused into living animals at rates slow enough to avoid perturbing endogenous pool sizes. Glycolytic, gluconeogenic, tricarboxylic acid (TCA) cycle and amino acid metabolism flux are assessed. Metabolites chosen for monitoring are ones identified by LC-MS analysis of extracts of neighboring tissue sections that have molecular ion masses which wouldn't overlap with other metabolites in mass spectral images (*i.e.* in the absence of LC separation). Chosen metabolites must also be detectable with sufficient sensitivity for quantification of low fractional labeling values (~20-40%). The authors provide open-source software for visualizing labeling patterns. The pattern for each metabolite includes a spatial display of the different isotopic forms of the molecule, the average extent of labeling, and labeling normalized to tracer labeling levels measured by LC-MS of serum. Using these techniques, the authors visualize the differential flux through glycolysis and gluconeogenesis between kidney cortex and medulla; the differential usage of glutamine, citrate and fatty acids as TCA cycle substrates across the kidney; and differential utilization of glucose and 3-hydroxybutyrate between different brain regions in response to a high carbohydrate or high fat diet. Brain nitrogen sources are also found to differ between brain regions: branched-chain amino acids contribute more in the midbrain and ammonia in the thalamus.

Yuan Z, Zhou Q, Cai L, Pan L, Sun W, Qumu S, Yu S, Feng J, Zhao H, Zheng Y, Shi M, Li S, Chen Y, Zhang X, Zhang M Q. SEAM is a spatial single nuclear metabolomics method for dissecting tissue microenvironment. *Nature Methods* 18;2021:1223-1232.

To improve spatial resolution of metabolomic mass spectral imaging for discrimination of subcellular structures, Yuan *et al.* employ a different ionization method – secondary ion mass spectrometry (SIMS) using a Bi liquid metal ion gun (LMIG), which produces Bi_3^+ projectiles. This primary ion beam is rastered over an area of 400 x 400 µm of tissue section and spectra are collected for 256 x 256 pixels with a pixel diameter of ~1.5 µm and a pixel acquisition rate of 150 µs/pixel. Fragmentation of desorbed metabolite ions in SIMS is prevalent, making identification difficult. The authors'

principle aim, however, is simply to distinguish nuclei from different cell types and metabolic states. For the purpose of spatial image construction, they report using a nonlinear dimensionality-reduction algorithm that deploys an unsupervised learning approach to compress the numerous mass spectral signals in each pixel to a small number of characters that preserve the main structure present in the data. In this way they are able to discriminate cells in images of human liver sections corresponding to known subgroups of endothelial cells, known subgroups of hepatocytes, and Kupffer cells on the basis of metabolomic profiling of their cell nuclei.

Crainiciuc G, Palomino-Segura M, Molina-Moreno M, Sicilia J, Aragones D G, Li J L Y, Madurga R, Adrover J M, Aroca-Crevillén A, Martin-Salamanca S, Del Valle A S, Castillo S D, Welch H C E, Soehnlein O, Graupera M, Sánchez-Cabo F, Zarbock A, Smithgall T E, Di Pilato M, Mempel T R, Tharaux P-L, González S F, Ayuso-Sacido A, Ng L G, Calvo G F, González-Díaz I, Díaz-De-María F, Hidalgo A. Behavioural immune landscapes of inflammation. *Nature* 601;2022:415-421.

In the setting of acute inflammation, leukocytes undergo rapid and continuous changes in motility and morphology as they encounter and adapt to local stimuli and to anatomical barriers during evolution of an inflammatory response. Crainiciuc *et al.* transfer into recipient mice leukocytes expressing cytoplasmic fluorescent proteins that enable precise measurement of the morphology and movement of hundreds of individual cells at sites of experimentally induced inflammation. Fluorescent markers that distinguish neutrophils, dendritic cells and macrophages are employed. Cells are tracked by 2-photon microscopy *in vivo*. The authors collect data for 118 different parameters describing cell shape and movement, and select sets of parameters for construction of *t*-distributed stochastic neighbor embedding (*t*-SNE) plots to distinguish subgroups of cells differing in behavior. In this way, they study inflammation associated with influenza infection of the trachea, ischemia-reperfusion injury of muscle, and laser burn injury of skin. The authors note distinct behaviors elicited by different stimuli among subgroups of cells which suggest that analyses of this kind will contribute to the identification of pathogenic and non-pathogenic immune states. They pursue the analysis by comparing wild-type cells with cells bearing mutations at 24 gene loci for the purpose of identifying regulators of the measured behaviors. Interestingly, they identify the Src kinase Fgr, a protein known to be involved in regulation of cell shape and migration, as a modulator of fibrosis and vascular inflammation mediated by neutrophils. The methodology is hoped to

contribute to knowledge of inflammation in many clinically important settings, including chronic vascular inflammation and immune activity in tumors.

CELL BIOLOGY

Schraivogel D, Kuhn T M, Rauscher B, Rodríguez-Martínez M, Paulsen M, Owsley K, Middlebrook A, Tischer C, Ramasz B, Ordoñez-Rueda D, Dees M, Cuylen-Haering S, Diebold E, Steinmetz L M. High-speed fluorescence image-enabled cell sorting. *Science* 375;2022:315-320.

In flow cytometry, the capability to sort cells based upon the expression of markers detected by fluorescence signals has made fundamental contributions in cell biology. Schraivogel *et al.* now describe instrumentation capable of high-speed sorting of cells in a flow cytometer based upon characteristics of subcellular structure detected by high-resolution fluorescence imaging, and they illustrate applications of this capability. The instrumentation incorporates an implementation of fluorescence imaging using radiofrequency-tagged emission (FIRE) technology. FIRE is a high-speed imaging technique that uses radiofrequency multiplexing methods derived from the field of communications technology to speed up image acquisition. FIRE is combined with a traditional cuvette-based droplet sorter and new low-latency signal processing and sorting technology in an instrument commercially available from Decton, Dickinson, and Co., Franklin Lakes, NJ. The instrument is capable of imaging at rates up to 15,000 cells/s. Its capabilities are illustrated by sorting cells differing in a variety of characteristics, including cell shape, numbers of nuclei or nucleoli, Golgi integrity, and translocation of the NF-κB pathway component RelA from cytoplasm to nucleus upon cell activation by TNFα. Cells at different stages of the cell cycle that have not previously been amenable to separation in flow cytometry are isolated. The authors also combine image-enabled cell sorting with CRISPR-pooled screening to identify regulators of the NF-κB pathway. The technology will provide opportunities for non-specialist labs to undertake analyses of molecular mechanisms underlying cellular physiology hitherto beyond their capacity to address.

DRUG DEVELOPMENT

Rurik J G, Tombácz I, Yadegari A, Fernández P O M, Shewale S V, Li L, Kimura T, Soliman O Y, Papp T E, Tam Y K, Mui B L, Albelda S M, Puré E, June C H, Aghajanian H, Weissman D, Parhiz H, Epstein J A. CAR T cells produced in vivo to treat cardiac injury. *Science* 375;2022:91-96.

T cells engineered to express chimeric antigen receptors (CARs) directed toward malignant cells have shown clinical efficacy in the treatment of various hematologic cancers. The use of such genetically modified cells has received FDA approval for immunotherapy of several malignancies. The possibility that CAR T cells might be employed to treat further, non-malignant conditions is of considerable interest. Rurik *et al.* here explore the use of CAR T cells to control cardiac fibrosis. They target fibroblast activation protein (FAP), a marker of activated fibroblasts, cells that deposit collagen which stiffens the myocardium. Conventionally, CAR T cells are made *ex vivo* by retroviral transduction with a gene encoding a suitable CAR. The pool of transduced T cells is then expanded in culture for injection of cells into the recipient. Instead, Rurik *et al.* accomplish their results by injecting into the recipient a chemically modified mRNA packaged into lipid nanoparticles in a fashion analogous to mRNA vaccination against COVID-19. The authors incorporate into the nanoparticles an antibody against the T cell marker CD5 to direct them for uptake by T cells specifically *in vivo*. In addition to the saving of time and cost with this innovation, the production of the chimeric T cell receptor is limited in duration to the intracellular lifetime of the RNA. This ensures that potential toxicity of the treatment is limited. The authors demonstrate the disappearance of anti-FAP CAR T cells in mouse recipients 1 week after injection. They show that the treatment successfully reduces fibrosis and preserves cardiac function after experimentally induced injury. These results suggest that *in vivo* generation of CAR T cells by administration of well-characterized mRNA-bearing lipid nanoparticles may reward investigation for the treatment of diverse diseases in the future.

Sadybekov A A, Sadybekov A V, Liu Y, Iliopoulos-Tsoutsouvas C, Huang X-P, Pickett J, Houser B, Patel N, Tran N K, Tong F, Zvonok N, Jain M K, Savych O, Radchenko D S, Nikas S P, Petasis N A, Moroz Y S, Roth B L, Makriyannis A, Katritch V. Synthon-based ligand discovery in virtual libraries of over 11 billion compounds. *Nature* 601;2022:452-459.

Discovery of pharmaceuticals on the basis of their ability to bind to a designated drug target is increasingly facilitated by virtual screening of ligands for complementarity to the target’s known or predicted 3-D structure. Although the synthesis of compounds identified by virtual screening might not be readily feasible, this difficulty has been overcome by screening virtual databases of “readily accessible” (REAL) compounds comprised of building blocks (“synthons”) with validated reactivity from which drug compounds can be assembled. However, the problem remains that REAL libraries are growing in size and have already become huge, sometimes exceeding 10 billion drug-

like compounds. The computational cost of searching such libraries with standard docking algorithms is becoming prohibitive. Sadybekov *et al.* propose a hierarchical screening approach that takes advantage of the modular nature of the building blocks in a REAL library. It begins with docking of a prebuilt set of reaction scaffolds and corresponding synthons. Only the best scaffold/synthon combinations are subsequently used for docking to identify fully elaborated hits for chemical synthesis and experimental testing. In a search for novel cannabinoid antagonists, the authors prescreen an 11-billion-compound library in this way, leaving just 2 million compounds for iterative docking computation. Of the compounds subsequently synthesized and tested, 33% were hits (twice as many as standard virtual library screening), including 14 with sub-micromolar affinity. Analogues of the best compounds yielded further improvement. The authors achieved one K_i value of 0.9 nM with this methodology, and a selectivity measured by CB_2/CB_1 ratio of 50-200-fold. They estimate a saving in computational resources of ~100x. In principle, the same general approach is applicable to any docking problem.