A spoonful of sugar

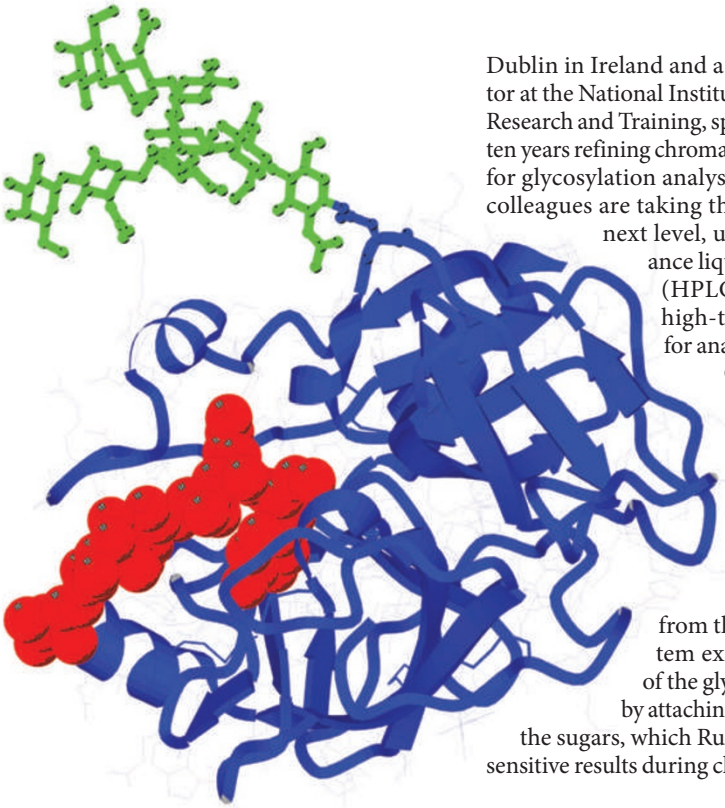
Carbohydrates are important in many biological processes, but the full extent of their distribution and function remains unclear. Advances in technology are now reveal those secrets. Nathan Blow reports.

Sugars occur in a variety of forms and locations throughout the human body. From those that are attached to proteins during glycosylation to the carbohydrates that decorate the surfaces of cells lining the lungs and digestive tract, the range of possible sugar conformations and glycoforms is tremendous. As a result, analysing carbohydrates is a tricky business for anyone interested in glycobiology.

The term glycobiology was coined in 1988 by biochemist Raymond Dwek at the University of Oxford, UK. Dwek used the phrase simply to emphasize the importance of relating sugars back to basic biology rather than just isolating and examining them outside of their biological context. Instead, he named a field that is thriving in its own right.

Today glycobiology is intertwined with fields such as immunology, virology, reproductive biology and drug discovery. "More people are starting to realize that sugars are not just there for protecting surfaces from proteolysis, but they have some functional role to play," says Ian Wilson, a structural biologist at the Scripps Research Institute in La Jolla, California.

But even as more researchers accept the importance of sugars in basic biology, many glycobiologists worry that the barrier to entry into their field remains too high, potentially delaying or hampering discovery and innovation. "The technical difficulty is so great now that many scientists are turned away," says Peter Seeberger, a chemist at the Swiss Federal Institute of Technology in Zurich. The solution, he adds, is to "lower the hurdle by providing access to technology more easily".



M. WORMALD, P. RUDD

A molecular model of a prostate-specific antigen with tumour-associated glycosylation (in green).

To those ends, Seeberger is trying to develop user-friendly automated solutions for complicated procedures such as the synthesis of complex carbohydrates.

Seeberger is not alone. Pauline Rudd, a professor of glycobiology at University College Dublin in Ireland and a principal investigator at the National Institute for Bioprocessing Research and Training, spent the better part of ten years refining chromatography approaches for glycosylation analysis. Now she and her colleagues are taking their approach to the next level, using high-performance liquid chromatography (HPLC) as the basis for a high-throughput pipeline for analysing glycans.

Glycoproteins featuring N-linked glycans are first immobilized either in gels or on membranes, and the glycans are then released using an enzyme that cleaves the sugars from the proteins. The system examines the patterns of the glycans on the proteins by attaching fluorescent labels to the sugars, which Rudd says offers highly sensitive results during chromatography.

Sweet analysis

The labelled sugars are run on a normal phase HPLC column and the resulting peaks are correlated to a pre-run dextran ladder, thereby assigning a 'glucose unit value' to each of the peaks. "We have a database that is automatically interrogated to give us a list of sugars that could have these particular glucose units," says Rudd. Using this information, a series of exoglycosidase digestions is performed and those data are fed back into the computer program to assign final structures.

The researchers recently installed an automated liquid-handling platform from Hamilton Robotics of Reno, Nevada, so that they could do their glycan analyses in 96-well plates. "One analysis will take about eight hours, so the aim is to get it done by the end of the shift," says Rudd.

Speed is important, Rudd notes, because the drug industry increasingly wants to monitor the glycosylation patterns of proteins. "When people want to achieve quality by design, they need to determine the optimal culture conditions and time for harvesting monoclonal antibodies," she says. "Therefore, they need to understand how the glycosylation changes over the course of production." Rudd says that her pipeline can test samples every hour, over a number of days or at different pH conditions to find those optimal points.

P. RUDD

Pauline Rudd is advancing high-throughput glycan analysis.

Elizabeth Higgins, chief executive and founder of GlycoSolutions in Worcester, Massachusetts, feels that a different factor is driving the drug industry's interest in glycan analysis. GlycoSolutions offers glycomics services and analyses, and last year worked on 20 different glycosylation analysis projects for various pharmaceutical companies. Higgins says that the analyses were largely done to meet regulatory requirements. "Most companies we work with are driven by getting data for the Food and Drug Administration," she says.

Another company working on high-throughput tools for analysing glycosylation patterns to aid drug development is Procognia, in Ashdod, Israel. Because many different glycoforms can exist, an extensive knowledge of glycosylation patterns and how they change during drug manufacturing is important for he development of biosimilar drugs, says Ilana Belzer, Procognia's vice-president of research and development. To tackle this issue, the firm has developed GlycoScope, a high-throughput workflow for glycosylation analysis platform that uses lectin (carbohydrate-binding protein) arrays and informatics tools to provide glycosylation fingerprints and glycan structures for glycoproteins.

By providing values of molecular weight that can be used to deduce initial structures, mass spectrometry (MS) is yet another approach to glycan analysis. "MS is very good at defining the sugar profiles of cell surfaces," says James Paulson, a glycobiologist at the Scripps Research Institute who studies glycan binding proteins that mediate cellular communication in the immune system. With additional isolation and fragmentation using either matrix-assisted laser desorption/ionization or electrospray ionization followed by tandem MS (MS/MS), researchers can deconvolute the configuration of sugars, a process very similar to protein sequencing.

The past few years have seen the arrival of many commercial programs and algorithms that assign glycan structures based on MS spectra. PREMIER Biosoft based in Palo Alto, California, sells SimGlycan, which uses MS/MS data to query a database of more than 8,000 theoretical glycan fragmentation patterns to generate a list of probable structures. Developers at the Palo Alto Research Center have designed new software packages that identify and annotate glycopeptides from a combina-

SURFACE SENSING

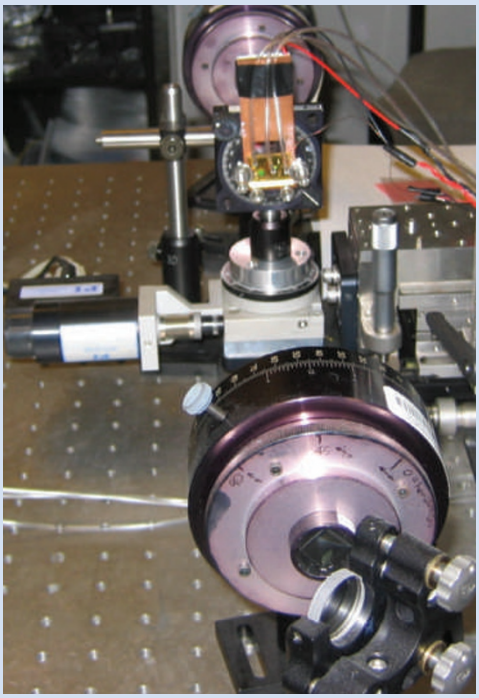
G. HWANG

In recent years, glycan arrays have been used to identify the cell-surface sugars bound to by pathogens such as the flu virus. In parallel, companies such as CombiMatrix in Mukilteo, Washington, have developed specialized diagnostic instruments to identify pathogens, including tools to differentiate flu strains. Now researchers at the MITRE Corporation in Bedford, Massachusetts, and the University of California, San Diego (UCSD), have developed a method that could blur the line between these tow types of tool.

The device was described in the December issue of *IEEE Sensors Journal* by MITRE researchers Grace Hwang and Elaine Mullen and UCSD researchers Lin Pang and Y. Shaya Fainman[4]. It features an array developed at the UCSD with a gold surface that is perforated with nanometre-wide holes. A glycoprotein is attached to the gold surface inside the hole and the pathogen or carbohydrate-binding lectin is added. The instrument detects binding events through surface plasmon resonance (SPR), measuring fluctuations in electron density at the boundary between the metal and a dielectric surface.

Hwang and Mullen use microfluidic delivery channels to place the glycoproteins on the gold surface. When the proteins attach to the gold there is a detectable change in plasmon resonance. Once this reaches equilibrium, the pathogen or lectin is introduced using the delivery channel and any binding that takes place further changes the plasmon resonance. The device is reusable as acids can be used to break the glycan–lectin bonds and clean the array.

"The reason we wanted to use a plasmonic device was because plasmons are very sensitive to perturbations at the metal–dielectric interface," says Hwang. For studying pathogens such as the flu virus, sensitivity can be an issue. The binding of the virus to different oligosaccharides occurs in the low millimolar range says Ian Wilson, a structural biologist at the Scripps Research Institute in La Jolla, California. As a result, he notes, glycan arrays often need to amplify the fluorescence signal, which requires additional antibodies.



Surface plasmon resonance provides label-free methods to look at carbohydrate interactions.

Hwang and Mullen's system avoids this problem, as plasmon detection does not require fluorescence to measure binding interactions — potentially opening the instrument up to a wide range of sensitive interactions. At the moment, however, the researchers are still working to improve the device's sensitivity for detecting flu viruses — their calculations suggest that it should be possible to identify up to one million influenza particles per millilitre.

Unlike other glycan arrays, the SPR system doesn't need printing or linkers to attach sugar targets to the gold surface. "The disulphide bonds in the glycoproteins will typically break and then bind to gold spontaneously," says Mullen. The is helpful because when the sulphide bonds form with the gold surface, the oligosaccharides of the glycoproteins are oriented properly with their bioactive sugars projecting towards the medium.

Nevertheless, using glycoproteins in this way means that Hwang and Mullen have to choose carefully and be confident in the glycosylation patterns of the glycoproteins they use as their target. To help them, Mullen and her colleagues built a database called SugarBindDB (http://sugarbinddb.mitre.org). "We know which glycoproteins to choose by looking at our own database of pathogens and their specific sugar sequences," says Mullen. "Then we go to the GlycoSuite, a database of oligosaccharides, to determine which glycoprotein it was attached to when it was isolated and what organisms it came from."

Hwang acknowledges that it is challenging to identify potential glycoproteins that present only the sugars required for selective pathogen sensing. If the glycoprotein displays a mixture of sugars, then it could bind to non-pathogens. It is even more difficult to identify potential pathogen targets displayed on glycoproteins in human tissue, but she and Mullen think this is a challenge not just for their device, but for glycobiologists in general.

"I realized from discussions with other researchers that predictive tools to compute binding affinities between sugars and lectins do not exist today," Hwang says, noting that this is a gap in glycan research tools that does not exist for nucleic acids and proteins. But she thinks in time, as more biophysical information is gained about glycan structures and properties, glycan arrays will catch up. N.B.

tion of single and tandem MS data.

MS analysis works well early on, says Higgins, but can be dangerous when it comes to working out the exact sugar structure because researchers often make assumptions based on mass alone. But the main challenge in using mass spectrometry for glycan analysis is figuring out the linkages between sugars. This is complex because the system needs consistently to fragment the sugars at the correct point to show one sugar is linked to a certain position on another sugar, says Paulson. He adds that such consistency remains an issue.

Rudd says that the HPLC approach along with enzyme digests can identify specific sugar linkages. HPLC columns can resolve the sugars on the basis of their conformations, and each monosaccharide contributes a specific incremental value to the retention time of an oligosaccharide. The pools of released sugars are treated with enzyme arrays in which each enzyme is highly specific for a particular monosaccharide in a particular linkage. The sequence and the linkage between sugars can be determined simultaneously for all the sugars in the pool. To help researchers interested in using this approach, Rudd's group recently made available the database GlycoBase and the analytical tool AutoGU to aid in the assignment of provisional structures based on HPLC profiles[1].

Glycan arrays and bird flu

The hunt for specific binding partners to various branched sugars or sugar-binding proteins called lectins requires a higher-throughput system. This can be achieved using glycan arrays. First described in 2002, these arrays feature different oligosaccharides or polysaccharides printed on slides or held in wells on a plate. "I think that glycan arrays have been a spectacular success over the past few years," says Paulson.

Initially, the arrays contained relatively small numbers of glycans and were designed mainly to study the specificity of antibodies and carbohydrate-binding proteins. But Wilson is one of a number of researchers who realized that some of these arrays would prove useful for diverse applications relevant to their own research. He uses glycan arrays for studying how the influenza virus binds to cells.

THERMO FISHER SCIENTIFIC

Advances in mass spectrometry are improving analyses of sugar composition.

Some viruses, such as flu and HIV, attach themselves to host cells during the early stages of infection by binding to sugars on the cells' surface. Paulson, in fact, discovered in the 1980s that avian flu viruses recognize different sugar receptors from their human virus counterparts. For Wilson, glycan arrays offered a way to look in detail at the specificity of different flu strains for various sugars, especially the H5N1 strain of bird flu that emerged in 1997 as a worldwide health concern, as well as the strain that caused the human pandemic in 1918. "We have analysed 50 to 60 or maybe even more influenza haemagglutinin mutants on the array to look for how the 1918 and H5N1 influenza strains can convert from human-to-avian or avian-to-human receptor specificity," says Wilson. Work has gone far in explaining how mutations can change the sugars to which influenza strains bind, thereby interconverting the receptor-binding characteristics of avian strains and human strains.

Wilson thinks that glycobiology is brought to the attention of a much wider audience when researchers use tools such as glycan arrays to work on well-known microorganisms such as the flu virus (see 'Surface sensing'). "There are a lot of other uses for these arrays, but everyone understands flu and the risks of bird flu," he says.

The Consortium for Functional Glycomics (CFG), an effort funded by the US National Institute of General Medical Sciences, aims to provide unique resources for glycobiology research. Headed by Paulson, the consortium has expanded the number of glycans available for arrays. "There are now 480 glycans in the consortium library," says Paulson, "and they all have amino-terminal linkers that allow them to be printed on slides using standard robotics." The ease of generating and analysing these arrays is opening the field to ever more researchers who can now submit samples to the CFG for rapid analysis.

Commercial developers also make high-content arrays with both glycans and carbohydrate-binding lectins attached to the surface. Robotic Labware Designs in Encinitas, California, offers printing services for glycan arrays as well as a series of preprinted glycan arrays. QIAGEN, headquartered in Hilden, Germany, provides the Qproteome GlycoArray kit for glycosylation analysis. This array and analysis software, developed by Procognia, contains a series of specific lectins that bind different monosaccharides, which allows researchers to determine the pattern and relative abundance of specific glycosylation epitopes in a glycoprotein.

Although 480 glycans on one array might not seem impressive compared with DNA microarrays, which can contain over a million features, Paulson is unperturbed. For carbohydrate-binding proteins, which usually recognize and interact with the tips of glycans, 480 represents a reasonable approximation of the options.

ETH-ZURICH

Automated carbohydrate synthesis could speed glycobiology research efforts.

P. RUDD

Robotics are proving crucial in several high-throughput glycosylation analysis approaches.

"If you only consider the tips, or the last six or seven sugars, then it is a very finite number of structures, in the order of 500," he says.

The problem for glycobiologists is how quickly carbohydrate diversity can grow when those 500 structures are attached to different branches on a single N-linked glycan. "If you allow any one of those structures to occur on any one of the four branches, you have this huge number of structures that could theoretically exist," says Paulson. And this is where glycan arrays run into a wall — researchers want this level of diversity on their arrays to help them understand how proteins and pathogens bind sugars, but generating such a diversity of glycans can be difficult.

ETH-ZURICH

Peter Seeberger is working on fresh approaches to carbohydrate synthesis.

Although many groups still try to isolate sugars from natural sources to use in their research, most agree that improving synthesis methods and technology is essential to obtaining large quantities of diverse carbohydrates.

"I think that up to now, carbohydrate synthesis has been restricted to a relatively small group of experts who bring considerable technical knowledge to the table," says Seeberger. Even for experts, such synthesis can take a long time — weeks or even years when it comes to making complex carbohydrates or glycoconjugates. Seeberger and his group, along with a handful of other labs around the world, have been working to improve carbohydrate synthesis methods. Ultimately they hope to develop automated instruments that can synthesize carbohydrates much like DNA synthesizers currently produce nucleic acids.

There are currently two main approaches to carbohydrate synthesis: solid-phase or one-pot synthesis. In 2001, Seeberger and his colleagues described an automated system that uses solid-phase synthesis for carbohydrates[2]. A programmable one-pot synthesis approach, meanwhile, has been advanced by Chi-Huey Wong, from the Scripps Research Institute and Academia Sinica in Taipei, Taiwan, and his colleagues.

Cooking up sugars

In solid-phase synthesis, sugar building blocks are attached to a surface or a bead, which can be moved during the synthesis process to allow other monosaccharides to be added. The one-pot approach uses a computer program to determine which monosaccharides to place in a flask; the next reagent is added and the mixture stirred. This process is repeated until an oligosaccharide is obtained. "What you save is the different work-up steps that often take much more time to achieve than the actual synthesis," says Seeberger of the one-pot approach. He adds that in this sense, both approaches cut down on the purification and separation steps in carbohydrate synthesis.

Seeberger sees the building blocks as a big issue for both approaches. Unlike DNA, which has four nucleotide bases, or the 20 amino acids that comprise peptides, there are 10 common monosaccharides in humans and many more in bacterial systems. Even more vexing when it comes to synthesis is the potential for branching of sugars. For example, glucose can link to another sugar at two points in its structure — a 1–6 linkage or a 1–4 linkage. This means that two different building blocks must be available for synthesis, which adds another level of complexity. The synthesis of monosaccharide building blocks was advanced recently when Shang-Cheng Hung in Taiwan and his colleagues reported a selective one-pot synthesis approach for the synthesis of highly functionalized, differentially protected monosaccharides[3].

Some commercial companies are producing monosaccharide building blocks for chemical syntheses. Dextra Laboratories in Reading, UK, offers monosaccharides as well as various glycoconjugates and more complex N-linked oligosaccharides. And other companies such as Omicron Biochemicals of South Bend, Indiana, and GLYCOTEAM in Hamburg, Germany, offer carbohydrate chemical synthesis services.

Chemical synthesis is not the only route to obtaining synthetic carbohydrates — researchers can also take advantage of nature's methods. "Enzymatic synthesis is one approach the CFG uses and that has enormously accelerated the rate at which you can synthesize complex natural sugars," says Paulson. But the approach is limited by the number of glycosyltransferase enzymes needed to synthesize all the carbohydrates researchers may be interested in. The number of glycosyltransferases needed for synthesis can be almost as daunting as the number of monosaccharide building blocks in chemical approaches. For example, GlycoGene, a company based in Ibaraki, Japan, offers enzymatic synthesis services to researchers through the use of more than 180 different glycosyltransferases. For this reason, the CFG has merged enzymology and chemistry in the production of many of the sugars on its glycan array.

Although he is keen to see automated chemical synthesis up and running, Paulson sees gaps when it comes to the carbohydrates that can be synthesized with existing methods. "You cannot make everything you want now, although you can make some carbohydrates quickly and easily," he says. "The gaps are the key things and these might be what people are really interested in looking at."

Despite this, Seeberger still sees access to tools as the greatest challenge in glycobiology at the moment. "When you think about genomics and proteomics, you can sequence and you can synthesize," he says. "But those two things are still not generally possible in glycobiology."

The field of glycobiology is still finding its way 20 years after the word was first printed. Although advances in the analysis and synthesis of carbohydrates are leading to fresh insights, much remains to be discovered. But Dwek can sit back and take comfort in the knowledge that his word has blossomed into a field that continues to grow. "I think the future of glycobiology is very exciting," he says. ■

Nathan Blow is the technology editor for *Nature* and *Nature Methods*.

1. Campbell, M. P., Royle, L., Radcliffe, C. M., Dwek, R. A. & Rudd, P. M. *Bioinformatics* **24,** 1214–1216 (2008).
2. Plante, O. J., Palmacci, E. R. & Seeberger, P. H. *Science* **291,** 1523–1527 (2001).
3. Wang, C.-C. *et al. Nature* **446,** 896–899 (2007).
4. Hwang, G. M., Pang, L., Mullen, E. H. & Fainman, Y. *IEEE Sens. J.* **8,** 2074–2079 (2008).