Journal of Cereal Science 50 (2009) 306–311

Contents lists available at ScienceDirect

Journal of Cereal Science

journal homepage: www.elsevier.com/locate/jcs

Review

Lectin and carbohydrate microarrays: New high-throughput methods for glycoprotein, carbohydrate-binding protein and carbohydrate-active enzyme analysis

Margherita Fais [a,b], Rositsa Karamanska [a,b], David A. Russell [a], Robert A. Field [b,*]

[a] School of Chemical Sciences and Pharmacy, University of East Anglia, Norwich NR4 7TJ, UK
[b] Department of Biological Chemistry, John Innes Centre, Colney Lane, Norwich NR4 7UH, UK

ARTICLE INFO

Article history:
Received 29 August 2008
Received in revised form
22 June 2009
Accepted 26 June 2009

Keywords:
Carbohydrate
Lectin
Antibody
Carbohydrate-active enzyme
Carbohydrate microarray
High throughput analysis

ABSTRACT

New technologies based on microarray methods have begun to feature widely in carbohydrate chemistry and biology. These 'glycoarray' techniques, which in a number of cases emulate what has been achieved with DNA microarrays, allow for high throughput, quantitative analysis of protein–carbohydrate interactions. Lectin, antibody and enzyme specificity have been evaluated with these new techniques, which also extend to the detection of viruses and bacteria, and serodiagnosis of infection. In the plant field, high throughput mapping of cell wall carbohydrate structures has been reported, giving information not only on the localisation of given glycans within a plant, but also allowing systematic comparison between mutants and species. This review outlines some of the basic principles of 'glycoarrays' and illustrates recent reports of their development and application.

1. Background

It has been noted that, in mammalian systems, "…proteins are often unduly portrayed as (the) decisive hardware. In contrast, and actually prominent among the biochemical systems to store information, carbohydrate epitopes of cellular glycoconjugates favourably combine high-density coding with strategic positioning, rendering them readily accessible for interactions with adaptor molecules" (Villalobo et al., 2006). In a similar sense, Varki asserted in the early 90s, concerning the 'biological roles of oligosaccharides – all of the theories are correct' (Varki, 1993). However, whilst carbohydrate structure and recognition are clearly central to biology (Ambrosi et al., 2005; Gabius et al., 2004; Sharon and Lis, 2004; Varki et al., 2009), in contrast to nucleic acids and proteins, oligo- and polysaccharide structures are not directly genome-encoded in a templated fashion (Turnbull and Field, 2007). In recent years the vision of glycobiologists (Dwek, 1996) has increasingly turned to the concept of the 'glycome' – the complete set of glycan structures expressed by specific cells, tissues or organisms (for instance, Kondo et al., 2006). In turn this has led to a need for analysis of larger numbers of glycan structures, and their cognate binding proteins, which has provided the impetus for development of new technologies with high-throughput potential. This presents a formidable challenge, given the inherent difficulties in analysis of glycan structure and function, which is reflected by the way that, to date, glycomics has lagged behind genomics and proteomics. A recent informatics analysis of the diversity of mammalian carbohydrate structures identifies that just 11 monosaccharide connections account for >75% of all linkages. Thus, the number of structural combinations found in mammalian 'glycospace' is much smaller than had been expected (Werz et al., 2007). In the plant kingdom, things are on an altogether grander scale of complexity: the pectic oligosaccharide rhamnogalacturonan II (RG-II) alone is composed of at least 12 different types of sugar residue linked together by more than 20 different glycosidic linkages (O'Neill et al., 2004).

Major advances in glycan analysis continue to be made with the aid of chromatography-based techniques, often in conjunction with mass spectrometry. However, sample preparation and separation times are still not optimal (Wada et al., 2007). For instance, the *N*-linked glycans present in 96 glycoprotein samples can be released and purified in 2 or 3 days and profiled in a further 2 days. Such techniques are state-of-the-art for biomarker applications in rheumatoid arthritis, for instance, using as little as 5 μl of patient serum (Royle et al., 2008). Many developments in glycobiology have been

* Corresponding author. Tel.: +44 1603 450720; fax: +44 1603 450018.
E-mail address: rob.field@bbsrc.ac.uk (R.A. Field).

0733-5210/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jcs.2009.06.010

driven by the largesse of the medical funding pot. Nonetheless, as with other omics and systems biology approaches, the field is now rapidly advancing on a number of fronts, including into plant science (Yuan et al., 2008) and its relationship to nutrition and health (Kussmann et al., 2008). In this mini-review we highlight recent developments in lectin and carbohydrate microarray technology, together with representative examples of recent applications and potential for future developments. Further details of consortia offering access to glycoarray screening, on a service or collaborative basis, can be found at: http://www.functionalglycomics.org/static/index.shtml and http://www.glycoarrays.org.uk/

2. Basic principles of lectin and glycan microarrays

2.1. Microarrays: why and how?

The draw of multiple parallel analysis of carbohydrate–protein interactions through the use of microarray-based methods requiring microgram quantities of sample (or on occasion much less) is attractive (for general reviews of this topic see: Horlacher and Seeberger, 2006; Laurent et al., in press; Liang et al., 2008; Park et al., 2008; Shin et al., 2005). In contrast to solution-based assays, where numerous parallel analyses are employed, the spatially defined location of ligands on an array allows for 10 s, 100 s, or potentially 1000 s of samples to be analysed in a single experiment. The array approach opens the way for discovering new carbohydrate-binding proteins, or protein domains, and for mapping the specificity of such proteins. Figs. 1 and 2 outline the format for the generation and interrogation of glycan and lectin microarrays. In most cases, the hardware required maps onto that already in common use for nucleic acid microarrays (*i.e.* commercial slides, arraying instruments and scanners) or for chemical biology/proteomics more generally (*i.e.* standard MALDI-tof mass spectrometry equipment).

2.2. The mechanics of glycoarray production

In contrast to lectin immobilisation, which can generally be achieved by direct covalent attachment to NHS-activated surfaces, a particular challenge for the development of glycan arrays lies in the need to be able to reliably immobilise a wide variety of chemically and structurally diverse glycans. Whilst typically only up to 10 s of lectins are arrayed at any one time, with glycans the objective is to print (at least) 100 s of different structures.

Of the readily available lectins, those from plants dominate. However, whilst these proteins can be powerful tools, many are themselves glycoproteins, which can complicate carbohydrate recognition studies. Efforts are being made to make recombinant, non-glycosylated lectins more accessible, as demonstrated by a recently reported route to the generation of lectin-fusion protein constructs suitable for arraying (Hsu et al., 2008).

A number of different approaches to surface attachment of glycans have been explored; this topic has been thoroughly reviewed (for instance, Culf et al., 2006; Park et al., 2008). In general, the methods fall into two categories: non-covalent or covalent immobilisation (Table 1).

Whilst direct physical adsorption presents a technically straightforward route to array production, its use is generally limited to polysaccharides, or multiply charged oligosaccharides such as pectins (Willats et al., 2002) and potentially glycosaminoglycans. One is therefore faced with the challenge of efficiently generating derivatives of reducing glycans that possess primary amines or other functional groups (for example Feizi and Chai, 2004; Song et al., 2008; Xia et al., 2005). In many cases, reductive amination is employed in this derivatization step, which then necessitates a further purification of derivatized glycans prior to array printing, which can add significantly to the sample preparation process. It should be noted that, in contrast to reductive amination (Feizi and Chai, 2004; Fukui et al., 2002; Song et al., 2008; Xia et al., 2005), where the reducing sugar ring is lost on derivatization, oxime formation appears to leave this ring intact. This is important where small glycans and/or proteins recognising glycan structure near the reducing terminus are being investigated (Liu et al., 2007). In order to remove the need for pre-derivatisation, surface-tethered *O*-alkyl hydroxlamines and acyl hydrazides have been investigated (Lee and Shin, 2005; Park et al., 2007; Zhi et al., 2008); such methods benefit from working with very high sample concentrations. A novel bi-functional spacer has been reported (Bohorov et al., 2006) that has selective reactivity towards free reducing glycans and a second functionality for coupling to NHS-ester activated surfaces. Glycans conjugated with this type of bi-functional linker have been arrayed and used to detect various lectins and antibodies. This methodology has been exemplified with naturally isolated *N*-glycans, *O*-glycans, milk oligosaccharides and bacterial polysaccharides at the microgram scale.

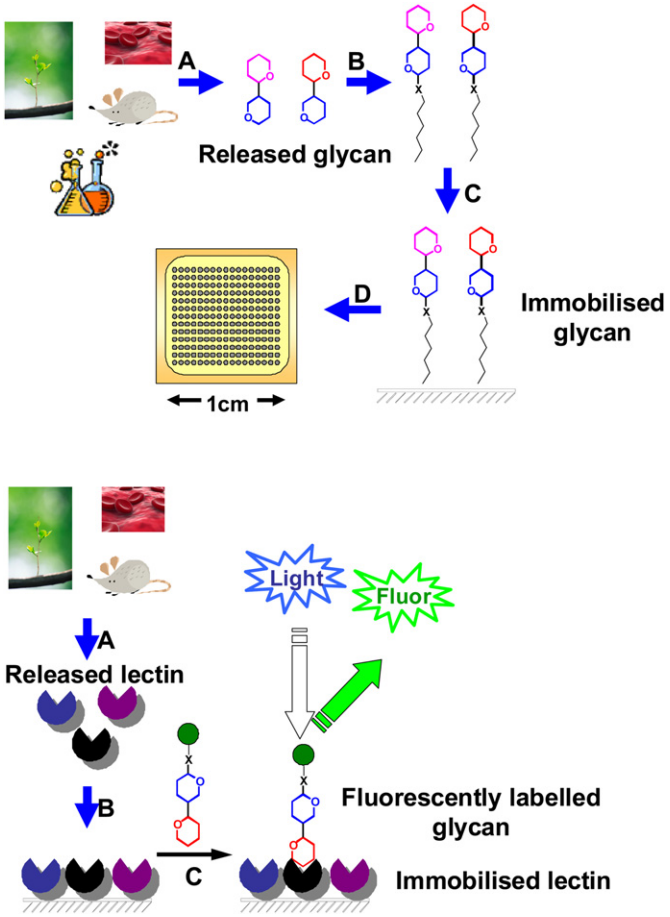

Fig. 1. Glycan and lectin microarrays. *Top* – Outline approach to the generation on a carbohydrate microarray. Naturally occurring glycan from tissues or cells, or materials arising from chemical synthesis are produced as reducing sugars (A) for derivatization (B) in a form that can be immobilised (C) in microarray format (D). Methods for glycan microarray analysis can be found in Fig. 2. *Bottom* – Outline approach to the generation of a lectin microarray and its use with fluorescently labelled glycans. Lectins, natural or recombinant, are generated in soluble form (A) for immobilisation in array format (B). Lectin array analysis is typically performed with the aid of fluorescently labelled glycans (C).