Carbohydrate Research 390 (2014) 9–19

Contents lists available at ScienceDirect

Carbohydrate Research

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

Minireview

Activity-based proteomics probes for carbohydrate-processing enzymes: current trends and future outlook

Keith A. Stubbs *

School of Chemistry and Biochemistry, The University of Western Australia, 35 Stirling Highway, Crawley, WA 6009, Australia

ARTICLE INFO

Article history:
Received 24 December 2013
Received in revised form 21 February 2014
Accepted 22 February 2014
Available online 4 March 2014

Keywords:
Activity-based proteomics probes
Carbohydrate-processing enzymes
Glycobiology
Glycoside hydrolases
Glycosyl transferases

ABSTRACT

Carbohydrate-processing enzymes are gaining more attention due to their roles in health and disease as these enzymes are involved in the construction and deconstruction of vast arrays of glycan structures. As a result, the development of methods to identify these enzymes in complex biological systems is of increasing interest. Activity-based proteomics probes (ABPPs) are increasingly being used in glycobiology to detect and identify functionally related proteins (and homologues) within a biological system. This review will describe the design of activity-based proteomics probes, provide examples of compounds that have been used to profile activity in the area of carbohydrate-processing enzymes, and give some future perspectives.

1. Introduction

One of the great scientific achievements of the 20th century was the completion of the sequencing of the human genome,[1] which has enabled scientists to further explore complex biological processes ranging from the substrates and products of biological reactions to the genes responsible for such activities. Furthermore, utilisation of this knowledge, especially in conjunction with animal models, has allowed for deeper investigations into the genetic causes of human diseases.[2–5]

On a more global scale, as genomes of other organisms are decoded, information regarding gene sequences has accumulated in such large quantities that scientists are now faced with the challenge of assigning functions to these genes in a systematic fashion. Bioinformatics is used to assign gene function through sequence homology, but currently there are limited methods for assigning the function of these genes at the biochemical level. This massive task has been handed to the field of proteomics and will allow scientists to gain a much greater understanding of biological systems and potentially uncover new enzymes involved in biochemical pathways.

One specific set of enzymes that is gaining more attention due to their role in health and disease are the carbohydrate-processing enzymes.[6] These enzymes are involved in constructing and deconstructing vast arrays of complex glycan structures. One point to note though is that, for example in humans, although the number of different monosaccharides found in glycans is actually small, it is a combination of the inherent type, number of units and the linkages between them that generates this complexity.[7] What further increases this complexity is that these structures are not static but are constantly being remodelled, degraded and built in response to different biological cues. In addition, the quantities of some glycans are quite small and so there are likely to be new glycan structures present in very low abundance, and their associated biochemical pathways still awaiting discovery.[8] As a result of this complex cellular environment, a large number of carbohydrate-processing enzymes have been found, and continue to be discovered, that play specific roles in various biological responses. The number of putative carbohydrate-processing enzymes that have been bioinformatically annotated from the vast number of potential gene products is enormous and this number continues to grow as evolutionarily diverse bacterial genomes are sequenced. The carbohydrate database CAZy[9–12] has been developed to display the genomic, structural and biochemical information of carbohydrate-processing enzymes. To date, there are 2434 bacterial, 73 eukaryotic, and 244 viral sequenced genomes that have been found to contain carbohydrate-processing enzymes, but many of these putative enzymes have only been identified through bioinformatics and have not had their bioinformatic annotation confirmed or their function identified. This is where activity-based proteomics (ABP) can attempt to fill the gap. This field utilises compounds known as activity-based proteomics probes (ABPPs)[13] to detect and identify functionally related proteins (and homologues) in a biological

* Tel.: +61 864882725.
E-mail address: keith.stubbs@uwa.edu.au

http://dx.doi.org/10.1016/j.carres.2014.02.023
0008-6215/

system, whether it be a cell extract, living cells or in an animal model. Such probes can be synthetically prepared and fine tuned for a protein activity of interest, which is critical to avoid undesirable biological hits. These probes also have the ability to address multiple scientific topics depending on the experiment being conducted,[14–16] such as:

(1) To confirm the annotation of putative enzymes and to discover new enzymes of similar function that have not been annotated as such.
(2) To identify a specific enzyme activity in organisms with sequenced or unsequenced genomes.
(3) In the absence of useful antibodies, ABPPs can be used to analyse and track the levels of active proteins of interest in a native system or in response to stimuli.

This review will describe the design of ABPPs, provide several examples of compounds that have been used to profile activity in the area of carbohydrate-processing enzymes, and offer some thoughts for the future.

2. Design of activity-based proteomics probes (ABPPs)

The design of ABPPs relies on three structural units, and on the way in which the probe is prepared as it is important to functional outcome.[14–17] The first unit, named the recognition element, is required for specific targeting of the probe to the protein(s) having the function of interest. In terms of carbohydrate-processing enzymes this unit usually resembles a substrate or inhibitor of the enzyme class being targeted (Fig. 1).[18,19] What has greatly helped the development of affinity-based probes for carbohydrate-processing enzymes is the wealth of research that has gone into the development of potent, selective inhibitors that are useful at the cellular level. It is desirable that the recognition element should also contain a reactive moiety that can form a covalent bond with the enzyme of interest. If the recognition element does not have such a reactive group, then another part of the overall probe should confer the ability to covalently label the protein of interest.

The second unit required, known as the linker, acts as a tether between the recognition element and the third unit, known as the reporter group. The linker usually functions to increase the solubility of the probe in water and to act as a spacer between the recognition element and the reporter group so that the latter does not interfere with binding to the target enzyme. However, the linker can have additional uses such as increasing selectivity of the probe[20] and as an aid in mass spectrometry analysis.[21,22]

The reporter group acts as a tool to aid the visualisation and enrichment of the proteins labelled by the probe. Early probes utilised radioisotopes as reporters, but today, commonly used reporter groups include biotin, rhodamine and fluorescein (Fig. 2A) with the latter two allowing for in-gel detection of labelled proteins or for studying the abundance and location of proteins in living cells. Biotin, due to its high affinity for streptavidin, is typically used as an enrichment and purification agent which allows for the removal of unlabelled proteins and the enriched proteins can then be identified by mass spectrometry.[23,24] The biotin unit can also be used as an epitope for Western blotting. There are some caveats surrounding the choice of the reporter group which should be considered, these include the possible attenuation of cellular penetration of the probe and also steric hindrance, as reporter groups tend to be cumbersome, which may be deleterious for binding of the probe to the proteins of interest. This latter point is especially critical for carbohydrate-processing enzymes. Many carbohydrate-processing enzymes act on the termini of glycoconjugates and have pocket-shaped active site architectures. This allows for these enzymes to retain an exquisite preference for their respective substrate, but this architecture restricts accommodation of the reporter group. One method for overcoming these problems lies in the success of bioorthogonal reactions in biological chemistry.[25,26] The probe design can include a bioorthogonal reactive group that can be subsequently tagged using a chemoselective reaction such as the Staudinger ligation[27], a copper(I) catalysed[28,29] or strain-promoted Huisgen

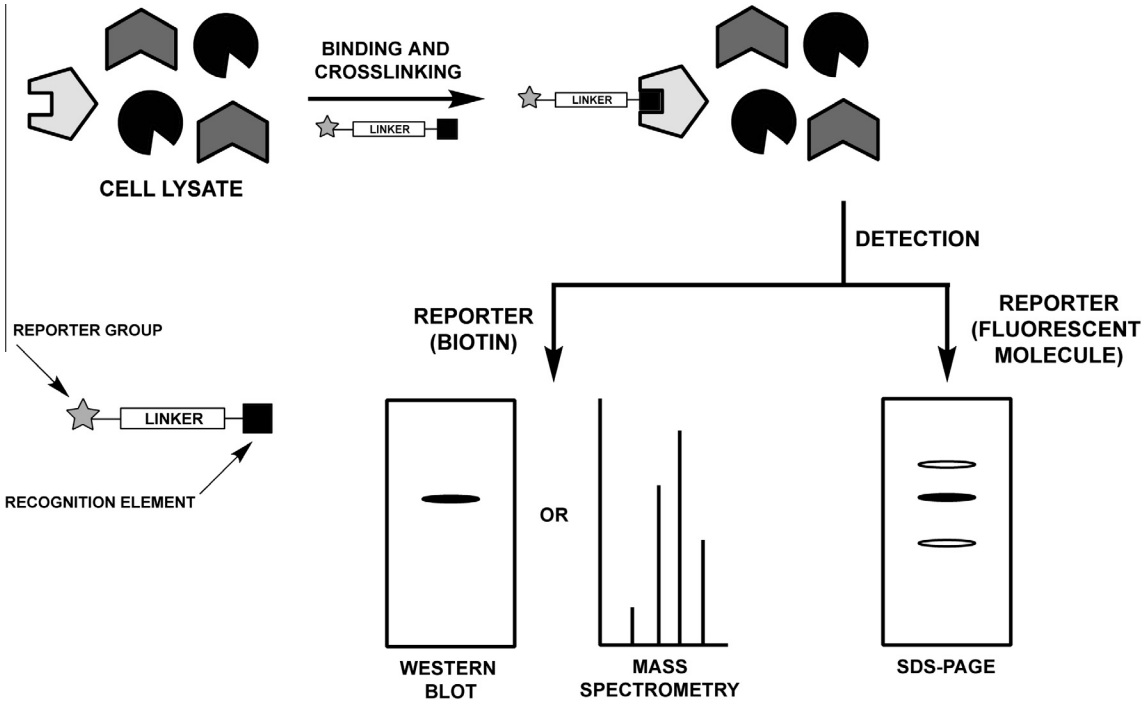

Figure 1. A typical experiment using an ABPP. A proteome from a cell lysate or live cells is treated with the activity-based probe. Key to the success of the probe is the covalent attachment of the probe to the enzyme(s) of interest. Depending on the reporter group, the labelled enzyme(s) can then be either directly visualised with SDS–PAGE followed by fluorescent imaging or can be purified and analysed by mass spectrometry, or the labelled enzymes can be detected by Western blot.