Biotechnology Advances 31 (2013) 17–37

Contents lists available at SciVerse ScienceDirect

Biotechnology Advances

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

Research review paper

Carbohydrate synthesis and biosynthesis technologies for cracking of the glycan code: Recent advances

Hynek Mrázek, Lenka Weignerová, Pavla Bojarová, Petr Novák, Ondřej Vaněk, Karel Bezouška *

Department of Biochemistry, Faculty of Science, Charles University Prague, Hlavova 8, CZ-12840 Praha 2, Czech Republic
Institute of Microbiology, Academy of Sciences of Czech Republic, Vídeňská 1083, CZ-14220 Praha 4, Czech Republic

ARTICLE INFO

Available online 28 March 2012

Keywords:
Cellular factories
Complex carbohydrates
Glycodrugs
Glycoforms
Glycoprotein hormones
Protein glycosylation
Structure activity studies
Vaccination

ABSTRACT

The glycan code of glycoproteins can be conceptually defined at molecular level by the sequence of well characterized glycans attached to evolutionarily predetermined amino acids along the polypeptide chain. Functional consequences of protein glycosylation are numerous, and include a hierarchy of properties from general physicochemical characteristics such as solubility, stability and protection of the polypeptide from the environment up to specific glycan interactions. Definition of the glycan code for glycoproteins has been so far hampered by the lack of chemically defined glycoprotein glycoforms that proved to be extremely difficult to purify from natural sources, and the total chemical synthesis of which has been hitherto possible only for very small molecular species. This review summarizes the recent progress in chemical and chemoenzymatic synthesis of complex glycans and their protein conjugates. Progress in our understanding of the ways in which a particular glycoprotein glycoform gives rise to a unique set of functional properties is now having far reaching implications for the biotechnology of important glycodrugs such as therapeutical monoclonal antibodies, glycoprotein hormones, carbohydrate conjugates used for vaccination and other practically important protein–carbohydrate conjugates.

© 2012 Elsevier Inc. All rights reserved.

Contents

1. Introduction

Protein glycosylation is among the most frequent and abundant protein modifications with numerous functional consequences on protein solubility, stability, folding, assembly into fully active complexes, and

* Corresponding author at: Department of Biochemistry, Faculty of Science, Charles University Prague, Hlavova 8, CZ-12840 Praha 2, Czech Republic. Tel.: +420 2 2195 1273; fax: +420 2 2195 1283.
E-mail address: karel.bezouska@natur.cuni.cz (K. Bezouška).

0734-9750/$ – see front matter © 2012 Elsevier Inc. All rights reserved.
doi:10.1016/j.biotechadv.2012.03.008

specific biological interactions. As such, it has proved to be of enormous importance in the biotechnological production of glycoprotein drugs and other commercially important protein carbohydrate conjugates. Despite all the progress in our knowledge on protein glycosylation and many practical achievements in this area, our understanding of the molecular code translating the effect of protein attached glycans into the functional protein space remains far from complete. In this review we summarize the most recent progress in this area. We particularly emphasize the practical implications of theoretical findings in the key directions of contemporary glycoprotein biotechnologies.

1.1. Defining the glycan code

The concept of the carbohydrate-based biological code as the third type of code in living cells has been around for more than three decades since it was introduced in the pioneering works by Nathan Sharon dealing with the structure of glycoproteins and their recognition by lectins (Sharon, 1980). Although the salient features of this unifying concept remain still true today, the term "glycocode" has not been well defined at chemical level. The term has been used in modern glycobiology and glycobiotechnology in three principally different contexts: (a) in the frame of the original concept as the third type of biological code besides those defined by nucleotides and amino acids (Jones and Larive, 2011; Ly et al., 2011); (b) as a distinct spatiotemporal expression patterns of glycans functionally important in certain biological processes such as the directing of olfactory sensory neurons during axonal pathfinding (Murrey et al., 2009) or at the fetomaternal interface supposed to be critical for pregnancy success or species hybridization (Jones and Aplin, 2009). The molecular nature of the glycocode described in (a) and (b) remains poorly defined (least understood), despite application of modern research technologies (El-Boubbou et al., 2010; Feizi and Chai, 2004; Pilobello and Mahal, 2007); (c) the "*bona fide*" third dimensional space in biological polymers occuring as a result of enzymatic or nonenzymatic post-translational modification of proteins, namely glycosylation or glycation. For the latter understanding of glycocode we coin a term "glycan code" in order to distinguish it from the "glycocode" described above, and from the "glycosylation code" controlling the mode of attachment of carbohydrate side chain to the polypeptide by the glycosylation biosynthetic machinery (e.g. the sequon Asn-Xxx-Ser/Thr provides a "code" for *N*-glycosylation). In contrast to the glycocode in (a) and (b), the glycan code can be conceptually well described as the sequence of structurally defined glycans attached to defined amino acids along the polypeptide chain. Such a code would then be different for each particular glycoform of the glycoprotein. Under this definition the analysis of the glycan code in most real world natural glycoproteins that are huge mixtures of individual glycoform species still remains extremely difficult. However, technologies are now emerging that allow to prepare, analyze and verify chemically defined glycoforms in sufficient amount enabling to address principal issue and questions concerning this well defined functional code (Pohl, 2004; Shental-Bechor and Levy, 2009). At the same time, however, the accummulating findings in this area of research are of an immense practical importance since they are critical for solving a number of biotechnological problems in areas such as glycoprotein folding, glycoprotein complex assembly, glycoprotein stability and solubility, and protection of glycoproteins against proteases, active radicals within the cell as well as to the number of pathologies related to these fundamental biochemical processes (neurodegenerative diseases, stress related diseases, aging, cancer etc.) (Martínek et al., 2010; Mitra et al., 2006; Petrescu et al., 2006). All these aspects are then of a particular relevance in biotechnologies related to glycoprotein drugs and practically important carbohydrate–protein conjugates.

The outstanding questions along the way to dissect the "third dimension" glycan code are fewfold. First, what is the minimum functionally relevant size of the glycan? Second, how can be functional importance of individual glycans influenced by the exact nature of protein–carbohydrate linkage? Third, what is the functional importance of the position of glycan along the polypeptide chain, often evolutionarily conserved, and is change in its position going to influence glycan function? Fourth, are carbohydrates functionally unique, or is it possible to replace them by other molecules of similar structure and/or chemical character? Using experiments performed so far with a limited number of model glycoproteins, we still do not have good answers to these principal questions. Considering the inherited complexity of the glycoproteins and their complexes, studies have been often limited to rather simple model systems, and we have to rely on molecular modeling as well as molecular dynamics to fill in for the missing data from the experiments. In order to reduce the complexity of the experimental system, the initial observations by Drickamer and Taylor (1998) proved seminal. Applying the evolutionary approach, they noticed that the function of protein attached glycans related to evolutionarily old functions such as protein folding, protein stability, and protection of the polypeptides are often confined to the biosynthetically conserved "core" region of the attached oligosaccharide while the peripheral extensions involving the evolutionarily newer complex glycan sequences are involved in cellular recognition or extracellular tagging. This hypothesis represents an important simplification for studies on glycoprotein glycan code driven by the suggestion that for the core region it "often does not matter which particular sugars are attached to the protein, the protein just has to be glycosylated on a particular amino acid residue" (Dwek, 1996). This phenomenon is probably due to the proximity of the core region of the oligosaccharide to the protein surface where it can most directly influence its chemical properties. By contrast, the terminal glycan elaborations are distally located in order to facilitate their interaction with the external environment (Drickamer and Taylor, 1998). Such a description of the glycan code has been supported by additional evolutionary considerations indicating that glycan diversity in complex multicellular organisms is driven by selection pressures of both endogenous and exogenous origin (Gagneux and Varki, 1999). An argument has been put forward that an exogenous selection pressures mediated by viral and microbial pathogens have played a predominant role leaving it difficult to appreciate and elucidate the specific endogenous roles of glycans within the organisms that synthesize them (Gagneux and Varki, 1999).

The recently published data now appear to reveal these endogenous roles much more clearly. A minimalist model based on the folding topology of a native protein in which each amino acid and sugar ring was represented by a single bead revealed that glycans can by means of entropic contributions modulate the free energy landscape of the polypeptide influencing important processes such as folding, stability, and protein complex assembly (Hoffmann and Flörke, 1998). More recently, Shental-Bechor and Levy investigated *in silico* the folding of 63 engineered SH3 domain variants that had been glycosylated with different number of polysaccharide chains at different sites on the protein's surface. The thermodynamic stabilization induced by glycosylation was coupled with kinetic stabilization (Shental-Bechor and Levy, 2008). A related experimental study concluded that specific, evolutionarily conserved protein–glycan contacts must play an important role in mediating the beneficial energetic effects on protein folding that glycosylation may confer (Price et al., 2010). Similar conclusions have been reached recently by Mrázek and colleagues using a more natural experimental system. A use has been made of fungal hexosaminidases, unique enzymes with a complex protein architecture in which the *N*-glycosylated catalytic units meets the *O*-glycosylated (namely *O*-mannosylated) propeptide in order to form the fully active tetramer (Plíhal et al., 2004; Ryšlavá et al., 2011). Hexosaminidase from *Penicillium oxalicum* together with some other fungal glycosidases has proved particularly useful in these studies since its propeptide is both N-terminally *N*-glycosylated

Download English Version:

https://daneshyari.com/en/article/10231528

Download Persian Version:

https://daneshyari.com/article/10231528

Daneshyari.com