Chemistry & Biology
Previews

Predictable Enzymatic Glycosylation

Cheng-Wei Tom Chang[1,*]
[1]Department of Chemistry and Biochemistry, Utah State University, 0300 Old Main Hill, Logan, UT 84322-0300, USA
*Correspondence: tom.chang@usu.edu
DOI 10.1016/j.chembiol.2009.06.001

An innovative approach for manipulating glycosyltransferase-catalyzed glycosylation has now been developed (Truman et al.). Created using a domain-swapping strategy, these chimeric glycotransferases have predictable substrate specificity and may lead to the breakthrough developments in the preparation of carbohydrate-containing molecules of biological interest.

The major advancements in biological sciences over the last century have been made in the protein and the nucleic acids fields of research, and, although continued contributions have been made to the understanding of carbohydrates biology, the field is now gaining momentum. Carbohydrates often play mysterious but indispensable roles in diverse biological events (Dwek, 1996); however, the structural complexity of carbohydrates has significantly hindered the progress. In order to reveal the insights of carbohydrates' function, the manipulation of sugar components on aglycons of interest, a strategy known as glycodiversification (Thibodeaux et al., 2007), glycorandomization (Yang et al., 2004), or neoglycorandomization (Langenhan et al., 2005), is often needed. Among the reactions involved in glycodiversification, glycosylation is the most pivotal step. Glycosylation is typically accomplished via chemical (Demchenko, 2008) or enzymatic (Melançon et al., 2006; Zhang et al., 2006) methods. While chemical glycosylation may be applicable to diverse natural and non-natural glycosyl donors and aglycons, the syntheses of these donors and aglycons could be difficult to conduct and accomplish. In addition, the scale-up production of desired glycoconjugates could be complicated by the additional protection and de-protection steps. In contrast, enzymatic glycosylation using glycosyltransferases (GTs) and the corresponding glycosyl donors appears to be more suitable for scale-up production of glycoconjugates, using, for example, bioengineering and fermentation processes. Nevertheless, the substrate specificity of GTs often significantly lowers the efficiency of incorporating foreign carbohydrates. Exploring intrinsic or engineered GTs' substrate promiscuity offers a possible step forward toward solving this problem, albeit impeded by the current inability to fully predict substrate acceptability of GTs.

A report focused on using a domain-swapping strategy based on the sequence comparison and the domain analysis of GTs now provides a rationale-based prediction of the substrate specificity of the designed chimeric GTs (Truman et al., 2009). In general, GTs contain two domains: the domain for binding of nucleoside diphospho-sugar (NDP-sugar), which functions as the glycosyl donors, and the domain for binding of aglycons (glycosyl acceptor). Truman et al. (2009) construct chimeric GTs by swapping the *N*- and *C*- domains of two natural GTs and evaluate their enzymatic behavior. Initial studies of these native and chimeric GTs confirm that the *N*-terminal domain controls acceptor (aglycon) binding, while the *C*-terminal domain controls sugar (from NDP-sugar) binding. The authors further explore the domain-swapping strategy by focusing on GtfAH1 as a powerful example of chimeric glycosyltransferases. The chimeric GtfAH1 contains *N*-terminal domain from GtfA and *C*-terminal domain from Orf1. GtfA and Orf1 transfer *epi*-vancosamine and *N*-acetylglucosamine (GlcNAc) onto desvancosaminyl vancomycin (DVV) and teicoplanin glucosaminyl-pseudoaglycon, respectively, and are highly specific toward their native substrates. The chimeric GtfAH1, however, shows much more tolerance with different substrates and transfer sGlcNAc onto various glycopeptides acceptors with comparable effiency in comparison to the parent GTs.

Glycodiversification using a biological approach is a power tool for elucidating the structure activity relationship (SAR) that is essential for making carbohydrate-containing biomolecules with practical applications. The reported domain-swapping strategy offers a rare opportunity of predicting the substrate specificity of the engineered GTs, which could serve as the foundation for biological glycodiversification. Coupled with X-ray structure studies, it may be even possible to introduce unnatural sugars onto the aglycons of interest. With all the exciting accomplishments, several technical aspects may need to be addressed prior to the development of mature technology. First, it requires a collection of *C*-terminal domains from various GTs in order to take full advantage of biological glycodiversification. For example, it would be extremely important and interesting to reproduce the same efficiency of chimeric GTs that can transfer L-sugars, deoxysugars, aminosugars, and sugars with branched-chain. Second, it is also critical to demonstrate the feasibility of glycosylating aglycons other than glycopeptides with the chimeric GTs. This aim, presumably, can be accomplished with the collection of *N*-terminal domains from various GTs. Finally, it will be crucial to examine the assumption that the "linker" sequences for the connection of *N*- and *C*- terminal domains can be applicable or universal to various classes of GTs.

Chemical glycodiversification may be more suitable for library construction of glycoconjugates with non-natural structural components, needed for further detailed SAR investigations. Nevertheless, the chemical syntheses represent a great hurdle for implementing such an approach in generating desired products on industrial scale. The work presented by Truman et al. (2009) is a significant step forward in developing methods for the large-scale preparation of carbohydrate-containing biomolecules. Further results regarding the above-mentioned

CHBIOL 1527_1528

technical issues of employing the domain swapping approach will certainly corroborate the merits of biological glycodiversification, and enable the practical and scale-up production of glycoconjugates free of the limitations imposed by nature.

REFERENCES

Demchenko, A.V., ed. (2008). Handbook of Chemical Glycosylation: Advances in Stereoselectivity and Therapeutic Relevance (Weinheim, Germany: Wiley-VCH Verlag GmbH & Co.).

Dwek, R.A. (1996). Chem. Rev. *96*, 683–720.

Langenhan, J.M., Peters, N.R., Guzei, I.A., Hoffmann, F.M., and Thorson, J.S. (2005). Proc. Natl. Acad. Sci. USA *102*, 12305–12310.

Melançon, C.E., Thibodeaux, C.J., and Liu, H.-w. (2006). ACS Chem. Biol. *1*, 499–504.

Thibodeaux, C.J., Melancon, C.E., and Liu, H.-w. (2007). Nature *446*, 1008–1016.

Truman, A.W., Dias, M.V.B., Wu, S., Blundell, T.L., Huang, F., and Spencer, J.B. (2009). Chem. Biol. *16*, this issue, 676–685.

Yang, J., Hoffmeister, D., Liu, L., Fu, X., and Thorson, J.S. (2004). Bioorg. Med. Chem. *12*, 1577–1584.

Zhang, C., Griffith, B.R., Fu, Q., Albermann, C., Fu, X., Lee, I.-K., Li, L., and Thorson, J.S. (2006). Science *313*, 1291–1294.

Oxygenase Catalyzed 5-Methylcytosine Hydroxylation

Christoph Loenarz[1] and Christopher J. Schofield[1,*]
[1]Chemistry Research Laboratory and The Oxford Centre for Integrative Systems Biology, University of Oxford, 12 Mansfield Road, Oxford, OX1 3TA, UK
*Correspondence: christopher.schofield@chem.ox.ac.uk
DOI 10.1016/j.chembiol.2009.06.002

Recent reports identify the oxygenase catalyzed production of 5-hydroxymethylcytosine as a modification to mammalian DNA (Kriaucionis and Heintz, 2009; Tahiliani et al., 2009). This discovery has potential far-reaching implications for epigenetic regulation and will stimulate efforts to identify new nucleic acid modifications.

Modifications to genomic DNA include methylation at the 5-position of cytosine (C) bases within cytosine-guanine dinucleotide (CpG) sites. Production of 5-methylcytosine (5mC) alters DNA structure without interfering with base pairing, and is recognized as an important epigenetic modification. Active gene promoters generally have a low abundance of 5mC, because promoter methylation can cause transcriptional repression. The formation of 5mC is catalyzed by DNA methyltransferases (DNMTs) and regulates gene expression patterns and initiates chromatin remodeling.

While plants employ pathways involving 5mC glycosylases and base excision repair to reverse cytosine methylation, attempts at identifying enzymes with 5mC reversal or modification activity in metazoans have been extensive but largely fruitless (reviewed in Ooi and Bestor [2008]). However, in recent years, other epigenetic modifications once believed to be static, such as histone lysyl methylation, have been demonstrated to be reversible.

The starting point that led Tahiliani and colleagues (Tahiliani et al., 2009) to the discovery of 5-hydroxymethylcytosine (hmC) as a modification to mammalian DNA was the identification of a previously unknown member of the 2-oxoglutarate (2OG) oxygenase family. 2OG oxygenases couple the two-electron oxidation of substrate to the oxidative decarboxylation of 2OG, and depend on Fe(II) as a cofactor. While bacterial 2OG oxygenases catalyze an extraordinarily wide range of oxidative chemistry including epimerization, ring closure, desaturation, and halogenation reactions, the activities of identified metazoan 2OG oxygenases have so far been limited to hydroxylation and demethylation via hydroxylation (Loenarz and Schofield, 2008).

Human 2OG oxygenases are involved in a diverse range of biological functions, including histone demethylation, collagen stabilization, DNA repair, hypoxia sensing, and fatty acid metabolism (Loenarz and Schofield, 2008). Although the catalytic domains of 2OG oxygenases share a common, albeit sometimes distorted, double-stranded β-helix (or jelly-roll) fold (first observed in antibiotic biosynthesis enzymes), they often have little overall sequence conservation, which impedes bioinformatic studies. However, approaches based on structurally informed sequence searches starting from 2OG oxygenases with known functions have led to the discovery of human orthologs, including enzymes with important biomedical functions.

In 1993, Borst and coworkers discovered that the DNA of African trypanosomes, protist parasites that cause African sleeping sickness, contained the unusual base β-D-glucosylhydroxymethyl-uracil, termed base J (Figure 1Ai). Their efforts to understand the nature of this modification and its role in expression led to the proposal of two trypanosomatid J-binding proteins JBP1 and JBP2 as thymidine hydroxylases of the 2OG oxygenase family (Yu et al., 2007). Starting from structurally informed iterative sequence profile searches using the JBP1 and JBP2 oxygenase domains, bioinformatic work from Tahiliani et al. (2009) led to the prediction of the three human enzymes TET1-3 as Fe(II) and 2OG oxygenases that catalyze 5-methylcytosine hydroxylation (Figure 1Aii). TET1-3 appear to be part of the wider TET/JBP fmaily, present minimally in metazoans, fungi, and algae.

CHBIOL 1527_1528