12. Ades, S.E., and Sauer, R.T. (1995). Biochemistry *34*, 14601–14608.
13. Connolly, J.P., Augustine, J.G., and Francklyn, C. (1999). Nucleic Acids Res. *27*, 1182–1189.
14. Banerjee-Basu, S., and Baxevanis, A.D. (2001). Nucleic Acids Res. *29*, 3258–3269.
15. Dovat, S., Ronni, T., Russell, D., Ferrini, R., Cobb, B.S., and Smale, S.T. (2002). Genes Dev. *16*, 2985–2990.
16. Jantz, D., and Berg, J.M. (2004). Proc. Natl. Acad. Sci. USA *101*, 7589–7593.
17. Eklund, E.A., Goldenberg, I., Lu, Y., Andrejic, J., and Kakar, R. (2002). J. Biol. Chem. *277*, 36878–36888.
18. Banerjee-Basu, S., Moreland, T., Hsu, B.J., Trout, K.L., and Baxevanis, A.D. (2003). Nucleic Acids Res. *31*, 304–306.
19. Marshall, S.A., Morgan, C.S., and Mayo, S.L. (2002). J. Mol. Biol. *316*, 189–199.

Chemistry & Biology, Vol. 11, July, 2004, DOI 10.1016/j.chembiol.2004.07.004

Cellular Addresses: Step One in Creating a Glycocode

In this issue of *Chemistry & Biology*, a library screening approach reveals at least four types of enzymes that attach galactosamine to build cell surface mucin-type glycoproteins [1]. A better molecular understanding of how these information-carrying oligosaccharides are created sets the stage for designing more selective inhibitors and potential therapeutics.

Cell surfaces are covered in diverse strings and branches of carbohydrate structures that create a kind of three-dimensional address system, or glycocode, which mediates interactions with a variety of biological components such as proteins or other cells [2, 3]. These addresses often change as cells grow and differentiate or become diseased [4]. Pathogens such as viruses and bacteria use such complex sugar structures to adhere to tissue for invasion of their hosts. The inhibition or promotion of these carbohydrate-based interactions serves as a new frontier in therapeutics for the treatment of conditions from cancer to viral and bacterial infections [5–7]. Unfortunately, very little is known about how these complex codes are assembled and regulated at the molecular level. In many cases, the actual changes to the cell surface architecture that occur with age or disease are not even known yet.

The most common protein-associated cell surface carriers of glycocodes in humans are the mucins or mucin-like proteins. *O*-glycosylation of the protein backbone at serine or threonine side chains with *N*-acetylgalactosamine (GalNAc) followed by addition of various other sugars creates densely clustered regions of carbohydrates that often eclipse the protein in size. The original hypothesis that these carbohydrate chains are initiated by only a few polypeptide GalNAc-transferases (ppGalNAcTs) is now being replaced with the realization that the system is far more complex (Figure 1) [8]. The human genome contains 24 putative ppGalNAcTs, and each isoform varies in its spatial and temporal regulation as well as tissue location [9]. A picture is emerging in which the highly glycosylated mucin domains are created by the action of several different ppGalNAcTs on the same protein backbone without simultaneous sugar chain initiation at every amino acid that is ultimately glycosylated [1, 8].

In this issue, the Bertozzi, Gerken, and Tabak groups report the first systematic study of the in vitro peptide or glycopeptide substrate requirements of eight members of the ppGalNAcT family [1]. These groups created a library of compounds based on a 13-amino acid segment of a rat mucin that includes every possible combination of sugar-modified threonine residues with up to four galactosamines. This library was incubated with each of eight glycosyltransferase isoforms and uridinediphospho-*N*-azidoacetylgalactosamine (UDP-GalNAz) under saturating conditions. The azide-labeled substrate was previously shown to undergo this enzymatic process and provided a convenient method to detect the newly added sugars after Staudinger ligation with a modified phosphine and standard signal amplification [10].

Analysis of the data reveals four basic types of ppGalNacTs. Some prefer peptides with no sugars or only one sugar already attached; glycosylation of nearby amino acids inhibits these so-called early transferases. The intermediate ppGalNAcTs prefer peptides with two, or to a lesser extent, three sugars attached. The late transferases glycosylate peptides that already have three or even four sugars attached nearby. Interestingly, the enzymes in these three categories have some redundant functions, in that two different transferases will glycosylate the same peptide. Therefore, the loss of function of one of these enzymes in vivo can perhaps be rescued in part by other isoforms. In contrast, two of the eight tested ppGalNAcTs form a fourth category, which contains very specialized functions that cannot be taken over by other isoforms. Indeed, the knockout of one of these latter transferases in a fruit fly is shown to be lethal [11, 12]; therefore, the effect of a knockout of the other specialized ppGalNAcT will be of particular interest. Should the other isoforms prove incapable of rescuing the function of this specialized transferase, the real power of the molecular approach to mucin biosynthesis studies reported in this issue will become apparent.

The next question is if the now known differences in substrate acceptance among these isoforms can be

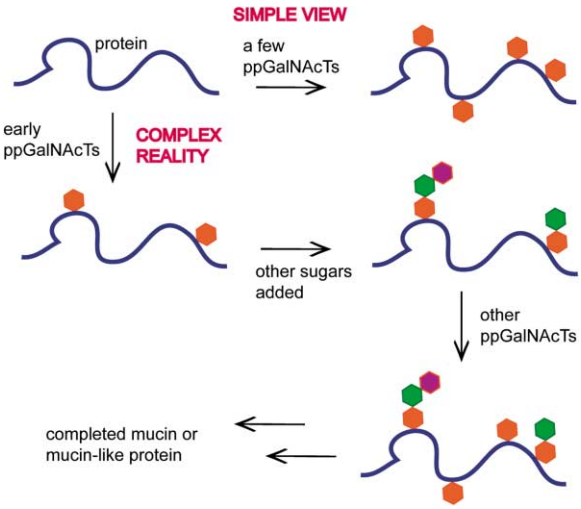

Figure 1. Proposed Mechanisms for the Modification of a Polypeptide by Isoforms of pGalNAcTs to Form Highly Glycosylated Proteins

used to design selective inhibitors of the enzymes involved in mucin biosynthesis. Selective inhibitors of each isoform of ppGalNAcT are valuable not only as research tools but as potential cancer therapeutics. Cell surface mucin expression is often altered in cancer cells and influences tumor metastasis [13]. In the March 2004 issue of *Chemistry & Biology*, the Bertozzi and Tabak groups reported screening of a 1338-member uridine-based library to find inhibitors of mouse ppGalNacTs [14]. Two compounds from the library could affect mucin-type *O*-glycosylation but not *N*-linked glycosylation in cells; however, the two library members did not show significant selectivity among ppGalNAcT isoforms. An understanding of the individual substrate specificities now provides the platform to build isoform-selective inhibitors of the enzymes that initiate carbohydrate chains on the mucin protein backbone.

Nicola Pohl
Department of Chemistry and
the Plant Sciences Institute
Gilman Hall
Iowa State University
Ames, Iowa 50011

Selected Reading

1. Pratt, M.R., Hang, H.C., Ten Hagen, K.G., Rarick, J., Gerken, T.A., Tabak, L.A., and Bertozzi, C.R. (2004). Chem. Biol. *11*, this issue, 1009–1016.
2. Gabius, H.-J., Siebert, H.-C., Andre, S., Jimenez-Barbero, J., and Ruediger, H. (2004). Chembiochem *5*, 740–764.
3. Davis, B.G. (2000). Chem. Ind. (London) 134–138.
4. Tong, L., Baskaran, G., Jones, M.B., Rhee, J.K., and Yarema, K.J. (2003). Biotechnol. Genet. Eng. Rev. *20*, 199–244.
5. Drinnan, N., and Ramsdale, T. (2003). In Molecular Pathomechanisms and New Trends in Drug Research, G. Keri and T. Istvan, eds. (London: Taylor & Francis), pp. 178–190.
6. Apostolopoulos, V., Plebanski, M., and McKenzie, I. (2003). In Immunobiology of Carbohydrates, S.Y.C. Wong and G. Arsequell, eds. (New York: Kluwer/Plenum), pp. 292–301.
7. Bertozzi, C.R., and Kiessling, L.L. (2001). Science *291*, 2357–2364.
8. Hassen, H., Bennett, E.P., Mandel, U., Hollingsworth, M.A., and Clausen, H. (2000). In Carbohydrates in Chemistry and Biology, Volume 3, B. Ernst, G.W. Hart, and P. Sinay, eds. (New York: Wiley-VCH), pp. 273–292.
9. Ten Hagen, K.G., Fritz, T.A., and Tabak, L.A. (2003). Glycobiology *13*, 1–16.
10. Hang, H.C., Yu, C., Pratt, M.R., and Bertozzi, C.R. (2004). J. Am. Chem. Soc. *126*, 6–7.
11. Ten Hagen, K.G., and Tran, D.T. (2002). J. Biol. Chem. *277*, 22616–22622.
12. Schwientek, T., Bennett, E.P., Flores, C., Thacker, J., Hollmann, M., Reis, C.A., Behrens, J., Mandel, U., Keck, B., Schafer, M.A., et al. (2002). J. Biol. Chem. *277*, 22623–22638.
13. McDermott, K.M., Crocker, P.R., Harris, A., Burdick, M.D., Hinoda, Y., Hayashi, T., Imai, K., and Hollingsworth, M.A. (2001). Int. J. Cancer *94*, 783–791.
14. Hang, H.C., Chong, Y., Ten Hagen, K.G., Tian, E., Winans, K.A., Tabak, L.A., and Bertozzi, C.R. (2004). Chem. Biol. *11*, 337–345.

Chemistry & Biology, Vol. 11, July, 2004, DOI 10.1016/j.chembiol.2004.07.003

Novel Mechanism for Priming Aromatic Polyketide Synthases

In this issue of *Chemistry & Biology*, a novel priming mechanism is proposed for aromatic polyketide biosynthesis, with an iterative type I polyketide synthase generating a starter unit primed for a type II polyketide synthase [6]. This novel priming system participates in hedamycin biosynthesis, a DNA alkylating agent.

Many natural products are bioactive, and polyketides constitute one of the most important families of natural products [1, 2]. Polyketides are widely distributed in plants, fungi, and bacteria and exhibit a wide range of biological activities of interest to the pharmaceutical industry (antibiotics, anticancer, antifungals, or immunosuppressive agents) and the agrochemical industry (insecticides or antiparasitic agents) [3]. From simply looking at the chemical structures of polyketides, there are no common structural features to suggest that these compounds all belong to the same family. However, when the various biosynthetic pathways for generating polyketide are considered, it is clear that they share a common thread: all polyketide biosynthesis involves the assembly of carbon chains from acyl precursors in a series of reactions catalyzed by a complex enzymatic system, the polyketide synthase (PKS). The reaction begins when PKS is primed by a starter molecule and then