



Published in final edited form as:

Nat Biotechnol. 2010 November ; 28(11): 1153–1156. doi:10.1038/nbt1110-1153.

Chinese hamster ovary cells can produce galactose- α -1,3-galactose antigens on proteins

Carlos J Bosques¹, Brian E Collins¹, James W Meador III¹, Hetal Sarvaiya¹, Jennifer L Murphy¹, Guy DelloRusso¹, Dorota A Bulik¹, I-Hsuan Hsu¹, Nathaniel Washburn¹, Sandra F Sipsy¹, James R Myette¹, Rahul Raman², Zachary Shriver², Ram Sasisekharan², and Ganesh Venkataraman¹

Ganesh Venkataraman: ganesh@momentapharma.com

¹Momenta Pharmaceuticals, Cambridge, Massachusetts, USA

²Harvard-MIT Division of Health Sciences and Technology, Koch Institute for Integrative Cancer Research, Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

To the Editor

Chinese hamster ovary (CHO) cells are widely used for the manufacture of biotherapeutics, in part because of their ability to produce proteins with desirable properties, including ‘human-like’ glycosylation profiles. For biotherapeutics production, control of glycosylation is critical because it has a profound effect on protein function, including half-life and efficacy. Additionally, specific glycan structures may adversely affect their safety profile. For example, the terminal galactose- α -1,3-galactose (α -Gal) antigen can react with circulating anti α -Gal antibodies present in most individuals¹. It is now understood that murine cell lines, such as SP2 or NSO, typical manufacturing cell lines for biotherapeutics, contain the necessary biosynthetic machinery to produce proteins containing α -Gal epitopes^{2–4}. Furthermore, the majority of adverse clinical events associated with an induced IgE-mediated anaphylaxis response in patients treated with the commercial antibody Erbitux (cetuximab) manufactured in a murine myeloma cell line have been attributed to the presence of the α -Gal moiety⁴. Even so, it is generally accepted that CHO cells lack the biosynthetic machinery to synthesize glycoproteins with α -Gal antigens⁵. Contrary to this assumption, we report here the identification of the CHO ortholog of *N*-acetyllactosaminide 3- α -galactosyltransferase-1, which is responsible for the synthesis of the α -Gal epitope. We find that the enzyme product of this CHO gene is active and that glycosylated protein products produced in CHO contain the signature α -Gal antigen because of the action of this enzyme. Furthermore, characterizing the commercial therapeutic protein abatacept (Orencia) manufactured in CHO cell lines, we also identified the presence of α -Gal. Finally, we find that the presence of the α -Gal epitope likely arises during clonal selection because different

© 2011 Nature America, Inc. All rights reserved.

COMPETING FINANCIAL INTERESTS

The authors declare competing financial interests: details accompany the full-text HTML version of the paper at <http://www.nature.com/naturebiotechnology/>

Note: Supplementary information is available on the Nature Biotechnology website.

subclonal populations from the same parental cell line differ in their expression of this gene. Although the specific levels of α -Gal required to trigger anaphylaxis reactions are not known and are likely product specific, the fact that humans contain high levels of circulating anti- α -Gal antibodies suggests that minimizing (or at least controlling) the levels of these epitopes during biotherapeutics development may be beneficial to patients. Furthermore, the approaches described here to monitor α -Gal levels may prove useful in industry for the surveillance and control of α -Gal levels during protein manufacture.

To determine whether CHO cells potentially express an *N*-acetyllactosaminide 3- α -galactosyltransferase-1 (Ggta1, E.C. 2.4.1.87)^{1,2}, we first sought to identify and, if present, subsequently clone this gene from Chinese hamster cDNA. Based on sequence alignment of mouse, rat, dog and cow Ggta1 orthologs, we were able to clone by PCR the full-length coding sequence for Ggta1 from cDNA pools derived from four different Chinese hamster tissues (brain, kidney, ovary and spleen) and confirm their homology to the rodent orthologs by direct DNA sequencing (**Supplementary Fig. 1**). The ovary and spleen consensus sequences are >99% identical to each other but do exhibit six single-nucleotide polymorphisms, resulting in one conserved amino acid difference. The remaining nucleotide differences between the orthologs are silent mutations. Molecular cloning of the Ggta1 gene from Chinese hamster cDNA also allowed us to clone directly from the CHO cell genome a critical section of the Ggta1 gene, including exons 8 and 9 and the 5-kb intervening intron adjoining them; exon 9 was of particular interest as it comprises the largest exon within the putative catalytic domain of the galactosyltransferase.

To determine whether the cloned gene product has the requisite features to be an active *N*-acetyllactosaminide 3- α -galactosyltransferase, we then completed a detailed analysis of the protein structure predicted by the gene sequence. A protein-BLAST search of the CHO Ggta1 exon 8–9 (CHOexo8–9) gene product showed highest sequence homology (>90%) with mouse, rat and bovine Ggta1. We further investigated the structural attributes of the putative enzyme through comparison to existing Ggta1 structures. Notably, the bovine form of Ggta1 is the most extensively characterized, including the availability of several X-ray co-crystal structures for both the wild-type and active site mutants of this enzyme in complex with the UDP-Gal donor and acceptor substrate⁶.

The structural model of the CHOexo8–9 gene product was constructed using the SWISS-MODEL interface (<http://swissmodel.expasy.org/>) in the ‘automated’ mode. The bovine α -1,3 galactosyltransferase (α 1,3GalT) was automatically selected as a template based on its high sequence homology to CHOexo8–9. The structural complex of the CHOexo8–9 with the donor and acceptor substrate (Fig. 1) was constructed with the help of the co-crystal structures of bovine α 1,3GalT with UDP-2F-galactose (UDP-2F-Gal; a donor inhibitor containing a fluorine atom at the 2-position of galactose sugar)⁶ and type II lactosamine acceptor substrate⁷. The critical active site residues in CHOexo8–9 are identical to those of bovine α 1,3GalT. Specifically, residues Asp225 and Asp227, along with the oxygen atoms of the disphosphate group of the donor UDP-Gal, are positioned to coordinate a critical divalent Mn²⁺ ion. Furthermore, highly conserved residues Trp249, Trp250, Trp314 and Trp356 are involved in positioning the acceptor substrate for galactosyl transfer. Notably, these residues have been proposed to define the acceptor substrate specificity of the α 1→3

galactosyltransferase family, given that the analogous residues are different in Gb3 synthase, Forssman antigen synthase and human blood group A and B synthases⁸. The presence of these residues in the CHOexo8–9 gene product strongly suggests that the cloned gene, if active, is truly an *N*-acetyllactosaminide 3- α -galactosyltransferase.

To confirm the enzymatic function of Ggta1, the catalytic domain of the coding sequence was recombinantly expressed in *Escherichia coli* as a fusion protein possessing an N-terminal maltose binding protein. The expression strategy for the catalytic domain was chosen based on previous work reported for the bovine Ggta1 ortholog⁹. Ggta1 substrate specificity was examined using structurally defined sugar acceptors as substrates including a galactose residue β -1,4 linked to *N*-acetylglucosamine (Gal β 1,4-GlcNAc), *N*-acetyllactosamine (LacNAc) as well as two naturally occurring N-linked glycans: asialo, monogalactosylated biantennary core substituted with fucose (NA2G1F); and asialo, digalactosylated biantennary core substituted with fucose (NA2G2F). The results are summarized in **Supplementary** Figure 2 and indicate that Ggta1 is enzymatically active. Ggta1 demonstrates a specific activity of ~ 16 nmol of product $\text{min}^{-1} \text{mg}^{-1}$ enzyme, using LacNAc as substrate, with an apparent K_M of 200 μM . Additionally, both N-glycan acceptors NA2G1F and NA2G2F are suitable substrates for Ggta1. Tandem mass spectrometry (MS/MS) fragmentation analysis of the product from the NA2G1F reaction unequivocally demonstrates the specific transfer of galactose from UDP-Gal onto the terminal galactose of NA2G1F (**Supplementary** Fig. 2); galactosyl transfer onto the terminal GlcNAc is not observed. Furthermore, detailed MS/MS analysis of the enzymatic products strongly suggests the presence of a Gal α 1 \rightarrow 3Gal moiety.

To determine the specificity of the Gal–Gal linkage in the product, we then analyzed ~ 150 μg of the di-galactosylated product (Gal–Gal β 1,4-GlcNA-2AB), generated from LacNAc, by gas chromatography/mass spectrometry (GC/MS; **Supplementary** Fig. 3). In this analysis, terminal galactose and 3-linked galactose are detected as the most prevalent signals observed by GC/MS. To further confirm the stereospecific arrangement of the linkage between the Gal1–3Gal linkage on the Gal1–3Gal β 1,4-GlcNA-2AB product, we subsequently treated the enzymatically generated, digalactosylated product with α -galactosidase and likewise analyzed it by LC-MS. The results clearly demonstrate the susceptibility of the Ggta1-catalyzed product to α -galactosidase (**Supplementary** Fig. 3). These data confirm the specific catalytic functionality of the cloned Chinese hamster Ggta1 gene as an α 1,3GalT necessary and sufficient for the biosynthesis of the α -Gal carbohydrate epitope in a CHO-based cell line.

To determine whether such antigenic species are actually present in pharmaceutical products, we examined in detail a representative biological product produced in CHO cells. Abatacept (cytotoxic T-lymphocyte antigen 4 (CTLA4-IgG)) was chosen because it is a prototypic biologic manufactured in CHO cells, resulting from the fusion of the ligand-binding domain of the CTLA4 and the human immunoglobulin heavy chain constant region.

In abatacept, we observed a glycan species with an abundance of $\sim 0.2\%$ of the mass of total N-linked glycans. The unknown species had a neutral mass of 2,070 Da, consistent with the glycan composition HexNAc₄ Hex₆ Fuc₁. Analysis of this structure by MS/MS sequencing

yielded fragment ions expected for a typical hybrid glycan structure (Fig. 2). However, we observed that the overall fragmentation pattern for this unknown was different from that typically observed for common hybrid structures. Specifically, examination of the ratio between the two B ions with m/z (mass-to-charge ratio) of 366 and 528 and the two Y ions with m/z of 1,542 and 1,704 suggests that the unknown is more likely to contain an α -Gal structure (Fig. 2b,c) than a hybrid structure. Liquid chromatography (LC)-MS/MS analysis before and after treatment with α -galactosidase showed a clear decrease in the peak containing the 2,070 Da species (retention time \sim 73.0 min) and a concomitant increase in a glycan with mass of 1,908 Da (retention time \sim 67.0 min), indicating loss of a single monosaccharide, ostensibly α -Gal. MS/MS analysis on the 1,908 Da product confirmed the expected HexNac₄ Hex₅ Fuc₁ glycan product (**Supplementary Fig. 4**). Furthermore, quantitative monosaccharide analysis after the treatment of the glycan mixture with α -galactosidase indicated that abatacept contains 370 pmol of terminal α -Gal per milligram of protein (30 mmol/mol of protein; **Supplementary Table 1**). Finally, to confirm and extend these results, we performed LC-MS/MS analysis of glycopeptides derived from a tryptic digest of abatacept, which revealed the presence of an α -Gal containing glycopeptide at one N-linked glycosylation site, namely asparagine 107 (**Supplementary Fig. 5**). The identification of this glycopeptide species demonstrates that the α -galactose glycan is linked to the CTLA4 domain of abatacept and is not derived from a contaminating glycoprotein.

To extend this analysis, the full-length gene encoding Ggta1 was transiently transfected into a previously developed CHO cell line clone that stably expressed CTLA4-IgG. The parental cell line was shown to be negative for expression of the endogenous Ggta1 gene. A highly functional copy of the murine ortholog of the Ggta1 gene was likewise transfected into the same cell line and used as a positive control. Recombinant CTLA4-IgG was purified from the spent media and the levels of α -Gal were determined by monosaccharide analysis as described in the **Supplementary Methods**. The results are summarized in **Supplementary Table 1**. Significantly, CTLA4-IgG expressed from CHO cells transfected with the Ggta1 expression vector contained the α -Gal structure, whereas the control cells mock-transfected with the 'empty' expression vector did not. The amount of α -Gal observed in the recombinantly expressed CTLA4-IgG was comparable to the amount measured for abatacept. Taken together, these results show that commercial biotherapeutics manufactured in CHO can, in fact, contain α -Gal and suggest that the enzyme product of the identified CHO Ggta1 has the appropriate activity to produce α -Gal-containing products.

Given the identification of the gene sequence and to better understand how the α -Gal epitope appears on CHO-based products, we sought to investigate what cell line backgrounds may have the potential to synthesize the product. To this end, we generated a series of stable clonal cell lines containing the CTLA4-IgG gene from different CHO backgrounds (CHOK1 and CHOdhr⁻, a line lacking dihydrofolate reductase activity) and used the primers described above to screen for Ggta1 transcript levels (Table 1). Interestingly, we observed clones both positive and negative for the Ggta1 transcript, even within the same cell line background, indicating that selection was critical to identify clones with low levels of Ggta1 activity. To correlate the Ggta1 transcript levels with actual output of the α -Gal structure, CTLA4-IgG product was isolated from each clonal population and

characterized for α -Gal content (Table 1). All clones that expressed Ggt1 at higher levels than background ($C_p < 38$) were also observed to contain α -Gal on the expressed protein product. Furthermore, the level of transcript trended with the level of α -Gal observed on the product. These data illustrate that the expression of the Ggt1 gene correlates to the expression of the α -Gal product which can arise during the course of clonal selection, and it suggests that any product expressed within these backgrounds has the potential to contain the α -Gal structure.

In conclusion, the data presented here show that, in contrast to what is currently assumed⁵, proteins manufactured in CHO cells may contain α -Gal epitopes. This may be the case especially for proteins with glycosylation sites outside the Fc domains of IgG molecules. The fact that humans contain circulating anti- α -Gal antibodies suggest that controlling the levels of this antigenic epitope during biotherapeutics development could have a positive impact on the clearance and safety profile of the drug. Although the levels of α -Gal in the CHO proteins analyzed in this study are lower in comparison to those typically observed in products derived from murine cell lines, it is important to monitor and control the levels of these species during CHO biotherapeutics development, especially because the specific levels of α -Gal required to induce anaphylaxis reactions are not well understood and are likely to be product dependent.

Acknowledgments

The Linkage analysis study was supported in part by the National Institutes of Health National Center for Research Resources (NIH/NCRR)-funded grant entitled 'Integrated Technology Resource for Biomedical Glycomics' (grant no. P41 RR018502-01) to the Complex Carbohydrate Research Center. We would like to thank E. Arevalo, S. Wudyka, T. Carbeau and S. Prabhakar for valuable technical assistance and J. Robblee for critical review of the manuscript.

References

1. Macher BA, Galili U. *Biochim Biophys Acta*. 2008; 1780:75–88. [PubMed: 18047841]
2. Larsen RD, et al. *Proc Natl Acad Sci USA*. 1989; 86:8227–8231. [PubMed: 2510162]
3. Sheeley DM, Merrill BM, Taylor LC. *Anal Biochem*. 1997; 247:102–110. [PubMed: 9126378]
4. Chung CH, et al. *N Engl J Med*. 2008; 358:1109–1117. [PubMed: 18337601]
5. Jenkins N, Parekh RB, James DC. *Nat Biotechnol*. 1996; 14:975–981. [PubMed: 9631034]
6. Jamaluddin H, Tumbale P, Withers SG, Acharya KR, Brew K. *J Mol Biol*. 2007; 369:1270–1281. [PubMed: 17493636]
7. Boix E, Zhang Y, Swaminathan GJ, Brew K, Acharya KR. *J Biol Chem*. 2002; 277:28310–28318. [PubMed: 12011052]
8. Zhang Y, et al. *Glycobiology*. 2004; 14:1295–1302. [PubMed: 15229192]
9. Shah PS, Bizik F, Dukor RK, Qasba PK. *Biochim Biophys Acta*. 2000; 1480:222–234. [PubMed: 11004566]
10. Domon B, Costello C. *Glycoconj J*. 1988; 5:397–409.

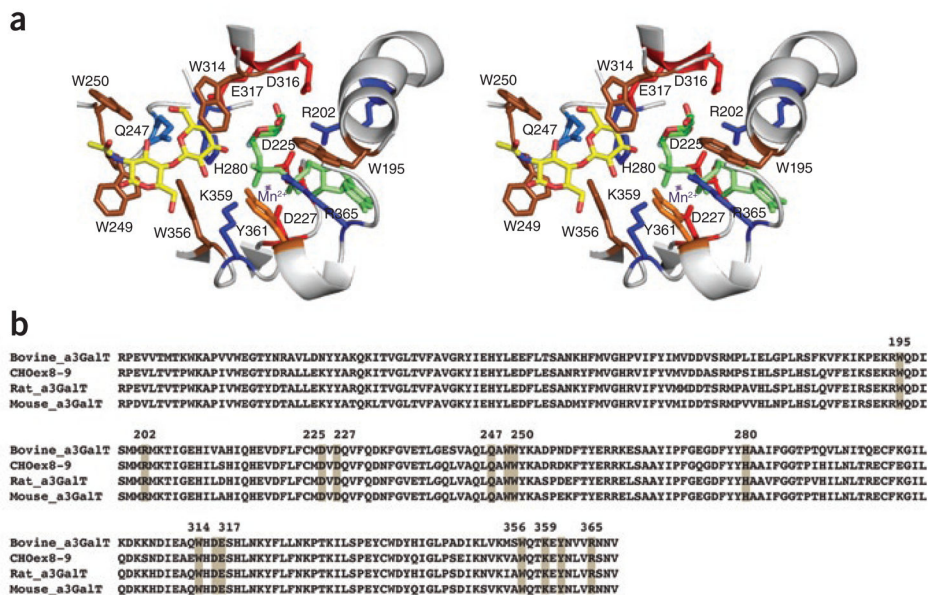


Figure 1.

Structural complex of CHO Ggt1 exons 8–9 gene product with UDP-2F-Gal donor and LacNAc acceptor substrate. **(a)** Shown in stereo view is the cartoon representation of the enzyme active site with the side chains of the key residues labeled using the single amino acid code and the numbering based on the bovine α 3GalT crystal structure (PDB: 1G93). The UDP-2F-Gal is shown in stick representation in green and the type II LacNAc acceptor substrate is shown in stick representation in yellow. The location of the Mn^{2+} cation (in purple) is obtained from the bovine α 3GalT crystal structure template used to construct the homology model. **(b)** Sequence alignment of the various α 3GalT with the highly conserved critical active site residues highlighted in gray.

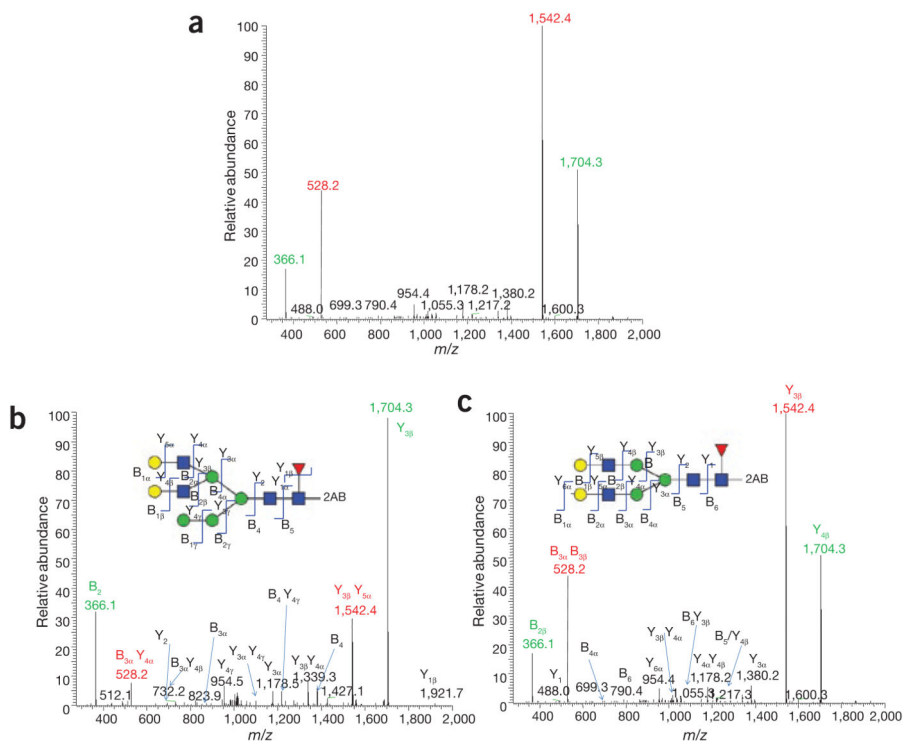


Figure 2.

MS/MS fragmentation analysis of N-glycan species observed in abatacept. (a–c) Comparison of the MS/MS fragmentation profile between an N-glycan species with a neutral mass of 2,070 Da observed in Orenca (a) and the fragmentation profile of an isobaric hybrid species (b) and an isobaric α -Gal-containing species (c) from our N-glycan fragmentation database.

Nomenclature for glycan structures: blue squares, GlcNAc; red triangles, fucose; green circles, mannose; yellow circles, galactose. The additional galactose is depicted in the lower branch to simplify the fragmentation nomenclature. Fragmentation nomenclature is based on ref. 10.

Table 1

Comparison of Ggta1 transcript levels from CHO clones expressing CTLA4-IgG and the corresponding amounts of terminal α -Gal expressed on the abatacept protein product

Clone	Average Cp	Picomol of α -gal/mg protein
Orencia (abatacept)	n.a	369.8
CHOK1 ⁻ clone 0003	31.5 \pm 0.4	404.1
CHOK1 ⁻ clone 1145	32.5 \pm 0.2	205.4
CHOdhfr ⁻ clone 0301	34.9 \pm 0.8	162.7
CHOdhfr ⁻ clone 1303	38.5 \pm 2.1	n.d.
CHOK1 ⁻ clone 1105	35.2 \pm 0.4	70.0

Average Cp values for controls (in the absence of reverse transcriptase) were Cp > 38. n.d., not detected.