REPORT

Knowing the glycosylation machinery of mammary gland cells: Potential and drawbacks for the expression of proteins of pharmaceutical interest

✉ José A Cremata, Raquel Montesino, Jorge R Toledo, Oliberto Sánchez, Elsa Rodríguez, María P Rodríguez

Center for Genetic Engineering and Biotechnology, CIGB
Ave. 31 e/ 158 y 190, AP 6162, CP 10600, Cubanacán,
Playa, Havana, Cuba
E-mail:jose.cremata@cigb.edu.cu

ABSTRACT

N-glycosylation analysis of recombinant human Erythropoietin (rhEPO) obtained by adenoviral transduction of goat mammary gland epithelial cells (GMGE) rhEPO-GMGE and in the milk of goats (rhEPO-Lc) was carried out by a combination of normal-phase HPLC (Amide-80) and ion exchange chromatography of the 4ABA labeled, enzymatically released N-glycans and further characterized by MALDI, ESI-MS and LC/MS. The most abundant N-glycans of rhEPO-GMGE are the monosialylated multiantennary core-fucosylated type, but fucosylation was also found in outer arms. However, rhEPO-Lc showed low branched, core-fucosylated, N-glycans. Here the charged N-glycans were found to be mostly a2,6-monosialylated with Neu5Ac or Neu5Gc at a ratio of 1:1, in contrast with the N-glycans from rhEPO produced in GMGE cells, where the charged glycans display the Neu5Ac. An important finding was the presence in rhEPO-Lc of biantennary N-glycans with lactosediamine (GalNAc-GlcNAc) ending arms that can be either neutral or sialylated, which is poorly represented in rhEPO-GMGE. This type of non-reducing terminal has not been found in rhEPO-CHO. These features differentiate the recombinant EPO expressed in the goat mammary gland from the classical EPO expressed in CHO cells, where the N-glycans are mostly fully sialylated multiantennary structures. These results emphasize that the difference between N-glycan populations of a given glycoprotein are sensitive to the cell-type and cell environment where they are cultivated.

Keywords: N-glycosylation, recombinant erythropoietin, mammary gland, goat milk, chromatographic profile, mass spectrometry

*I*ntroduction

The large-scale production of recombinant biopharmaceutical glycoproteins in the milk of transgenic animals is becoming more widespread due to very promising economic production processes. To date several proteins have been expressed in the milk of transgenic animals such as human lactoferrin in mice [1] and cows [2]; a1-antitrypsin [3] and C1 inhibitor [4] in rabbits; human antithrombin [5] and human tissue-type plasminogen activator [6, 7] in goats; human Factor VIII [8] and human Protein C [9] in pigs; and human erythropoietin in rabbits [10] and pigs [11]. However, the functionality and specificity of mammary gland glycosylation is, as yet, not completely understood compared to other mammalian cell lines. Recently, our group has expressed hEPO in the milk of goats (*Capra hircus*) by the adenoviral transduction of the mammary secretory epithelial cells [12].

Human erythropoietin (hEPO) is a glycoprotein hormone with three potential N-glycosylation sites at Asn-24, -38, and -83, and one O-glycosylation site at Ser-126. The glycosylation on hEPO is essential for *in vivo* activity [13], as the unglycosylated hormone shows full *in vitro* but no *in vivo* biological activity [14]. Thus, an alternative system for the production of biologically active hEPO requires the biosynthetic machinery of mammalian cells such as Chinese hamster ovary (CHO) cell cultures, currently used for the production of recombinant hEPO that is able to assemble the necessary tetrasialylated tetraantennary N-glycans to maintain the half-life of the protein [15, 16].

✉ Author of correspondence

An adenoviral vector, carrying the hEPO cDNA, enabled the expression of the recombinant glycoprotein at levels over 2 g/L in goat milk. However, the milk-derived recombinant hormone (rhEPO-Lc), showed a lower molecular weight, more basic isoforms and a very low *in vivo* hematopoietic activity compared to the described homologous rhEPO produced in CHO cell cultures [17]. Moreover, we have also established a continuous, non-transformed epithelial cell line (GMGE) from the primary culture of goat mammary gland and expressed rhEPO in this system (rhEPO-GMGE). The recombinant protein expressed in GMGE cells showed an N-glycosylation pattern that significantly differed from that of the classical EPO-CHO, with core and outer arm fucosylation and low sialic acid content. The glycoprotein showed a reduced hematopoietic activity [18]. The low hematopoietic activity of both rhEPO-GMGE and rhEPO-Lc can be presumably associated to differences in glycosylation with respect to rhEPO produced in the CHO cell line.

*R*esults and discussion

N-glycosylation of epithelial goat mammary gland cells in vitro *and* in vivo*: Human Erythropoietin as a model*

Purification and preliminary characterization

The hEPO expressed in the GMGE cell line was purified to homogeneity as described by Toledo *et al*. [19].

1. Nuijens JH, Van Berkel PH, Geerts ME, Hartevelt PP, De Boer HA, Van Veen HA, *et al*. Characterization of recombinant human lactoferrin secreted in milk of transgenic mice. J Biol Chem 1997;272:8802-7.

2. Van Berkel PH, Welling MM, Geerts M, Van Veen HA, Ravensbergen B, Salaheddine M, *et al*. Large scale production of recombinant human lactoferrin in the milk of transgenic cows. Nat Biotechnol 2002; 20:484-7.

3. Massoud M, Bischoff R, Dalemans W, Pointu H, Attal J, Schultz H, *et al*. Expression of active recombinant human alpha 1-antitrypsin in transgenic rabbits. J Biotechnol 1991;18:193-203.

4. Koles K, Van Berkel PH, Pieper FR, Nuijens JH, Mannesse ML, Vliegenthart JF, *et al*. N- and O-glycans of recombinant human C1 inhibitor expressed in the milk of transgenic rabbits. Glycobiology 2004; 14:51-64.

5. Edmunds T, Van Patten SM, Pollock J, Hanson E, Bernasconi R, Higgins E, *et al*. Transgenically produced human antithrombin: structural and functional comparison to human plasma-derived antithrombin. Blood 1998;91(12):4561-71.

6. Denman J, Hayes M, O'Day C, Edmunds T, Bartlett C, Hirani S, *et al*. Transgenic expression of a variant of human tissue-type plasminogen activator in goat milk: purification and characterization of the recombinant enzyme. Biotechnology 1991;9: 839-43.

This procedure ensures an efficient recovery of the different glycoforms of the rhEPO. The rhEPO-Lc and rhEPO-CHO were also purified by the same protocol. rhEPO-CHO was used as reference in this study (Figure1A)

Enzymatic deglycosylation with PNGaseF followed by SDS-PAGE showed hEPO-CHO with an apparent molecular weight of 20 kDa, corresponding to the polypeptide chain that still has the O-glycosylation site (Ser 126) occupied; rhEPO-GMGE showed similar behavior. However a band at 18 kDa, which is attributable to the completed deglycosylated protein backbone, was observed when analyzing the rhEPO-Lc (Figure 1C).

A comparison of the IEF patterns of the rhEPO produced in the three different systems but purified using the same procedure (Figure 1B), showed that rhEPO-GMGE contains a larger number of less acidic glycoforms compared to rhEPO-CHO as reference, while in the rhEPO-Lc the less acidic glycoforms prevail.

7. Ebert KM, Di Tullio P, Barry CA, Schindler JE, Ayres SL, Smith TE, *et al*. Induction of human tissue plasminogen activator in the mammary gland of transgenic goats. Biotechnology 1994;12(7):699-702.

8. Paleyanda RK, Velander WH, Lee TK, Scandella DH, Gwazdauskas FC, Knight JW, *et al*. Transgenic pigs produce functional human factor VIII in milk. Nat Biotechnol 1997;15:971-5.

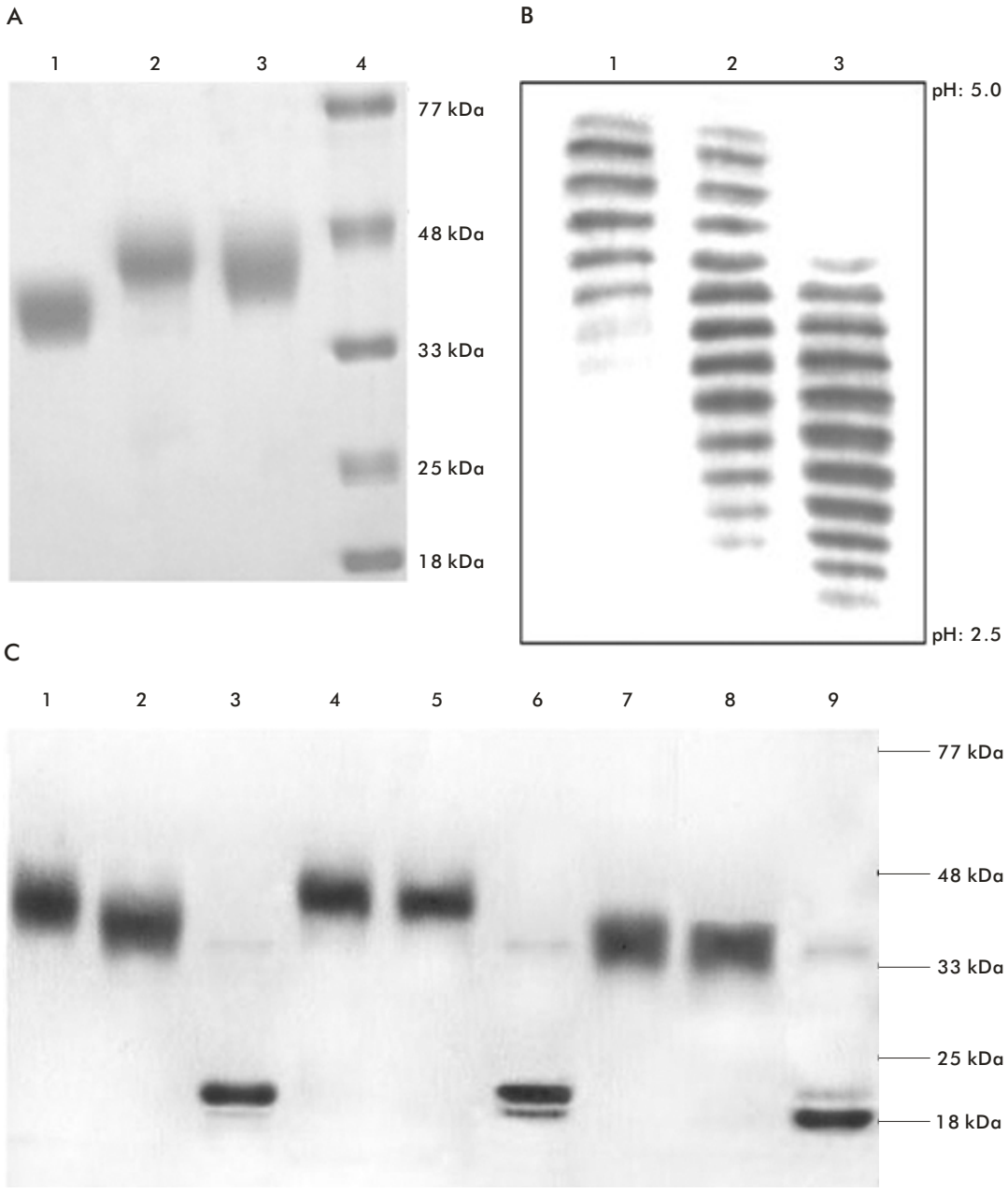


Figure 1. Electrophoretic migration variation and isoform distribution of rhEPO expressed in goat mammary gland epithelial cells *in vitro* and *in vivo*. (A) SDS-PAGE (12.5%) of different rhEPO samples (Coomassie blue-stained) Lane 1: rhEPO-Lc; 2: rhEPO-GMGE and 3: rhEPO-CHO, (B) Isoelectric focusing and immuno-identification with anti-EPO-HRP conjugate mAb. Isoelectrofocusing was carried out in the pH range 2.5-5.0. Lane 1: rhEPO-Lc; 2: rhEPO-GMGE and 3: rhEPO-CHO, (C) SDS-PAGE (12.5%) with Coomassie blue-stainning. Lanes 1, 4 and 7 rhEPO-CHO, rhEPO-GMGE and rhEPO-Lc respectively. Lanes 2, 5 and 8 rhEPO-CHO, rhEPO-GMGE and rhEPO-Lc desialylated with sialidase from *Salmonella typhimurium*. Lanes 3, 6 and 9 rhEPO-CHO, rhEPO-GMGE and rhEPO-Lc N-deglycosylated with PNGase F.

Desialylation with *Salmonella typhimurium* sialidase evidenced a lower sialic acid content in rhEPO-GMGE and even less in rhEPO-Lc compared to rhEPO-CHO (Figure 1C) as suggested by minor shifts SDS-PAGE corroborating the IEF results on the higher proportion of acidic isoforms in rhEPO-CHO.

The analysis of sialic acid content by reverse-phase HPLC C18 column with the fluorescence detection of a specific derivatization of a ketoacid with 1,2-diamine-4,5-methylen-dioxybencene (DMB-derivatives) of the N-glycans from rhEPO-Lc showed the presence of N-Acetyl- and N-Glycolyl-neuraminic acids (Neu5Ac:Neu5Gc) in a 1:1 ratio, in contrast to a major presence of Neu5Ac linked to the non-reducing terminal arms in glycans from rhEPO-GMGE and the reference rhEPO-CHO [17, 18].

Additionally, no changes in the protein backbone were detected by mass spectrometric sequencing. The disulfide bonds between Cys 7-161 and 29-33 were confirmed. This experiment also revealed that the three potential N-glycosylation sites (Asn 24, 38 and 83) were occupied [17]. Thus, the differences in apparent molecular weight and IEF isoform patterns between the three rhEPO molecules are consequences of differential N-glycans processing and modification through the transit by the Golgi apparatus in the secretory pathway.

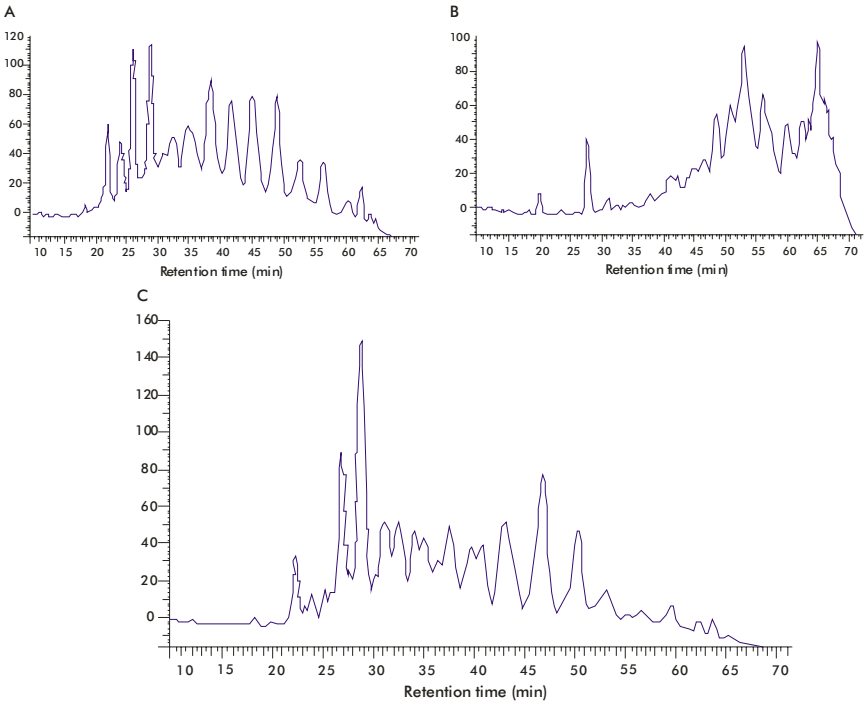


Figure 2. Amide-80 HPLC (normal-phase) profiles of the 4ABA-labeled N-glycan pool present in rhEPO expressed in mammary gland epithelial cells. A) rhEPO-CHO, B) rhEPO-GMGE and C) rhEPO-Lc. The profile of rhEPO-CHO is used as a reference.

Study of the N-linked oligosaccharides from hEPO expressed in epithelial goat mammary gland cells

N-glycans from rhEPO-Lc and rhEPO-GMGE were enzymatically released by PNGase F and labeled with 4-aminobenzoic acid and their profiles obtained in normal phase-HPLC in an Amide-80 column (Figure 2). Anion exchange (DEAE-HPLC) fractions were further analyzed by MALDI and ESI-MS as well as by LC/MS [17, 18] and the major N-glycan structures were characterized in each fraction.

Analysis of neutral and sialylated N-glycans from EPO-Lc

The 4ABA-derived *N*-glycans, neutral and charged fractions from DEAE-HPLC, were analyzed by LC-ESI-MS. Both fractions are composed by heterogenic N-glycan populations. The most abundant neutral 4ABA-labeled N-glycan ([M+H]+ ion at *m/z* 1991.0 was assigned to a monosaccharide relative composition Hex3HexNAc6dHex (Hex=Hexose, HexNAc=N-Acetyl Hexosamine y dHex= deoxy-Hexose). Collison Induced Dissociation of the precursor ion gave fragment ions which facilitated the N-glycan preliminary structure proposition, corresponding to a complex biantennary type oligosaccharide with GalNAc-GlcNAc non-reducing terminals, which are absent in rhEPO-CHO and definitely of low frequency in human glycoproteins. The most abundant oligosaccharides from the charged fraction have a similar monosaccharide composition but with a single Neu5Gc residue (Hex3HexNAc6dHex1Neu5Gc1). The molecular ion at *m/z* 2298.1 gave fragment ions that suggested a core fucosylated biantennary structure with an atypical GalNAc-GlcNAc disaccharide in each antenna, singly capped with one Neu5Ac or Neu5Gc residue. Thus, the relative monosaccharide composition of the charged N-glycan in rhEPO-Lc is described by the general formula Hex3-8HexNAc3-8dHex1-2Neu5Gc/Ac1 (Figure 3, [17]).

Analysis of neutral and sialylated N-glycan from hEPO-GMGE

Normal phase HPLC analysis on an Amide-80 column of the 4ABA-labelled N-glycan from rhEPO-GMGE give an unexpected profile with fractions eluting at a high retention time that are poorly represented in rhEPO-Lc and rhEPO-CHO. Sialylation was discarded as the main factor that produced such a high retention time, since the content of Neu5Ac is three fold lower than in the reference rhEPO-CHO. N-glycan structure analysis was carried out as described for rhEPO-Lc.

It was determined that the rhEPO-GMGE, the most abundant neutral N-glycan (Figure 4), can be described by the general formula Hex7HexNAc6dHexx where x=0-4. From this general formula we can assume that the N-glycans are mostly of the tetraantennary type. Certain compositions, e.g., Hex3HexNAc6dHexx and Hex4Hex NAc5dHexx suggested the presence of GaINAc-GlcNAc instead of N-Acetyl-lactosamine (Gal-GlcNAc) elements. The negative-ion ESI MS/MS spectrum of the N-glycan with a single charged precursor ion at *m/z* 1988.7 [M-H]- confirms the composition Hex3HexNAc6dHex3, which is similar to rhEPO-Lc. Hence, the N-glycan can be assumed to be an asialo, core fucosylated biantennary glycan with two *N,N*-diacetyllactosediamine arms but differing from rhEPO-Lc, this structure is not very abundant in the N-glycan population of rhEPO-GMGE. Additionally, Gala(1-3)Gal glycoepitopes could be associated to N-glycans with a monosaccharide relative

9. Van Cott KE, Lubon H, Gwazdauskas FC, Knight J, Drohan WN, Velander WH. Recombinant human protein C expression in the milk of transgenic pigs and the effect on endogenous milk immunoglobulin and transferrin levels. Transgenic Res 2001;10:43-51.

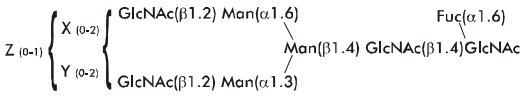
10. Massoud M, Attal J, Thépot D, Pointu H, Stinnakre MG, Théron MC, *et al*. The deleterious effects of human erythropoietin gene driven by the rabbit whey acidic protein gene promoter in transgenic rabbits. Reprod Nutr Dev 1996;36:555-63.

11. Park JK, Lee YK, Lee P, Chung HJ, Kim S, Lee HG, *et al*. Recombinant human erythropoietin produced in milk of transgenic pigs. J Biotechnol 2006;122: 362-71.

12. Toledo JR, Sánchez O, Montesino R, Fernández García Y, Rodríguez MP, Cremata JA. Differential in vitro and in vivo glycosylation of human erythropoietin expressed in adenovirally transduced mouse mammary epithelial cells. Biochim Biophys Acta 2005;1726:48-56.

13. Yuen CT, Storring PL, Tiplady RJ, Izquierdo M, Wait R, Gee CK, *et al*. Relationships between the N-glycan structures and biological activities of recombinant human erythropoietins produced using different culture condi-tions and purification procedures. Br J Haematol 2003;121:511-26.

14. Misaizu T, Matsuki S, Strickland TW, Takeuchi M, Kobata A, Takasaki S. Role of antennary structure of N-linked sugar chains in renal handling of recombinant human erythropoietin. Blood 1995;86: 4097-104.



	Hex	HexNAc	dHex	Neu5Ac	Neu5Gc
Neutral	3	6	1	-	-
	4	5	1	-	-
	5	4	1	-	-
Charged	3	6	1	1	-
	3	6	1	-	1
	4	5	1	1	-
	4	5	1	-	1
	5	4	1	1	-
	5	4	1	-	1

Figure 3. A schematic representation of the most abundant neutral and charged N-glycans from rhEPO-Lc. Above, a multi-structure proposition of the highly represented biantennary N-glycans. Below, the relative monosaccharide composition in terms of Hex, HexNAc, dHex, Neu5Ac and Neu5Gc.

composition Hex8-HexNAc6-dHexx. The analysis of the N-glycan with *m/z* 1609.9 [M-2H]2, from the charged fraction gives a tetra-antennary compound having the composition of Hex7HexNAc6dHex3 Neu5Ac1. It is fucosylated not only in the inner GlcNAc residue but also in the outer antennas as well, and monosialylated with a single Neu5Ac residue. In these N-glycans antennae fucosylation occurred at the GlcNAc residue but not on the Gal terminal residue thus giving glycoepitopes like Lewis X or Sialyl Lewis X. This event has not yet been observed in CHO cells. Thus, the N-glycan population of rhEPO-GMGE can be described by the general formula Hex3-9HexNAc2-8dHex1-5Neu5Ac1 [18].

The substantial functional difference of the epithelial mammary gland cell glycosylation machinery cultured *in vitro* or *in vivo* were summarized by the results described above and shown in Table 1.

The lactosediamine (GalNAc-GlcNAc) motive, which is not widely spread in vertebrates, is present on both rhEPO-Lc and rhEPO-GMGE, but in a lower amount in the latter. Another distinctive structural feature is the polyfucosylation of the outer arm N-glycan in rhEPO-GMGE, not yet described for recombinant proteins expressed in the milk of rabbits, goats, and cows or in the classical CHO system. The low sialic acid content is common of both the GMGE and the milk expressed proteins but Neu5Ac prevails in GMGE cells and a similar proportion of Neu5Ac and Neu5Gc is found in the rhEPO-Lc. A drawback of the epithelial mammary gland cell system for the expression of recombinant biopharmaceuticals is the presence of Gala(1-3)Gal non-reducing terminal epitope, which, even at low frequencies, is highly immunogenic in humans. The presence of this glycoepitope would lead to a faster blood stream clearance of the recombinant glycoprotein.

Determination of the hematopoietc activity of rhEPO expressed in epithelial goat mammary gland cells in vitro and in vivo

The *in vivo* hematopoietic activity of rhEPO-Lc and rhEPO-GMGE were compared with rhEPO-CHO. The rhEPO-Lc resulted hematopoietically inactive. Although, rhEPO-GMGE showed lower hematopietic

15. Jacobs K, Shoemaker C, Rudersdorf R, Neill SD, Kaufman RJ, Mufson A, *et* al. Isolation and characterization of genomic and cDNA clones of human erythropoietin. Nature 1985;313:806-10.

16. Lin FK, Suggs S, Lin CH, Browne JK, Smalling R, Egrie JC, *et al*. Cloning and expression of the human erythropoietin gene. Proc Natl Acad Sci U S A 1985;82: 7580-4.

17. Montesino R, Toledo JR, Sánchez O, Sánchez A, Harvey DJ, Royle L, *et* al. Monosialylated biantennary N-glycoforms containing GalNAc-GlcNAc antennae predominate when human EPO is ex-pressed in goat milk. Arch Biochem Biophys 2008; 470:163-75.

18. Sánchez O, Montesino R, Toledo JR, Rodríguez E, Díaz D, Royle L, *et* al. The goat mammary glandular epithelial (GMGE) cell line promotes polyfucosylation and N,N'-diacetyllactosediaminylation of N-glycans linked to recombinant human erythropoietin. Arch Biochem Biophys 2007;464: 322-34.

19. Toledo JR, Sánchez O, Montesino R, García G, Montañez M, Zamora PA, *et al*. High expression level of recombinant human erythropoietin in the milk of non-transgenic goats. J Biotechnol 2006;123: 225-35.

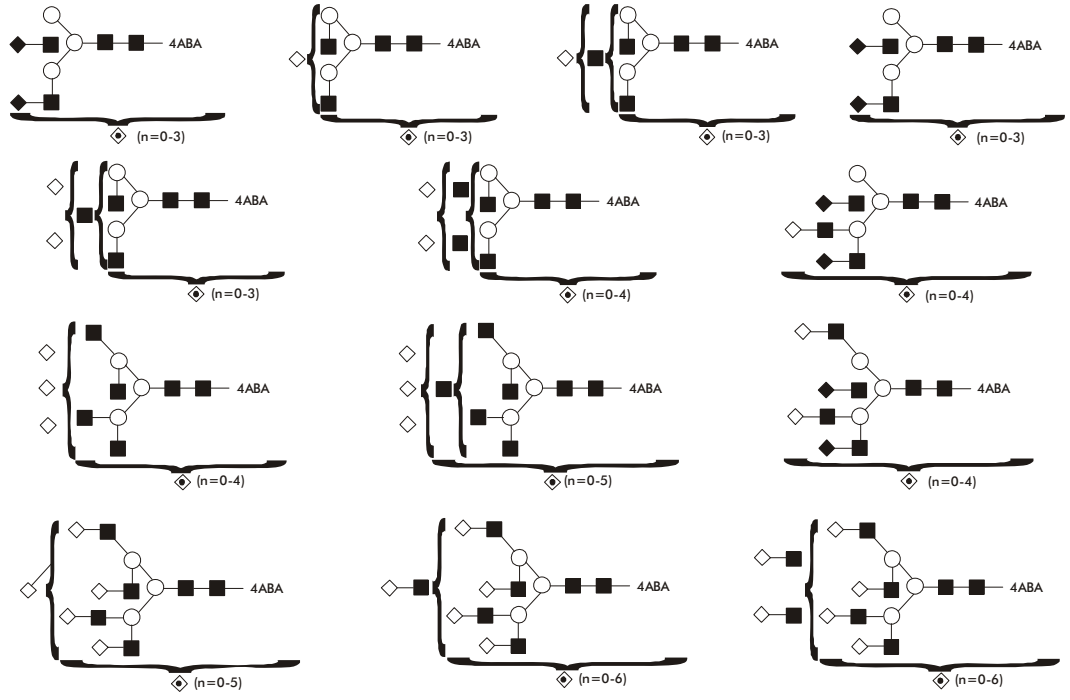


Figure 4. Diagrams of several possible neutral N-glycan structures, linked to rhEPO-GMGE. Structures were deduced from the monosaccharide composition determined by MS. The symbols for the glycan structures are as follows: GlcNAc, black square; GalNAc, black diamond; Man, white circle; Gal, white diamond; Fuc, diamond with a dot inside; α-linkage, dotted line, β-linkage, solid line; -, 1-4 linkage; /, 1-3 linkage; l, 1-2 linkage, «n» is the number of Fuc residues a-linked to the inner-most core and outer antennae GlcNAc residues.

Table 1. N-glycosylation features of the rhEPO expressed in mammary epithelial cells cultured *in vitro* or *in vivo*

Structures	rhEPO-CHO	rhEPO-Lc	rhEPO-GMGE
Source	CHO cells	Milk of goats	GMGE cell line from no transformed goats
Number of antennae	Tetrasialylated tetraantennary	Monosialylated diantennary	Monosialylated, polyfucosylated Di, Tri and Tetraantennary
Sialic acid content	Mean, >10 mol/mol proetin, over 95% Neu5Ac	Mean, <1 mol/mol protein, Nue5Ac:Neu5Gc 1:1	Mean, 6mol/mol protein, over 95% Neu5Ac
Neutral	Only Galβ1-4GlcNAc- terminals	Abundant GalNAcβ1-4GlcNAc- and Galβ1-4GlcNAc- terminals	Abundant Galα1-4(Fuc)GlcNAc- terminals
Charged	Neu5Acα2-3Galβ1-4 GlcNAc-	Neu5Gc/Acα2-6GalNAcβ1-4GlcNAc- and Neu5Gc/Acα2-6Galβ1-4GlcNAc-	Neu5Acα2-6Galβ1-4 (Fuc)GlcNAc-

activity than rhEPO-CHO, a significant increase of *in vivo* activity evidenced that the concentration of rhEPO-GMGE was increased in the assay (Figure 5). This result perfectly agrees with the sialic acid content of each molecule, since the decisive role this residue plays in the *in vivo* biological activity [13] is well known. Glycoproteins N-glycans capped with sialic acid residues are prevented from clearance through abolished or diminished interactions with the kidney Mannose Receptors and liver Galactose Repectorsm which explains the differences observed in the biological activity among the three recombinant hEPO molecules.

The difference in glycosylation observed in the rhEPO expressed in the epithelial mammary cell gland *in vitro* and *in vivo* is not yet clear and needs an exhaustive biological interpretation. The mammary gland during lactation functions like a bio-factory that secretes a huge amount of proteins to the cistern lumen. Thus, it is likely that during lactation there is a regulation of the glycosylation machinery that includes several glycosidases and glycosyltransferases that determine the final modifications of the N-glycans linked to the polypeptide backbone during the transit by the secretory pathway. These results, besides confirming the cell-type glycosylation dependence, reinforce the statement that changes of the host cell environment can lead to dramatic changes in the functionality of the overall glycosylation machinery.

Relevance of the study

The present report gives data on N-glycosylation glycoproteins expressed in mammary gland epithelial cells either *in vivo* or *in vitro*; the potential and drawbacks of this system with regard to the glycosylation machinery are discussed emphasizing those molecules of pharmaceutical interest. For the first time a model recombinant protein was expressed in mammary gland epithelial cells cultured *in vitro* or produced *in vivo* in the milk of goats and the N-glycosylation was fully characterized.

The milk of goats for the production of recombinant proteins is an attractive system due the very high expression levels reached (over 2 g/L). Thus, understanding the glycosylation features of the mammary gland pathway is essential when choosing the protein to be expressed, mainly those in which biological activity depends on a multiantennary and highly sialylated N-glycan pattern.

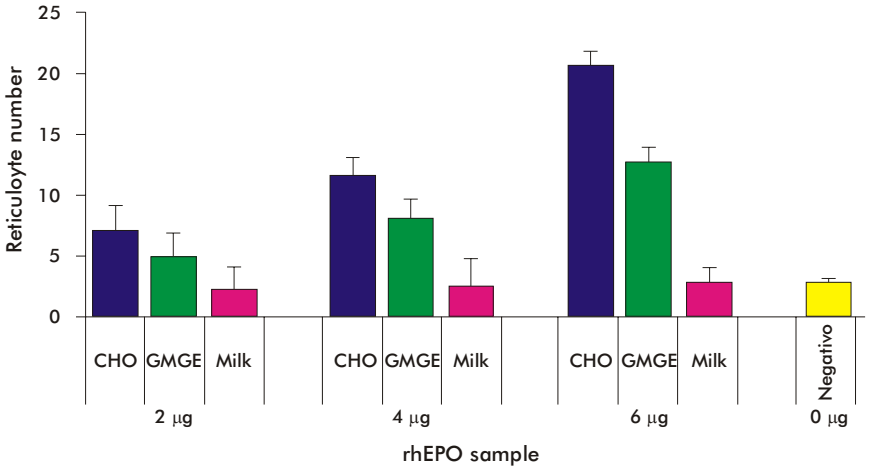


Figure 5. *In vivo* hematopoietic activity assay of rhEPO expressed in goat mammary gland epithelial cells *in vitro* and *in vivo*, rhEPO-GMGE and rhEPO-Lc respectively. Activity was determined by the increase in reticulocytes after subcutaneous inoculation into B6D2F1 normocytic mice, with doses of 2, 4 and 6 μg. The reticulocytes were counted 4 days after inoculation of the recombinant hormones. The figure shows the mean and S.D. of three experiments.

The results described above lead to a more rational design of the genetic manipulation of the goat mammary gland glycosylation machinery in order to have a more human-like glycosylation that may ensure the requirements for a structure-function relationship.

Conclusions

The characterization of protein model N-glycosylation expressed in goat mammary gland epithelial cells, either in cell culture (cell line) or body fluid (milk), gave significant differences in the glyco-sylation pattern that might be associated to the regulation of glycosyltransferases / glycosidases during lactation.

By knowing the molecular events that might modulate the glycoforms profile of glycoproteins expressed in the mammary gland it is possible to:

1. Decide whether the mammary gland is an appropriate bioreactor for the expression of a given glycoprotein if the glycosidic requirements for its biological function are already known.

2. Design a more precise glycosyltransferases/glycosidases genetic manipulation for the «humanization» of the mammary gland that ensures the requirements for a structure-function relationship in the case of glycoproteins for pharmaceutical use.