Rapid and sensitive methods for the analysis and identification of O-glycans from glycoproteins

Kozak, R.P.

Chapter 1

General introduction

Glycosylation

Glycosylation is the enzymatic process in which oligosaccharides (known as glycans) are added to proteins, lipids or other organic molecules. Protein glycosylation is the most common post-translational modification found in all domains of life. It has been estimated that more than 50% of human proteins are glycosylated.[1] Proteins are made up of combinations of 20 different amino acids and glycans are built from several monosaccharides including: galactose (Gal), glucose (Glc), fucose (Fuc), mannose (Man), *N*-acetylgalactosamine (GalNAc), *N*-acetylglucosamine (GlcNAc), and *N*-acetylneuraminic acid (NeuAc).[2] In this context protein glycosylation can display structural heterogeneity with respect and with respect to the glycan's structure. There are two main types of protein glycosylation, N-linked glycosylation and O-linked glycosylation (Figure 1).

The N-linked glycans are covalently attached to the amino group in the asparagine side chain18 in the sequence Asn-X-Ser/Thr (where X is any amino acid except proline). The glycan is later modified in the ER and Golgi apparatus, creating the three types of N-glycans found in mature glycoproteins: high mannose, complex and hybrid with the core structure $Man_3GlcNAc_2$.[3,4]

The O-linked glycans are covalently attached to the hydroxyl group of serine or threonine within a peptide backbone. A range of monosaccharides linked to serine or threonine have been found (e.g. *N*-acetylglucosamine, *N*-acetylgalactosamine, xylose, mannose and fucose). Extended structures from a core *N*-acetyloglucosamine also can be found and these are called mucin type O-glycans.[5-7] The focus of this thesis is on the development of analytical methods for analysis of mucin type O-glycosylation.

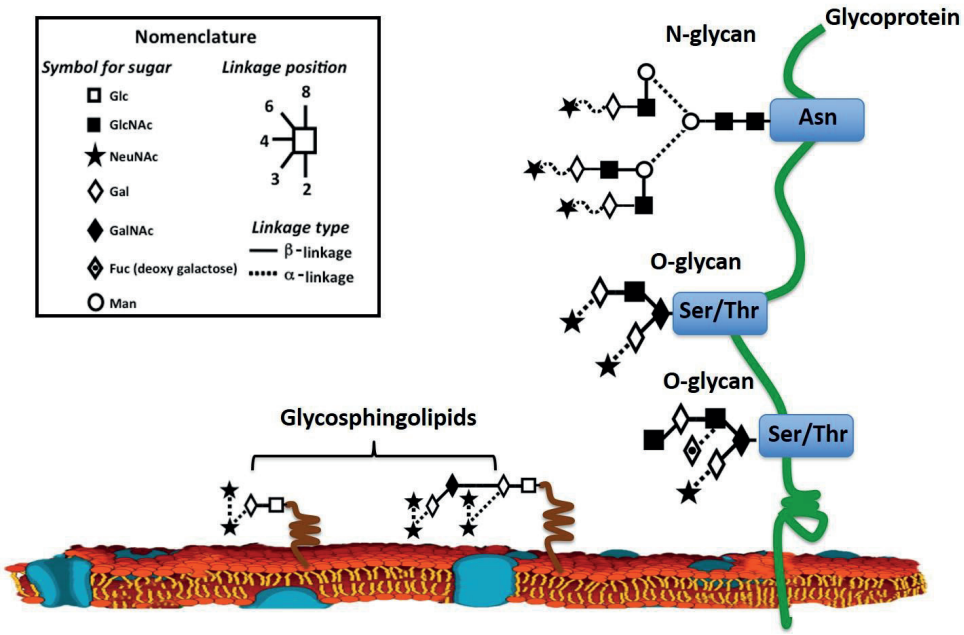

Figure 1. Types of glycosylation. Glycans are schematically represented according to Harvey et al.[8]

Biopharmaceuticals

Over the last decade the rate of approved protein therapeutics (biopharmaceuticals or biologic drugs) in the United States and European Union has been rapidly growing.[9] The cost of developing and bringing biopharmaceuticals to the market is much higher (when compared to classic small molecule drugs), offering increased profit margins. The biopharmaceutical market is currently estimated at US$199.7 billion globally and it is expected to reach US$ 497.9 billion by 2020, growing with an approximate annual rate of 13.5 %.[10]

Biotherapeutic development has been focused on treating a wide variety of chronic diseases including diabetes, cancer, hepatitis C and B and chronic renal failure – as well as less common ones such as hemophilia, Fabry's disease, growth deficiency, multiple sclerosis and Crohn's disease. The appeal of biopharmaceuticals

is that they harness the power of natural therapeutically active biomolecules and often have greater efficacy and specificity than small molecule drugs. The large majority of biologic drugs have been produced in cell culture and they are highly complex glycoproteins composed of various polypeptide domains, some of which are glycosylated.[11] There are two main types of glycosylation present in biopharmaceuticals: N- and O-glycosylation and both can impact safety, efficacy, immunogenicity, solubility, protein folding and half-life of the drug [12,13] (Table 1).

Given the potential impact of glycosylation on the safety and efficacy profiles of biologic drugs biopharma companies must optimise, measure and control the glycosylation patterns of their therapeutics. Unfortunately, there are three aspects that make these tasks very difficult. Firstly glycans are extremely complex molecules and most glycoprotein therapeutics have more than one glycosylation site. Secondly, glycans cause structural and functional diversity of the proteins to which they are attached. Most glycoprotein drugs exist not as a single molecule but as a heterogeneous mixture of glycoforms which share the same protein but which differ in the glycans. Each of these glycoforms might have its own distinct clinical safety and efficacy profile. Thirdly, changes in glycosylation patterns are the major source of batch to batch variability for glycoprotein therapeutics.[13] To overcome these challenges biopharma companies have to employ analytical technologies for detailed characterisation of glycosylation during drug design and biomanufacturing. In parallel with this is a drive from health agencies such as the US Food and Drug Administration (FDA), the European Medicines Evaluation Agency (EMEA), and the International Conference on Harmonization (ICH) which have produced guidance

Generic Designation	Therapeutic Protein	Expression System	Glycosylation
Algulcosidase alfa	Acid glucosidase	CHO cells	N-links[14]
Anti-hemophilic factor	Factor VIII	BHK & CHO cells	N-links & O-links[15]
Chorionic Gonadotropin alfa	Chorionic Gonadotropin	CHO cells	N-links & O-links[16,17]
Darbepoietin alfa	Engineered erythropoietin with 2 extraN-glycans (long-acting)	CHO cells	N-links & O-links[16,17]
Epoetin alfa	Erythropoietin	CHO cells	N-links & O-links[18,19]
Etanercept	A dimeric fusion protein consisting of the extracellular ligandbinding portion of the human 75 kilodalton (p75) tumor necrosis factor receptor (TNFR) linked to the Fc portion of human IgG1	CHO cells	N-links & O-links[18-20]
Factor VIIa	Factor VIIa	BHK cells	N-links & O-links[21]
Factor IX	Factor IX	CHO cells	N-links & O-links[22,23]
Hepatitis B vaccine	Surface antigen of Hepatitis B	*S. cerevisiae*	N-links & O-links[24]
Hepatitis B vaccine	S, pre-S and pre-S2	Murine cell line	N-links & O-links[24]
Lenograstim	Granulocyte colony stimulating factor (G-CSF)	CHO cells	O-links[25]
Thyrotropin alfa	Thyroid stimulating hormone	CHO cells	N-links[26,27]

Table 1. Examples of approved glycosylated therapeutics.

documents relating to regulatory requirements for measurement and control of drug glycosylation.[28,29] The variability of drug glycosylation is a major concern to biopharma companies and regulators as it can cause inconsistencies in the clinical performance of the drug product. The consequences are potential risks to patients being treated with drugs that are either ineffective or too strong. This also exposes companies to litigation for producing unreliable therapeutics. Given this, there is considerable interest in mitigating these risks by using a Quality by Design (QbD) to deal with glycosylation. In practice, this entails identifying, measuring and optimising

Glycosylation Critical Quality Attributes (GCQAs) starting at an early stage in drug development and continuing through the product's life cycle.[29,30] To be successful, QbD of drug glycosylation requires use of a range of orthogonal methods for glycosylation structure analysis. In practice, these are drawn from three main and complementary analytical platforms: high-performance liquid chromatography (HPLC), mass spectrometry (MS) and capillary electrophoresis (CE).

Biosynthesis and functions of O-linked glycans

Biosynthesis of mucin type O-glycans is very complex and is initiated in the Golgi apparatus by a *N*-acetyl galactosaminyltransferase (UDP-GalNAc) that transfers a *N*-acetylgalactosamine (GalNAc) residue to the side chain of a serine or a threonine in a glycoprotein via α-linkage to form GalNAcα1- Ser/Thr, which is also known as the Tn antigen.[31,32] Stepwise enzymatic elongation by specific transferases produces several core structures, which are further elongated or modified by sialylation, sulfatation, acetylation, fucosylation, and polylactosamine-extension.[33,34]

Synthesis of Tn antigen is followed by transfer of galactose (Gal), *N*-acetylglucosamine (GlcNAc) or *N*-acetylgalactosamine (GalNAc) to the Tn antigen to form eight core O-glycan structures (Figure 2). In addition core 1–4 structures are extended to form various structures including polyLacNAc chains[35], Lewis antigens[36], and various blood group antigens including the well-known ABO blood groups.[36]

The glycans associated with proteins modulate protein function and signaling. They can also alter the dynamics of glycoprotein endocytosis and cell surface half-life through binding to multivalent lectins.[37] Glycans play important roles in biological processes such as protein secretion (as they influence protein folding)[38], protein

clearance [39], receptor binding and activation [40] and cell adhesion.[41] Glycans also contribute to the regulation of cytosolic and nuclear functions, immune surveillance, inflammatory reactions, auto-immunity, hormone action and tumour metastasis.[42,43] Specific glycan patterns can change when the physiological state changes – for example, during disease or aging.[44,45] Protein glycosylation also plays very important role during absorption, distribution, metabolism and excretion of a drug in an organism.[46,47]

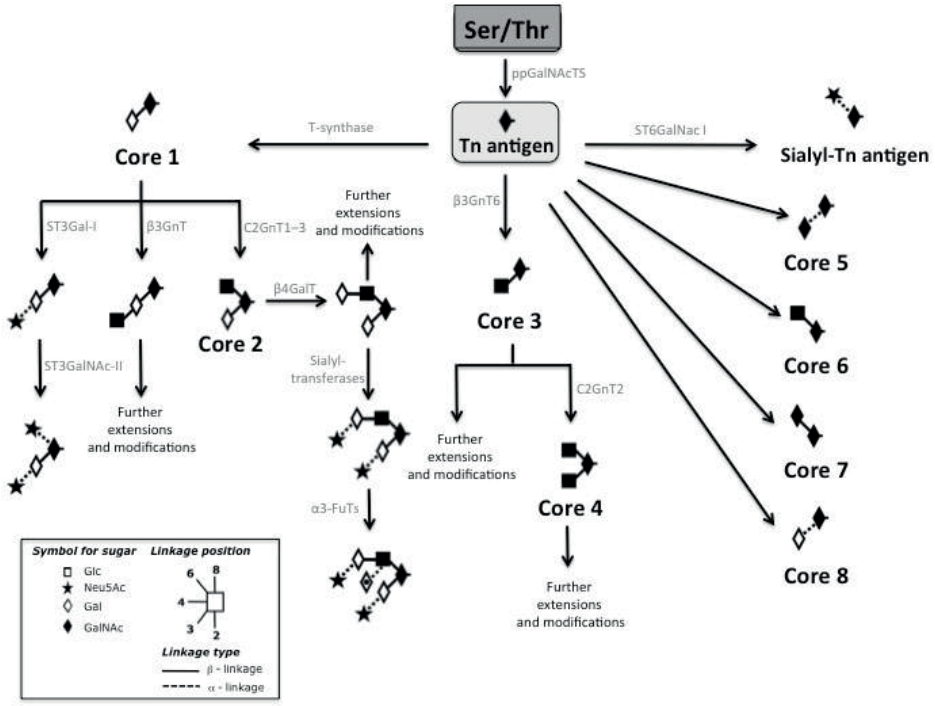

Figure 2. O-glycan biosynthesis pathway.

Functions of O-linked glylcans

During the last decade or so, many aspects of the biological roles of mucin type O-glycans have been reported. Functions of O-glycans are very diverse and include among others: (i) they maintain the highly extended and rigid structure of mucins, as

clustered O-glycans induce the peptide core to adopt a stiff and extended conformation that prevents folding into a globular structure;[48-50] (ii) they contribute to preventing the desiccation of the ocular surface by providing a hydrophilic character to mucins;[51] (iii) they play a role in the processing and expression of glycoproteins; (iv) they are involved in recognition in different processes (e.g. selecting binding in leukocyte circulation, fertilization, glycoprotein clearance);[50,52] (v) they are also known to restrict adenoviral vector access to apical receptors expressed in respiratory epithelium.[53]

Changes in O-glycosylation have been associated with many diseases such as inflammatory bowel disease, cystic fibrosis, IgA nephropathy and cancer. Inflammatory bowel diseases and cystic fibrosis are characterized by alterations in sulfation and sialylation of terminal O-glycan chains, which may result in a detrimental effect on the physico-chemical properties of the mucus, its protective barrier function, which would in turn impact bacterial colonization.[54,55] IgA nephropathy is an immune disorder characterized by reduced galactosylation of the Tn-antigen in IgA1.[56-58]

Mucins produced in cancer cells have an altered expression of glycosyltranferases that results in short and highly sialylated *O*-glycan chains, such as the Tn and sialyl-Tn antigens. Changes in O-glycan structures in cancer have many consequences for the function of epithelial cells, altering their antigenic and adhesive properties, as well as their potential to invade and metastasize.[59-62] Some of these functional properties of mucin type O-glycans might be used to design new biological drugs.

Aspects of glycosylation analysis will be extensively described in Chapter 2.

Scope of the thesis

Over the last decade the number of techniques for O-glycan analysis have been developed (described in Chapter 2), however release of O-glycans remains very challenging. The main reason for this is that there is no enzyme available for universal O-glycan release from proteins. Therefore, chemical release is the most effective method for release full range of O-glycans. Unfortunately all of the O-glycan release methods show a side reaction known as “peeling”, which is a stepwise degradation of the polysaccharide starting at the reducing end and removing one sugar residue at a time. Peeling is well known problem when performing the O-glycan release and often results in poor repeatability with variable amounts of the small peeled glycans. This is a major problem for comparability studies or quality control. The scope of this thesis was to develop or improve methods for the release of O-glycans with free reducing termini with high yield and limited degradation for the more detailed characterization of biological samples and biopharmaceuticals. In Chapters 3 and 4 the development of sample preparation techniques for O-glycan release (in their nonreduced form) by hydrazinolysis are described. In Chapter 5, the use of procainamide for fluorescent labelling of glycans is introduced. The procainamide labelled glycans gave higher fluorescent response and significantly higher ion intensity in positive ESI-MS when compared to most widely used 2-AB. Chapter 6 describes the application of the improved protocols for O-glycan release by hydrazinolysis combined with developed procainamide labelling system followed by HILIC-UHPLC-ESI-MS/MS analysis for evaluation of O-glycosylation changes in human saliva. The closing chapter of this thesis, Chapter 7, contains a general discussion which places the individual chapters into context, is included.

Reference List:

(1) Shental-Bechor, D.; Levy, Y. *Curr. Opin. Struct. Biol* **2009**, *19*, 524-533.
(2) Moremen, K. W.; Tiemeyer, M.; Nairn, A. V. *Nat. Rev. Mol. Cell Biol* **2012**, *13*, 448-462.
(3) Kornfeld, R.; Kornfeld, S. *Annu. Rev. Biochem* **1985**, *54*, 631-664.
(4) Nilsson, I. M.; von, H. G. *J. Biol. Chem* **1993**, *268*, 5798-5801.
(5) I., B.; H., S.; P., S. In *Essentials of Glycobiology. 2nd edition*; Cold Spring Harbor Laboratory Press, 2009.
(6) Spiro, R. G. *Glycobiology* **2002**, *12*, 43R-56R.
(7) Tran, D. T.; Ten Hagen, K. G. *J. Biol. Chem* **2013**, *288*, 6921-6929.
(8) Harvey, D. J.; Merry, A. H.; Royle, L.; Campbell, M. P.; Dwek, R. A.; Rudd, P. M. *Proteomics* **2009**, *9*, 3796-3801.
(9) Walsh, G. *Nat. Biotechnol* **2010**, *28*, 917-924.
(10) *Biopharmaceuticals - A Global Market Overview*, 2013.
(11) Higgins, E. *Glycoconj. J* **2010**, *27*, 211-225.
(12) Hossler, P.; Khattak, S. F.; Li, Z. J. *Glycobiology* **2009**, *19*, 936-949.
(13) Dotz, V.; Haselberg, V.; Shubhakar, A.; Kozak, R. P.; Falck, D.; Rombouts, Y.; Reusch, D.; Somsen, G. W.; Fernandes, D. L.; Wuhrer, M. *Trends in Analytical Chemistry* **2015**, *73*, 1-9.
(14) Jongen, S. P.; Gerwig, G. J.; Leeflang, B. R.; Koles, K.; Mannesse, M. L.; van Berkel, P. H.; Pieper, F. R.; Kroos, M. A.; Reuser, A. J.; Zhou, Q.; Jin, X.; Zhang, K.; Edmunds, T.; Kamerling, J. P. *Glycobiology* **2007**, *17*, 600-619.
(15) Hironaka, T.; Furukawa, K.; Esmon, P. C.; Fournel, M. A.; Sawada, S.; Kato, M.; Minaga, T.; Kobata, A. *J. Biol. Chem* **1992**, *267*, 8012-8020.
(16) Kessler, M. J.; Mise, T.; Ghai, R. D.; Bahl, O. P. *J. Biol. Chem* **1979**, *254*, 7909-7914.
(17) Weisshaar, G.; Hiyama, J.; Renwick, A. G. *Glycobiology* **1991**, *1*, 393-404.
(18) Takeuchi, M.; Kobata, A. *Glycobiology* **1991**, *1*, 337-346.
(19) Lai, P. H.; Everett, R.; Wang, F. F.; Arakawa, T.; Goldwasser, E. *J. Biol. Chem* **1986**, *261*, 3116-3121.
(20) Hokke, C. H.; Bergwerff, A. A.; Van Dedem, G. W.; Kamerling, J. P.; Vliegenthart, J. F. *Eur. J. Biochem* **1995**, *228*, 981-1008.

(21) Pennica, D.; Lam, V. T.; Weber, R. F.; Kohr, W. J.; Basa, L. J.; Spellman, M. W.; Ashkenazi, A.; Shire, S. J.; Goeddel, D. V. *Biochemistry* **1993**, *32*, 3131-3138.
(22) Fischer, B.; Mitterer, A.; Dorner, F.; Eibl, J. *J. Thromb. Thrombolysis* **1996**, *3*, 57-62.
(23) Agarwala, K. L.; Kawabata, S.; Takao, T.; Murata, H.; Shimonishi, Y.; Nishimura, H.; Iwanaga, S. *Biochemistry* **1994**, *33*, 5167-5171.
(24) Schmitt, S.; Glebe, D.; Tolle, T. K.; Lochnit, G.; Linder, D.; Geyer, R.; Gerlich, W. H. *J. Gen. Virol* **2004**, *85*, 2045-2053.
(25) Kubota, N.; Orita, T.; Hattori, K.; Oh-eda, M.; Ochi, N.; Yamazaki, T. *J. Biochem* **1990**, *107*, 486-492.
(26) Zhou, Q.; Park, S. H.; Boucher, S.; Higgins, E.; Lee, K.; Edmunds, T. *Anal. Biochem* **2004**, *335*, 10-16.
(27) Cole, E. S.; Lee, K.; Lauziere, K.; Kelton, C.; Chappel, S.; Weintraub, B.; Ferrara, D.; Peterson, P.; Bernasconi, R.; Edmunds, T.; . *Biotechnology (N. Y.)* **1993**, *11*, 1014-1024.
(28) Endo, T. *Biol. Pharm. Bull* **2009**, *32*, 765-766.
(29) Schiestl, M.; Stangler, T.; Torella, C.; Cepeljnik, T.; Toll, H.; Grau, R. *Nat. Biotechnol* **2011**, *29*, 310-312.
(30) Fernandes, D. L. In *Biopharmaceutical Production Technology, Volume 1 & Volume 2 (ed G. Subramanian)*; Wiley-VCH Verlag GmbH & Co.KGaA,Weinheim,Germany., 2012.
(31) Ju, T.; Wang, Y.; Aryal, R. P.; Lehoux, S. D.; Ding, X.; Kudelka, M. R.; Cutler, C.; Zeng, J.; Wang, J.; Sun, X.; Heimburg-Molinaro, J.; Smith, D. F.; Cummings, R. D. *Proteomics. Clin. Appl* **2013**, *7*, 618-631.
(32) Ju, T.; Otto, V. I.; Cummings, R. D. *Angew. Chem. Int. Ed Engl* **2011**, *50*, 1770-1791.
(33) Hounsell, E. F.; Davies, M. J.; Renouf, D. V. *Glycoconj. J* **1996**, *13*, 19-26.
(34) Yamashita, Y.; Chung, Y. S.; Horie, R.; Kannagi, R.; Sowa, M. *J. Natl. Cancer Inst* **1995**, *87*, 441-446.
(35) Liu, J.; Jin, C.; Cherian, R. M.; Karlsson, N. G.; Holgersson, J. *J. Biotechnol* **2015**, *199*, 77-89.
(36) Stanley, P.; Cummings, R. D. **2009**.
(37) Dennis, J. W.; Lau, K. S.; Demetriou, M.; Nabi, I. R. *Traffic* **2009**, *10*, 1569-1578.

(38) Hoseki, J.; Ushioda, R.; Nagata, K. *J. Biochem* **2010**, *147*, 19-25.
(39) Fukuda, M. N.; Sasaki, H.; Lopez, L.; Fukuda, M. *Blood* **1989**, *73*, 84-89.
(40) Marth, J. D.; Grewal, P. K. *Nat. Rev. Immunol* **2008**, *8*, 874-887.
(41) Gu, J.; Sato, Y.; Kariya, Y.; Isaji, T.; Taniguchi, N.; Fukuda, T. *J. Proteome. Res* **2009**, *8*, 431-435.
(42) Varki, A. *Glycobiology* **1993**, *3*, 97-130.
(43) Hart, G. W.; Slawson, C.; Ramirez-Correa, G.; Lagerlof, O. *Annu. Rev. Biochem* **2011**, *80*, 825-858.
(44) Freeze, H. H.; Aebi, M. *Curr. Opin. Struct. Biol* **2005**, *15*, 490-498.
(45) Ohtsubo, K.; Marth, J. D. *Cell* **2006**, *126*, 855-867.
(46) Sola, R. J.; Griebenow, K. *BioDrugs* **2010**, *24*, 9-21.
(47) Li, H.; d'Anjou, M. *Curr. Opin. Biotechnol* **2009**, *20*, 678-684.
(48) Jentoft, N. *Trends Biochem. Sci* **1990**, *15*, 291-294.
(49) Hanisch, F. G. *Biol. Chem* **2001**, *382*, 143-149.
(50) Van den Steen, P.; Rudd, P. M.; Dwek, R. A.; Opdenakker, G. *Crit Rev. Biochem. Mol. Biol* **1998**, *33*, 151-208.
(51) Argueso, P.; Gipson, I. K. *Exp. Eye Res* **2001**, *73*, 281-289.
(52) Fukuda, M.; Tsuboi, S. *Biochim. Biophys. Acta* **1999**, *1455*, 205-217.
(53) Stonebraker, J. R.; Wagner, D.; Lefensty, R. W.; Burns, K.; Gendler, S. J.; Bergelson, J. M.; Boucher, R. C.; O'Neal, W. K.; Pickles, R. J. *J. Virol* **2004**, *78*, 13755-13768.
(54) Xia, B.; Royall, J. A.; Damera, G.; Sachdev, G. P.; Cummings, R. D. *Glycobiology* **2005**, *15*, 747-775.
(55) Corfield, A. P.; Carroll, D.; Myerscough, N.; Probert, C. S. *Front Biosci* **2001**, *6*, D1321-D1357.
(56) Barratt, J.; Smith, A. C.; Feehally, J. *Nephrology. (Carlton.)* **2007**, *12*, 275-284.
(57) Allen, A. C.; Bailey, E. M.; Barratt, J.; Buck, K. S.; Feehally, J. *J. Am. Soc. Nephrol* **1999**, *10*, 1763-1771.
(58) Smith, A. C.; de Wolff, J. F.; Molyneux, K.; Feehally, J.; Barratt, J. *J. Am. Soc. Nephrol* **2006**, *17*, 1192-1199.
(59) Brockhausen, I. *Biochim. Biophys. Acta* **1999**, *1473*, 67-95.
(60) Brockhausen, I. *EMBO Rep* **2006**, *7*, 599-604.
(61) Tsuboi, S.; Hatakeyama, S.; Ohyama, C.; Fukuda, M. *Trends Mol. Med* **2012**, *18*, 224-232.

(62) Kudelka, M. R.; Ju, T.; Heimburg-Molinaro, J.; Cummings, R. D. *Adv. Cancer Res* **2015**, *126*, 53-135.