FEBS Letters 546 (2003) 154–158 FEBS 27402

Minireview

O-GlcNAc turns twenty: functional implications for post-translational modification of nuclear and cytosolic proteins with a sugar

Lance Wells, Gerald W. Hart*

Johns Hopkins School of Medicine, Department of Biological Chemistry, 725 N. Wolfe St., Baltimore, MD 21205, USA

Received 17 March 2003; revised 23 April 2003; accepted 23 April 2003

First published online 6 June 2003

Edited by Richard Marais

Abstract O-linked β-*N*-acetylglucosamine (O-GlcNAc) is a dynamic nucleocytoplasmic post-translational modification more analogous to phosphorylation than to classical complex O-glycosylation. A large number of nuclear and cytosolic proteins are modified by O-GlcNAc. Proteins modified by O-GlcNAc include transcription factors, signaling components, and metabolic enzymes. While the modification has been known for almost 20 years, functions for the monosaccharide modification are just now emerging. In this review, we will focus on the cycling enzymes and emerging roles for this post-translational modification in regulating signal transduction and transcription. Finally, we will discuss future directions and the working model of O-GlcNAc serving as a nutrient sensor.

Key words: O-GlcNAc; Post-translational modification; Glycosylation; Phosphorylation; Transcription; Insulin signaling; Nutrient sensing; Diabetes

1. Introduction

O-linked β-*N*-acetylglucosamine (O-GlcNAc), which was described in 1984, is a monosaccharide modification abundant on serine and threonine residues of a multitude of nucleocytoplasmic proteins [1]. Over the intervening years, the number of modified nuclear and cytoplasmic proteins identified has reached nearly 100 and includes transcription factors, cytoskeletal components, metabolic enzymes, and signaling components (reviewed in [2,3]). The enzymes responsible for O-GlcNAc addition (O-GlcNAc transferase, OGT) and removal (neutral β-*N*-acetylglucosaminidase, O-GlcNAcase) have been cloned and partially characterized in the last 6 and 2 years, respectively (reviewed in [4]). Work showing the dynamic and inducible nature of the modification has suggested that O-GlcNAc is a regulatory modification, but only work in the last few years has begun to elucidate O-GlcNAc's biological roles [5,6]. Thus, the 'O-GlcNAc field' is still relatively immature compared to the much better understood phosphorylation field. It is attractive to compare O-GlcNAc to serine/threonine phosphorylation as they share several key features including the residues and proteins they modify, their dynamic and inducible nature, and the existence of cycling enzymes [7].

In 1992, Edwin G. Krebs described the lengthy time that it took for phosphorylation/dephosphorylation to be accepted as a mechanism for regulating proteins, especially outside the glycogen metabolism field [8]. In particular, Krebs and Fischer's groundbreaking work in the mid to late 1950s came 25 years after the discovery of serine/threonine phosphorylation in 1933. After the pioneering phosphorylase kinase work by Krebs and Fischer, it would take another decade before it was recognized that phosphorylation could be regulated by extracellular stimuli (cAMP and protein kinase A, 1968, [9,10]) and another decade after that (1976) before the number of enzymes known to undergo phosphorylation/dephosphorylation exceeded 20 [8]. O-GlcNAc will soon be celebrating its 20th anniversary and likewise only recent work has begun to shed light on the biological function of this abundant modification [11,12]. A significant portion of the time and effort in the last 20 years in this small but growing field has focused on generating the necessary methods and tools to study this monosaccharide (reviewed in [13]). In this review, we will briefly discuss the cycling enzymes (reviewed in [4]) and then focus on the accumulating data that O-GlcNAc is involved in regulating protein function. We will conclude by describing the working model that O-GlcNAc is serving as a nutrient-sensor master switch that attenuates cellular responses to extracellular stimuli.

2. The cycling enzymes for O-GlcNAc modification of nucleocytoplasmic proteins (see Fig. 1)

OGT (uridine diphospho-*N*-acetylglucosamine:polypeptide β-*N*-acetylglucosaminyltransferase, EC 2.4.1) was originally purified to near homogeneity in 1992 from rat liver [14]. At that time, it was determined that it had a native molecular weight of approximately 340 kDa [14]. The rat, human, and *Caenorhabditis elegans* genes were cloned almost simultaneously in 1997 [15,16]. Following fine mapping of the human OGT gene to generate a full-length sequence [17], it was found that the human and rat sequence are nearly 100% homologous. Attempted knock-out studies in mice showed that the gene for OGT resides on the X chromosome and is necessary for embryonic stem cell viability [18]. In most tissues, the protein is a trimer composed of three 110 kDa polypeptides

*Corresponding author. Fax: (1)-401-614 8804.
E-mail address: gwhart@jhmi.edu (G.W. Hart).

0014-5793/03/$22.00 © 2003 Published by Elsevier Science B.V. on behalf of the Federation of European Biochemical Societies.
doi:10.1016/S0014-5793(03)00641-0

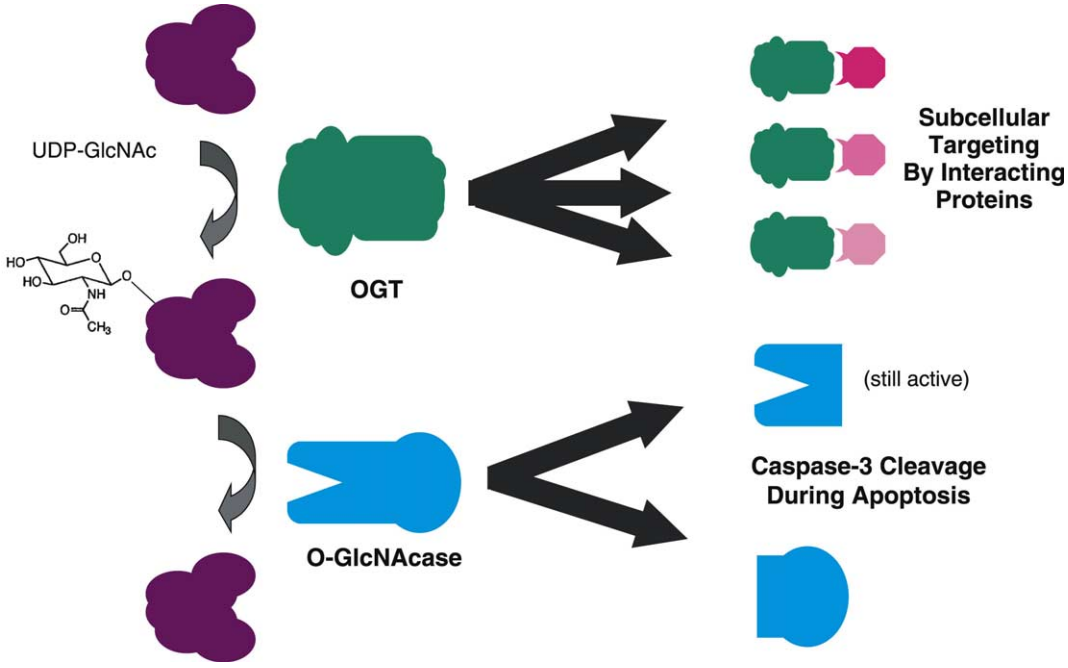

Fig. 1. Enzymes of O-GlcNAc cycling. OGT adds O-GlcNAc to nucleocytoplasmic proteins and O-GlcNAcase removes the dynamic modification. Recent data suggest that OGT is targeted through various protein–protein interactions and splicing. O-GlcNAcase is a substrate for caspase-3, an executioner protease in apoptosis.

and each polypeptide appears to be composed of two domains. The N-terminus of the mammalian polypeptide contains 11.5 tetratricopeptide repeats. These repeats are thought to be involved in mediating protein–protein interactions [19] and it has been shown that these repeats in OGT impact on protein trimerization and substrate recognition [20–22]. The C-terminus of the protein appears to be the catalytic domain of the protein with a putative UDP-GlcNAc binding site and weak homology to the glycogen phosphorylase/glycosyl transferase family [23]. Furthermore, a 100 amino acid deletion from the C-terminus results in a catalytic inactive enzyme [21,24]. Recent work has identified proteins that interact with OGT and these proteins may be targeting OGT to subcellular locations (Fig. 1, [22,24]). Also, Hanover and colleagues have recently shown that there are multiple isoforms of OGT and at least one splice variant is targeted to the mitochondria [25,26]. OGT has also been shown to be O-GlcNAc-modified itself as well as tyrosine-phosphorylated [15]. Understanding the regulation of this enzyme will be key to future investigations of the O-GlcNAc modification.

The existence of a neutral β-*N*-acetylglucosaminidase (termed hexC) has been known for over 25 years [27]. Dong et al. partially purified and characterized this nucleocytoplasmic enzyme and termed it O-GlcNAcase in 1994 [28]. Cloning and characterization of the enzyme in the last couple of years [29,30] suggest that O-GlcNAcase most likely represents the previously described hexC activity. The primary sequence of O-GlcNAcase had previously been identified as a meningioma-expressed antigen and showed weak hyaluronidase activity [31]. Furthermore, while the putative hyaluronidases and O-GlcNAcase activity are in the N-terminus, the C-terminus shows weak homology to histone acetyltransferases in the C-terminus [32]. The possibility that this one polypeptide has multiple enzymatic activities is currently being explored. The 916 amino acid enzyme is quite distinct from the lysosomal hexosaminidases [29,30]. Most notable is O-GlcNAcase's localization primarily to the cytosol and to a lesser extent the nucleus. O-GlcNAcase is also not a general hexosaminidase but instead is a specific β-*N*-acetylglucosaminidase, and thus it is not inhibited by *N*-acetylgalactosamine. Compared to OGT and due to its recent cloning, little information is currently known about interacting proteins, post-translational modification, and regulation of the enzyme. It is interesting to note that while the recombinant enzyme migrates at the expected molecular weight for a monomer (~140 kDa), the purified enzyme from brain migrates at ~600 kDa using size-exclusion chromatography [30]. This suggests that O-GlcNAcase is in a complex and in fact O-GlcNAcase co-purified with several identified proteins including the heat-shock proteins HSP110 and HSC70 [29]. One intriguing observation is that O-GlcNAcase is processed by caspase-3 to generate a 64 kDa C-terminal fragment [30]. Determining whether this proteolytic processing is separating the O-GlcNAcase activity from regulatory domains or from the putative acetyltransferase activity is an area of future investigation (see Fig. 1). The exact role that the O-GlcNAc modification plays in and the importance of the cleavage of O-GlcNAcase to apoptosis are currently areas of active research in our group.

3. The hexosamine biosynthetic pathway, diabetes, and O-GlcNAc (see Fig. 2)

The hexosamine biosynthetic pathway (HSP) generates the sugar nucleotide UDP-GlcNAc, which is the donor for O-GlcNAc addition to nucleocytoplasmic proteins (reviewed in [33]). The first and rate-limiting step in this pathway is the conversion of fructose-6-phosphate to glucosamine-6-phos-

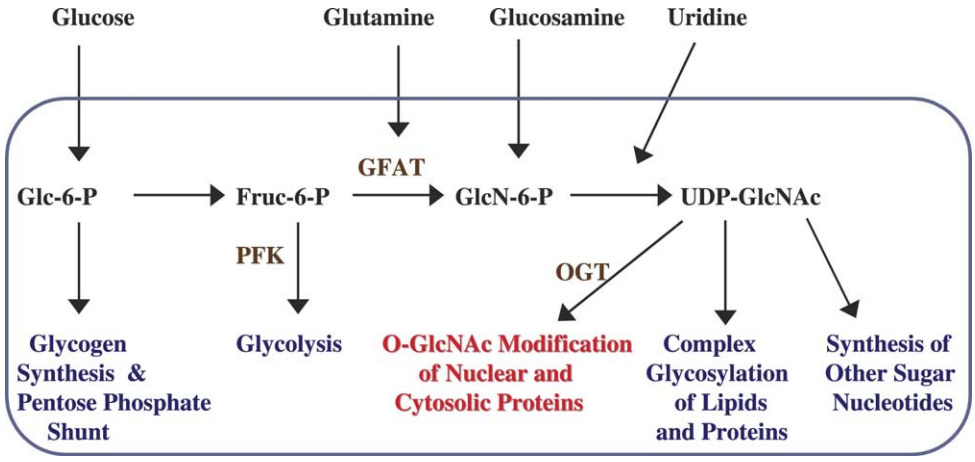

Fig. 2. Glucose metabolism in peripheral insulin-responsive tissues. Glucose can be metabolized through several pathways including the hexosamine biosynthetic pathway (from fruc-6-p to UDP-GlcNAc). The enzyme GFAT is shown since it is the rate-limiting step in UDP-GlcNAc formation. UDP-GlcNAc serves as a high-energy sugar nucleotide donor for several processes including OGT-catalyzed modification of nucleocytoplasmic proteins with O-GlcNAc.

phate with the concomitant conversion of glutamine to glutamate (see Fig. 2). This reaction is catalyzed by the enzyme glutamine:fructose-6-phosphate amidotransferase (GFAT, Fig. 2). Glucosamine-6-phosphate is then rapidly converted through a series of steps to UDP-GlcNAc that can then be used in glycosylation of lipids and proteins or be acted on by other enzymes for the generation of other UDP-sugars. Approximately 4% of the glucose in the cell branches into this pathway from glycolysis (see Fig. 2).

Marshall and colleagues were able to link increased flux through the HSP to insulin resistance in peripheral tissues [34]. Insulin resistance is the hallmark of diabetes. Hyperglycemia, glucosamine treatment, free fatty acid elevation, as well as overexpression of GFAT have all been shown to increase flux through the HSP and induce varying degrees of insulin resistance (reviewed in [33]). Since O-GlcNAc modification is dependent on flux through the HSP (see Fig. 2), it has been proposed as a possible mechanism for mediating insulin resistance. In fact, elevated O-GlcNAc levels were found in diabetic rats, insulin resistance 3T3-L1 adipocytes, and atherosclerotic plaques of diabetic patients (reviewed in [33]). We were able to establish a causal relationship between elevated O-GlcNAc levels and insulin resistance in 3T3-L1 adipocytes by pharmacological inhibition of the O-GlcNAcase [35]. Further we were able to demonstrate that elevated levels of O-GlcNAc inhibit proper insulin-dependent activation of Akt. The finding that the defect in insulin signaling was at or above Akt is consistent with the work of others looking at increased flux through the HSP and insulin resistance [36]. Subsequently, transgenic mice overexpressing OGT in skeletal muscle and fat were found to have lowered glucose disposal rates [37]. Furthermore, endothelial nitric oxide synthase is not properly activated when there are elevated O-GlcNAc levels and glycogen synthesis appears to be impaired with elevation of O-GlcNAc levels [38,39]. Since endothelial nitric oxide synthase and glycogen synthase are both downstream of Akt and appear to be O-GlcNAc-modified themselves, it appears that O-GlcNAc is either directly or indirectly modulating their activity. All of these data taken together suggest that O-GlcNAc modification of nucleocytoplasmic proteins is serving as a negative feedback system for insulin signaling, but elucidating the impact of O-GlcNAc at specific sites on key proteins remains to be investigated.

Another link to the diabetic state and elevated O-GlcNAc levels is the 'glucose toxicity' observed in diabetes that is believed to be responsible for many of the microvascular and macrovascular complications. For example, mouse embryonic fibroblasts with impaired synthesis of UDP-GlcNAc showed a marked increase in their ability to withstand apoptotic stimuli [40]. Consistent with the above studies, increased Akt activity (known to be anti-apoptotic) was seen in these cells and lowered O-GlcNAc levels. Interestingly, while O-GlcNAc levels were very responsive to the change in UDP-GlcNAc levels, there were no detectable changes in complex O- and N-linked glycosylation. Further, in retinal neurons, increased flux through the HSP was found to block the protective effect of insulin against apoptosis [41]. Also, O-GlcNAcase is a substrate for the executioner caspase-3, even though the effect of proteolysis is not clear since cleavage has no effect on in vitro O-GlcNAcase activity [30]. Finally, β-cells of the pancreas undergo apoptosis in response to streptozotocin treatment. While the toxic effects of this compound are though to be mediated by its alkylating capability [42], Kudlow and colleagues have shown that streptozotocin, a GlcNAc mimetic, is a weak inhibitor of O-GlcNAcase [43]. All of these data together suggest that O-GlcNAc may be playing a major role in diabetic complications and more generally in apoptosis. The precise mechanism by which O-GlcNAc modification is exerting its effect however remains to be determined.

4. O-GlcNAc and transcription

The C-terminal domain of RNA polymerase II is modified by O-GlcNAc [44]. Recent work has demonstrated a reciprocal and inhibitory relationship between phosphorylation and O-GlcNAc modification of this C-terminal domain [45]. Since phosphorylation is required for elongation, the possibility exists that the O-GlcNAc-modified form of RNA polymerase II is involved in the pre-initiation complex and may serve as a checkpoint for premature elongation.

To date, more than 10 transcription factors have been

shown to be modified by O-GlcNAc [46]. In more than one case the site of O-GlcNAc modification is the same as phosphorylation. For example, c-myc is O-GlcNAc-modified at Thr58, a known site of phosphorylation and a mutational hot spot in lymphomas [47]. Recent work suggests that glycosylation of c-myc plays a role in protein stability and subcellular localization of the transcription factor [48]. The Sp-1 transcription factor is one of the most well studied O-GlcNAc-modified proteins. Work has shown that O-GlcNAc modification of Sp-1 appears to modulate its transactivation capability by affecting Sp-1's interaction with TAF110 [49]. Recent work looking at OGT interacting proteins has identified OIP-106 and mSin3A, both of which may target OGT to transcriptional complexes (reviewed in [4]). Interestingly, mSin3A is part of a histone deacetylase complex and OGT activity was shown to be required for maximal transcriptional repression. Furthermore, increased O-GlcNAc modification of p53 and NF-κβ are associated with increased transcriptional activity (reviewed in [6]). Thus, there is compelling evidence that O-GlcNAc modification of the transcriptional machinery is modulating activity. It remains to be elucidated how O-GlcNAc is modulating activity at the molecular level and how O-GlcNAc modification of the transcription machinery is itself being regulated.

5. Conclusions and future directions

O-GlcNAc modification of nuclear and cytosolic proteins was described in the 1980s. The majority of the work in the 1990s centered on identifying more modified proteins, the enzymes for its addition and subtraction, and the development of tools to make the study of this modification less cumbersome. Work at the turn of the century has begun to focus on the function of this enigmatic modification. Thus, we believe that O-GlcNAc, in an analogous manner to the history of phosphorylation, is on the verge of an explosion in understanding. As more and more functional data concerning this modification appear in the literature and as tools for its study become available, more investigators are beginning to study O-GlcNAc. We believe this has also been greatly facilitated by the development of an O-GlcNAc-specific antibody, the cloned genes for OGT and O-GlcNAcase and a mass spectrometry-based technique for site-mapping that makes investigation more straightforward (reviewed in [6]). While some functional studies have emerged, the precise roles of O-GlcNAc in regulating signal transduction cascades, transcription, translation, and the cell cycle remain to be elucidated. There is also compelling evidence for O-GlcNAc being involved in disease states such as diabetes, cancer, and age-related neurological disorders but much work remains to be performed to understand mechanisms. Our laboratory is using a working model that O-GlcNAc modification of proteins is acting as a nutrient sensor (reviewed in [50]). In this model, cells are taking into account their energy levels to modulate what proteins are produced in that cell. Also, modulation of O-GlcNAc levels is impinging on extracellular-stimulated signal transduction events such that the cell is not acting as a slave to the stimuli but instead is taking into account its metabolic state and responding to the signal appropriately. While this is an attractive model, it remains to be rigorously tested. In conclusion, we believe that O-GlcNAc is now emerging as an important regulatory modification.

References

[1] Torres, C.R. and Hart, G.W. (1984) J. Biol. Chem. 259, 3308–3317.
[2] Hart, G.W., Haltiwanger, R.S., Holt, G.D. and Kelly, W.G. (1989) Ciba Found. Symp. 145, 102-12, discussion 112-8.
[3] Wells, L., Whalen, S.A. and Hart, G.W. (2003) Biochem. Biophys. Res. Commun. 302, 435–441.
[4] Iyer, S.P. and Hart, G.W. (2003) Biochemistry 42, 2493–2499.
[5] Wells, L., Vosseller, K. and Hart, G.W. (2001) Science 291, 2376–2378.
[6] Vosseller, K., Sakabe, K., Wells, L. and Hart, G.W. (2002) Curr. Opin. Chem. Biol. 6, 851–857.
[7] Comer, F.I. and Hart, G.W. (2000) J. Biol. Chem. 275, 29179–29182.
[8] Krebs, E.G. (1992) Nobel Lecture Physiology or Medicine.
[9] Walsh, D.A., Perkins, J.P. and Krebs, E.G. (1968) J. Biol. Chem. 243, 3763–3765.
[10] Miyamoto, E., Kuo, J.F. and Greengard, P. (1968) Science 165, 63–65.
[11] Hanover, J.A. (2001) FASEB J. 15, 1865–1876.
[12] Zachara, N.E. and Hart, G.W. (2002) Chem. Rev. 102, 431–438.
[13] Zachara, N.E., Cole, R.N., Hart, G.W. and Gao, Y. (2002) Current Protocols in Protein Science 17, 6.1-22.
[14] Haltiwanger, R.S., Blomberg, M.A. and Hart, G.W. (1992) J. Biol. Chem. 267, 9005–9013.
[15] Kreppel, L.K., Blomberg, M.A. and Hart, G.W. (1997) J. Biol. Chem. 272, 9308–9315.
[16] Lubas, W.A., Frank, D.W., Krause, M. and Hanover, J.A. (1997) J. Biol. Chem. 272, 9316–9324.
[17] Nolte, D. and Muller, U. (2002) Mamm. Genome 13, 62–64.
[18] Shafi, R., Iyer, S.P., Ellies, L.G., O'Donnell, N., Marek, K.W., Chui, D., Hart, G.W. and Marth, J.D. (2000) Proc. Natl. Acad. Sci. USA 97, 5735–5739.
[19] Blatch, G.L. and Lassle, M. (1999) Bioessays 21, 932–939.
[20] Kreppel, L.K. and Hart, G.W. (1999) J. Biol. Chem. 274, 32015–32022.
[21] Lubas, W.A. and Hanover, J.A. (2000) J. Biol. Chem. 275, 10983–10988.
[22] Iyer, S.P., Akimoto, Y. and Hart, G.W. (2003) J. Biol. Chem. 278, 5399–5409.
[23] Wrabl, J.O. and Grishin, N.V. (2001) J. Mol. Biol. 314, 365–374.
[24] Yang, X., Zhang, F. and Kudlow, J.E. (2002) Cell 110, 69–80.
[25] Hanover, J.A., Yu, S., Lubas, W.B., Shin, S.H., Ragano-Caracciola, M., Kochran, J. and Love, D.C. (2003) Arch. Biochem. Biophys. 409, 287–297.
[26] Love, D.C., Kochran, J., Cathey, R.L., Shin, S.H. and Hanover, J.A. (2003) J. Cell Sci. 116, 647–654.
[27] Beutler, E. and Kuhl, W. (1977) Ann. Hum. Genet. 41, 163–167.
[28] Dong, D.L. and Hart, G.W. (1994) J. Biol. Chem. 269, 19321–19330.
[29] Gao, Y., Wells, L., Comer, F.I., Parker, G.J. and Hart, G.W. (2001) J. Biol. Chem. 276, 9838–9845.
[30] Wells, L., Gao, Y., Mahoney, J.A., Vosseller, K., Chen, C., Rosen, A. and Hart, G.W. (2002) J. Biol. Chem. 277, 1755–1761.
[31] Heckel, D., Comtesse, N., Brass, N., Blin, N., Zang, K.D. and Meese, E. (1998) Hum. Mol. Genet. 7, 1859–1872.
[32] Schultz, J. and Pils, B. (2002) FEBS Lett. 529, 179–182.
[33] McClain, D.A. (2002) J. Diabet. Complications 16, 72–80.
[34] Marshall, S., Bacote, V. and Traxinger, R.R. (1991) J. Biol. Chem. 266, 4706–4712.
[35] Vosseller, K., Wells, L., Lane, M.D. and Hart, G.W. (2002) Proc. Natl. Acad. Sci. USA 99, 5313–5318.
[36] Patti, M.E., Virkamaki, A., Landaker, E.J., Kahn, C.R. and Yki-Jarvinen, H. (1999) Diabetes 48, 1562–1571.
[37] McClain, D.A., Lubas, W.A., Cooksey, R.C., Hazel, M., Parker, G.J., Love, D.C. and Hanover, J.A. (2002) Proc. Natl. Acad. Sci. USA 99, 10695–10699.
[38] Du, X.L., Edelstein, D., Dimmeler, S., Ju, Q., Sui, C. and Brownlee, M. (2001) J. Clin. Invest. 108, 1341–1348.
[39] Parker, G.J., Lund, K.C., Taylor, R.P. and McClain, D.A. (2003) J. Biol. Chem. 278, 10022–10027.
[40] Boehmelt, G. et al. (2000) EMBO J. 19, 5092–5104.

[41] Nakamura, M., Barber, A.J., Antonetti, D.A., LaNoue, K.F., Robinson, K.A., Buse, M.G. and Gardner, T.W. (2001) J. Biol. Chem. 276, 43748–43755.
[42] Bhuyan, B.K. (1970) Cancer Res. 30, 2017–2023.
[43] Roos, M.D., Xie, W., Su, K., Clark, J.A., Yang, X., Chin, E., Paterson, A.J. and Kudlow, J.E. (1998) Proc. Assoc. Am. Physicians 110, 422–432.
[44] Kelly, W.G., Dahmus, M.E. and Hart, G.W. (1993) J. Biol. Chem. 268, 10416–10424.
[45] Comer, F.I. and Hart, G.W. (2001) Biochemistry 40, 7845–7852.
[46] Comer, F.I. and Hart, G.W. (1999) Biochim. Biophys. Acta 1473, 161–171.
[47] Chou, T.Y., Hart, G.W. and Dang, C.V. (1995) J. Biol. Chem. 270, 18961–18965.
[48] Kamemura, K., Hayes, B.K., Comer, F.I. and Hart, G.W. (2002) J. Biol. Chem. 277, 19229–19235.
[49] Roos, M.D., Su, K., Baker, J.R. and Kudlow, J.E. (1997) Mol. Cell Biol. 17, 6472–6480.
[50] Wells, L., Vosseller, K. and Hart, G.W. (2003) Cell Mol. Life Sci. 60, 222–228.