

UCSF

UC San Francisco Previously Published Works

Title

Human RNA Polymerase II Promoter Recruitment in Vitro Is Regulated by O-Linked N-Acetylglucosaminyltransferase (OGT)*

Permalink

<https://escholarship.org/uc/item/3tx5h2jz>

Journal

Journal of Biological Chemistry, 291(27)

ISSN

0021-9258

Authors

Lewis, Brian A
Burlingame, Alma L
Myers, Samuel A

Publication Date

2016-07-01

DOI

10.1074/jbc.m115.684365

Copyright Information

This work is made available under the terms of a Creative Commons Attribution License, available at <https://creativecommons.org/licenses/by/4.0/>

Peer reviewed

Human RNA Polymerase II Promoter Recruitment *in Vitro* Is Regulated by O-Linked N-Acetylglucosaminyltransferase (OGT)*

Received for publication, August 10, 2015, and in revised form, March 28, 2016. Published, JBC Papers in Press, April 27, 2016, DOI 10.1074/jbc.M115.684365

Brian A. Lewis^{†1}, Alma L. Burlingame[§], and Samuel A. Myers[§]

From the [†]Transcriptional Regulation and Biochemistry Unit, Metabolism Branch, Center for Cancer Research, NCI, National Institutes of Health, Bethesda, Maryland 20892 and [§]Department of Pharmaceutical Chemistry, University of California, San Francisco, California 94158

Although the O-linked N-acetylglucosamine (O-GlcNAc) modification of the RNA polymerase II C-terminal domain was described 20 years ago, the function of this RNA polymerase II (pol II) species is not known. We show here that an O-GlcNAcylated pol II species (pol II γ) exists on promoters *in vitro*. Inhibition of O-GlcNAc-transferase activity and O-GlcNAcylation prevents pol II entry into the promoter, and O-GlcNAc removal from pol II is an ATP-dependent step during initiation. These data indicate that O-GlcNAc-transferase activity is essential for RNA pol II promoter recruitment and that pol II goes through a cycling of O-GlcNAcylation at the promoter. Mass spectrometry shows that serine residues 2 and 5 of the pol II C-terminal domain are O-GlcNAcylated, suggesting an overlap with the transcription factor IIH (TFIIH)-dependent serine 5 phosphorylation events during initiation and P-TEFb (positive transcriptional elongation factor b) events during elongation. These data provide unexpected and important insights into the role of a previously ill-defined species of RNA polymerase II in regulating transcription.

The current model of the RNA polymerase II (pol II)² transcription cycle proposes that a preinitiation complex (PIC) forms on promoters using pol IIA, a species of pol II in which its C-terminal domain (CTD; consensus heptad repeat Y¹S²P³T⁴S⁵P⁶S⁷) is hypophosphorylated. The CTD is then extensively phosphorylated at several residues during elongation (pol IIO species) (1–3). However, pol IIA is also O-GlcNAcylated (4, 5), thus casting doubt on whether pol IIA or O-GlcNAcylated pol II is the initiation-specific species. This

has remained an unresolved issue as has the general function of the O-GlcNAc modification of pol II.

The O-GlcNAcylation of proteins is a common post-translational modification but one that is severely underappreciated and poorly understood. O-GlcNAcylation is nearly as common as phosphorylation and in many cases is found in a reciprocal relationship with phosphorylation on those same serine/threonine residues. This mutual exclusivity is likely the source of regulation, switching proteins between phosphorylated and O-GlcNAcylated species. O-GlcNAc addition to serine or threonine residues is catalyzed by O-GlcNAc-transferase (OGT) and UDP-GlcNAc, and its removal is catalyzed by β -N-acetylglucosaminidase (OGA). UDP-GlcNAc is synthesized by the hexosamine biosynthetic pathway using fructose 6-phosphate as the starting sugar compound (6). That O-GlcNAc is produced by an influx of glucose and is actively added to and removed from proteins suggests an intriguing relationship between nutrient state and the proteome (6, 7).

We previously showed that both OGT and OGA activities were required for transcription *in vitro* using HeLa nuclear extracts and that inhibition of O-GlcNAc addition or removal resulted in a defect in PIC assembly (5). Serine-to-alanine mutations suggested that serine residues 5 and 7 of the CTD were O-GlcNAcylated, and an O-GlcNAcylated CTD could block phosphorylation by TFIIH, suggesting that these are mutually exclusive modifications (5, 8). *In vivo*, OGT shRNA reduced transcription, O-GlcNAcylation, and pol II promoter occupancy (5). These data suggested that O-GlcNAcylation played a role in transcription, but it was not clear what role, if any, O-GlcNAcylated pol II played in this process.

Here we explored further the nature of the PIC defect we reported previously (5). We found that OGT inhibition blocks not only the O-GlcNAcylation of pol II but also prevents pol II entry into the PIC. Furthermore, GlcNAc removal from pol II is an ATP-dependent step during initiation of transcription. Lastly, we provide the first mass spectrometry analysis of an O-GlcNAcylation CTD and show that the CTD serine 5 residue is O-GlcNAcylated by OGT. These data formulate a model of O-GlcNAc-dependent transcription initiation and provide evidence that O-GlcNAcylated pol II is a functional species at promoters. Overall, these data support the idea that regulation of O-GlcNAcylation occurs as part of the transcriptional process in humans.

* This work was supported by the NCI Intramural Program (to B. A. L.) and by the Biomedical Technology Research Centers Program of the National Institutes of Health NIGMS, National Institutes of Health NIGMS Grant 8P41GM103481, and Howard Hughes Medical Institute (to S. A. M. and A. L. B.). The authors have no conflicts of interest to declare. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

¹ To whom correspondence should be addressed. Tel.: 301-435-8323; E-mail: lewisbri@mail.nih.gov or phdbiochemist@gmail.com.

² The abbreviations used are: pol II, RNA polymerase II; CTD, C-terminal domain; TF, transcription factor; PIC, preinitiation complex; OGT, O-GlcNAc-transferase; OGA, β -N-acetylglucosaminidase; PUGNAc, O-(2-acetamido-2-deoxy-D-glucopyranosylidene)amino-N-phenylcarbamate; WGA, wheat germ agglutinin; IP, immunoprecipitation; AMP-PNP, adenosine 5'-(β , γ -imino)triphosphate; pol IIO, unmodified pol II; rGAC, ribonucleotides GTP/ATP/CTP.

Materials and Methods

In Vitro PIC Assays—PICs were formed by adding nuclear extract, BC buffer (20 mM Tris, pH 7.9 at 4 °C, 10% glycerol, 0.2 mM EDTA, 0.2 mM DTT, 100 mM KCl) containing 100 mM KCl (BC100), 2 μ g of *Escherichia coli* DNA to E3 promoter DNA immobilized on M280 Dynabeads that were blocked for 30 min using 100 mg/ml BSA in BC100 (5). PICs were incubated as indicated with 0.1 mM STO45849 or DMSO (0.5 μ l of 10 mM STO45849 stock in DMSO or 0.5 μ l of DMSO control) or 3 mM PUGNAc. For the wheat germ agglutinin affinity purification assays, assembled PICs on the adenovirus E3 promoter were washed in BC100, 0.05% Nonidet P-40 and eluted with 100 μ l of BC buffer containing 500 mM KCl (BC500). This was diluted 2 \times with H₂O and incubated for 1 h at room temperature with the equivalent of 20 μ l of WGA-agarose slurry (Vector Laboratories) or for 1 h with 1 μ l of 8WG16 mAb (Covance) and protein G beads (Roche Applied Science). STO45849 was obtained from TimTec.

PICs using the CMV promoter were formed as above but in the absence of *E. coli* DNA. 100 ng of CMV promoter DNA was bound to M280 beads as described previously (9). WGA affinity purification assays and 8WG16 immunoprecipitations were as described above. Western blotting using 8WG16 and 110.6 were as described previously (5). To detect *O*-GlcNAcylated pol II using 110.6 Western blotting (Fig. 1D, right panel), the PIC assay used 400 ng of DNA template.

For nucleotide additions, all concentrations were 0.5 mM final. PICs were washed with BC100, 0.05% Nonidet P-40, and then bound material was eluted with BC500. WGA assays were done as above; IPs and blotting were done using 8WG16 antibody (Covance) or anti-MED6 (Santa Cruz Biotechnology). IPs were washed with BC100, 0.05% Nonidet P-40 and eluted with sample buffer, subjected to SDS-PAGE, and transferred to nitrocellulose for Western blotting.

Liquid Chromatography-Tandem Mass Spectrometry (LC-MS/MS) and Data Analysis—OGT labeling of GST-CTD was done as described (5). GST-CTD fusions containing CTD repeats 27–52 were digested with a 1:50 trypsin:substrate ratio at 37 °C for 3 h. Peptides were desalted and analyzed as described previously (16) except the 90-min LC gradient only reached 20% acetonitrile. Data were searched using Protein Prospector v 5.10.0. against the Swiss-Prot database (March 21, 2012) where all 535,248 entries, including the user-added protein, GST-CTD amino acid sequence, was searched. Manual inspection for 204.087 *m/z* detection in higher energy collisional dissociation spectra was used to identify *N*-acetylglucosamine-modified peptides. In these instances, data were interpreted manually.

Results

Pol II γ Resides at Promoters in Vitro—Our previous work showed that both OGT and OGA were functionally necessary for PIC formation on promoters and that pol II was *O*-GlcNAcylated in nuclear extracts (pol II γ) (5). These findings raised the question of whether pol II γ is physically associated with the promoter. To address this question, we formed PICs in the presence or absence of the OGT inhibitor

STO45849 or the OGA inhibitor PUGNAc on the adenovirus E3 promoter bound to magnetic beads (10, 11). We eluted PICs with 500 mM KCl (to disrupt PIC-DNA and protein-protein interactions within the PIC) and mixed the eluate with WGA-agarose beads to isolate *O*-GlcNAcylated proteins from the PICs. WGA-bound proteins were eluted in sample buffer and analyzed by Western blotting for pol II (Fig. 1A). We found that pol II γ was in the PIC but that STO45849 largely blocked recruitment of pol II γ during PIC assembly. In contrast, the pol II remained *O*-GlcNAcylated after PUGNAc treatment.

We next asked whether total pol II recruitment to the PIC was affected by STO45849, and we found significant losses of pol II on the E3 promoter after treatment (Fig. 1, B and C). In contrast, the addition of the OGA inhibitor PUGNAc did not affect total pol II recruitment to the PIC (Fig. 1B). These experiments show that OGT activity is necessary for the optimal recruitment of pol II to the promoter.

We next asked whether other promoters might also contain pol II γ . We chose to analyze the CMV immediate early promoter as conditions for its activity have been worked out previously (12). We compared levels of pol II γ and total pol II in the PIC in the presence and absence of STO45849 using the WGA and pol II IP/Western blotting assays in Fig. 1, A and B. WGA failed to bind any pol II after treatment of the PIC with STO45849, indicating that pol II γ synthesis had been blocked (Fig. 1D, left panel). Additionally, there was no pol II present in the PICs after STO45849 treatment as determined by immunoprecipitating and detecting pol II with a pol II antibody, indicating that, as with the E3 promoter, total pol II recruitment into the PIC was blocked (Fig. 1D, middle panel).

To more accurately determine whether pol II γ was in the PIC, we immunoprecipitated pol II eluted from the PIC in BC500 and then asked whether anti-GlcNAc Western blotting would detect a signal where pol II migrated. This would eliminate any possibilities that the WGA experiments are purifying pol II indirectly via another *O*-GlcNAcylated protein (although our dissociation of the PIC in 500 mM KCl likely precludes this). Importantly, we found that the pol II IP contains a *O*-GlcNAcylated species that comigrates with pol II. This species was in the untreated PICs but was absent from the STO45849-treated PICs (Fig. 1D, right panel). This experiment shows that pol II γ is a component of the PIC and that it is formed during PIC formation.

We then asked whether removal of *O*-GlcNAc from pol II occurred during transcription. This might be expected if pol II γ is an active species in the PIC because CTD serine 5 residues must be available for phosphorylation by TFIIF, and previous experiments have indicated that *O*-GlcNAcylation of the CTD reduced its phosphorylation by TFIIF (5, 8). Also, the elongating species of pol II, pol IIO, is not *O*-GlcNAcylated (4), which again suggests that GlcNAc is removed from the CTD at initiation.

We formed PICs and then added NTPs to initiate transcription, eluted the PICs with 500 mM KCl, and assayed for *O*-GlcNAcylated pol II (WGA assay) and total pol II (pol II IP). The addition of NTPs significantly reduced the amount of pol II γ on the PIC, consistent with pol II promoter escape and GlcNAc removal during that phase of transcription (Fig. 2A).

Human RNA Pol II Promoter Recruitment Is Regulated by OGT

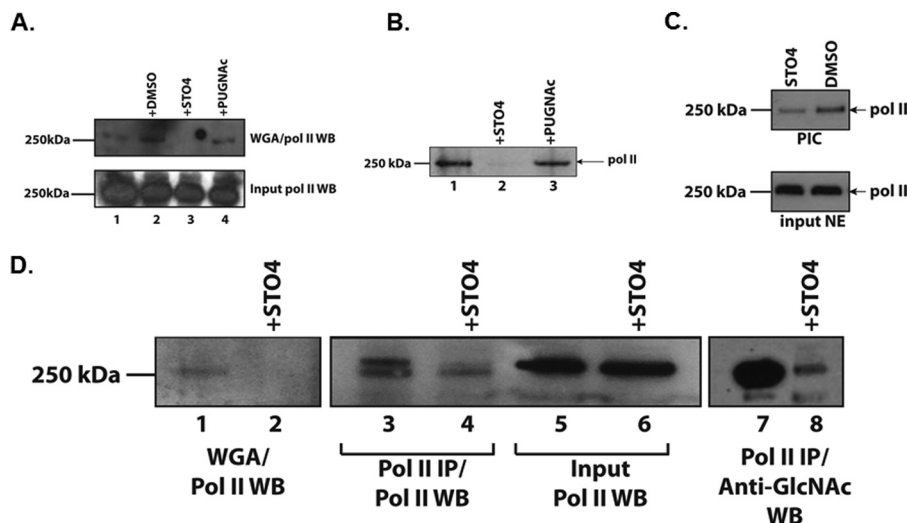


FIGURE 1. O-GlcNAcylated pol II localizes to promoters *in vitro*. *A*, WGA analysis for detection of pol II γ in PICs. PICs on the adenovirus E3 promoter were formed for 30 min in the presence or absence of the indicated inhibitor or DMSO control, washed in 100 mM KCl, and then eluted with 500 mM KCl. O-GlcNAcylated proteins were purified with WGA resin and analyzed by Western blotting for pol II Rpb1 subunit/8WG16 antibody (*WGA/pol II WB*). The *bottom panel* shows the amount of input pol II in the nuclear extract added to the PICs. *B*, analysis of total pol II content in PIC assays in the presence of STO45849 or PUGNac. PICs were assembled and washed in BC100, 0.05% Nonidet P-40, and pol II was eluted from the PIC in sample buffer, separated by 4–12% SDS-PAGE, transferred to nitrocellulose, and analyzed by Western blotting with the 8WG16 anti-pol II antibody. *C*, total pol II in the PIC is reduced by OGT inhibitor STO45849. PICs were assembled for 30 min with either STO45849 in DMSO or DMSO alone; then washed in BC100, 0.05% Nonidet P-40; eluted in BC500; and immunoprecipitated with 8WG16 anti-pol II antibody. IPs were subjected to SDS-PAGE, transferred to nitrocellulose, and probed with 8WG16. The *bottom panel* shows equivalent amounts of input nuclear extract added to the PICs as assayed by Western blotting for pol II. *D*, analysis of O-GlcNAcylated and total pol II in the PICs formed on the human CMV immediate early promoter. PICs were formed on the CMV promoter in the presence or absence of STO45849 (*STO4*). PICs were eluted in BC500. O-GlcNAcylated proteins were collected on WGA resin, washed, and loaded for SDS-PAGE. Western blots were probed using the anti-pol II antibody 8WG16 (*lanes 1 and 2, WGA/Pol II WB*). Total pol II was determined by immunoprecipitating and blotting with 8WG16 (*lanes 3 and 4, Pol II IP/Pol II WB*). 20% of nuclear extract input is also shown (*lanes 5 and 6, Input*; note that the input represents the amount of nuclear extract (NE) added to either the control or STO45849-containing assay tube). O-GlcNAcylated pol II was detected by pol II IP, SDS-PAGE, and subsequent Western blotting using the anti-O-GlcNAc antibody 110.6 (*right panel, Pol II IP/Anti-O-GlcNAc WB*). Data shown are representative of two to three experiments.

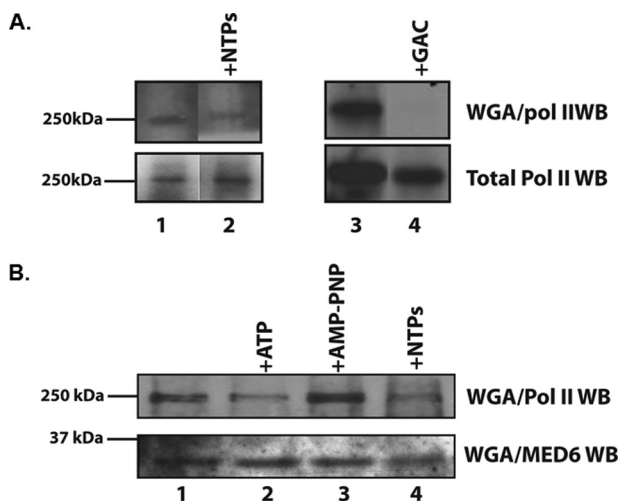


FIGURE 2. O-GlcNAc is removed from pol II upon initiation of transcription. *A*, addition of NTPs or rGAC promotes removal of GlcNAc from pol II. PICs were formed as in Fig. 1*A*. Transcription was initiated using either NTPs or the three ribonucleotides rGAC. PICs were washed in BC100, 0.05% Nonidet P-40 and then eluted in BC buffer containing 500 mM KCl (BC500). The BC500 eluate was diluted 1:1 with BC buffer and incubated with WGA-agarose beads. Total pol II was isolated from a duplicate PIC by eluting in BC500. The eluate was then immunoprecipitated with 8WG16. IPs were eluted in sample buffer, subjected to SDS-PAGE, and transferred to nitrocellulose. Pol II was detected by Western blotting (*WB*) with 8WG16 anti-CTD antibody. The figure was assembled from the same Western blots. *B*, ATP stimulates GlcNAc removal from pol II. PICs were formed using the E3 promoter as described above, and the indicated nucleotides were added. PICs were washed (BC100, Nonidet P-40) and eluted in BC500. The BC500 eluate was mixed with WGA-agarose beads to purify O-GlcNAcylated proteins. WGA-bound material was eluted with sample buffer, separated by 4–12% SDS-PAGE, transferred to nitrocellulose, and analyzed by Western blotting with anti-pol II or -MED6 antibodies. Data shown are representative of two to three experiments.

Similar results were obtained adding only rGAC, which allows elongation to approximately +10 (Fig. 2*A*). Additionally, ATP by itself stimulated removal of O-GlcNAc, whereas the nonhydrolyzable ATP analog AMP-PNP did not (Fig. 2*B*, *lanes 2 and 3*). In no case did we observe a significant loss of total pol II (*bottom panels*). As another control, although Mediator also bound to the WGA resin, indicating that it was O-GlcNAcylated, we did not observe any loss in Mediator from promoters upon NTP addition, indicating that it is not subject to OGA activity (Fig. 2*B*). The Mediator control is important because the Mediator complex can associate directly with pol II (13), and several Mediator subunits are O-GlcNAcylated (14). This experiment suggests that it is specifically pol II that is undergoing GlcNAc removal. These experiments confirm that the pol II γ is a promoter-localized species of pol II and suggest that the pol II γ is an active species of pol II and that the O-GlcNAc modification is specifically removed from the CTD upon initiation of transcription.

OGT Modifies Serine Residues 2 and 5 of the RNA Pol II CTD—OGT modifies serine and threonine residues (6). There are four such potential residues in the CTD repeat ($\gamma^1S^2P^3T^4S^5P^6S^7$): serine residues 2, 5, and 7 and threonine 4. Analysis of calf thymus pol II by Edman degradation found that both Thr⁴ and Ser⁵ were O-GlcNAcylated (4). However, serine residues are particularly labile in Edman reactions (15), which may have thus underrepresented serine 5 O-GlcNAcylation. Serine-to-alanine substitution experiments *in vitro* suggested that serine residues 5 and 7 are modified by OGT; the T4A mutant did not show any reduction in O-GlcNAcylation (5).

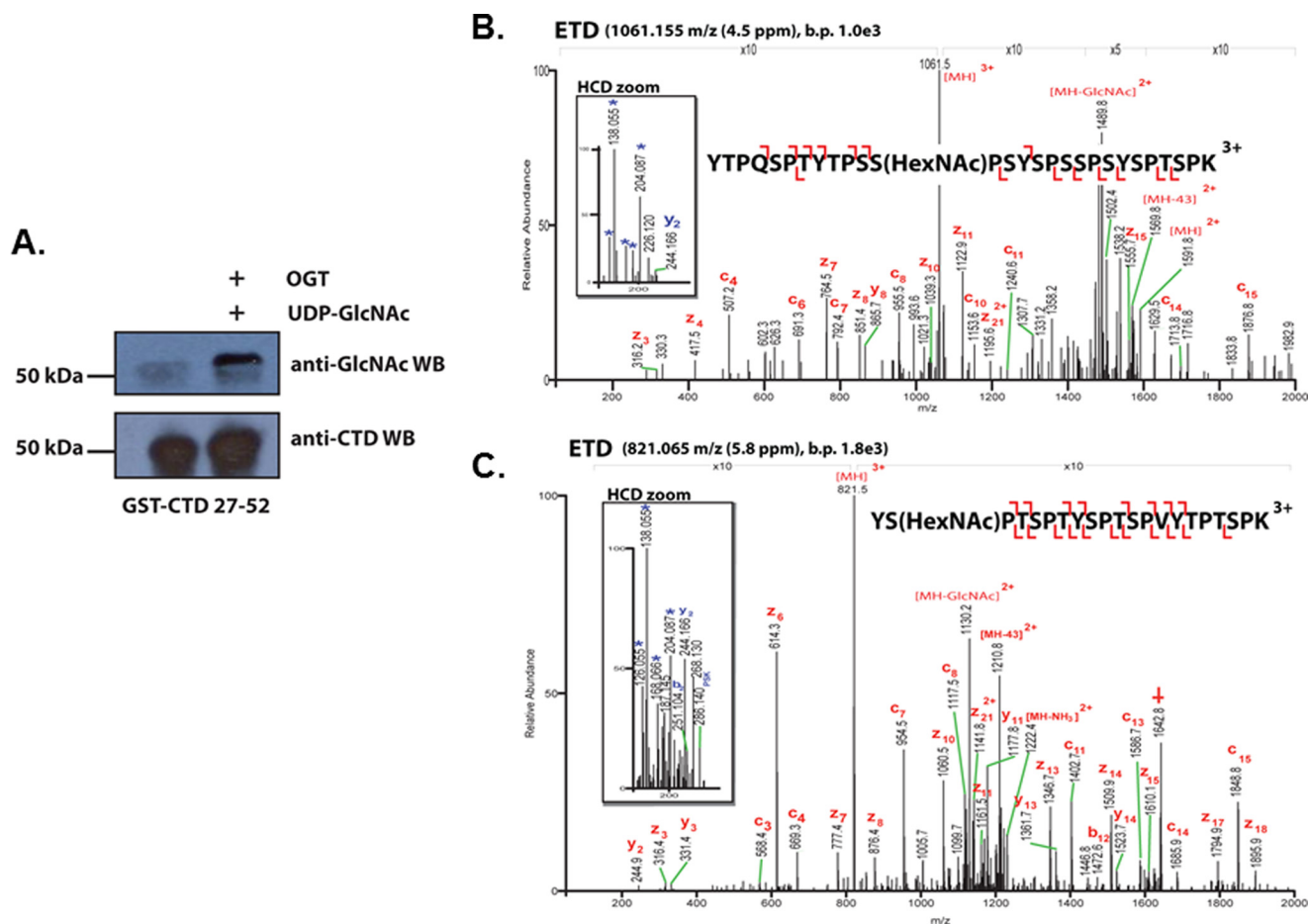


FIGURE 3. CTD O-GlcNAcylation occurs on serine residues 2 and 5. *A*, GST fused to the C-terminal nonconsensus half of human CTD (lane 1) was O-GlcNAcylated using recombinant OGT and UDP-GlcNAc (lane 2). O-GlcNAc was detected by Western blotting (WB) using the 110.6 anti-O-GlcNAc antibody. Equal loading of GST-CTD in both lanes was confirmed by a duplicate blot using the CTD antibody 8WG16. *B* and *C*, electron transfer dissociation (ETD) mass spectra GST-CTD repeats 34 (*B*) and 44 (*C*) modified by *N*-acetylglucosamine (HexNAc) at serine 5 (*B*) or serine 2 (*C*). The inset labeled “HCD zoom” indicates the detection with high mass accuracy of the GlcNAc oxonium ion and fragments thereof (denoted by *) by higher energy collisional dissociation of the same precursor ion. *b.p.*, base peak. *MH* refers to precursor mass. \perp indicates co-isolated contaminating ion.

The potential indirectness of our alanine mutant analysis and the observed differences between native calf thymus pol II and recombinant CTD substrate required that we pursue this further.

The main difficulty in interrogating the pol II CTD by mass spectrometry is a lack of proteolytic cleavage sites in the first 26 repeats. We thus focused on analysis of human repeats 27–52 where the lysine residues serve as trypsin cleavage sites. We treated recombinant GST-CTD protein, containing human repeats 27–52, with recombinant OGT and UDP-GlcNAc (Fig. 3*A*) and then performed LC-MS/MS analysis (16). Collisional activation and electron transfer dissociation MS/MS of GST-CTD tryptic peptides showed O-GlcNAcylation of serine 5 in CTD repeat 34, indicating that serine 5 is a direct target of OGT (Fig. 3*B*). In contrast, we did not observe any O-GlcNAcylation of Thr⁴ *in vitro*. These data suggest that the previously observed block to TFIIF-dependent phosphorylation of the serine 5 residue of the pol II CTD (5, 8) is due to O-GlcNAcylation of serine 5. Additionally, we noted, unexpectedly, that the serine 2 residue of the 44th CTD repeat was O-GlcNAcylated (Fig. 3*C*). A serine 2 modification was not predicted from the original calf thymus pol II analysis. However, as noted previously, serine

residues are particularly labile in Edman degradation reactions (15, 17).

Discussion

The Pol II γ Model of PIC Formation—The model for pol II species involved in PIC formation, initiation, and elongation was developed through the late 1980s and early 1990s. In systems containing purified human factors, pol IIA preferentially formed PICs as compared with pol IIO (3). An even more striking result was obtained with partially purified fractions from a whole cell extract that were devoid of pol II. The subsequent addition of calf thymus pol IIA or pol IIO clearly showed retention of pol IIA but not pol IIO in PICs that were transcriptionally active (2). One criticism of the pol IIA model is that the calf thymus pol IIA was likely O-GlcNAcylated as well (4), and thus it is ambiguous as to which pol II species are compatible with PIC formation.

In contrast, in highly purified systems using recombinant general transcription factors and native pol II, no difference between pol IIA and O was seen in abortive initiation assays (18). Because these systems lack OGT, OGA, or UDP-GlcNAc, it is likely that they do not functionally distinguish among A, O,

Human RNA Pol II Promoter Recruitment Is Regulated by OGT

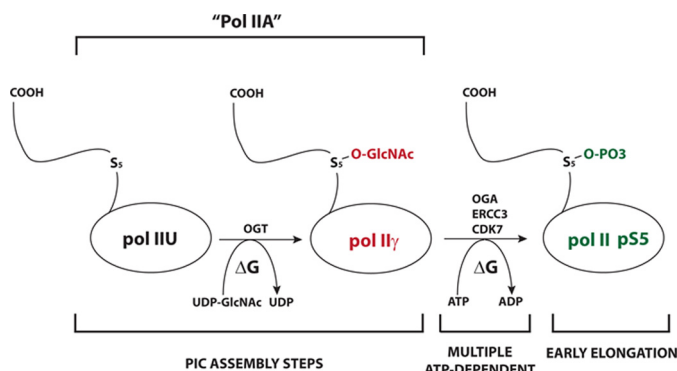


FIGURE 4. Model of RNA polymerase II modification during PIC formation and transcription initiation. Our work herein suggests that pol II γ is at least one of the final species of pol II that is stably associated with a PIC. We hypothesize a model where pol IIU associates with the promoter and is O-GlcNAcylated by OGT during PIC assembly to form pol II γ (because OGT is in the PIC (5)). Then pol II γ becomes a target of regulation of initiation events where GlcNAc must be removed for subsequent CTD phosphorylation (Fig. 2 and Ref. 21). Transcription initiation commences with several ATP-dependent steps: the activation of OGA, the ERCC3 helicase, and CDK7 kinase subunits within TFIIF (22). OGA then removes GlcNAc from the pol II CTD, permitting CDK7 phosphorylation of serine 5 in the CTD. The data and model also suggest that there are several distinct kinetic steps during PIC formation defined by OGT and OGA activity. In this model, pol IIA is likely a mixture of pol IIU and pol II γ as shown previously (4).

and γ species. In contrast, crude *in vitro* systems contain factors that distinguish between pol IIA and pol IIO (2), and Dignam-based (23) crude nuclear extract systems are O-GlcNAc-dependent (5).

Our data herein offer evidence that pol II γ is a promoter-specific species of pol II. We have detected pol II γ on PICs formed with crude nuclear extracts that are OGT- and OGA-dependent transcription systems (5). These experiments also show a requirement for OGT in PIC assembly and recruitment of pol II as an OGT inhibitor blocks the formation of pol II γ and reduces the amount of pol II in the PIC (Fig. 1). This result is supported by OGT shRNA data showing reductions in pol II promoter occupancy *in vivo* (5). Furthermore, the addition of ATP (or NTPs) shows that the removal of GlcNAc from pol II occurs concomitantly with the initiation of transcription as would be expected for a pol II participating in initiation and elongation. Note that our use of OGT and OGA inhibitors does not comment on what proteins in the PIC, other than pol II, are substrates for these enzymes. It would not be surprising if other PIC components are O-GlcNAcylated.

Our data suggest either that non-GlcNAcylated pol II associates with the promoter and is O-GlcNAcylated by OGT during PIC assembly to form pol II γ or that pol II is recruited to the PIC as pol II γ (Fig. 4). Based on the stable association of pol II γ with the PIC, we suggest that pol II γ , not the unmodified pol II, represents the final pol II species in the assembly of the PIC (Fig. 4). In principle, unmodified pol II (pol IIU) exists in either case as previous experiments indicate that GlcNAc must be removed for phosphorylation to occur. The only question is whether pol IIU exists transiently or stably. Thereafter, OGA is activated by an ATP-dependent step during initiation, catalyzing GlcNAc removal and permitting subsequent CTD phosphorylation (Figs. 1, 2, and 4). However, we cannot exclude the possibility that there are OGT-dependent and -independent

pathways of PIC formation. In addition, the lack of complete removal of GlcNAc suggests that it may continue to exist on an elongating pol II.

O-GlcNAcylation of CTD Serine Residues 2 and 5—Previous Edman degradation data showed both Thr⁴ and Ser⁵ O-GlcNAcylation in calf thymus pol II (4). Furthermore, a GlcNAcylated CTD is refractory to TFIIF phosphorylation (5, 8). Lastly, CTD alanine substitutions did not show a requirement for Thr⁴ in a Thr-to-Ala mutant but did show a requirement for Ser⁵ in reactions using recombinant human OGT (5). MS analysis found that serine 5 was in fact O-GlcNAcylated (Fig. 3). Our conclusion then is that serine 5 can be modified by either TFIIF or OGT. These experiments all suggest that O-GlcNAcylation of serine 5 can regulate the amount of phosphorylation of serine 5. We also found the first evidence that serine 2 residues can be O-GlcNAcylated. This suggests that both serine 2 and 5 O-GlcNAcylation are regulated events during the initiation and elongation processes and may affect TFIIF and PTEFb (positive transcriptional elongation factor b phosphorylation) of the CTD. It is reasonable to suggest that these O-GlcNAcylation events prevent premature or aberrant phosphorylation of the CTD or even serve to regulate those phosphorylation events during the transcription cycle. Further experimentation is required to examine the causal effects of these O-GlcNAcylation events. Additionally, techniques must be invented to access the distribution of O-GlcNAcylated residues in the N-terminal half of the human CTD, which is refractory to mass spectrometry.

In vivo, mutational analysis of the CTD to determine functional consequences of lack of O-GlcNAcylation is of limited use. For example, because serine 5 is both phosphorylated and O-GlcNAcylated, a mutated serine would represent two defects, a loss of both phosphorylation and O-GlcNAcylation, and thus it would be impossible to assign any functional consequences to the absence of one modification *versus* the other.

Additional Functions of Pol II γ —The discovery of pol II γ in the PIC indicates the existence of another level of regulation in the PIC mediated by O-GlcNAc cycling. We hypothesize several functions of pol II γ . First, the O-GlcNAcylation of the CTD is necessary for pol II entry into the PIC. Second, CTD O-GlcNAcylation might also serve to prevent or promote factors binding to the CTD or may prevent premature, aberrant CTD phosphorylation. Lastly, the hydrolysis of UDP-GlcNAc, which is a high energy donor, may make free energy contributions (similar to ATP hydrolysis (19)) to promote kinetic steps in PIC formation (Fig. 4).

Summary—The existence of pol II γ on promoters and the transcriptional requirement for O-GlcNAc and its cycling are both unexpected and outside the paradigms of transcriptional regulation. Our work sheds new light on the mechanism and regulation of transcription by RNA polymerase II; suggests a more physiological definition of a PIC; and draws direct connections between the cellular nutrient state, RNA polymerase II function, and genome-wide transcriptional regulation in diseases such as cancer and diabetes where elevated O-GlcNAcylation is part of a transformed and insulin resistance phenotype (6, 20).

Author Contributions—B. A. L. conceived the study, designed the experiments, performed the experiments in Figs. 1 and 2, and wrote the manuscript. A. L. B. and S. A. M. performed and interpreted the mass spectrometry data in Fig. 3.

Acknowledgments—GST-CTD 27–52 expression vector was kindly provided by David Bentley. We also thank Jason Piotrowski for expert technical assistance, Jerry Hart for 110.6 antibody, and David Levens and Dinah Singer for comments and critical reading of the manuscript.

References

1. Sims, R. J., 3rd, Belotserkovskaya, R., and Reinberg, D. (2004) Elongation by RNA polymerase II: the short and long of it. *Genes Dev.* **18**, 2437–2468
2. Chesnut, J. D., Stephens, J. H., and Dahmus, M. E. (1992) The interaction of RNA polymerase II with the adenovirus-2 major late promoter is precluded by phosphorylation of the C-terminal domain of subunit IIa. *J. Biol. Chem.* **267**, 10500–10506
3. Lu, H., Flores, O., Weinmann, R., and Reinberg, D. (1991) The nonphosphorylated form of RNA polymerase II preferentially associates with the preinitiation complex. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 10004–10008
4. Kelly, W. G., Dahmus, M. E., and Hart, G. W. (1993) RNA polymerase II is a glycoprotein. Modification of the COOH-terminal domain by O-GlcNAc. *J. Biol. Chem.* **268**, 10416–10424
5. Ranuncolo, S. M., Ghosh, S., Hanover, J. A., Hart, G. W., and Lewis, B. A. (2012) Evidence of the involvement of O-GlcNAc-modified human RNA polymerase II CTD in transcription *in vitro* and *in vivo*. *J. Biol. Chem.* **287**, 23549–23561
6. Bond, M. R., and Hanover, J. A. (2013) O-GlcNAc cycling: a link between metabolism and chronic disease. *Annu. Rev. Nutr.* **33**, 205–229
7. Ruan, H. B., Singh, J. P., Li, M. D., Wu, J., and Yang, X. (2013) Cracking the O-GlcNAc code in metabolism. *Trends Endocrinol. Metab.* **24**, 301–309
8. Comer, F. I., and Hart, G. W. (2001) Reciprocity between O-GlcNAc and O-phosphate on the carboxyl terminal domain of RNA polymerase II. *Biochemistry* **40**, 7845–7852
9. Cheng, B., and Price, D. H. (2009) Isolation and functional analysis of RNA polymerase II elongation complexes. *Methods* **48**, 346–352
10. Gross, B. J., Kraybill, B. C., and Walker, S. (2005) Discovery of O-GlcNAc transferase inhibitors. *J. Am. Chem. Soc.* **127**, 14588–14589
11. Macauley, M. S., and Vocadlo, D. J. (2010) Increasing O-GlcNAc levels: an overview of small-molecule inhibitors of O-GlcNAcase. *Biochim. Biophys. Acta* **1800**, 107–121
12. Adamson, T. E., Shore, S. M., and Price, D. H. (2003) Analysis of RNA polymerase II elongation *in vitro*. *Methods Enzymol.* **371**, 264–275
13. Jones, K. A., Kadonaga, J. T., Rosenfeld, P. J., Kelly, T. J., and Tjian, R. (1987) A cellular DNA-binding protein that activates eukaryotic transcription and DNA replication. *Cell* **48**, 79–89
14. Hsu, J. Y., Juven-Gershon, T., Marr, M. T., 2nd, Wright, K. J., Tjian, R., and Kadonaga, J. T. (2008) TBP, Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-dependent transcription. *Genes Dev.* **22**, 2353–2358
15. Reason, A. J., Dell, A., Romero, P. A., and Herscovics, A. (1991) Specificity of the mannosyltransferase which initiates outer chain formation in *Saccharomyces cerevisiae*. *Glycobiology* **1**, 387–391
16. Myers, S. A., Daou, S., Affar el, B., and Burlingame, A. (2013) Electron transfer dissociation (ETD): The mass spectrometric breakthrough essential for O-GlcNAc protein site assignments—a study of the O-GlcNAcylated protein host cell factor C1. *Proteomics* **13**, 982–991
17. Hart, G. W., Housley, M. P., and Slawson, C. (2007) Cycling of O-linked β -N-acetylglucosamine on nucleocytoplasmic proteins. *Nature* **446**, 1017–1022
18. Goodrich, J. A., and Tjian, R. (1994) Transcription factors IIE and IIH and ATP hydrolysis direct promoter clearance by RNA polymerase II. *Cell* **77**, 145–156
19. Qian, H. (2007) Phosphorylation energy hypothesis: open chemical systems and their biological functions. *Annu. Rev. Phys. Chem.* **58**, 113–142
20. Ma, Z., and Vosseller, K. (2013) O-GlcNAc in cancer biology. *Amino Acids* **45**, 719–733
21. Lewis, B. A. (2013) O-GlcNAcylation at promoters, nutrient sensors, and transcriptional regulation. *Biochim. Biophys. Acta* **1829**, 1202–1206
22. Thomas, M. C., and Chiang, C. M. (2006) The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**, 105–178
23. Dignam, J. D., Lebovitz, R. M., and Roeder, R. G. (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. *Nucleic Acids Res.* **11**, 1475–1489