Distributive O-GlcNAcylation on the Highly Repetitive C-terminal Domain of RNA Polymerase II

*Lei Lu†, Dacheng Fan†, Chia-Wei Hu†, Matthew Worth‡, Zhi-Xiong Ma†, Jiaoyang Jiang†**

†Pharmaceutical Sciences Division, School of Pharmacy, and ‡Department of Chemistry,

University of Wisconsin-Madison, Madison, Wisconsin 53705

Supporting Information

Construction of Plasmids. CTD constructs: The DNA sequence of CTD52 was synthesized with codon optimization for *E. coli* expression. The synthetic DNA fragment was amplified by PCR followed by restriction enzyme digestion with BsaI and HindIII. The digested CTD52 fragment was ligated into pE-SUMOstar vector (LifeSensors) that was digested with the same restriction enzymes. The final His-SUMO-CTD52 plasmid has an N-terminal His_6-tag and SUMO tag, following by CTD52 sequence. Plasmids of His-SUMO-CTD7, 10, 15 and 20 were similarly cloned into pE-SUMOstar vector between BsaI and HindIII sites. Below is the DNA sequence of synthesized CTD52:

AGGTTACTCCCCGACCAGCCCTGCGTACGAGCCGCGTTCTCCGGGTGGCTATACCC
CACAATCTCCGTCGTATTCCCCGACTTCTCCGTCCTACAGCCCTACTAGCCCATCGTA
CAGCCCAACGAGCCCGAACTACAGCCCGACTTCCCCGAGCTATAGCCCAACTAGTC
CGTCGTACTCCCCTACGAGCCCTAGCTACTCCCCTACCTCCCCGTCGTACAGCCCAA
CGAGCCCTTCCTATTCCCCGACGAGCCCAAGCTACTCTCCTACCTCGCCTTCCTATAG
CCCGACGTCCCCAAGCTACAGCCCGACGAGCCCTAGCTACTCGCCGACCAGCCCGA
GCTACTCCCCAACCTCCCCTAGCTACAGCCCGACCAGCCCGTCTTATTCCCCGACCT
CTCCGTCTTACAGCCCTACTTCCCCGTCTTACTCTCCAACGTCCCCGAGCTACAGCCC
GACTAGCCCAAATTACAGCCCAACCAGCCCGAATTACACGCCGACCAGCCCGAGCT
ATAGCCCTACCAGCCCATCTTACAGCCCAACCTCCCCAAATTACACCCCAACGTCTC
CGAATTACTCTCCAACCTCTCCGAGCTATTCGCCTACCAGCCCGTCCTACTCCCCGAC
CTCGCCGAGCTACTCGCCGAGCTCCCCGCGTTACACGCCGCAGAGCCCGACCTACAC
CCCGTCCTCGCCGTCGTACTCTCCGTCTAGCCCGTCGTATAGCCCTGCGAGCCCTAA
GTACACGCCTACGAGCCCGTCTTATAGCCCGTCTAGCCCAGAATATACCCCTACCAG
CCCAAAGTACTCCCCAACCTCCCCGAAATACAGCCCGACGAGCCCGAAATACAGCC

CAACGAGCCCGACCTATTCTCCGACGACCCCGAAATACAGCCCGACGAGCCCGACT
TATTCCCCTACTTCTCCGGTGTACACCCCTACGAGCCCGAAGTATAGCCCGACGAGC
CCGACCTACTCTCCGACCAGCCCAAAGTACAGCCCGACGAGCCCAACGTACAGCCC
TACCTCTCCTAAAGGCAGCACGTATTCTCCGACCTCGCCTGGTTATTCCCCAACTAG
CCCAACCTACAGCCTGACGAGCCCGGCTATCAGCCCGGATGACAGCGATGAAGAAA
ACTAA

OGT constructs: The ncOGT (full-length human OGT), $OGT_{6.5}$ (spanning residues 245-1031 based on the numbering of the full-length protein), $OGT_{5.5}$ (spanning residues 279-1031), and $OGT_{4.5}$ (spanning residues 313-1031) constructs were kindly provided by the Suzanne Walker lab at Harvard Medical School. All of these OGT plasmids were cloned into pET24b vector with OGT protein fused to an N-terminal T7 tag, followed by a His_8-tag and an HRV3C protease cleavage site (LEVLFQGP). The ncOGT and $OGT_{4.5}$ constructs were reported previously.[1]

CTD52 and ncOGT co-expression construct: CTD52 sequence was amplified by PCR and ligated into pETDuet-1 vector (Novagen) between restriction sites of NcoI and NotI in MCS1 region to generate plasmid pETDuet-CTD52. This plasmid and the PCR amplified ncOGT sequence were both digested with restriction enzymes of NdeI and XhoI, and then the digested ncOGT fragment was ligated into the MCS2 region of pETDuet-CTD52 to generate final plasmid of pETDuet-CTD52-ncOGT.

OGA construct: cDNA of human OGA (isoform 1, Uniprot identifier #O60502-1) was amplified by PCR and cloned into pET21b vector between SalI and XhoI sites. The resulting OGA protein contains a C-terminal His_6-tag.

Expression and Purification of Proteins. OGTs: The plasmid of each OGT construct (ncOGT, $OGT_{6.5}$, $OGT_{5.5}$, and $OGT_{4.5}$) was transformed into BL21(DE3) competent cells. OGT proteins were expressed and purified as previously reported.[1]

CTDs: Each His-SUMO-CTD plasmid (CTD7, 10, 15, 20, and 52) was transformed into BL21(DE3)CodonPlusRIL competent cells. CTD proteins were expressed and purified similarly as OGT.

O-GlcNAcylated CTD52: The plasmid of pETDuet-CTD52-ncOGT was transformed into BL21(DE3)CodonPlusRIL competent cells for co-expression of ncOGT (without His-tag) and His-SUMO-CTD52. The purification of O-GlcNAcylated His-SUMO-CTD52 protein was conducted similarly as OGT.

OGA: The plasmid of OGA was transformed into BL21(DE3) competent cells. Protein purification was conducted similarly as OGT.

REFERENCES

(1) Lazarus, M. B., Nam, Y., Jiang, J., Sliz, P., and Walker, S. (2011) Structure of human O-GlcNAc transferase and its complex with a peptide substrate. *Nature 469*, 564–567.

Figure S1

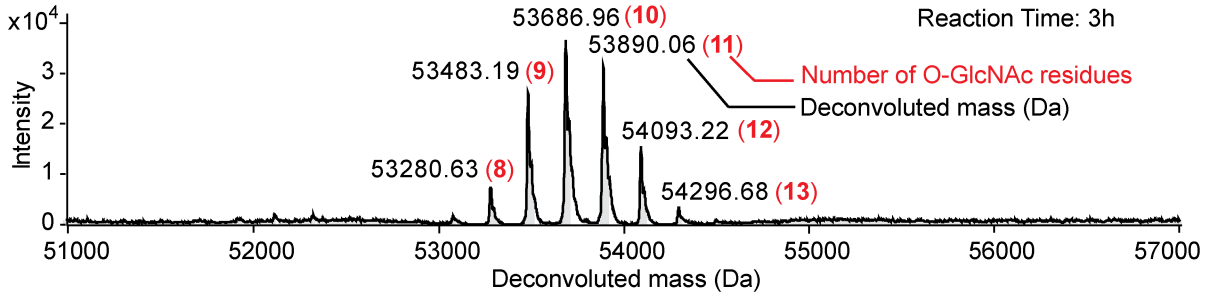

Figure S1. Intact protein MS of O-GlcNAcylated SUMO-CTD52 at a higher molar ratio. This experiment was conducted for 3 h with an increased ratio of protein concentration (SUMO-CTD52:OGT = 10:1). The numbers of O-GlcNAc residues are highlighted in red and bold.

Figure S2

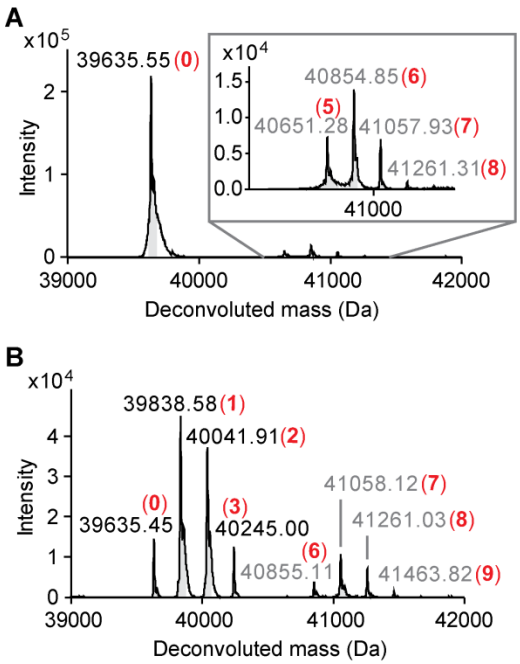

Figure S2. Deconvoluted mass spectra from distraction assay. (A) CTD52 (without tag) was partially glycosylated for 1 h (deconvoluted masses are enlarged in the insert and shown in gray) when the second batch of unmodified CTD52 (deconvoluted mass is shown in black) was added. Reaction was immediately stopped without further incubation. (B) Deconvoluted mass spectrum of the same reaction following addition of the second batch of CTD52 and incubation for another 15 min (same labeling as above). The numbers of O-GlcNAc residues are highlighted in red and bold.

Figure S3

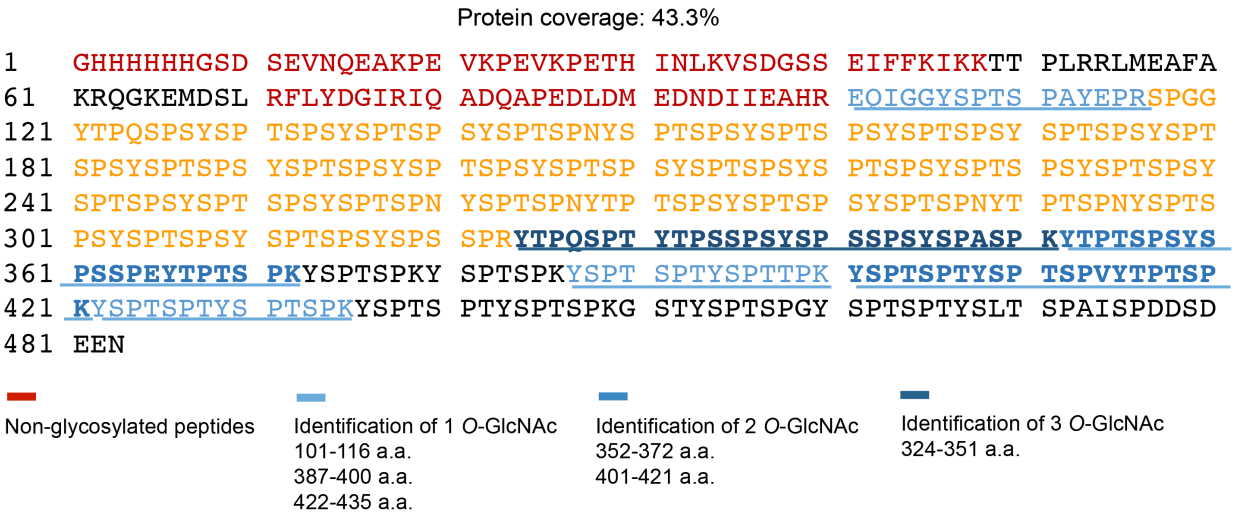

Figure S3. Sequence coverage and detected peptides from in-gel trypsin digestion of glycosylated SUMO-CTD52 after 60 min incubation with ncOGT and UDP-GlcNAc. Non-glycosylated, singly, doubly, and triply glycosylated peptides are highlighted in red, light blue, medium blue, and dark blue, respectively. The peptide in yellow was detected from in-solution trypsin digestion followed by intact protein MS, and the O-GlcNAcylation level was summarized in Figure 5.

Figure S4

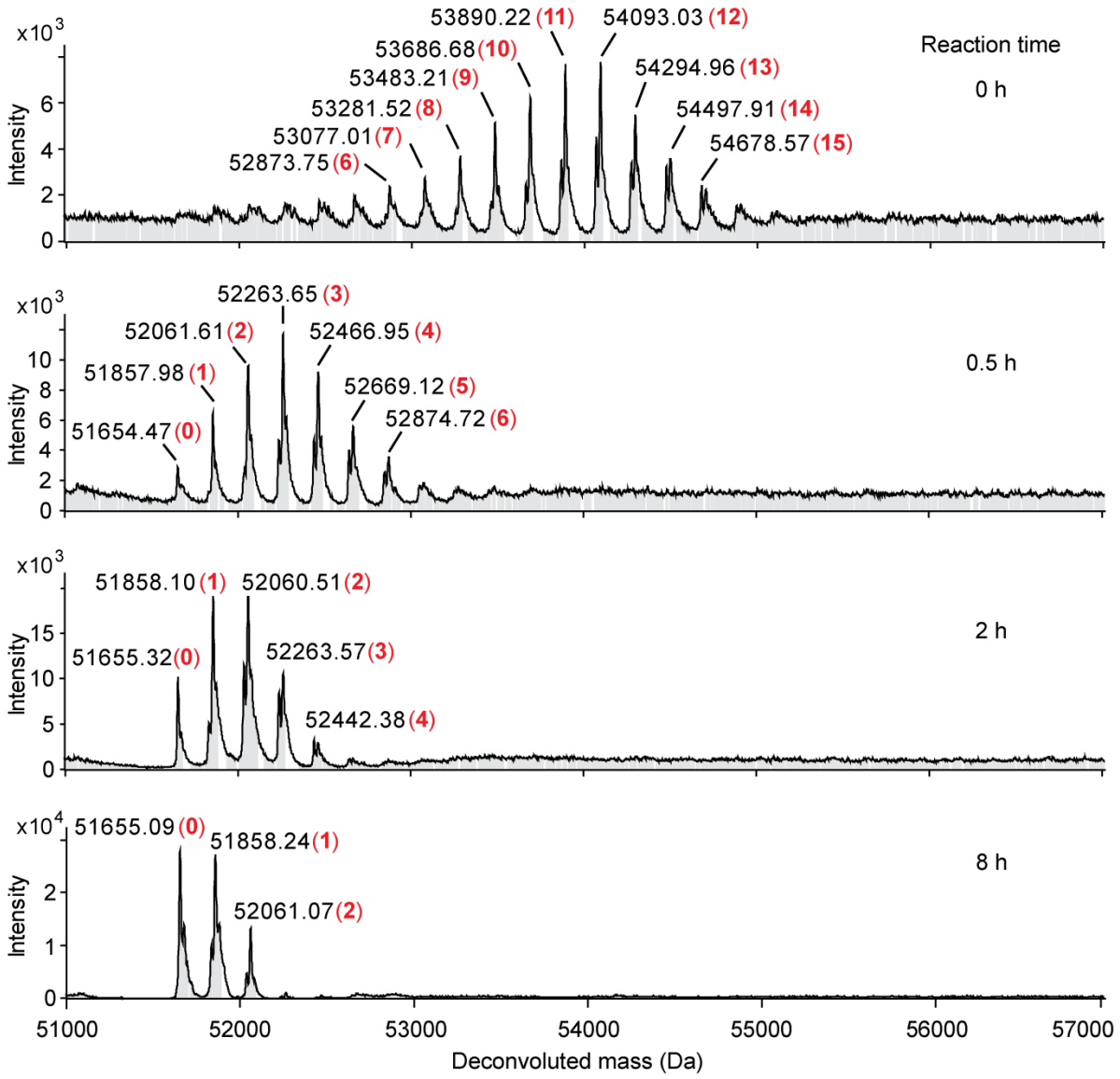

Figure S4. Intact protein MS showing that OGA almost completely removed O-GlcNAc from glycosylated CTD. An increased molar ratio of OGA to O-GlcNAcylated SUMO-CTD52 of 1:4 was applied. The numbers of O-GlcNAc residues are highlighted in red and bold.