Rumi functions as both a protein *O*-glucosyltransferase and a protein *O*-xylosyltransferase

Hideyuki Takeuchi[a], Rodrigo C. Fernández-Valdivia[b], Devin S. Caswell[a], Aleksandra Nita-Lazar[a,1], Nadia A. Rana[a], Thomas P. Garner[c], Thomas K. Weldeghiorghis[c], Megan A. Macnaughtan[c], Hamed Jafar-Nejad[b], and Robert S. Haltiwanger[a,2]

[a]Department of Biochemistry and Cell Biology, Institute of Cell and Developmental Biology, Stony Brook University, Stony Brook, NY 11794; [b]The Brown Foundation Institute of Molecular Medicine , Center for Metabolic and Degenerative Diseases, Department of Biochemistry and Molecular Biology, Medical School, University of Texas Health Science Center at Houston, Houston, TX 77030; and [c]Department of Chemistry, Louisiana State University, Baton Rouge, LA 70803

Edited* by Armando Parodi, Fundacion Instituto Leloir, Buenos Aires, Argentina, and approved August 18, 2011 (received for review June 17, 2011)

Mutations in *rumi* result in a temperature-sensitive loss of Notch signaling in *Drosophila*. *Drosophila* Rumi is a soluble, endoplasmic reticulum-retained protein with a CAP10 domain that functions as a protein *O*-glucosyltransferase. In human and mouse genomes, three potential Rumi homologues exist: one with a high degree of identity to *Drosophila* Rumi (52%), and two others with lower degrees of identity but including a CAP10 domain (KDELC1 and KDELC2). Here we show that both mouse and human Rumi, but not KDELC1 or KDELC2, catalyze transfer of glucose from UDP-glucose to an EGF repeat from human factor VII. Similarly, human Rumi, but not KDELC1 or KDELC2, rescues the *Notch* phenotypes in *Drosophila rumi* clones. During characterization of the Rumi enzymes, we noted that, in addition to protein *O*-glucosyltransferase activity, both mammalian and *Drosophila* Rumi also showed significant protein *O*-xylosyltransferase activity. Rumi transfers Xyl or glucose to serine 52 in the *O*-glucose consensus sequence (^{50}CASSPC^{55}) of factor VII EGF repeat. Surprisingly, the second serine (S53) facilitates transfer of Xyl, but not glucose, to the EGF repeat by Rumi. EGF16 of mouse Notch2, which has a diserine motif in the consensus sequence (^{587}CYSSPC^{592}), is also modified with either *O*-Xyl or *O*-glucose glycans in cells. Mutation of the second serine (S590A) causes a loss of *O*-Xyl but not *O*-glucose at this site. Altogether, our data establish dual substrate specificity for the glycosyltransferase Rumi and provide evidence that amino acid sequences of the recipient EGF repeat significantly influence which donor substrate (UDP-glucose or UDP-Xyl) is used.

development | glycobiology | mass spectrometry

The Notch signaling pathway plays a central role in developmental processes throughout the animal kingdom, controlling cell fate decisions in numerous tissues (1, 2). Its diverse functions are reflected in the wide range of human diseases linked to defects in the pathway, including Alagille syndrome, spondylocostal dysostosis, cerebral autosomal dominat arteriopathy with subcortical infarcts and leukoencephalopathy, multiple sclerosis, and several types of cancers (3). The extracellular domains of Notch receptors contain up to 36 EGF repeats, some of which can be modified with three unusual types of *O*-linked glycosylation: *O*-fucose, *O*-glucose, and *O*-GlcNAc (4, 5). Although many proteins contain EGF repeats, the consensus sequences for both *O*-fucose and *O*-glucose modifications are abundant and conserved in the EGF repeats of Notch proteins (4, 6). Work from several laboratories has shown the biological importance of both the *O*-fucose and *O*-glucose modifications on Notch (7–14).

Although glucose has long been known to exist as a biosynthetic intermediate on *N*-glycans, less is known about protein *O*-glucosylation, where glucose is linked to the hydroxyl of a distinct serine residue in the C^1-X-S-X-P-C^2 consensus amino acid sequence of an EGF repeat (15). The glucose is added by protein *O*-glucosyltransferase (Poglut) (16) and is known to be extended to a trisaccharide (Xyl-α1,3-Xyl-α1,3-Glc) on all mammalian proteins where it has been found, such as Notch (4, 17, 18), coagulation factors VII and IX, protein Z, Delta-like protein, and thrombospondin (19–21). Although Notch proteins carry many evolutionarily conserved *O*-glucosylation sites (4, 6), the biological role of *O*-glucosylation has until recently been unknown. We recently showed that *rumi*, the gene encoding Poglut, is essential for Notch signaling in *Drosophila*. Rumi contains a signal peptide, a CAP10 domain, and a Lys-Asp-Glu-Leu (KDEL)-like endoplasmic reticulum (ER) retention signal at its C terminus (12). The CAP10 gene had been suggested to be involved in the formation of capsule and virulence of *Cryptococcus neoformans*, providing a clue that the protein may be a glycosyltransferase (22). We showed that Rumi is a Poglut that modifies Notch. Moreover, we found that the Poglut activity of Rumi is essential for Notch function by analyzing a mutant allele with a point mutation (G189E) that eliminates the enzymatic activity of Rumi without affecting protein levels (12).

Here, we report that, of the three mammalian proteins containing CAP10 domains and ER retention signals, only the closest homologue to *Drosophila* Rumi (dRumi) rescues the *Notch* phenotype in *rumi* mutant flies in vivo and has Poglut activity in vitro. Surprisingly, biochemical analysis of the enzymatic activity of Rumi revealed that both mammalian and dRumi transfer xylose (Xyl) as well as glucose to some of their target EGF repeats. Furthermore, *O*-xylosylation of specific Notch EGF repeats occurs on Notch proteins produced in cells, suggesting that this modification is likely to be biologically relevant.

Results

To examine the importance of *O*-glucosylation for mammalian Notch, mammalian homologues of Rumi were identified using BLAST searches. In human and mouse genomes, three potential homologues were identified: one with a high degree of identity (54% identity to dRumi, also known at hCLP46/KTELC1; refs. 23 and 24), designated as Rumi, and two others (KDELC1/C2) with lower degrees of identity but with the clear presence of a CAP10 domain (25) (Fig. 1 *A* and *B*, and Fig. S1). All three genes encode a signal peptide, a CAP10 domain and a KDEL-like ER retention signal at the C terminus. Unlike Rumi, KDELC1 and KDELC2 contain a filamin-like domain (26). To examine whether these can regulate Notch signaling in vivo, we took a

Author contributions: H.T., R.C.F.-V., H.J.-N., and R.S.H. designed research; H.T., R.C.F.-V., D.S.C., A.N.-L., N.A.R., T.P.G., T.K.W., M.A.M., H.J.-N., and R.S.H. performed research; H.T., R.C.F.-V., D.S.C., N.A.R., M.A.M., H.J.-N., and R.S.H. analyzed data; and H.T., R.C.F.-V., H.J.-N., and R.S.H. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]Present address: National Institute of Allergy and Infectious Diseases, National Institutes of Health, Bethesda, MD 20892.

[2]To whom correspondence should be addressed. E-mail: rhaltiwanger@ms.cc.sunysb.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1109696108/-/DCSupplemental.

Downloaded by guest on January 23, 2022

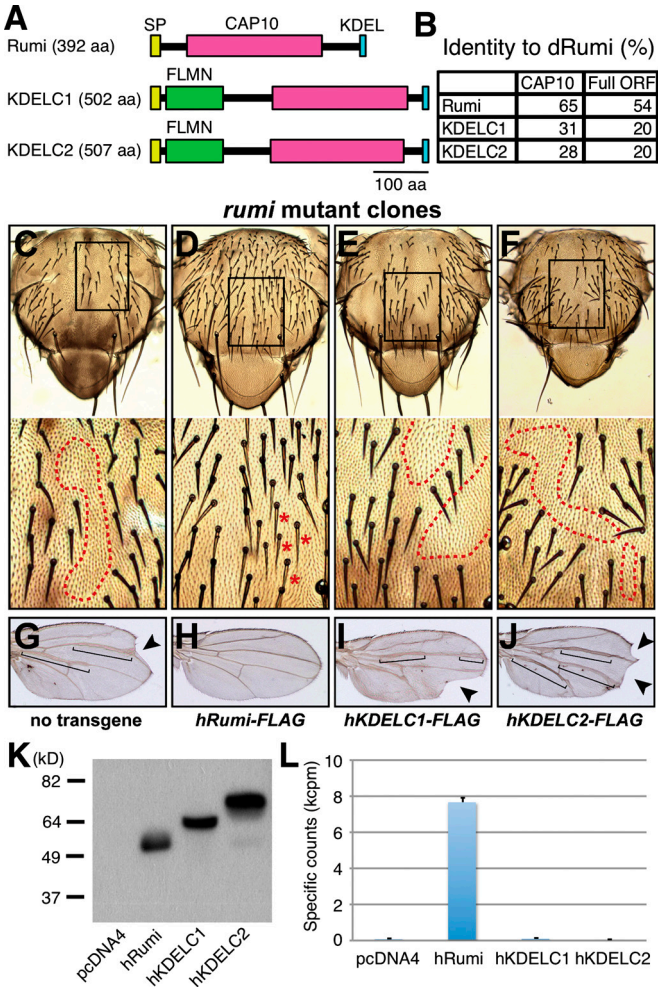

Fig. 1. Human Rumi is the functional and biochemical homologue of *Drosophila* Rumi. (*A*) Molecular structure of human proteins containing a CAP10 domain. SP, signal peptide; KDEL, KDEL ER-retention signal; FLMN, filamin-like domain. (*B*) Percent identity of amino acid sequences of human Rumi, KDELC1, and KDELC2 with dRumi. (*C*–*F*) Thoraces of adult flies harboring randomly located *rumi*$^{\Delta 26}$ mutant clones. The lower panels show close-ups of the boxes in the upper panels. The *rumi*$^{\Delta 26}$ mutant clones result in loss of sensory organs on thorax (*C*, dotted lines in the close-up). Simultaneous overexpression of a FLAG-tagged version of hRumi restores the lost sensory organs in *rumi*$^{\Delta 26}$ mutant clones (*D*). Note that the rescued sensory organs—four of which are marked by asterisks—lose the y^+ transgene on the reciprocal 3R chromosome and are therefore yellow, as opposed to *rumi*$^{+/-}$ heterozygous and *rumi*$^{+/+}$ sensory organs, which look darker. Human KDELC1–FLAG (*E*) or hKDELC2–FLAG (*F*) were not able to rescue the phenotype (dotted lines in the close-ups). (*G*–*J*) Wings of adult flies harboring randomly located *rumi*^{44} mutant clones. Loss of *rumi* results in the thickening of wing veins (brackets) and "Notches" of tissue loss in the wing margin (arrowhead) (*G*). Only hRumi–FLAG is able to rescue these phenotypes (compare *H* with *I* and *J*). All animals were raised at 30 °C after the second instar larval stage. (*K*) Protein expression detected by Western blot analysis using anti-Myc antibody in the culture media from 293T cells transfected transiently with an empty vector control (lane 1), human Rumi (lane 2), KDELC1 (lane 3), or KDELC2 (lane 4), all with a C-terminal Myc-His$_6$ tag. (*L*) The Poglut activity in the culture media from the same cells as Fig. 1*K*. The values indicate the mean ± SEM.

genetic approach and asked whether any of the human homologues can rescue *rumi* mutant phenotypes in *Drosophila*. We generated transgenic *Drosophila* strains, which allow the expression of a FLAG-tagged version of hRumi, hKDELC1, or hKDELC2 in *Drosophila* tissues. When raised at 30 °C, mitotic clones of *rumi* null alleles result in the loss of sensory bristles on *Drosophila* thorax and margin loss and vein expansion in *Drosophila* wing, consistent with loss of Notch signaling (12) (Fig. 1 *C* and *G*). Overexpression of the human proteins in mitotic clones homozygous for *rumi* alleles shows that hRumi–FLAG rescues the sensory organ and wing phenotypes of *rumi* clones (Fig. 1 *D* and *H*). However, neither hKDELC1–FLAG nor hKDELC2–FLAG is able to rescue the *rumi* loss-of-function phenotypes (Fig. 1 *E*, *F*, *I*, and *J*). To ensure that the lack of rescue by hKDELC1–FLAG and hKDELC2–FLAG is not because of protein instability or lack of expression, we performed anti-FLAG staining on third instar *Drosophila* wing imaginal discs (Fig. S2 *A*–*C*). We find that all three human proteins are expressed at relatively high levels in the mitotic clones (Fig. S2 *A′*–*C′*). The Notch protein accumulates in mutant clones of the fly *rumi* (12). Antibody staining indicates that this phenotype is also rescued by hRumi–FLAG but not by hKDELC1–FLAG or hKDELC2–FLAG (Fig. S2 *A″*–*C″*). Although all three proteins were expressed and secreted into the culture media after transient transfection in 293T cells (Fig. 1*K*), only hRumi showed significant Poglut activity (Fig. 1*L*). We conclude that hRumi is the only homologue which shows biochemical and functional conservation with dRumi.

Mammalian Rumi, like dRumi, has a C-terminal KDEL-like sequence, suggesting that it resides and functions in the ER. To test this notion, a truncated version of mouse Rumi lacking its KDEL-like sequence (Lys-Thr-Glu-Leu, KTEL) was expressed in 293T cells. Compared with wild-type, deletion of KTEL caused a decrease in the Poglut activity in cell lysates and a significant increase in the activity in the culture media (Fig. S3*A*). Furthermore, most of the endogenous Poglut activity of mammalian cell extracts was detected in fractions from a sucrose gradient corresponding to the ER (Fig. S3*B*) and was detected in soluble fractions, but not in membrane fractions (Fig. S3*C*). These results indicate that mammalian Rumi is a soluble ER-localized protein, consistent with what is known about dRumi (12).

During our initial characterization of its biochemical activity, we examined donor substrate specificity of Rumi using a variety of nucleotide sugars in in vitro assays with human factor VII (hFVII) EGF repeat as acceptor substrate. Surprisingly, we found that both UDP-Xyl and UDP-glucose were utilized by mouse Rumi (Fig. 2*A*). Because Xyl has the same orientation of hydroxyl groups but lacks the extracyclic methoxy group of glucose, UDP-Xyl could conceivably fit into the same active site as UDP-glucose (Fig. 2*B*). Further kinetic analysis showed that both activities of Rumi were dependent on the amount of Rumi (Fig. 2*C*), and the concentrations of hFVII EGF repeat (Fig. 2*D*), UDP-glucose, or UDP-Xyl (Fig. 2*E*). Both Poglut (14) and protein *O*-xylosyltransferase activity (Poxylt) activity (Fig. S3*D*) decreased approximately 50% in the extracts of the neonatal livers of Rumi$^{+/-}$ mice compared with those from wild-type mice. Furthermore, a glycine to glutamate mutation at G169, which is thought to be involved in binding UDP-glucose (27), made Rumi incapable of transferring glucose as well as Xyl (Fig. 2*F*). A similar result was observed with dRumi (Fig. S3*E*). These results suggest that Rumi transfers both glucose and Xyl to hFVII EGF repeat and that the dual activity of Rumi is evolutionarily conserved.

We examined the structures of the Poglut and Poxylt reaction products extensively. The products eluted as a single peak slightly earlier than the unglycosylated substrate on reverse phase HPLC, suggesting that the sugars were covalently attached to the substrate and made the products more hydrophilic (Fig. S4 *A*–*C*). Analysis by electrospray ionization-MS/MS (Fig. S4 *D*–*F*) confirmed that the product of the Poglut reaction is hFVII EGF modified with a hexose, whereas the product of the Poxylt reaction is the EGF repeat modified with a pentose. To confirm the sugar structure on the products, we performed in vitro reactions using UDP-[^{3}H]glucose or UDP-[^{14}C]Xyl and purified the radioactive [^{3}H]glucosylated- or [^{14}C]xylosylated-hFVII EGF repeats by reverse phase HPLC (Fig. S4*G*). The *O*-glycans were released from the purified products by alkali-induced β-elimination and

Downloaded by guest on January 23, 2022

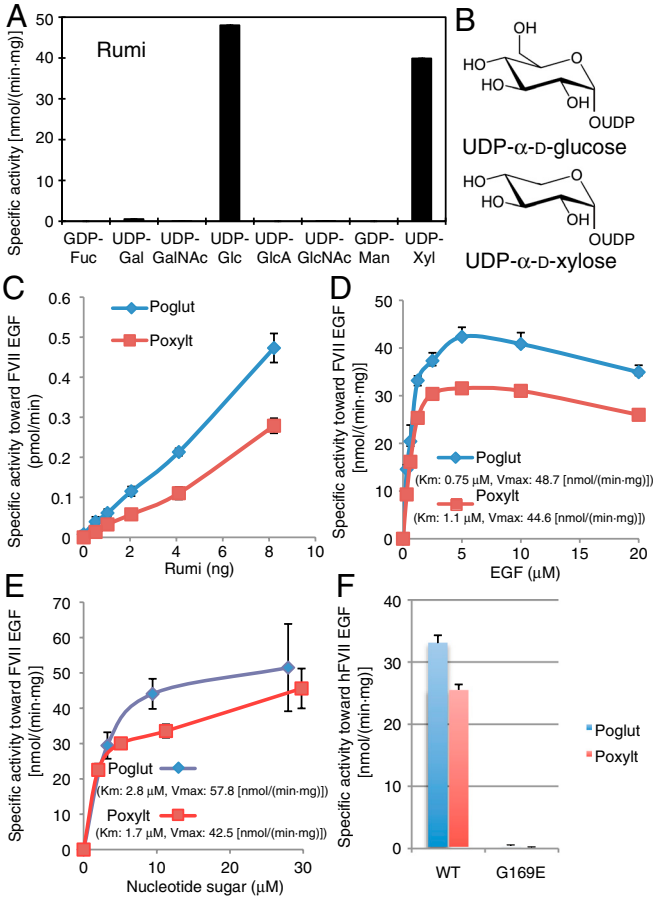

Fig. 2. Rumi has Poxylt activity as well as Poglut activity. (*A*) Donor substrate specificity of *N*-terminal His-tagged mouse Rumi (His-mRumi). A representative dataset from two independent assays is shown. (*B*) Structure of UDP-glucose and UDP-Xyl. (*C*) Enzyme dependence of Poglut (blue) or Poxylt (red) activities of His-mRumi toward hFVII EGF repeats (5 μM). (*D*) EGF dependence of Poglut (blue) or Poxylt (red) activities of His-mRumi in the presence of 10 μM UDP-sugars. Note, kinetic parameters were calculated using the data points below 5 μM of EGF repeats by Lineweaver–Burk plot because the decreased activities at higher range may be due to substrate inhibition. (*E*) UDP-sugar dependence of Poglut (blue) or Poxylt (red) activities of His-mRumi toward hFVII EGF repeats (5 μM). (*F*) Poglut (blue) or Poxylt (red) activities of FLAG-tagged versions of wild-type or G169E mutant hRumi toward hFVII EGF repeats (5 μM). The values indicate mean ± SEM.

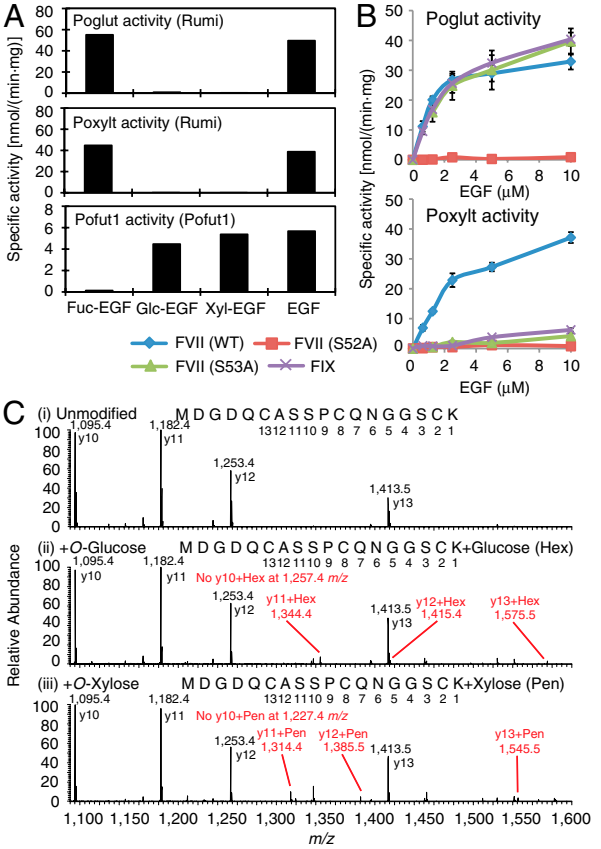

Fig. 3. The diserine motif in the *O*-glucose consensus sequence is an indicator for *O*-Xyl transfer by Rumi. (*A*) Poglut (*Top*) or Poxylt activity (*Middle*) of His-mRumi, or Pofut1 activity of human Pofut1 toward hFVII EGF repeats modified with *O*-fucose (Fuc-EGF), *O*-glucose (Glc-EGF), *O*-xylose (Xyl-EGF), or unmodified hFVII EGF repeats (5 μM). Data are representative of two independent assays. (*B*) EGF dependence of Poglut (*Upper*) or Poxylt (*Lower*) activity of hRumi–FLAG toward hFVII EGF repeat (WT, blue), S52A (red), or S53A (green) mutants, and hFIX EGF repeat (purple). The values indicate mean ± SEM. (*C*) HCD analysis after trypsin digestion of the hFVII EGF repeats unmodified or modified with *O*-glucose or *O*-Xyl by hRumi. Collision-induced dissociation fragmentation confirmed the identity of the peptide and the modifications (Fig. S6). (*i*) HCD fragmentation of the doubly charged form of the peptide MDGDQCASSPCQDGGSCK (*m/z* 980.3) from unmodified hFVII EGF repeats. Numerous sequence fragment y ions are observed that confirm the identity of the peptide. (*ii*) HCD fragmentation of the doubly charged form of the *O*-glucosylated peptide (*m/z* 1,061.3) from *O*-glucosylated hFVII EGF repeats. Numerous sequence fragment y ions are observed that not only confirm the identity of the peptide, but also the attachment of a hexose (glucose) to the first serine within the diserine motif. Although y10 + Hex (*m/z* 1,257.4) was not detected, y11 + Hex (*m/z* 1,344.4), y12 + Hex (*m/z* 1,415.4), and y13 + Hex (*m/z* 1,575.5) ions were detected, confirming the attachment of hexose to the first serine. (*iii*) HCD fragmentation of the doubly charged form of the *O*-xylosylated peptide (*m/z* 1,046.3) from *O*-xylosylated hFVII EGF repeats. Numerous sequence fragment y ions are observed that not only confirm the identity of the peptide, but also the attachment of a pentose (xylose) to the first serine within the diserine motif. Although y10 + Pen (*m/z* 1,227.4) was not detected, y11 + Pen (*m/z* 1,314.4), y12 + Pen (*m/z* 1,385.5), and y13 + Pen (*m/z* 1,545.5) ions were detected.

subjected to gel filtration chromatography, where they migrated at a size consistent with a monosaccharide (Fig. S4*H*). High-pH anion exchange chromatography coupled with pulsed amperometric detection analysis (Fig. S4*I*) demonstrated that the Poglut product is glucitol and the Poxylt product is xylitol, the theoretical products from *O*-glucosylated or *O*-xylosylated EGF repeats after alkali-induced β-elimination, respectively. NMR analysis of glucosylated or xylosylated human factor IX (hFIX) EGF repeats synthesized with hRumi in vitro confirmed the product identification and clearly showed that the anomericity of both *O*-glucose and *O*-Xyl is beta (Fig. S5). These results indicate that Rumi transfers a single *O*-glucose or *O*-Xyl monosaccharide to EGF repeats in a β-linkage.

To examine whether *O*-Xyl is added to the *O*-glucose consensus site of the EGF repeats, we performed the glycosyltransferase assays using the hFVII EGF repeats, which had been modified with prior addition of *O*-fucose by protein *O*-fucosyltransferase 1 (Pofut1) (28), *O*-glucose, or *O*-Xyl by Rumi (Fig. 3*A*). Pofut1 added *O*-fucose to *O*-glucosylated, *O*-xylosylated, or unmodified EGF repeats, but not to *O*-fucosylated EGF, whereas Rumi could add Xyl or glucose to either unmodified or *O*-fucosylated EGF repeat. These results suggest that the *O*-fucose site is distinct from both the *O*-glucose and *O*-Xyl sites on the EGF repeat. In contrast, neither *O*-glucosylated nor *O*-xylosylated EGF repeats served as acceptor for addition of *O*-Xyl or *O*-glucose by Rumi, suggesting that these modifications occur at the same site. Although the hFVII EGF repeat is known to be *O*-glucosylated at serine 52 (20), it contains a second serine residue (serine 53) in the *O*-glucose consensus sequence (^{50}CASSPC^{55}). To examine which serine is the attachment site for *O*-Xyl, we used EGF repeats in

Downloaded by guest on January 23, 2022

which Ser52 or Ser53 were mutated to Ala (S52A or S53A). Our prior studies using crude cell extracts as enzyme source showed that *O*-glucose is added to the S53A mutant, but not to the S52A mutant, consistent with the predicted modification site (16). As expected, purified Rumi was not able to glucosylate the S52A mutant, but the S53A mutant was still an excellent acceptor substrate (even slightly better than wild type), confirming that Rumi transfers *O*-glucose only to the predicted site, Ser52 (Fig. 3*B*). Similarly, the Poxylt activity of Rumi was completely lost with the S52A mutant, suggesting Ser52 is the site of modification. Interestingly, we detected a dramatic decrease in Poxylt activity of Rumi toward the S53A mutant compared with the wild type (Fig. 3*B*). This result suggests that either both Ser52 and Ser53 are modified with *O*-Xyl, or Ser53 is important for *O*-xylosylation at Ser52. To differentiate between these possibilities, we examined the specific amino acid modified using mass spectrometry. Collision-induced dissociation fragmentation of tryptic peptides from glucosylated or xylosylated EGF repeats confirmed the attachment of hexose or pentose, respectively, to the peptide containing the consensus sequence (MDGDQCAS[52]S[53]PCQNGGSCK) (Fig. S6). Higher-energy C-trap dissociation (HCD) analysis showed that both *O*-glucose and *O*-Xyl are attached to Ser52 (Fig. 3*C*). These results suggest that, unlike the Poglut activity that modifies the consensus motif regardless of the identity of the X amino acids surrounding the S, the Poxylt activity of Rumi is sensitive to the flanking amino acid sequences, and that the Ser residue next to the *O*-glucose site facilitates *O*-xylosylation.

To examine the importance of the diserine motif in more detail, we used the hFIX EGF repeat as acceptor substrate. Factor IX is known to be modified by *O*-glucose (20), but lacks the second serine within the *O*-glucose consensus sequence ([51]CESNPC[56]). Interestingly, although hFIX EGF repeats can be both *O*-glucosylated and *O*-xylosylated by Rumi, it is a poor acceptor substrate for *O*-Xyl transfer, similar to the hFVII S53A mutant (Fig. 3*B*). These results confirm that the presence of the diserine motif significantly promotes the transfer of Xyl to EGF repeats by Rumi.

To examine whether *O*-Xyl exists on EGF repeats produced in cells, we examined the glycosylation of EGF16 from mouse Notch2 (mNotch2), which also contains a diserine motif within an *O*-glucose consensus sequence ([587]CYSSPC[592]). Initially we confirmed that Rumi could add *O*-glucose or *O*-Xyl to the first serine within the motif of bacterially expressed mNotch2 EGF16 in vitro (Fig. S7). To determine whether *O*-Xyl occurs on EGF16 produced in cells, we performed mass spectral analysis of the peptide containing the consensus sequence resulting from Asp-N digestion of EGF13–18 from mNotch2 expressed in 293T cells: [582]DQIDECYSSPCLN[594]. Interestingly, we observed both an *O*-pentose trisaccharide and a canonical *O*-glucose-pentose-pentose trisaccharide on this peptide (Fig. 4*A* and Fig. S8). Extracted ion chromatograms of the ions corresponding to the unmodified, *O*-glucose trisaccharide, and *O*-pentose trisaccharide forms of this peptide are shown in Fig. 4*A*. Assuming that both glycoforms of the peptide ionize to the same degree, the *O*-pentose trisaccharide form appears to be present at approximately 5% of the level of the *O*-glucose trisaccharide form. This ratio is consistent with what is known about relative levels of UDP-glucose and UDP-Xyl within cells (17), although it could change in a cell-dependent manner because nucleotide sugar levels are known to vary from cell to cell (29). *O*-glucose-pentose-pentose trisaccharide has been detected on several other peptides corresponding to *O*-glucose motifs on other EGF repeats derived from EGF13–18 of mNotch2 (14). However, we have not found evidence for *O*-pentose trisaccharide at any of these other sites, none of which contain the diserine motif. When the second serine residue in the EGF16 site was mutated to Ala ([587]CYSAPC[592]), only the peptide modified with *O*-glucose trisaccharide was detected (the *O*-pentose form and unmodified form could not be detected, Fig. 4*B* and Fig. S8). These results demonstrate that EGF repeats within Notch can be modified by *O*-Xyl or *O*-glucose in cells (with endogenous levels of UDP-glucose and UDP-Xyl), and that the diserine motif appears to be an indicator for where Xyl modifications may occur.

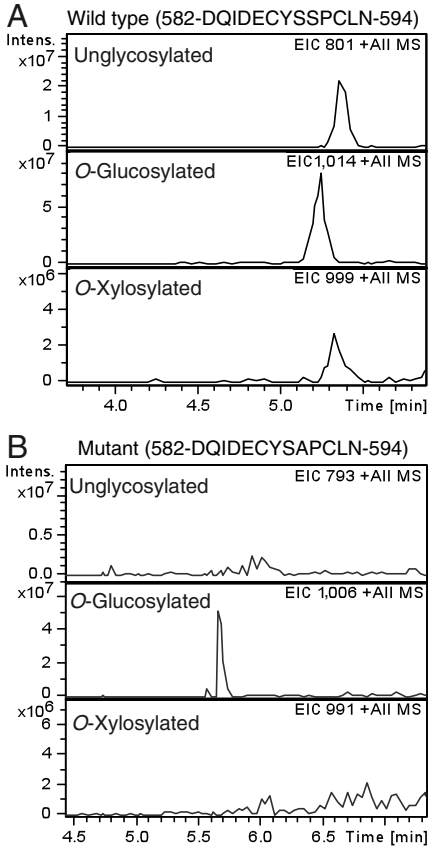

Fig. 4. O-Xyl trisaccharide on EGF16 of mNotch2 is dependent on the diserine motif. Extracted ion chromatograms (EICs) of the ions corresponding to (glyco)peptides containing the *O*-glucose consensus sequence from EGF16 of mNotch2 derived from Asp-N digests of wild-type (*A*) or the mutated (*B*) mNotch2 proteins (see Fig. S8 for mass spectra confirming the identity of these ions). (*A*) EICs of the ions corresponding to the unglycosylated (*Top*), *O*-glucosylated (*Middle*), and *O*-xylosylated (*Bottom*) forms of [582]DQIDECYSSPCLN[594] from wild-type mNotch2 EGF16. The MS data were searched for the doubly charged form of these (glyco)peptides: *m/z* 801.0 for the unglycosylated, *m/z* 1,014.0 for the *O*-glucosylated, and *m/z* 999.0 for the *O*-xylosylated forms. (*B*) EICs of the ions corresponding to the unglycosylated (*Top*), *O*-glucosylated (*Middle*), and *O*-xylosylated (*Bottom*) forms of [582]DQIDECYSAPCLN[594] from the mutated mNotch2 EGF16. The MS data were searched for the doubly charged form of these (glyco)peptides: *m/z* 793.0 for the unglycosylated (*Top*), *m/z* 1,006.0 for the *O*-glucosylated (*Middle*), and *m/z* 991.0 for the *O*-xylosylated (*Bottom*) forms.

Discussion

Mammalian CAP10-Like Genes. Recently, we identified dRumi as an essential component for Notch signaling in flies and showed that dRumi is a Poglut (12). In the present study, we analyzed three candidate mammalian genes as dRumi homologues and showed that mammalian Rumi has Poxylt activity as well as Poglut activity, whereas the other two homologues, KDELC1 and KDELC2, have neither of those activities. Furthermore, mammalian Rumi rescues defects in Notch signaling in *rumi* clones, indicating that mammalian Rumi can regulate Notch signaling in vivo. Loss of one copy of the Rumi gene reduced approximately 50% of both Poglut and Poxylt activities in neonatal mouse livers (14) (Fig. S3*D*). Together, these results strongly suggest that Rumi is the only Poglut/Poxylt for EGF repeats in mammals.

Downloaded by guest on January 23, 2022

The molecular function(s) of KDELC1 and KDELC2 is still unknown. We have shown that the CAP10 domain of dRumi has two DXD-like motifs and that the latter one is functionally important in vitro and in vivo (12, 30). Mammalian Rumi also has two DXD-like motifs in the CAP10 domain, but KDELC1 and KDELC2 do not have the latter one (Fig. S1). Moreover, both dRumi and hRumi have a WXGG motif which is essential for their enzymatic activity (12, 14) (Fig. 2 and Fig. S3) and is implicated in interaction with UDP-sugar donors (27). However, KDELC1 and KDELC2 lack this motif (Fig. S1). Unlike Rumi, KDELC1 and KDELC2 have a filamin-like domain which is known to be involved in interaction with other molecules such as Actin (26) (Fig. 1). To examine the potential roles of modulation of *O*-glucosylation and/or *O*-xylosylation through interaction with Rumi, we performed the glycosyltransferase assays with Rumi in the presence of the purified KDELC1 or KDELC2. Neither KDELC1 nor KDELC2 affected the dual activities of Rumi. Thus, further work is needed to determine the biochemical function of KDELC1 and KDELC2.

How Does Rumi Recognize Donor Substrates? Surprisingly, Rumi utilized UDP-glucose and UDP-Xyl with almost the same affinity (Fig. 2*E*), but not UDP-glucuronic acid. The only structural difference between UDP-glucose, UDP-Xyl, and UDP-glucuronic acid is at the C6 position of the sugar (Fig. 2*B*). The larger, charged C6 group of UDP-glucuronic acid probably precludes significant binding to the active site, but the lack of a C6 group in UDP-Xyl must allow binding. *O*-glucose and *O*-Xyl are β-linked to the serine in EGF repeats (18, 21) (Fig. S5), but are in an α-linkage in UDP-glucose and UDP-Xyl, suggesting that Rumi is an inverting enzyme. Because a point mutation at G169 abrogated both Poglut and Poxylt activities of Rumi, Rumi is likely to use the same active site to transfer *O*-glucose and *O*-Xyl to EGF repeats.

How Does Rumi Recognize Acceptor Substrates? Rumi only modifies properly folded EGF repeats containing the *O*-glucose consensus sequences (C^1-X-S-X-P-C^2) and does not transfer glucose or Xyl to EGF repeats in which this serine (S) is mutated to alanine (Fig. 3). Surprisingly, efficient transfer of Xyl (but not glucose) required the presence of a second serine, C^1-X-S-S-P-C^2. This result suggests that the specificity of the enzyme for the donor substrate is influenced by the structure of the acceptor. It is known that single amino acid mutations in the active sites of some glycosyltransferases can alter their donor substrate specificity such as β4Gal-T and GlcAT-I (31), but altering donor substrate specificity by changes in acceptor substrate is highly unusual. Although structural work will need to be done to fully explain this observation, it suggests that the nucleotide sugar binds at the interface between the enzyme and the EGF repeat, and that sequences within the EGF repeat modulate the affinity of the tertiary complex. We have seen differential efficiency of *O*-glucosylation among the Notch EGF repeats containing the *O*-glucose consensus sequences (14), which is most likely due to the differences in the amino acid sequence of the EGF repeats. Interestingly, among the EGF repeats examined in the EGF13–18 of mNotch2, EGF16 is the only one which has this diserine motif (^{587}CYSSPC^{592}) and the only one which is modified with *O*-Xyl glycans. Database searches reveal the presence of a number of other EGF repeats with a diserine motif in other potential Rumi targets such as mNotch1, mNotch3, Delta-like1, Crumbs1, and Regeneration-associated muscle protease. We are currently examining the glycosylation status of those proteins.

What Does the Dual Specificity of Rumi Mean Biologically? The G189E missense mutation in dRumi and its equivalent G169E mutation in the mammalian Rumi abolish the enzymatic activity of the Rumi proteins without affecting their expression level or stability. The G189E mutation results in temperature-sensitive loss of Notch signaling in flies, similar to null alleles of *rumi* (12). Moreover, our recent study shows that the G169E mutant fails to rescue the Notch-dependent neurite outgrowth phenotypes caused by Rumi knockdown in the Neuro2A mouse neuroblastoma cells (14). These results indicate that Rumi primarily regulates Notch signaling through its enzymatic activity. Based on the ability of Rumi to function both as a Poglut and as a Poxylt, and also given the presence of both of these carbohydrates on Notch EGF repeats, we propose that regulation of Notch signaling by Rumi depends on the addition of *O*-glucose or *O*-Xyl to Notch pathway components by Rumi.

Our results indicate that in some EGF repeats, *O*-Xyl glycans are attached to the serine residue in the C^1-X-S-X-P-C^2 consensus sites, which was previously thought to be exclusively modified by *O*-glucose (Fig. S9). We found *O*-pentose trisaccharides on mNotch2. Because Xyl is the only pentose known to modify proteins in mammals, the *O*-pentose trisaccharide is most likely an *O*-Xyl elongated with two Xyls (e.g., Xyl-α1,3-Xyl-α1,3-Xyl-β*O*-Ser). If the α1,3-xylosyltransferases that extend the *O*-Xyl were down-regulated in a tissue, the β-linked *O*-Xyl could serve as a primer for glycosaminoglycan biosynthesis. Analysis of data from EST databases suggests that the α1,3-xylosyltransferases are indeed expressed in tissue-specific, developmentally regulated patterns, suggesting that, in certain tissues with low levels of these enzymes, the β-linked *O*-Xyl monosaccharides are exposed and therefore glycosaminoglycan addition may be favored (Fig. S9). Interestingly, Kamimura et al. showed that heparan sulfate 3-*O*-sulfotransferase regulates Notch signaling in *Drosophila* (32), suggesting that glycosaminoglycans might be involved in Notch pathway regulation.

Experimental Procedures

Materials. A cell line, called 293T cells, from the American Type Culture Collection were grown in Dulbecco's modified Eagle's medium (Invitrogen) supplemented with 10% fetal bovine serum. GDP-[^{3}H]fucose, UDP-[^{3}H]galactose, UDP-[^{3}H] *N*-acetylgalactosamine, UDP-[^{3}H]glucose, UDP-[^{3}H]glucuronic acid, UDP-[^{3}H] *N*-acetylglucosamine, and UDP-[^{14}C]Xyl were purchased from American Radiolabeled Chemicals. GDP-[^{3}H]mannose was a kind gift from Neta Dean (Stony Brook University). UDP-α-D-Xyl was purchased from Complex Carbohydrate Research Center at the University of Georgia. All other reagents were of the highest quality available.

Constructs for Expression of Mammalian Rumi, KDELC1, and KDELC2. The cDNAs of human Rumi, KDELC1, and KDELC2 were purchased from Open Biosystems and their ORFs were inserted into pcDNA4/TO/myc-His vector so that the recombinant proteins were expressed with a C-terminal myc-His_6 tag. FLAG-tagged version of wild-type and G169E mutant human Rumi was constructed in pTracer vector (14). The cDNA of mouse Rumi was purchased from RIKEN and inserted into pcDNA4/TO/myc-His vector. The His_6 tag sequence was inserted immediately after the signal peptide. DNA sequences of the constructs were confirmed by sequencing. All the primers are available upon request.

Purification of the His_6- or FLAG-Tagged Proteins by Ni-Nitrilotriacetate Chromatography. His_6-tagged Rumi proteins or EGF13–18 of mouse Notch2 from 293T cells transfected with the expression vectors were purified as described in ref. 33. FLAG-tagged version of human Rumi (wild type and G169E mutant) was purified as described in ref. 12. Protein concentration was determined by Coomassie staining using BSA as a standard.

Poglut/Poxylt Assay. Poglut assays were performed as previously described (12, 16). A 10-μL reaction mixture contained 50 mM Hepes (pH 7.0), 10 mM $MnCl_2$, the indicated amounts of acceptor substrates (e.g., hFVII EGF repeats), 0.16 μM (0.01 mCi/ml;

Downloaded by guest on January 23, 2022

1 Ci = 37 GBq) UDP-[^{3}H]glucose, 10 μM UDP-glucose, 0.5% Nonidet P40, and purified proteins. For Poxylt assay, 10 μM UDP-[^{14}C]Xyl was used as donor substrate. The reaction was performed at 37 °C for 20 min and stopped by adding 900 μL of 100 mM EDTA (pH 8.0). The sample was loaded onto a C18 cartridge (100 mg, Agilent). After the cartridge was washed with 5 mL of H_2O, the EGF repeat was eluted with 1 mL of 80% methanol. Incorporation of [^{3}H]glucose or [^{14}C]Xyl into the EGF repeats was determined by scintillation counting of the eluate. Reactions without substrates were used as background control. The mouse and human proteins are very similar biochemically. Most biochemical analyses shown in the manuscript were done using both mouse and human Rumi proteins but due to space limitations, only the mouse Rumi data are shown. Some data using human Rumi (Fig. 2*F* and Fig. 3*B*) are included to demonstrate the biochemical similarity.

Mass Spectrometric Analysis. Notch proteins purified by Ni-nitrilotriacetate affinity chromatography were reduced, alkylated, digested with Asp-N protease (Sigma), and analyzed by nano-liquid chromatography/tandem mass spectrometry using an Agilent 6340 ion-trap mass spectrometer with a nano-HPLC CHIP-cube interface auto sampler as reported previously (34, 35). HCD analysis was performed at Stony Brook University Proteomics Center using LTQ Orbitrap from Thermo Fisher Scientific (36).

ACKNOWLEDGMENTS. We thank members of the Haltiwanger, Jafar-Nejad, and Macnaughtan labs for helpful discussions and Megumi Takeuchi and Shinako Kakuda for technical assistance. We also thank Antonius Koller at Stony Brook University Proteomics Center for technical assistance. This work was supported by National Institute of Health Grants GM61126 (to R.S.H.) and GM084135 (to H.J.-N.), Basil O'Connor Award 5-FY07-654 and Research Grant 1-FY10-362 (to H.J.-N.), a research grant from Mizutani Foundation for Glycoscience (H.T.) and a fellowship of Astellas Foundation for Research on Metabolic Disorders (H.T.).

1. Kopan R, Ilagan MX (2009) The canonical Notch signaling pathway: Unfolding the activation mechanism. *Cell* 137:216–233.
2. Fortini ME (2009) Notch signaling: The core pathway and its posttranslational regulation. *Dev Cell* 16:633–647.
3. Rampal R, Luther KB, Haltiwanger RS (2007) Notch signaling in normal and disease states: Possible therapies related to glycosylation. *Curr Mol Med* 7:427–445.
4. Moloney DJ, et al. (2000) Mammalian Notch1 is modified with two unusual forms of *O*-linked glycosylation found on epidermal growth factor-like modules. *J Biol Chem* 275:9604–9611.
5. Matsuura A, et al. (2008) *O*-linked *N*-acetylglucosamine is present on the extracellular domain of notch receptors. *J Biol Chem* 283:35486–35495.
6. Haines N, Irvine KD (2003) Glycosylation regulates notch signaling. *Nat Rev Mol Cell Biol* 4:786–797.
7. Okajima T, Irvine KD (2002) Regulation of notch signaling by *O*-linked fucose. *Cell* 111:893–904.
8. Moloney DJ, et al. (2000) Fringe is a glycosyltransferase that modifies Notch. *Nature* 406:369–375.
9. Bruckner K, Perez L, Clausen H, Cohen S (2000) Glycosyltransferase activity of Fringe modulates Notch-Delta interactions. *Nature* 406:411–415.
10. Sasamura T, et al. (2003) neurotic, a novel maternal neurogenic gene, encodes an *O*-fucosyltransferase that is essential for Notch-Delta interactions. *Development* 130:4785–4795.
11. Shi S, Stanley P (2003) Protein *O*-fucosyltransferase I is an essential component of Notch signaling pathways. *Proc Natl Acad Sci USA* 100:5234–5239.
12. Acar M, et al. (2008) Rumi is a CAP10 domain glycosyltransferase that modifies Notch and is required for Notch signaling. *Cell* 132:247–258.
13. Takeuchi H, Haltiwanger RS (2010) Role of glycosylation of Notch in development. *Semin Cell Dev Biol* 21:638–645.
14. Fernandez-Valdivia R, et al. (2011) Regulation of mammalian Notch signaling and embryonic development by the protein *O*-glucosyltransferase Rumi. *Development* 138:1925–1934.
15. Harris RJ, Spellman MW (1993) *O*-linked fucose and other post-translational modifications unique to EGF modules. *Glycobiology* 3:219–224.
16. Shao L, Luo Y, Moloney DJ, Haltiwanger R (2002) *O*-glycosylation of EGF repeats: Identification and initial characterization of a UDP-glucose: Protein *O*-glucosyltransferase. *Glycobiology* 12:763–770.
17. Bakker H, et al. (2009) Functional UDP-xylose transport across the endoplasmic reticulum/Golgi membrane in a Chinese hamster ovary cell mutant defective in UDP-xylose Synthase. *J Biol Chem* 284:2576–2583.
18. Whitworth GE, Zandberg WF, Clark T, Vocadlo DJ (2010) Mammalian Notch is modified by D-Xyl-alpha1-3-D-Xyl-alpha1-3-D-Glc-beta1-O-Ser: Implementation of a method to study *O*-glucosylation. *Glycobiology* 20:287–299.
19. Hase S, et al. (1988) A new trisaccharide sugar chain linked to a serine residue in bovine blood coagulation factors VII and IX. *J Biochem* 104:867–868.
20. Nishimura H, et al. (1989) Identification of a disaccharide (Xyl-Glc) and a trisaccharide (Xyl_2-Glc) *O*-glycosidically linked to a serine residue in the first epidermal growth factor-like domain of human factors VII and IX and protein Z and bovine protein Z. *J Biol Chem* 264:20320–20325.
21. Hase S, Nishimura H, Kawabata S, Iwanaga S, Ikenaka T (1990) The structure of $(xylose)_2$glucose-*O*-serine 53 found in the first epidermal growth factor-like domain of bovine blood clotting factor IX. *J Biol Chem* 265:1858–1861.
22. Okabayashi K, Hasegawa A, Watanabe T (2007) Microreview: Capsule-associated genes of Cryptococcus neoformans. *Mycopathologia* 163:1–8.
23. Teng Y, et al. (2006) Cloning, expression and characterization of a novel human CAP10-like gene hCLP46 from CD34(+) stem/progenitor cells. *Gene* 371:7–15.
24. Wu ZL, Ethen CM, Prather B, Machacek M, Jiang W (2011) Universal phosphatase-coupled glycosyltransferase assay. *Glycobiology* 21:727–733.
25. Kimata Y, et al. (2000) Identification of a novel mammalian endoplasmic reticulum-resident KDEL protein using an EST database motif search. *Gene* 261:321–327.
26. Popowicz GM, Schleicher M, Noegel AA, Holak TA (2006) Filamins: Promiscuous organizers of the cytoskeleton. *Trends Biochem Sci* 31:411–419.
27. Busch C, Hofmann F, Gerhard R, Aktories K (2000) Involvement of a conserved tryptophan residue in the UDP-glucose binding of large clostridial cytotoxin glycosyltransferases. *J Biol Chem* 275:13228–13234.
28. Wang Y, et al. (2001) Modification of epidermal growth factor-like repeats with *O*-fucose: Molecular cloning of a novel GDP-fucose: Protein *O*-fucosyltransferase. *J Biol Chem* 276:40338–40345.
29. Nakajima K, et al. (2010) Simultaneous determination of nucleotide sugars with ion-pair reversed-phase HPLC. *Glycobiology* 20:865–871.
30. Lee TV, Takeuchi H, Jafar-Nejad H (2010) Regulation of notch signaling via *O*-glucosylation insights from Drosophila studies. *Methods Enzymol* 480:375–398.
31. Qasba PK, Ramakrishnan B, Boeggeman E (2005) Substrate-induced conformational changes in glycosyltransferases. *Trends Biochem Sci* 30:53–62.
32. Kamimura K, et al. (2004) Regulation of Notch signaling by Drosophila heparan sulfate 3-*O* sulfotransferase. *J Cell Biol* 166:1069–1079.
33. Du J, et al. (2010) *O*-fucosylation of thrombospondin type 1 repeats restricts epithelial to mesenchymal transition (EMT) and maintains epiblast pluripotency during mouse gastrulation. *Dev Biol* 346:25–38.
34. Leonhard-Melief C, Haltiwanger RS (2010) *O*-fucosylation of thrombospondin type 1 repeats. *Methods Enzymol* 480:401–416.
35. Nita-Lazar A, Haltiwanger RS (2006) Methods for analysis of *O*-linked modifications on epidermal growth factor-like and thrombospondin type 1 repeats. *Methods Enzymol* 417:93–111.
36. Olsen JV, et al. (2007) Higher-energy C-trap dissociation for peptide modification analysis. *Nat Methods* 4:709–712.

Downloaded by guest on January 23, 2022