J|A|C|S ARTICLES

Published on Web 09/30/2010

Chemical Synthesis, Folding, and Structural Insights into *O*-Fucosylated Epidermal Growth Factor-like Repeat 12 of Mouse Notch-1 Receptor

Kazumi Hiruma-Shimizu,[†,‡,||] Kensaku Hosoguchi,[†,||] Yan Liu,[§] Naoki Fujitani,[†] Takashi Ohta,[‡] Hiroshi Hinou,[†] Takahiko Matsushita,[†,‡] Hiroki Shimizu,[‡] Ten Feizi,[§] and Shin-Ichiro Nishimura*[,†,‡]

Graduate School of Life Science, Frontier Research Center for Post-Genomic Science and Technology, Hokkaido University, N21, W11, Sapporo 001-0021, Japan, National Institute of Advanced Industrial Science and Technology (AIST), Sapporo 062-8517, Japan, and Glycosciences Laboratory, Imperial College London, Northwick Park Campus, Harrow, Middlesex, HA1 3UJ, U.K.

Received June 15, 2010; E-mail: shin@glyco.sci.hokudai.ac.jp

Abstract: Notch receptors are cell surface glycoproteins that play key roles in a number of developmental cascades in metazoa. The extracellular domains of Notch-1 receptors are composed of 36 tandem epidermal growth factor (EGF)-like repeats, many of which are modified at highly conserved consensus sites by an unusual form of *O*-glycan, with *O*-fucose. The *O*-fucose residues on certain EGF repeats may be elongated. In mammalian cells this can be a tetrasaccharide, Siaα2,3Galβ1,4GlcNAcβ1,3Fucα1→. This elongation process is initiated by the action of *O*-fucose-specific β1,3 *N*-acetylglucosaminyltransferases of the Fringe family. There is evidence that the addition of GlcNAc by Fringe serves as an essential modulator of the interaction of Notch with its ligands and the triggering of activation. Here we describe the efficient synthesis, folding, and structural characterization of EGF repeat 12 (EGF 12) of a mouse Notch-1 receptor bearing different *O*-fucose glycan chains. We demonstrate that the three disulfide bonds, Cys^{456}-Cys^{467} (C1−C3), Cys^{461}-Cys^{476} (C2−C4), and Cys^{478}-Cys^{487} (C5−C6) were correctly formed in the nonglycosylated as well as the *O*-fucosylated forms of EGF 12. Three-dimensional structural studies by NMR reveal that the methyl group of fucose is in close contact with ILe^{475}, Met^{477}, Pro^{478} residues and this stabilizes the conformation of the antiparallel β-sheet of EGF 12. The addition of the GlcNAc residue on *O*-fucosylated EGF 12 induces a significant conformational change in the adjacent tripeptide sequence, $Gln^{462}Asn^{463}Asp^{464}$, which is a motif involved in the natural, enzymatic *O*-fucosylation at the conserved site ($Cys^{461}X_4Ser/ThrCys^{467}$).

Introduction

The Notch signaling pathway is a highly conserved cell signaling system present in metazoans and involved in a wide variety of developmental processes.[1] Glycosylation of the Notch receptor is essential for the regulation of the Notch signaling pathway.[2,3] The extracellular domain of a mouse Notch-1 receptor contains 36 tandem epidermal growth factor (EGF)-like repeats, many of which are glycosylated with unusual forms of *O*-glycans, namely *O*-fucose (Fuc) and *O*-glucose glycans,[4] as well as the more recently described *O*-*N*-acetylglucosamine (GlcNAc)[5] glycans. The modification by *O*-fucose glycans occurs at the putative consensus sequences $C2X_{4-5}Ser/ThrC3$[6] (where Ser/Thr is the modified amino acid, X can be any amino acid, and C2, C3 are the second and third conserved cysteines in the EGF repeat) and is believed to be an important modulator of Notch signaling.[7,8] One of the most evolutionarily conserved *O*-fucose sites is in EGF 12, which is considered to be a portion of the Delta-Serrate-Lag2 (DSL) ligand-binding site.[9,10] Although the role of *O*-fucose glycans on EGF 12 appears to be crucial for the direct interaction with DSL ligands, the structural and functional roles of the glycosylation in EGF 12 have not been clarified in detail.[8,10,11] This is in part due to the disulfide-rich nature of the EGF repeats which precludes large-scale expression of native material for biophysical and biochemical studies. The *O*-fucose moieties on EGF domains of Notch are

[†] Hokkaido University.
[‡] National Institute of Advanced Industrial Science and Technology.
[§] Imperial College London.
[||] These authors contributed equally.

(1) Bray, S. J. *Nat. Rev. Mol. Cell Biol.* **2006**, *7*, 678–689.
(2) Luther, K. B.; Haltiwanger, R. S. *Int. Biochem. Cell Biol.* **2009**, *41*, 1011–1024.
(3) Stanley, P. *Curr. Opin. Struct. Biol.* **2007**, *17*, 530–535.
(4) Moloney, D. J.; Shair, L. H.; Lu, F. M.; Xia, J.; Locke, R.; Matta, K. L.; Haltiwanger, R. S. *J. Biol. Chem.* **2000**, *275*, 9604–9611.
(5) Matsuura, A.; Ito, M.; Sakaidani, Y.; Kondo, T.; Murakami, K.; Furukawa, K.; Nadano, D.; Matsuda, T.; Okajima, T. *J. Biol. Chem.* **2008**, *283*, 35486–35495.
(6) Shao, L.; Moloney, D. J.; Haltiwanger, R. S. *J. Biol. Chem.* **2003**, *278*, 7775–7782.
(7) Okajima, T.; Irvine, K. *Cell* **2002**, *111*, 893–904.
(8) Rampal, R.; Arboleda-Velasquez, J. F.; Nita-Lazer, A.; Kosik, K. S.; Haltiwanger, R. S. *J. Biol. Chem.* **2005**, *280*, 32133–32140.
(9) Rebay, I.; Fleming, R. J.; Fehon, R. G.; Cherbas, L.; Cherbas, P.; Artavanis-Tsakonas, S. *Cell* **1991**, *67*, 687–699.
(10) Xu, A.; Lei, L.; Irvine, K. D. *J. Biol. Chem.* **2005**, *280*, 30158–30165.
(11) Lei, L.; Xu, A.; Panin, V. M.; Irvine, K. D. *Development* **2003**, *130*, 6411–6421.

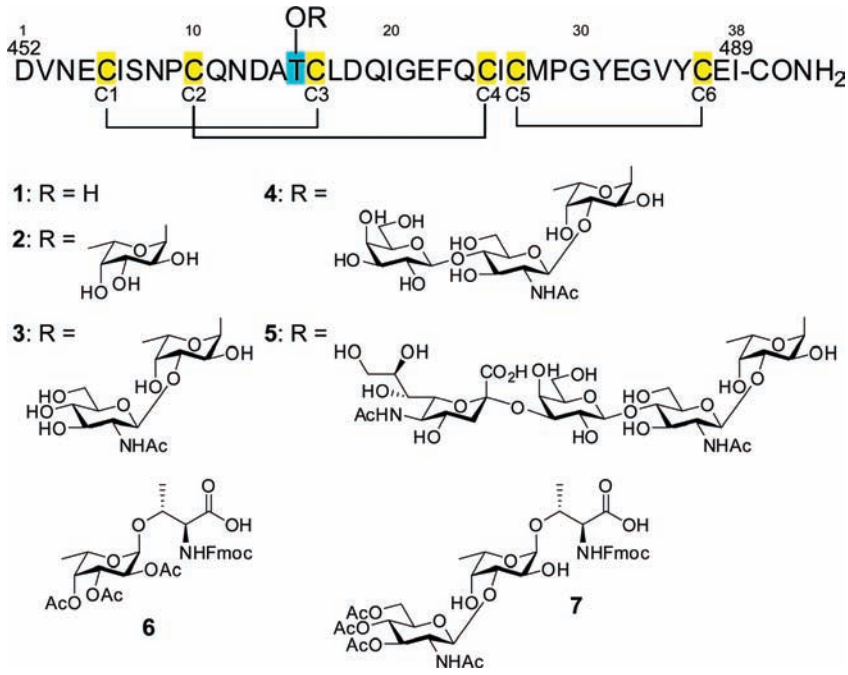


Figure 1. Glycosylated EGF 12 of mouse Notch-1 (amino acids 452−489) and key synthetic materials. *O*-Fucose site Thr^{466} is in cyan, and the six cysteine residues are in yellow.

often further elongated by fucose-specific β-1,3-*N*-acetylglucosaminyltransferases[12,13] of the Fringe family. It is well documented that by adding GlcNAc to *O*-fucose moieties Fringe regulates ligand-mediated Notch signaling.[15] In mammals the action of Fringe is the cue to the further elongation and formation of a mature tetrasaccharide structure, Siaα2,3Galβ1,4GlcNAcβ1,3Fucα1→.[4,12,15]

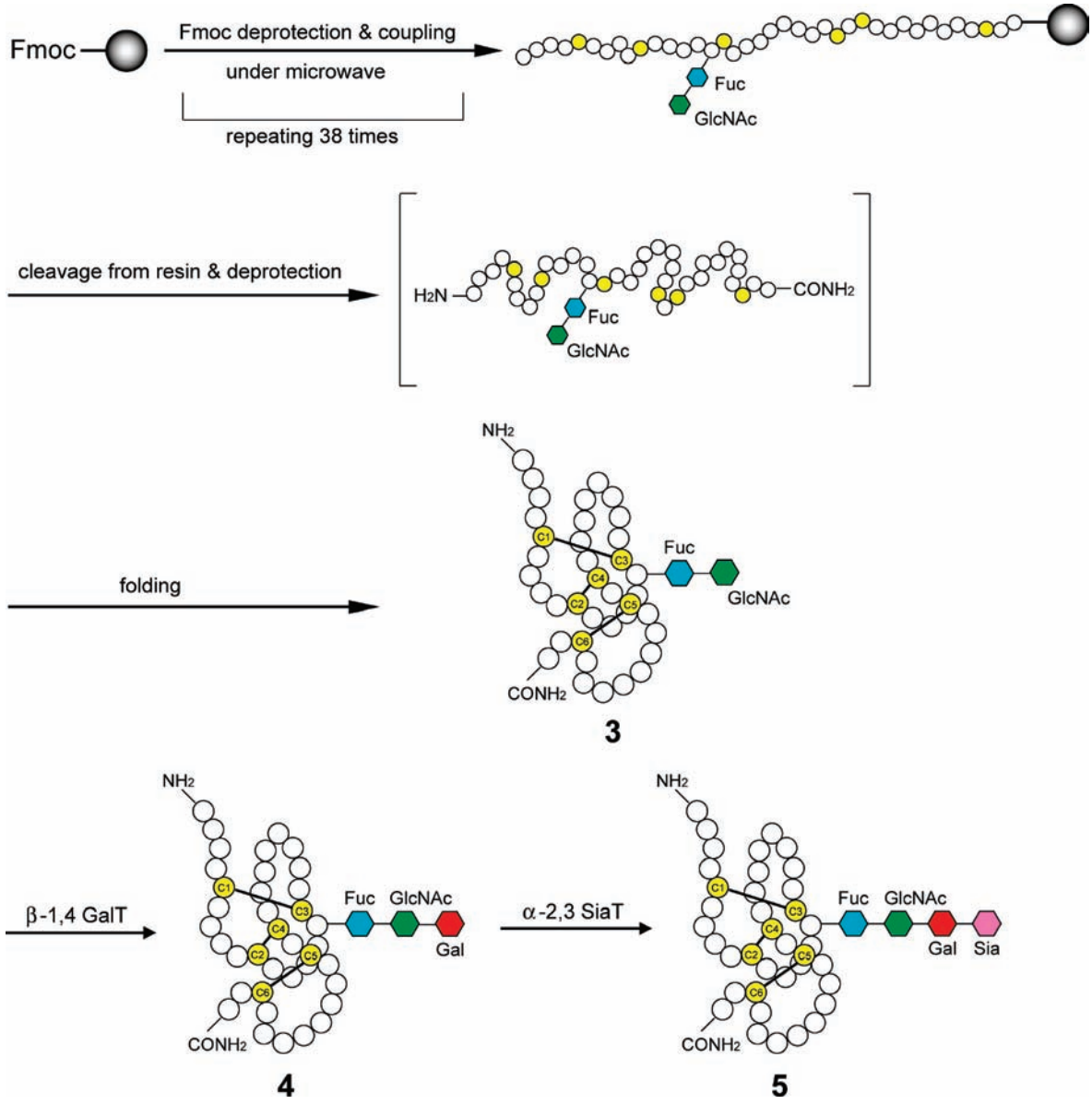
As part of a program to investigate the roles of *O*-fucose glycans in Notch−ligand interaction, we have prepared the fucosylated amino acid building blocks Fucα1-Thr(Fmoc) and GlcNAcβ1,3Fucα1-Thr(Fmoc). We considered that chemical and enzymatic synthesis of EGF 12 peptides (**1**, **2**, **3**, **4**, and **5**) listed in Figure 1 would be of interest in studies of the significance of the glycosylation in Notch signaling as well as the folding mechanism of EGF repeats bearing an *O*-fucosylated motif. The present study has been mostly focused on the chemical synthesis, *in vitro* folding, and structural characterization of EGF 12 having the disaccharide GlcNAcβ1,3Fucα1→ moiety (**3**), this being the glycopeptide structure that is formed in nature, which appears to be critical in the regulation of Notch−ligand interaction. We show that the addition of GlcNAc by Fringe to Fucα1→EGF 12 induces a conformational change in EGF 12, and we hypothesize that this tunes the binding strength with Delta-like and Jagged Notch ligands. We suggest that the key intermediate **3** is a useful general synthetic tool for the construction of variously glycosylated Notch analogs by further enzymatic sugar elongation with Gal and Sia residues.

Results and Discussion

Chemical Syntheses and Folding of EGF 12 (1−3). The synthetic approach for glycosylated EGF 12 (**3**) is shown in Scheme 1. It is based on the microwave-assisted solid-phase protocol established for the synthesis of tumor-associated mucin glycopeptides.[16−21] In previous studies we have demonstrated that Fmoc-Thr/Ser derivatives having *O*-acetylated sugars such as Fmoc-Thr/Ser(Ac$_3$GalNAcα1→)−OH (Tn antigen), Fmoc-Thr/Ser(Ac$_3$Galβ1,3Ac$_2$GalNAc→)−OH (core 1), Fmoc-Thr/Ser(Ac$_3$GlcNAcβ1,6Ac$_2$GalNAc→)−OH (core 6), and Fmoc-Thr/Ser[Ac$_3$Galβ1,3(Ac$_3$GlcNAc)β1,6GalNAc→]−OH (core 2) are versatile building blocks for the construction of a highly complex MUC1-related compound library.[20,21] Indeed, this strategy using a fully *O*-acetylated Fmoc-Thr(Ac$_3$Fuc→)−OH (**6**) permits rapid and efficient synthesis of *O*-fucosylated EGF 12 (**2**) which is composed of 38 amino acid residues (Swiss-Prot **Q01705**). Deprotection of the fucose residue proceeded smoothly under general conditions [1 N NaOH (pH 12.5)/MeOH, DTT, 25 °C, 2 h] and gave the desired linear glycopeptide of compound **2**. For the synthesis of compound **3**, we used partially protected building blocks **7** [Fmoc-Thr(Ac$_3$GlcNAclβ1,3Fuc→)−OH], as *O*-acetyl protection of the axial hydroxyl group at the C-4 position of the fucose residue was found to be difficult to remove under the above conditions when the adjacent equatorial hydroxyl group at the C-3 position is substituted by the bulky GlcNAc residue (see Supporting

(12) Moloney, D. J.; Panin, V. M.; Johnston, S. H.; Chen, J.; Shao, L.; Wilson, R.; Wang, Y.; Stanley, P.; Irvine, K. D.; Haltiwanger, R. S.; Vogt, T. F. *Nature* **2000**, *406*, 369–375.
(13) Jinek, M.; Chen, Y.-W.; Clausen, H.; Cohen, S. M.; Conti, E. *Nat. Struct. Mol. Biol.* **2006**, *13*, 945–946.
(14) Haines, N.; Irvine, K. D. *Nat. Rev. Mol. Cell Biol.* **2003**, *4*, 786–797.
(15) Chen, J.; Moloney, D. J.; Stanley, P. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 13716–13721.
(16) Matsushita, T.; Hinou, H.; Kurogochi, M.; Shimizu, H.; Nishimura, S.-I. *Org. Lett.* **2005**, *7*, 877–880.
(17) Matsushita, T.; Hinou, H.; Fumoto, M.; Kurogochi, M.; Fujitani, N.; Shimizu, H.; Nishimura, S.-I. *J. Org. Chem.* **2006**, *71*, 3051–3063.
(18) Fumoto, M.; Hinou, H.; Ohta, T.; Ito, T.; Yamada, K.; Takimoto, A.; Kondo, H.; Shimizu, H.; Inazu, T.; Nakahara, Y.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2005**, *127*, 11804–11818.
(19) Naruchi, K.; Hamamoto, T.; Kurogochi, M.; Hinou, H.; Shimizu, H.; Matsushita, T.; Fujitani, N.; Kondo, H.; Nishimura, S.-I. *J. Org. Chem.* **2006**, *71*, 9609–9621.
(20) Ohyabu, N.; Hinou, H.; Matsushita, T.; Izumi, R.; Shimizu, H.; Kawamoto, K.; Numata, Y.; Togame, H.; Takemoto, H.; Kondo, H.; Nishimura, S.-I. *J. Am. Chem. Soc.* **2009**, *131*, 17102–17109.
(21) Matsushita, T.; Sadamoto, R.; Ohyabu, N.; Nakata, H.; Fumoto, H.; Fujitani, N.; Takegawa, Y.; Sakamoto, T.; Kurogochi, M.; Hinou, H.; Shimizu, H.; Ito, T.; Naruchi, K.; Togame, H.; Takemoto, H.; Kondo, H.; Nishimura, S.-I. *Biochemistry* **2009**, *48*, 11117–11133.

Scheme 1. Synthetic Route of Sugar-Modified EGF 12 of Mouse Notch-1 (Amino Acids 452−489)



Information, Figure S1 and Scheme S2). As anticipated, use of the building block **7** allowed for the construction of full-length EGF 12 with the disaccharide GlcNAcβ1,3Fuc→moiety (**3**) by microwave-assisted solid-phase synthesis on PEGA resin. After de-*O*-acetylation, the crude product was directly subjected to oxidative folding in glutathione-based redox buffer at pH 8.[22] The folding process was monitored with reversed-phase high performance liquid chromatography (RPHPLC) analysis of aliquots of the sample solution, and distinct profiles were obtained before and after folding (Figure S2). The latter gave a sharp peak at 24 min as the major component (Figure S2b), of which the precursor ion is at *m*/*z* 4560.763 (Figure S2c). This is consistent with the predicted molecular mass of the folded glycopeptide **3** which has three disulfide bonds. Moreover, no other peaks corresponding to the same molecular mass were observed in this HPLC profile, and this indicates that it is unlikely any glycopeptide isomer with a different disulfide bonding pattern had been produced in the folding process. The nonglycosylated ('naked') EGF 12 peptide **1** and its *O*-fucosyl analog **2** were synthesized and subjected to the folding process under similar conditions. The folding products of **1** and **2** were analyzed by RP-HPLC and MS, and the results (data not shown) are similar to those observed with peptide **3**. Further studies were performed with the peptide fragments generated and isolated from HPLC purified fractions of folded naked EGF 12 peptide **1** and glycopeptides **2** and **3** after thermolysin digestion. These peptide fragments were characterized by MALDI-TOFMS and amino acid sequence analyses. The results of peptide **1** showed the three fragments, C1−C3 (*m*/*z* 870.2), C2−C4 (*m*/*z* 1284.9), and C5−C6 (*m*/*z* 1379.2), which demonstrated the correct folding of the synthetic EGF 12 peptide **1** (Table S1). In contrast, although the fragment of C5−C6 (*m*/*z* 1379.2) was observed in glycopeptides **2** and **3**, we did not detect any small fragments containing C1−C3, C2−C4, or those corresponding to peptides with mismatched disulfide bonds. This suggests that glycosylation at the Thr[466] residue may possibly afford steric protection to susceptible sites against the enzymatic hydrolysis.

(22) Zamborelli, T. J.; Dodson, W. S.; Harding, B. J.; Zhang, J.; Bennett, B. D.; Lenz, D. M.; Young, Y.; Haniu, M.; Liu, C.-F.; Jones, T.; Jarosinski, M. A. *J. Pept. Res.* **2000**, *55*, 359–371.

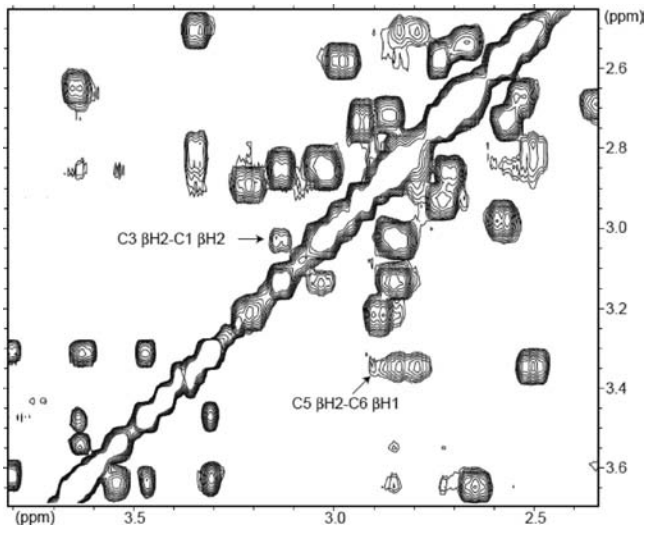


Figure 2. NOESY spectrum (mixing time 400 ms) of EGF 12 (**3**). NOEs between two β-protons of cross-linked cysteine residues were observed in C1−C3 and C5−C6.

Enzymatic Sugar Elongation of EGF 12 (3). In order to obtain a series of glycosylated EGF 12 peptides, we carried out sequential enzymatic elongation of the disaccharide on folded glycopeptide **3** by recombinant β-1,4-galactosyltransferase (β-1,4-Gal-T) and α-2,3-sialyltransferase (α-2,3-Sia-T) (see Scheme 1). These enzymatic reactions efficiently allowed the production of trisaccharide glycopeptide **4** and the tetrasaccharide glycopeptide **5**. Both peptides were purified by RP-HPLC and were corroborated by MS analyses. These compounds will be applied to future structural and functional studies in Notch−ligand interactions.

Structural Analysis of EGF 12 Peptides. In order to gain insights into the structures of the folded products, we performed NMR studies of purified naked EGF 12 (**1**) and glycosylated EGF 12 peptides **2** and **3**. The proton assignments and the structure calculation were made for all three peptides. We describe those for EGF 12 (**3**) as an example. The amide protons of the 38 amino acid backbone were fully assigned, except for those of the *N*-terminus Asp which were designated by comparison with the published data on human Notch-1 EGF 10−13 (BMRB Entry 6031).[23] NOESY spectra of peptide **3** showed two nuclear Overhauser enhancements (NOEs) between β-protons of the cross-linked cysteine residues in C1−C3 and C5−C6 (Figure 2). Although the NOE related to the C2−C4 cross-linkage could not be unambiguously assigned due to the overlapping with other signals, the two NOEs observed are sufficient to define the correct disulfide bonding pattern of glycopeptide **3**. The sequential connectivities between H^{α} and NH ($d_{\alpha N}$) were observed with strong intensity over the full length of this peptide, whereas much weaker NOEs were observed with the sequential connectivities between amide protons (d_{NN}). Furthermore, 21 amino acid residues of glycosylated EGF 12 (**3**) were shown to have a coupling constant between H^{α} and the NH proton ($^3J_{HN\alpha}$) of more than 8.0 Hz. These data suggest that glycosylated EGF 12 (**3**) is rich in β-sheet or extended structures without helical conformation (Figure S3). The hydrogen−deuterium exchange experiment using total correlation spectroscopy (TOCSY) measurement, which was recorded in D_2O solvent, revealed that a total of 11 backbone amide protons had slowly exchanging rates [Figure S4 (**3**)]. This indicates that the formation of hydrogen bonding in these amino acids may provide EGF 12 with a stable conformation. From calculations using a NOE network, coupling constants, and hydrogen bond information, we concluded that synthetic EGF-like glycopeptide **3** contains a double-stranded antiparallel β-sheet, which is typical for the EGF family [Table 1 and Figure 3A (**3**) and 3B (**3**)].

Table 1. Structural Statistics of Average Conformational Energies and RMSD for the 20 Accepted Structures of Glycosylated Mouse Notch-1 EGF 12 (**3**)

Average potential energies (kcal mol^{-1})[a]	
E_{total}	18.446 ± 0.528
E_{bonds}	0.864 ± 0.043
E_{angle}	8.155 ± 0.166
E_{impr}	0.915 ± 0.049
E_{VDW}[b]	3.008 ± 0.552
E_{NOE}[b]	4.276 ± 0.384
E_{cdih}[b]	0.0054 ± 0.004
RMSD from idealized geometry	
bonds (Å)	0.0012 ± 0
angles (deg)	0.346 ± 0.0031
impropers (deg)	0.1647 ± 0.0034
Pairwise rmsd of 20 structures from Cys^{456} to Cys^{467}, Gly^{472}, vPhe^{474} and Cys^{476} to Ile^{489} (Å)	
Backbone atoms (N, Cα, C′)	0.62 ± 0.14
All heavy atoms	1.06 ± 0.15

[a] E_{impr}, E_{VDW}, E_{NOE}, and E_{cdih} are the energy of improper torsion angles, the van der Waals repulsion energy, the square-well NOE potential energy, and the dihedral potential energy, respectively. [b] The force constants for the calculations of E_{VDW}, E_{NOE}, and E_{cdih} were 4.0 kcal mol^{-1} Å^{-4}, 50 kcal mol^{-1} Å^{-1}, and 200 kcal mol^{-1} rad^{-2}, respectively.

All energies and rmsd values were calculated using the CNS1.1 and MOLMOL programs, respectively.

To explore the effect of *O*-glycosylation on the structure of the EGF-like domain, the solution structures of both the naked EGF 12 (**1**) and the *O*-fucosylated EGF 12 (**2**) were calculated in the same manner as that of EGF 12 (**3**) after full assignment of proton signals. From the NOE patterns observed and results of hydrogen−deuterium exchange experiments, we conclude that peptides **1** and **2** also have the characteristic structures of the EGF family [Supporting Information, Figure S4 (**1** and **2**)]. The superposition of 20 final structures shown in Figure 3A and the energy-minimized structures shown in Figure 3B indeed supported the structural similarities of the three EGF 12 domains. For each peptide compound, the 20 final structures (Figure 3A) were well converged except for the flexible region at the *N*-terminus and the hinge region of the two antiparallel β-sheet strands. These tendencies were in good agreement with the solution structure of the EGF domain of blood coagulation factors.[24] The superposition of the backbone C, N, Cα atoms of the most energy-minimized structures of EGF 12 (**1**) and the glycosylated analogs **2** and **3** are shown in Figure 4. These structures were superimposed by minimizing the difference for the backbone atoms of residues 456−468, 473−489 to give an rmsd value of 1.33 ± 0.32 Å in backbone. The results suggest that *O*-fucosylation does not cause a major conformational change in the overall backbone structure of EGF 12 of the mouse

(23) Hambleton, S.; Valeyev, N. V.; Muranyi, A.; Knott, V.; Werner, J. M.; McMichael, A.; Handford, P. A.; Downing, A. K. *Structure* **2004**, *12*, 2173–2183.

(24) Kao, Y.-H.; Lee, G. F.; Wang, Y.; Starovasnik, M. A.; Kelley, R. F.; Spellman, M. W.; Lerner, L. *Biochemistry* **1999**, *38*, 7097–7110.

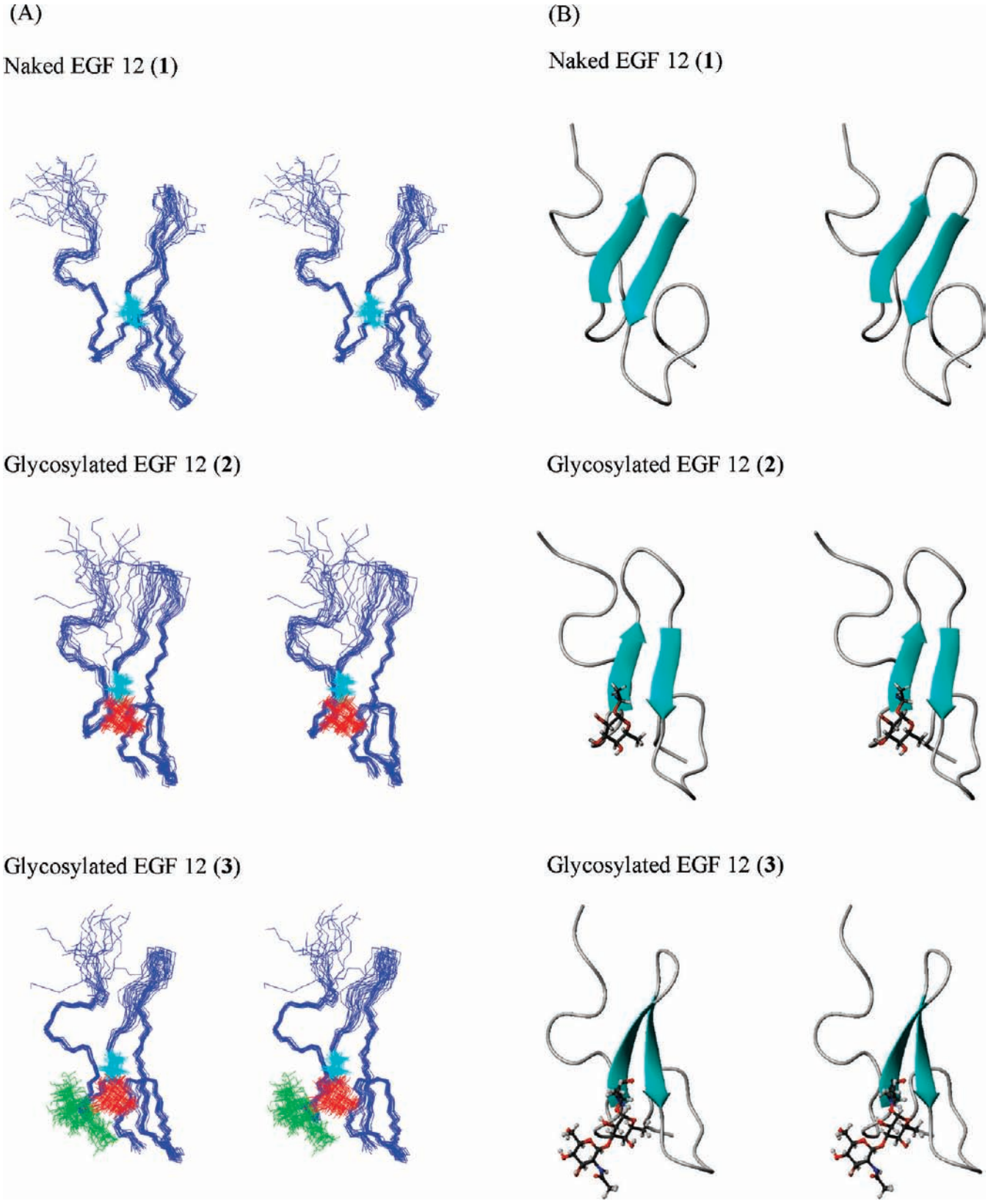


Figure 3. Three dimensional structures of naked and *O*-glycosylated EGF 12 analogs. These figures were generated with the program MOLMOL. (A) A stereoview of the backbone C, N, Cα atoms of the 20 final structures for EGF 12 analogs. These structures were superimposed by minimizing the difference for the backbone atoms of residues 456, 459−467, 476−488 in naked EGF 12 (**1**); 456, 460−467, 476−489 in *O*-fucosylated EGF 12 (**2**); and 456−467, 472, 474, 476−489 in *O*-glycosylated EGF 12 (**3**). The structures are oriented with the N terminus at the top. The heavy atoms of Thr^{466} residues are shown in cyan, and sugar moieties and Fuc and GlcNAc residues (all their atoms), in red and green, respectively. (B) A stereoview of the schematic representation of the most energy-minimized structures. The *β*-sheet strands are predicted by arrows, and the sugar moieties, including sugar-linked threonine, are shown in ball and stick. Oxygen is shown as a red ball, nitrogen as blue, hydrogen as white, and carbon as gray, respectively.

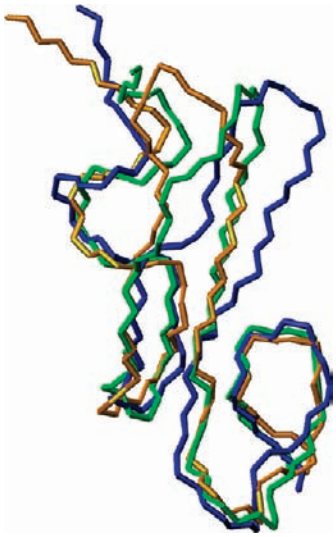


Figure 4. A comparison of the energy-minimized structures of the three EGF 12 peptides (**1**, **2**, and **3**). Each EGF 12 was superimposed by minimizing the difference for the backbone atoms of residues 456−468 and 473−488. The blue line is for the backbone of naked EGF 12 (**1**); the orange line, for that of the *O*-fucosylated EGF 12 (**2**); and the green line, for that of the GlcNAc-Fuc-modified EGF 12 (**3**). The rmsd values for the backbone atoms are 1.33 ± 0.32 Å. This figure was generated with the program PyMOL.

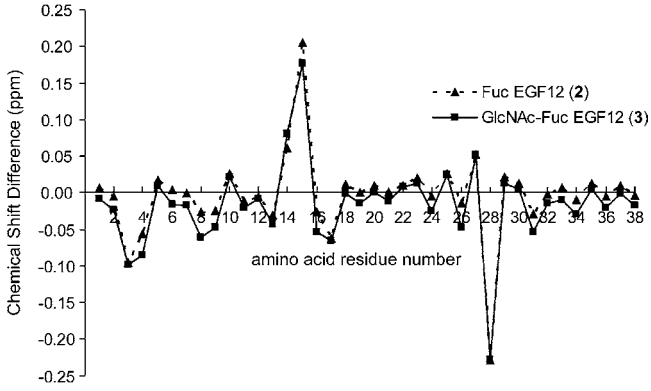


Figure 5. Chemical shift differences (in ppm) of CH^{α} protons of *O*-glycosylated EGF 12 (**2**) and (**3**) from naked EGF 12 (**1**).

Notch-1 receptor. Several differences were observed, however, in more detailed NMR studies on the structures of the region close to the glycosylated residue Thr^{466}: there were clear differences in chemical shift and the strength of sequential NOEs connectivities between H^{α} and NH ($d_{\alpha N}$), and changes were detected in slowly exchanging amide protons (Figures 5 and S4). This indicates that the structures of glycosylated EGF were changed in part. In both *O*-fucosylated EGF 12 (**2**) and (**3**), NOEs were observed between the ring protons of the fucose residues and Thr^{466} as expected. Interestingly, several NOEs were also observed between the methyl protons (C6) of the fucose residue and the methionine residue (Met^{477}) which is opposite to Thr^{466} on the other strand of the antiparallel β-sheet (Figure 6). Some additional NOEs between the fucose methyl group and the neighboring amino acids of Met^{477}, namely ILe^{475} and Pro^{478}, were also observed (Figure 6).

These findings, together with the change in the chemical shift of the α proton of Met^{477} (amino acid position 28 in Figure 5) in the glycosylated EGF peptides, demonstrated the close contact of the fucose residue with these amino acids (Figures 5, 6). In addition, the atomic radius molecular modeling showed that the

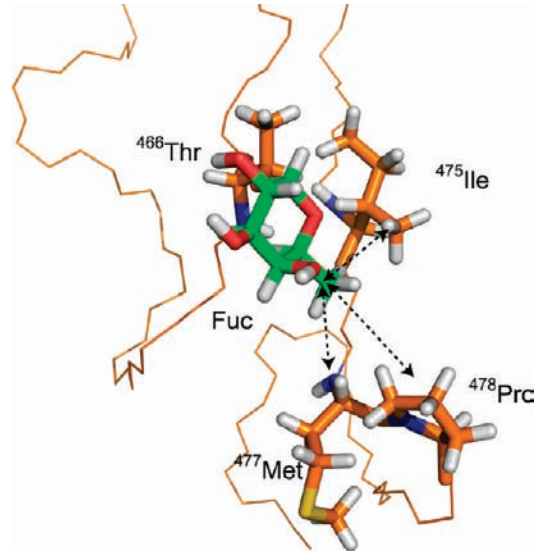


Figure 6. A schematic representation of *O*-fucosylated EGF 12 (**2**) focusing on Fucose, Thr^{466}, and ^{475}Ile-Cys-Met-Pro^{478}. This figure was generated with the program PyMOL.

methyl group of the fucose residue could fit in the intersheet spacing of the antiparallel β-sheet (data not shown). This further supports that the fucose moiety might directly function as a "bridge" in the formation of the antiparallel β-sheet to stabilize the structure through the interaction between fucose and the peptide backbone.

Next we compared all structures of this consensus region to probe the structural effect of *O*-fucosylation. The 20 final ensembles of energy-minimized structures of the *O*-fucose consensus sequence, Cys^{461}-Gln-Asn-Asp-Ala-Thr-Cys^{467}, are shown in Figure 7. Interestingly, while the conformation of this region was well converged in **3**, which has the disaccharide GlcNAcβ1,3Fuc→ moiety, the structures of the same region in nonglycosylated **1** and fucosylated **2** seemed to be somewhat disordered. The addition of the fucose residue resulted in some change in the conformation of glutamine and asparagine. Moreover, the GlcNAc residue that followed also caused a significant structural change in asparagine and aspartic acid as deduced from the different strengths of the sequential NOE connectivities between H^{α} and NH ($d_{\alpha N}$) (Figure S4). This is likely to be caused by the interaction of the carboxyl group of the Asp^{464} residue with the GlcNAc residue. Some of the Asp^{464} carboxyl groups in the 20 refined structures of (**3**) in Figure 7 were shown to be located close to *O*-6 of the GlcNAc residue, suggesting a potential hydrogen bonding between these moieties. Thus, although *O*-fucosylation does not cause an overall conformational change in EGF 12 of the mouse Notch-1 receptor, there are considerable differences around the fucosylation site including the *O*-fucose consensus sequence region.

The *O*-fucose consensus site in EGF-like repeats is considered to be $C2X_{4-5}Ser/ThrC3$,[6] where X can be any amino acid. In a previous structural study of the EGF-like domain of the blood coagulation factor[24] on which the *O*-fucosylated region has the sequence C2QNGGTC3, a conformational change was not observed to be associated with the introduction of the fucose residue. A possible explanation for this is the lack of amino acid side chains in this region which can interact with the fucose residue. C2QNGGTC3 also occurs as the most frequent sequence among the 21 *O*-fucosylated EGF-like repeats on mouse Notch-1. However, the sequence C2QNDATC3 is unique

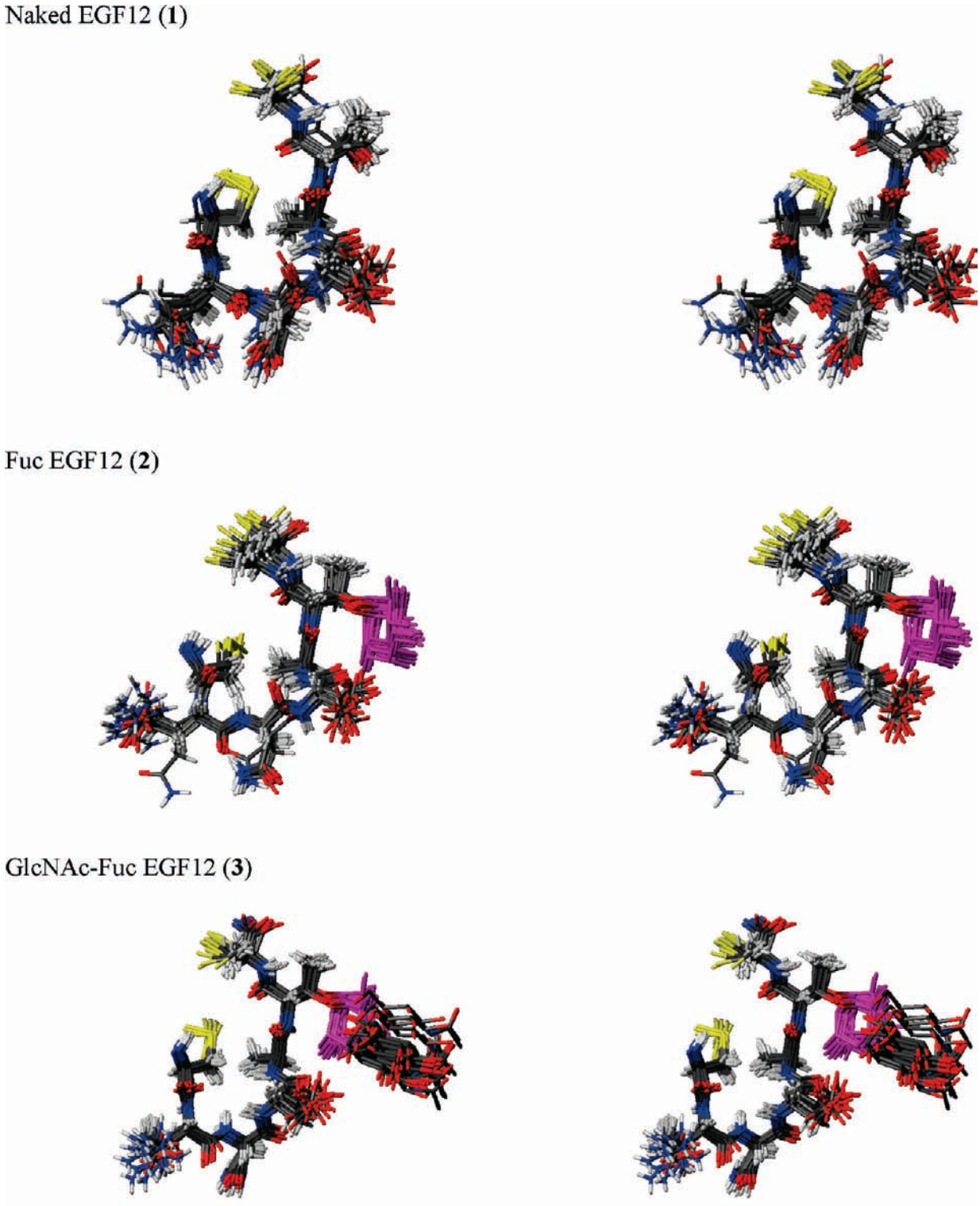


Figure 7. 20 final ensembles of structures of the *O*-fucose consensus region (^{461}Cys-Gln-Asn-Asp-Ala-Thr-Cys^{466}) for EGF 12's. In all images corresponding to each EGF 12 (**1**−**3**), hydrogen atoms are shown in gray; carbon atoms, black; oxygen atoms, red; sulfur atoms, yellow; and nitrogen atoms, blue. In the figures of glycosylated EGF 12 (**2**, **3**), the fucose residue is shown in magenta. These images were generated with the program MOLMOL.

to the EGF 12 motif, and as described above, the presence of Fuc and GlcNAc has a clear influence on the structure of this sequence. Thus, the rare C2QNDATC3 sequence in the *O*-fucosylated site of EGF 12 might have a special role in Notch function.[1,2] This is yet to be elucidated, as it has been predicted from NMR titration experiments that the ligand binding site is on the face of EGF 12 that is opposite to (away from) the *O*-fucose,[25] with the inference being that the sugar chain on threonine may not be directly involved in ligand binding.

Conclusion

In summary, five analogs of the EGF 12 repeat of a mouse Notch-1 receptor (compounds **1**−**5**) have been synthesized and correctly folded. By comparing the structures of the two

synthetic *O*-fucosylated forms of EGF 12 (**2** and **3**) with the nonglycosylated form (**1**), we provide evidence for conformational changes induced by glycosylation. Although *O*-fucosylation does not confer a perceptible change in the overall backbone structure of EGF 12, our results suggest that the fucose residue may serve to stabilize the conformation of the antiparallel β-sheet structure of EGF 12. More importantly, we have shown that the addition of the GlcNAc residue on the fucose induces a significant structural change of the *O*-fucosylated region. This provides structural insights into possible mechanisms of the regulation of Notch−ligand interaction by Fringe. Our findings are analogous to the reported poly-L-proline type II helix-like conformation of the tandem repeating polypeptide, antifreeze glycoproteins [AFGPs, poly(Ala-Ala-Thr)$_n$][26] that result from the addition of multiple *O*-GalNAc moieties of threonine residues. The report is to our knowledge the first to describe the structure determination of an EGF-like module having an extended disaccharide *O*-fucose modification. This opens the way to investigating the contribution of this conformational change to the Notch mediated signal transduction.

Experimental Section

Reagents and General Methods. All commercially available solvents and reagents used for peptide synthesis were A grade and used without purification. Recombinant human β-1,4-GalT (EC 2.4.1.22) was purchased from TOYOBO Co. Ltd. α-2,3-(*N*)-SiaT (EC 2.4.99.6) was purchased from Calbiochem Co. Ltd. Thermolysin was purchased from Sigma. All mixing was performed by using a vortex mixer. MALDI-TOFMS spectra were recorded with a Bruker AutoFLEX mass spectrometer in linear positive mode using matrix (2,5-dihydroxybenzoic acid). Samples for MALDI-TOFMS were desalted and concentrated using 10-μL C_{18} ZipTips (Millipore) according to the manufacturer's instructions. Typically, samples were dissolved in 1 μL of 30% (v/v) aqueous acetonitrile and mixed with the same volume of a saturated solution of 2,5-dihydroxybenzoic acid in water. The above mixtures (1 μL) were applied to a stainless steel target MALDI plate and air-dried before analysis in the mass spectrometer. All 1D and 2D ^{1}H NMR spectra for the identification of the synthetic peptides were collected with a Bruker AV-600 spectrometer at 600.13 MHz for proton frequency equipped with a Cryo-probe at 300 K. Two-dimensional homonuclear double-quantum-filtered scalar-correlated spectroscopy (DQF-COSY), TOCSY with an MLEV-17 sequence, and nuclear Overhauser enhancement spectroscopy (NOESY) spectra were recorded in the indirect dimension using States-TPPI phase cycling.

Solid-Phase Synthesis of EGF 12 (1−3). The solid-phase synthesis of the glycosylated EGF 12 analogs was carried out using PEGA resin (0.03 mmol) functionalized with a Rinkamide linker (0.24 mmol/g) according to our previous report.[16] After removing the Fmoc group on resin by 20% piperidine in DMF, each Fmoc-amino acid (0.09 mmol) except for the cysteine residue was coupled with resin in the presence of HBTU (0.09 mmol), HOBt (0.09 mmol), and DIEA (0.18 mmol) in DMF with microwave irradiation. Fmoc-cysteine was introduced using FmocCys(Trt)COOPfp (0.09 mmol) in the presence of HOBt (0.09 mmol) to prevent undesired epimerization. As for the Thr^{466} residue, the building block **6** or **7** was used instead of Fmoc-Thr(*t*But) in each case. Removal of N^{α}-Fmoc protection and the coupling reaction were repeated until the *N*-terminal amino acid residue. Upon completion of the synthesis, each peptide resin was treated with a mixture of TFA/EDT/H_2O/TIS = 95:2.5:2.5:1 for 1 h at room temperature. The solution was filtered and concentrated by a flow of nitrogen gas, and the crude peptide was precipitated using cold *tert*-butylmethylether. To synthesize naked EGF 12 (**1**) this crude precipitate was involved in a folding process. To deprotect the *O*-acetyl group of the sugar moiety the solid material was dissolved in basic methanol (pH 12.5) containing DTT (6.5 mM), and the solution was stirred. After 2 h the solution was neutralized and concentrated by a flow of nitrogen gas at 40 °C. The crude material underwent preparative RP-HPLC to remove excess DTT.

Folding of Glycopeptides. The linear glycopeptides (1 mg) were dissolved in a redox buffer (20 mL) containing 100 mM ammonium acetate (pH 8.0), with 1 mM reduced and 0.1 mM oxidized glutathione. The mixtures were routinely stirred for 3 days at room temperature, although all oxidation reactions were complete within 48 h as evidenced by analytical RP-HPLC (data not shown). After neutralization by treating with 30% aqueous acetic acid, the solutions were lyophilized and subjected to purification by RP-HPLC (see Supporting Information, Figure S2).

Sugar Elongation by Glycosyltransferases. Glycosylated EGF 12 (**3**) was treated with a mixture of 2.2 mU of β-1,4-GalT and UDP-Gal (1.79 μmol) in a total volume of 890 μL of 50 mM HEPES/NaOH buffer, pH 7.0, and 10 mM $MnCl_2$. After incubation for 2 h at 25 °C, the reaction mixture was directly subjected to preparative RP-HPLC to purify compound **4**, whose transfer ratio of galactose residue was estimated to be 100% based on the HPLC profile. Subsequently, compound **4** (0.95 μmol) was treated with a mixture of 98 μU of α-2,3-SiaT and CMP-Sia (1.89 μmol) in a total volume of 950 μL of 50 mM HEPES/NaOH buffer, pH 7.0, and 10 mM $MnCl_2$. After incubation overnight at 25 °C, the reaction mixture was subjected to preparative RP-HPLC. The resulting HPLC profile of the purified material clearly indicated that sialylation of the trisaccharide moiety of **4** proceeded to completion to give the EGF 12 derivative bearing tetrasaccharide **5**.

Purification of EGF 12 Derivatives. Synthetic naked peptides and glycopeptides were purified using a preparative C-18 reversed-phase column (Inertsil ODS-3 10 mm × 250 mm) on a HITACHI liquid chromatography system (HPLC) with an L7150 pump, at a flow rate of 4 mL/min. The column temperature was 40 °C, and UV monitoring was carried out at 220 nm. Buffer A was distilled water containing 0.1% TFA, and buffer B was acetonitrile containing 0.1% TFA. A linear gradient of 10−50% of B over 60 min was used unless otherwise stated. In the case for the purification of the peptides containing sialic acid, another buffer was used: A was 25 mM ammonium acetate buffer pH 5.8, and B was acetonitrile containing 10% of 25 mM ammonium acetate buffer pH 5.8. A linear gradient of 2−6% of B over 60 min was used. Analytical runs were performed on a HITACHI liquid chromatography system (ELITE Lachrom) with an Inertsil ODS-3 (4.6 mm × 250 mm) reversed-phase column. UV absorbance was measured at 220 nm, and the column temperature was 25 °C. The flow rate was 1 mL/min, and elution conditions were described as follows: the elution buffer was the same as that described above. A linear gradient of 20−45% of B over 45 min was used to elute peptides **1**−**4**. Purification of sialylated compound **5** was performed using 10−55% of B over 45 min.

EGF 12 (**1**): Retention time of analytical RP-HPLC was 28.2 min. MALDI-TOFMS (*m*/*z*) calcd for $C_{175}H_{262}N_{45}O_{62}S_7$ $[M + H]^+$ 4209.678, found 4209.574.

EGF 12 (**2**): Retention time of analytical RP-HPLC was 28.9 min. MALDI-TOFMS (*m*/*z*) calcd for $C_{181}H_{271}N_{45}O_{66}S_7$ $[M + H]^+$ 4358.813, found 4358.88.

EGF 12 (**3**): Retention time of analytical RP-HPLC was 24.5 min. MALDI-TOFMS (*m*/*z*) calcd for $C_{189}H_{284}N_{46}O_{71}S_7$ $[M + H]^+$ 4562.006, found 4560.763.

EGF 12 (**4**): Retention time of analytical RP-HPLC was 23.5 min. MALDI-TOFMS (*m*/*z*) calcd for $C_{195}H_{294}N_{46}O_{76}S_7$ $[M + Na]^+$ 4746.128, found 4747.35.

(25) Cordle, J.; Johnson, A.; Tay, J. Z. Y.; Roversi, P.; Wilkin, M. B.; de Madrid, B. H.; Shimizu, H.; Jensen, S.; Whiteman, P.; Jin, B.; Redfield, C.; Baron, M.; Lea, S. M.; Handford, P. A. *Nat. Struct. Mol. Biol.* **2008**, *15*, 849–857.
(26) Tachibana, Y.; Fletcher, G. L.; Fujitani, N.; Tsuda, S.; Monde, K.; Nishimura, S.-I. *Angew. Chem., Int. Ed.* **2004**, *43*, 856–862.

EGF 12 (**5**): Retention time of analytical RP-HPLC was 25.3 min. MALDI-TOFMS (*m*/*z*) calcd for $C_{206}H_{311}N_{47}O_{84}S_7$ $[M + H]^+$ 5015.401, found 5016.11.

Thermolytic Digestion. Thermolysin was dissolved in 10 mM phosphate buffer, and the concentration was adjusted to 0.03 mM before use. After the purification, each folded EGF (100 μg) was digested with diluted thermolysin (10 μL) in 25 mM ammonium acetate buffer pH 5.8 (100 μL) at 37 °C for 24 h and separated by RP-HPLC. Each fraction was analyzed by MALDI-TOFMS and amino acid analysis. The fragments derived from compound **1** were shown in the text. Although the fragment due to C5−C6 (^{477}Ile-Cys-Met-Pro-Gly-Tyr-Glu-Gly-Val-Tyr-Cys-Glu^{488}, theoretical *m*/*z* as $[M + H]^+$ 1379.575) was identified at *m*/*z* 1379.19 for **2** and *m*/*z* 1379.27 for **3**, respectively, ion peaks due to the fragments containing C1−C3 or C2−C4 could not be detected.

NMR Spectroscopy. Sequential assignments were achieved according to the standard methods for small proteins established by Wüthrich and co-workers.[27] Naked EGF12 (**1**) was dissolved at a final concentration of 0.75 mM, glycosylated EGF12 (**2**) was 0.67 mM, and disaccharide EGF12 (**3**) was 0.55 mM in 300 μL containing 90% H_2O/10% D_2O at pH 5.3, respectively. The pH was adjusted to 5.3 by addition of aq NaOH and aq HCl. TOCSY experiments were applied for a spin-locking time of 100 ms, and NOESY experiments were carried out with mixing times of 100, 200, and 400 ms. The suspension of water signal was performed by presaturation during the relaxation delay (2 s) and by a 3-9-19 WATERGATE pulse sequence with a field gradient.[28,29] TOCSY and NOESY spectra which were acquired with 2048 × 2048 data matrices were zero-field to yield 2048 × 2048 data matrices. DQF-COSY spectra with 8192 × 512 frequency data points were also recorded and zero-field to yield 8192 × 8192 matrices in order to measure the coupling constants. The sweep width of 8192 Hz was applied. Time domain data in both dimensions were multiplied by a sine bell window function with a 90° phase shift prior to Fourier transformation. All NMR data were processed by NMRPipe software[30] and analyzed using the Sparky program.[31]

The chemical shift assignments have been deposited in the BioMagResBank database (http://www.bmrb.wisc.edu) under BMRB accession number 11006 (**3**). The final constraints used in the structure calculations have been deposited in the Protein Data Bank with accession number 2RR0 (**1**), 2RR2 (**2**), 2RQZ (**3**).

Structure Calculation. Three-dimensional structures of synthesized peptides **1**−**3** were calculated using the CNS 1.1 program[32] with standard protocols for distance geometry-simulated annealing and refinement. In order to calculate the family of structures, we used a total of 377 distance restraints and 27 dihedral angle ϕ restraints for EGF (**1**), a total of 374 distance restraints and 30 dihedral angle ϕ restraints for Fuc EGF12 (**2**), and a total of 517 distance restraints and 39 dihedral angle ϕ restraints used for glycosylated EGF12 (**3**), respectively. Distance restraints for calculations were estimated from the cross-peak intensities in NOESY spectra with a mixing time of 400 ms; the estimated restraints were then classified into five categories (very strong, strong, medium, weak, and very weak) and assigned upper limits of 2.6, 2.9, 3.5, 5.0, and 6.0 Å, respectively. In the first stage of structure determination, synthesized peptides **1**−**3** were calculated using only interproton distance information. After the validation of fulfilling distance restraints for the obtained structure, the restraints of the dihedral angle ϕ were adopted for structure calculation. When the coupling constant $^3J_{HN\alpha}$ was more than 8.0 Hz and less than 6.0 Hz, the dihedral angle ϕ was constricted to $-120 \pm 30°$ and $-60 \pm 30°$, respectively.[27] In the final step of structure determination, hydrogen bond restraints were used as distance constraints of 1.5−2.5 Å between amide protons and carbonyl oxygen and 2.5−3.5 Å between amide nitrogen and amide protons. Amide protons protected from exchange with solvent were identified by amide hydrogen/deuterium exchange experiments using a series of 1D ^{1}H NMR and TOCSY spectra recorded after reconstituting the EGF 12 in D_2O. These protected amide protons are presumably either involved in hydrogen bonds or buried in the interior of the protein. When a hydrogen bond acceptor could be clearly identified in the initial structures for a protected amide proton, two distance constraints were then included in the refinement stage of the structure calculation for each hydrogen bond. The conformation of the sugar rings was held fixed to the chair conformation. All analyses of rmsd values and the solution structures of EGF 12 derivatives were performed with the PROCHECK[33] and MOLMOL[34] programs.

Acknowledgment. This work was partly supported by grants for "Innovation Project for Future Drug Discovery and Medical Care" from the Ministry of Education, Culture, Science and Technology and for "Development of Novel Diagnostic and Medical Applications through Elucidation of Sugar Chain Functions" from New Energy and Industrial Technology Development Organization (NEDO), UK Research Councils' Basic Technology Grant (GR/S79268), and UK Engineering and Physical Research Councils Translational Grant EP/G037604/1. We thank Mr. T. Hirose at the Center for Instrumental Analysis, Hokkaido University, for amino acid analysis.

Supporting Information Available: Experimental details for the chemical synthesis of building blocks **6**, **7** and the supplemental data of structural analysis of peptides. This material is available free of charge via the Internet at http://pubs.acs.org.

JA105216U

(27) Wüthrich, K. *NMR of Proteins and Nucleic Acids*; John Wiley and Sons: New York, 1986.
(28) Piotto, M.; Saudek, V.; Sklenár, V. *J. Biol. NMR* **1992**, *2*, 661–665.
(29) Sklenár, V.; Piotto, M.; Leppik, R.; Saudek, V. *J. Magn. Reson., Ser. A* **1993**, *102*, 241–245.
(30) Delaglio, F.; Grzesiek, S.; Vuister, G.; Zhu, G.; Pheifer, J.; Bax, A. *J. Biomol. NMR* **1995**, *6*, 277–293.
(31) Sparky 3.111, a NMR assignment and integration software, was downloaded on February 2, 2006 from http://www.cgl.ucsf.edu/home/sparky/ and used. This site was updated on October 13, 2006: Goddard, T. D. Kneller, D. G. *SPARKY 3*; University of California, San Francisco, 1999.
(32) Brünger, A. T.; Adams, P. D.; Clore, G. M.; DeLano, W. L.; Gros, P.; Grosse-Kunstleve, R. W.; Jiang, J. S.; Kuszewski, J.; Nilges, M.; Pannu, N. S.; Read, R. J.; Rice, L. M.; Simonson, T.; Warren, G. L. *Acta Crystallogr., Sect. D* **1998**, *54*, 905–921.
(33) Laskowski, R. A.; Rullmann, J. A. C.; MacArthur, M. W.; Kaptein, R.; Thornton, J. M. *J. Biomol. NMR* **1996**, *8*, 477–486.
(34) Koradi, R.; Billeter, M.; Wüthrich, K. *J. Mol. Graph* **1996**, *14*, 51–55.