Regulation of signal transduction pathways in development by glycosylation

Robert S Haltiwanger

Recent studies from several laboratories have provided evidence that cell surface complex carbohydrates play key roles in the regulation of developmentally relevant signal transduction events. The demonstration that Fringe, a known modifier of Notch function, is a fucose-specific N-acetylglucosaminyltransferase provided strong evidence that the Notch signaling pathway could be regulated by alterations of O-fucose structures. More recently, the demonstration that O-fucose modification of Cripto is essential for Nodal-dependent signaling provides further evidence of a role for glycosylation in signal transduction. These and other examples provide a new paradigm for the regulation of signal transduction events by glycosylation.

Address
Department of Biochemistry and Cell Biology, Institute for Cell and Developmental Biology, State University of New York at Stony Brook, Stony Brook, New York 11794-5215, USA

Current Opinion in Structural Biology 2002, 12:593–598

0959-440X/02/$ – see front matter

Abbreviations

CFC	Cripto, FRL-1, Cryptic
CSL	CBF1, Suppressor of hairless, Lag-1
EGF	epidermal growth factor
GlcNAc	N-acetylglucosamine
GPI	glycosyl phosphatidyl inositol
O-FucT-1	GDP-fucose: protein O-fucosyltransferase 1
TGFβ	transforming growth factor β
uPA	urinary-type plasminogen activator

Introduction

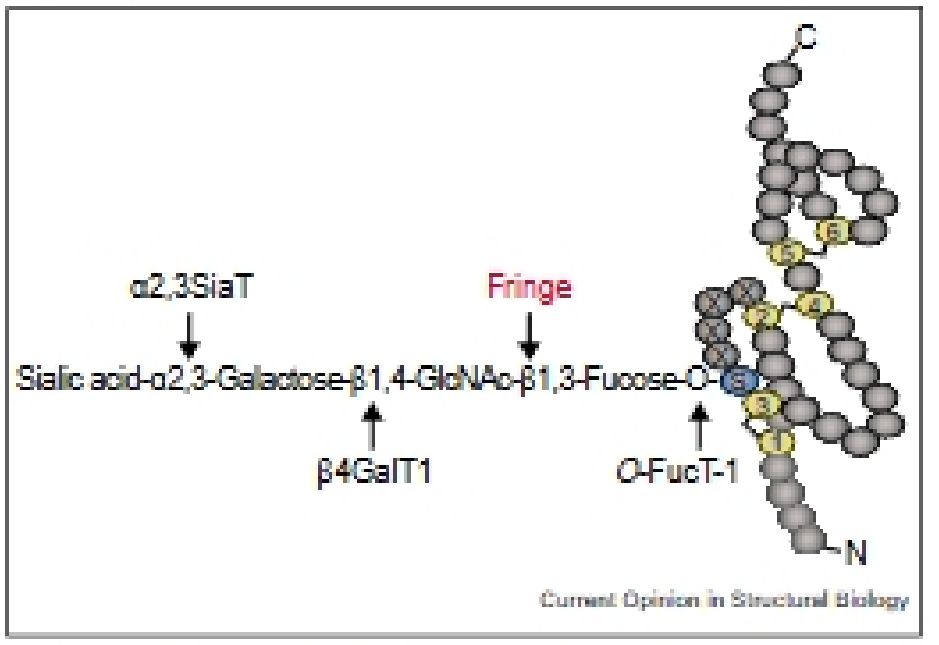
The discovery of the complexity and diversity of complex carbohydrates on the cell surface led researchers over 30 years ago to hypothesize that glycoconjugates play roles in communication between cells and in the transfer of information from the outside of the cell to the inside [1]. As this type of communication is essential for numerous stages of development, specific carbohydrate modifications were proposed to play roles in particular biological events at the cell surface. Over the years, numerous observations have supported this concept. Early on, many of the stage-specific embryonic antigens (e.g. SSEA-1, SSEA-3, SSEA-4, HNK-1) were demonstrated to be specific carbohydrate structures [2]. For instance, SSEA-1 is the Lewis x oligosaccharide, SSEA-3 and SSEA-4 are glycolipids of the globo series, and HNK-1 is a sulfated glycan. The expression of unique glycan structures at specific stages implied a particular function for that structure. The demonstration that the presence or absence of polysialic acid alters homotypic interactions of the neural cell adhesion molecule (NCAM) added further support [2]. More recently, the demonstration of embryonic lethality resulting from the genetic ablation of several glycosyltransferases has revealed that particular carbohydrate structures are essential for development to proceed past certain stages [3]. Nonetheless, with the exception of NCAM, identifying examples of particular carbohydrate structures on specific proteins mediating such effects was, for many years, elusive. Recent work has begun to identify some of these molecules. The first examples came with the demonstration that cell surface heparan sulfate proteoglycans play an essential role in Wnt, hedgehog, FGF (fibroblast growth factor) and TGFβ (transforming growth factor β) superfamily pathways. More recently, *O*-fucose modifications of epidermal growth factor (EGF)-like repeats have been shown to modulate Notch, TGFβ family (Nodal) and urinary-type plasminogen activator (uPA) signal transduction. Several excellent recent reviews have been written concerning the role of heparan sulfate proteoglycans in development [4–6] and thus will not be considered further here. This review will focus on the recent studies of the *O*-fucose modifications of EGF repeats of Notch, Cripto and uPA.

Involvement of O-fucose modifications of EGF repeats in signal transduction

O-linked carbohydrate modifications of EGF repeats

Fucose *O*-linked to serine or threonine was first observed over 25 years ago as amino acid fucosides isolated from human urine [7]. The first protein reported to bear *O*-fucose was uPA [8], quickly followed by tissue-type plasminogen activator and several clotting factors (factors VI, IX and XII) [9]. Comparison of the sequences surrounding the sites of *O*-fucose modification on these proteins showed the fucose to be localized to a putative consensus sequence within EGF repeats. EGF repeats are small (approximately 40 amino acid) protein motifs originally observed in epidermal growth factor. They are defined by the presence of six conserved cysteine residues that form three disulfide bonds (Figure 1; [10]). EGF repeats occur in dozens of cell surface and secreted proteins, and are known to play roles in protein–protein interactions. The *O*-fucose modification occurs between the second and third conserved cysteines of the putative consensus sequence $C^2XXGG\underline{S/T}C^3$ [9]. Proteins predicted to be modified with *O*-fucose based on the presence of this consensus sequence have been demonstrated to bear the modification [11•–14•], indicating that it can be used to make accurate predictions about whether a protein will bear *O*-fucose (Table 1). Recent work has suggested that the originally proposed consensus site is too narrow and that *O*-fucose modifications occur more broadly than predicted [14•]. As a result, a broader consensus site, $C^2X_{3-5}\underline{S/T}C^3$, has recently been proposed.

Figure 1



EGF repeat modified with the O-fucose tetrasaccharide.
A representation of an EGF repeat (based on EGF1 from factor VII [18]) modified with the O-fucose tetrasaccharide is shown. The conserved cysteines of the EGF repeat are in yellow and are numbered, and the disulfide bonds between them (C^1–C^3, C^2–C^4 and C^5–C^6) are shown. The serine (shown) or threonine modified with O-fucose is in blue, and the other amino acids between C^2 and C^3 are shown as X. The number of Xs can vary between 3 and 5 [14•]. The enzymes responsible for the addition of each saccharide are indicated: O-FucT-1 [15–17], Fringe [30••,37••], β4GalT1 [39•] and either α2,3SiaT (shown) or α2,6SiaT [11•,30••]. Adapted from [30••].

Table 1

O-fucose-modified proteins*.

Proteins known to be modified with O-fucose.
Blood clotting/dissolution
Urokinase (uPA) [8]
Tissue-type plasminogen activator (tPA) [46]
Desmodus (bat) salivary plasminogen activator [47]
Factor VII [48]
Factor IX [49]
Factor XII [50]
Notch signaling
Notch [11•,30••]
Delta [14•]
Serrate [14•]
Jagged [14•]
TGFβ signaling
Cripto [12••,13•]
Other
Fetal antigen-1/Delta-like protein (FA1/DLK) [51]
Proteins predicted to be modified with O-fucose based on the presence of the consensus sequence (C-X-X-G-G-S/T-C).
Slit
Crumbs
Cryptic
LDL receptor-related protein (LRP)
Acrogranin/epithelin
Brevican (PCCB)
Neurocan (PGCN)
Agrin
Hepatocyte growth factor activator
One-eyed pinhead
FRL-1
Versican (PGCV)
Multimerin
Fibrillin
Fibropellin
Perlecan (PGBM)

*Proteins that have been identified to contain O-fucose are listed along with proteins that contain the consensus sequence for O-fucose addition and may be modified. Proteins containing a consensus sequence within an EGF module were found by searching the sequence databases Swiss-Prot and PIR (Protein Information Resource) at the MOTIF web site (www.motif.genome.ad.jp) using the query pattern C-x-x-G-G-[ST]-C. Note that searches with the broader consensus site (C'X$_{3,5}$S/TC') are not informative due to the large number of proteins identified. Adapted from [52].

The enzyme responsible for the addition of *O*-fucose to EGF repeats, GDP-fucose: protein *O*-fucosyltransferase 1 (*O*-FucT-1), has been identified and cloned [15,16]. *O*-FucT-1 appears to be a type II membrane glycoprotein like most glycosyltransferases involved in the addition of sugars to proteins. It will not fucosylate synthetic peptides containing the consensus sequence for *O*-fucose addition, but requires a properly folded EGF repeat containing the consensus sequence (Figure 1; [15,17]). Homologs have been identified in species from *Caenorhabditis elegans* to humans, an expression pattern that is consistent with the distribution of proteins containing EGF repeats.

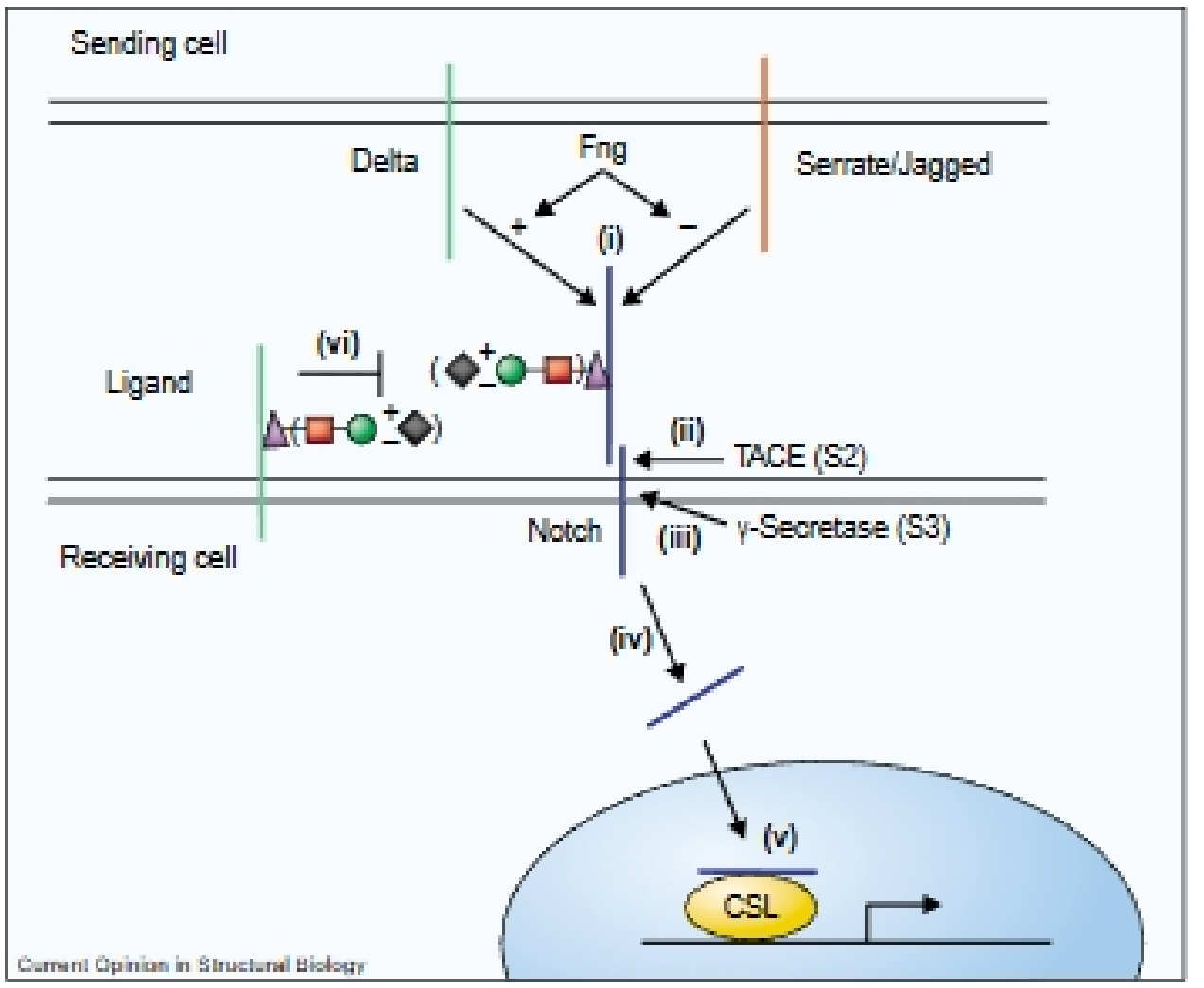
The structure of the EGF repeat from factor VII with and without the *O*-fucose has been determined using NMR [18]. Although the *O*-fucose presents a significant epitope on one face of the EGF repeat, no major conformational change in the polypeptide structure was observed. Thus, alterations in function due to changes in glycosylation are more likely to result from direct interactions with the carbohydrates or from steric blocking of a protein–protein interaction by the carbohydrate.

Early work on the function of *O*-fucose modifications indicated a role in modulating receptor–ligand interactions. Binding of uPA to the uPA receptor results in the activation of several signaling cascades within cells [19]. The EGF repeat of uPA is necessary and sufficient to activate the uPA receptor. Interestingly, removal of the *O*-fucose from the EGF repeat (either chemically or by synthesis in bacteria in which fucosylation does not occur) abrogates uPA receptor activation, although it has no effect on the binding of the EGF repeat to the receptor [20]. The mechanism by which the *O*-fucose activates the receptor is unknown, but these results demonstrate that the presence or absence of a simple sugar on an EGF repeat can regulate a signal transduction event.

Fringe is an O-fucose: β1,3-N-acetylglucosaminyl-transferase that modulates Notch signaling

Another example of how a signaling event can be regulated by alterations of carbohydrate structures on EGF repeats has been revealed in the Notch pathway. Notch is a large cell surface receptor protein that plays an essential role in numerous developmental events (for recent reviews on

Figure 2



Fringe modulates ligand-mediated activation of Notch (for a recent review on Notch activation, see [22]). Notch activation proceeds through five steps. (i) Notch in the signal-receiving cell becomes activated upon binding to Notch ligands (Delta and Serrate/Jagged) expressed on the cell surface of the sending cell. (ii) Ligand binding stimulates proteolytic cleavage of the extracellular domain of Notch at site 2 (S2) catalyzed by TACE. (iii) Release of the extracellular domain of Notch stimulates intramembranous cleavage at site 3 (S3) catalyzed by γ-secretase. (iv) γ-Secretase cleavage releases the intracellular domain from the membrane as a soluble protein in the cytoplasm. (v) The Notch intracellular domain translocates to the nucleus, where it interacts with the CSL family of transcriptional regulators to activate transcription. (vi) Ligand expressed in the same cell as Notch can inhibit activation through cell autonomous inhibition [53,54]. Fringe (Fng) modifies O-fucose residues on both Notch [30••,37••] and Notch ligands [14•], resulting in potentiation of signaling from Delta (Fng+) or inhibition of signaling from Serrate/Jagged (Fng−). O-fucose is represented by a triangle; β1,3-GlcNAc added by Fringe is a square; β1,4-galactose is a circle; α2,3/6-sialic acid is a diamond. The brackets indicate the carbohydrates that are added in response to Fringe action. The addition of the sialic acid is not necessary for Fringe to exert its effects on Jagged1-dependent Notch activation and, thus, is presented as +/− [39•]. Adapted from [30••].

Notch, see [21,22]). A variety of human diseases are caused by defects in Notch signaling, including T-cell leukemias, CADASIL, spondylocostal dysostosis and Alagille syndrome [23–25]. The extracellular domain of Notch is composed largely of tandem EGF repeats (36 in *Drosophila* Notch and mammalian Notch1 and Notch2), several of which contain consensus sites for *O*-fucose modification [11•]. The recent finding of a broader consensus site for *O*-fucose modification has greatly expanded the number of potential sites on Notch [14•]. Notch activation is initiated by ligand binding. Interestingly, Notch ligands (Delta and Serrate/Jagged) are also transmembrane proteins. Thus, Notch activation requires the expression of a ligand on an adjacent cell (Figure 2). Binding of ligand initiates a proteolytic event catalyzed by TACE, a member of the ADAMS family of metalloproteases, whereby the extracellular domain of Notch is released from the membrane. This is followed by another proteolytic event, just inside the membrane, catalyzed by γ-secretase. After the cytoplasmic domain of Notch is released from the membrane, it translocates to the nucleus, where it binds to members of the CSL (CBF1, Suppressor of hairless, Lag-1) family of transcriptional regulators, resulting in the activation of several downstream genes.

Notch activation can be regulated at several points within the signaling pathway. The Fringe protein was identified in *Drosophila* as a modulator of Notch activation, capable of potentiating signaling from Delta while inhibiting signaling from Serrate/Jagged (reviewed in [26,27]; Figure 2). Mutations in *fringe* result in defects in the proper development of several tissues in *Drosophila*, including wings, eyes and legs. Three Fringe homologs exist in mammals (Lunatic fringe, Manic fringe and Radical fringe) that are all capable of complementing *fringe* mutants in *Drosophila* [28,29]. Cell-based signaling assays have confirmed that Lunatic and Manic fringe can inhibit Jagged1-dependent Notch activation and stimulate Delta1-dependent activation [30••,31•,32•]. Genetic ablation of *Lunatic fringe* in mice results in fusion of somites, skeletal defects and perinatal death [33,34], although no obvious phenotype was observed in mice lacking *Radical fringe* [35,36]. No data are yet available on phenotypes of *Manic fringe* knockouts. Thus, Fringe carries out a variety of essential developmental functions by modulating Notch activity.

Fringe modulates Notch activity by altering the *O*-fucose carbohydrate structures on EGF repeats. Fringe is an