Structure and function of a Notch-like proteolytic switch domain in the extracellular matrix receptor dystroglycan

Michael J.M. Anderson, Amanda N. Hayward, Eva Grant, Lauren Greenberg, Robert Evans III, and Wendy R. Gordon*

Department of Biochemistry, Molecular Biology, and Biophysics, University of Minnesota, Minneapolis, MN, USA
* Please address correspondence to: wrgordon@umn.edu

Abstract:
The adhesion receptor dystroglycan provides a critical mechanical link between the extracellular matrix (ECM) and the actin cytoskeleton to help muscle cells withstand contraction and neural cells maintain the blood brain barrier. Disrupting the link is associated cancer and muscular dystrophy. Proteolysis of dystroglycan by matrix metalloproteinases (MMPs) provides a mechanism to break the mechanical anchor and is amplified in several pathogenic states, yet how MMP cleavage is regulated is largely unexplored. We solved the crystal structure of the extracellular membrane-adjacent domains of dystroglycan containing the MMP site and find the tandem Ig-like and SEA-like domains structurally interact in an L-shape to facilitate a permissive conformation for **α/β** auto-catalytic processing. The MMP cleavage site is located in a structural motif C-terminal to the SEA-like domain that folds back and nestles into a pocket within the SEA-like domain to protect the cleavage site from MMPs. We then show while the intact proteolytic switch domain is resistant to MMP cleavage, structurally disruptive muscular dystrophy-related mutations sensitize dystroglycan to proteolysis in secreted constructs and the full-length receptor. Intriguingly, previously uncharacterized cancer-associated mutations that map to the proteolytic switch domain similarly lead to increases in proteolysis. Taken together, this study demonstrates dystroglycan proteolysis is conformationally regulated and mutations in the domain lead to enhanced proteolysis.

Introduction:

Dystroglycan is an extracellular matrix (ECM) receptor expressed on the surface of cells that adjoin basement membranes, such as epithelial, neural, and muscle cells and plays important roles during development and adult homeostasis [1]. Dystroglycan is encoded by the *Dag1* gene and translated as a single polypeptide precursor that undergoes post-translational autolytic processing to generate two non-covalently interacting subunits at the cell-surface: the extracellular and highly glycosylated α-dystroglycan that binds laminin and other ligands in the ECM and the transmembrane β-dystroglycan that binds to the cytoskeleton via adaptor proteins such as dystrophin [2,3]. Dystroglycan's most well-understood function in cells is to provide a mechanical link between the extracellular matrix and the cytoskeleton [4]. In its function as a core member of the dystrophin-glycoprotein-complex (DGC) in muscle cells, dystroglycan provides a critical anchor between the ECM and the actin cytoskeleton to protect the sarcolemma from the forces of muscle contraction [5,6]. Similarly, dystroglycan expressed on astrocyte endfeet in the brain links to the basement membrane of endothelial cells lining blood vessels to help maintain the blood brain barrier [7]. Thus, loss of functional dystroglycan or breaking of the mechanical link it provides is generally detrimental to cells. Whole animal dystroglycan deletion leads to embryonic lethality [8], and conditional knockouts in muscle and brain tissues result in drastic tissue disorganization [9,10]. Hypoglycosylation in the laminin binding region of dystroglycan has been linked to cancer and muscular dystrophy. Mutation or truncation of dystrophin causes Duchenne muscular dystrophy. Recently, similar dystrophin truncations and mutations of dystrophin have been identified in cancers, suggesting the DGC plays a tumor and metastasis suppressor role in cancer.

One underexplored mechanism for breaking the DGC mechanical link is cleavage of dystroglycan by sheddases such as MMP2 and 9. Matrix metalloproteinase (MMP) cleavage likely plays a physiologic role to enable processes such as cell migration, wound repair, or function of neural synapses, but is apparently dysregulated in many diseases. Proteolysis of β-dystroglycan to a 31 kDa MMP cleaved fragment is enhanced in several pathogenic states, such as cancer [11–13], muscle diseases [14,15], and autoimmune disorders [7]. Additionally, high levels of MMPs are found in biopsies of muscular dystrophy patients [16], and muscular dystrophy phenotypes are ameliorated in mice treated with metalloprotease inhibitors [17–19]. Moreover, MMP-2 / MMP-9 double-knockout mice in which MMP cleavage of dystroglycan is blocked are resistant to the onset of autoimmune encephalomyelitis, underscoring the relevance of this cleavage in disease pathogenesis [7].

The non-covalent association of α- and β-dystroglycan occurs within a putative Sea urchin-Enterokinase-Agrin-like (SEA-like) domain, located extracellularly between the ECM-binding and transmembrane regions [20]. Importantly, this SEA-like domain also houses the predicted MMP cleavage sites and harbors one of the patient-derived muscular dystrophy mutations (C669F) identified in dystroglycan [21]. Several other cell surface receptors contain SEA-like domains just extracellular to the transmembrane region[22], most notably Notch[23,24]. In Notch, the SEA-like domain cooperates with an adjacent N-terminal domain to structurally occlude a protease site that becomes exposed by forces associated with cell to cell contact[25]. We recently used a Synthetic Notch Assay for Proteolytic Switches (SNAPS) to show that the region of dystroglycan comprising the Ig-like and SEA-like domains functions as a Notch-like proteolytic switch when used to replace the Notch NRR in the context of the full-length Notch receptor [26]. That is, dystroglycan shedding did not occur until forces associated with ligand binding unfolded the domain to expose the sites. This suggests dystroglycan's MMP sites may be conformationally regulated like Notch's ADAM site and that disease-related mutations in the domain might dysregulate proteolysis. Moreover, if dystroglycan proteolysis is regulated by the conformation of the domain like Notch, proteolysis could potentially be modulated by therapeutic antibodies. However, since the Notch transmembrane domain was present in the chimeric Notch signaling assay, it is likely the ADAM proteases were the sheddases and not MMP2/9 which are known to cleave dystroglycan in its physiologic context. Moreover, secondary structure prediction programs suggested the MMP sites mapped to an unstructured region C-terminal to the putative SEA-like domain. Thus, we solved the crystal structure of the putative dystroglycan proteolytic switch domain in order to understand if dystroglycan proteolytic cleavage is conformationally regulated, how disease-relevant mutations might alter dystroglycan's function, and how dystroglycan's juxtamembrane SEA-like domain compares to other cell surface receptors.

Results

Dystroglycan α/β-processing requires both Ig-like and SEA-like domains

We first aimed to map the domain boundaries of the putative dystroglycan proteolytic switch domain. Like several other cell surface receptors, dystroglycan contains a SEA-like domain that that houses both α/β-processing and putative MMP sites just extracellular to the transmembrane domain, and an adjacent N-terminal Ig-like domain that might play a role in stabilizing the SEA-like fold (Figure 1A).[15,27]. Previous studies of SEA domains in mucins revealed the autoproteolytic α/β-processing event can only occur if the domain is correctly folded due to the necessity of a precisely strained loop to permit a serine in the cleavage sequence to nucleophilically attack a neighboring glycine [28]. Thus, we utilized SEA-domain α/β-processing as an indication of correct protein folding to map the minimal functional unit of dystroglycan's proteolytic switch domain and devised a series of constructs of human dystroglycan including the SEA-like domain and its neighboring Ig-like domain. Dual-epitope tagged constructs of dystroglycan's proteolytic switch domain and sequential truncations were designed using predicted secondary structure as a guide (Fig 1B).

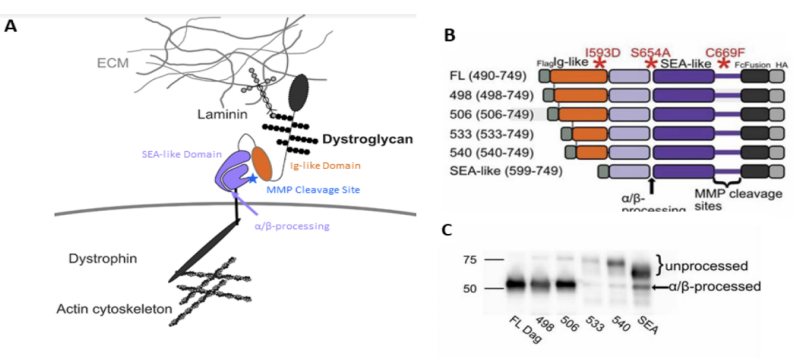

FIG 1. Dystroglycan α/β-processing requires both Ig-like and SEA-like domains. (A) Extracellular matrix receptor dystroglycan **(B)** Infographic of constructs used to map the dystroglycan proteolytic switch domain. Constructs consisted of Ig-like (orange) and SEA-like (purple) domains containing an N-terminal FLAG tag and C-terminal Fc and HA epitope tags. Sites of α/β-processing, MMP cleavage, and disease-related mutations (red asterisks) are depicted. FL = full-length. **(C)** Anti-HA western blot comparing immunoprecipitated protein levels of unprocessed and α/β-processed dystroglycan and associated truncation constructs depicted in panel (A). The % processed is quantified in the graph below using densitometry**.**

The constructs were expressed in HEK-293T cells, secreted, and subsequently immunoprecipitated. Two potential α-HA-tagged protein species were expected: a band of ~50 kDa corresponding to the correctly α/β-processed C-terminus in all constructs and a variable-sized larger band corresponding to the full-length unprocessed protein. The full-length (FL) construct containing the entire Ig-like and SEA-like domains showed a dominant 50 kDa band and a faint larger band, reflecting close to 100% α/β-

Space group	P 21 21 21
Cell Dimensions	
a, b, c (Å)	44.864 86.494 127.914
α, β, γ (°)	90 90 90
Resolution (Å)	51.43 - 2.432 (2.519 - 2.432
Number of observed reflections	137993 (12051)
Number of unique reflections	19288 (1838)
Redundancy	7.2 (6.6)
Completeness (%)	98.92 (94.33)
R-merge	0.09652 (1.799)
Mean I/σ_I	10.62 (1.14)
Refinement statistics	
Resolution (Å)	51.43 - 2.432 (2.519 - 2.432
R-work/R-free	0.2661/0.3474
Number of atoms	3521
Protein	3507
Ligands	2
Water	12
B-factors	80.88
Protein	80.92
Ligands	90.33
Water	67.89
Root mean square deviations	
Bond length (Å)	0.009
Bond angles (°)	1.34

Table 1. Crystallographic data on the solved proteolytic switch domain structure

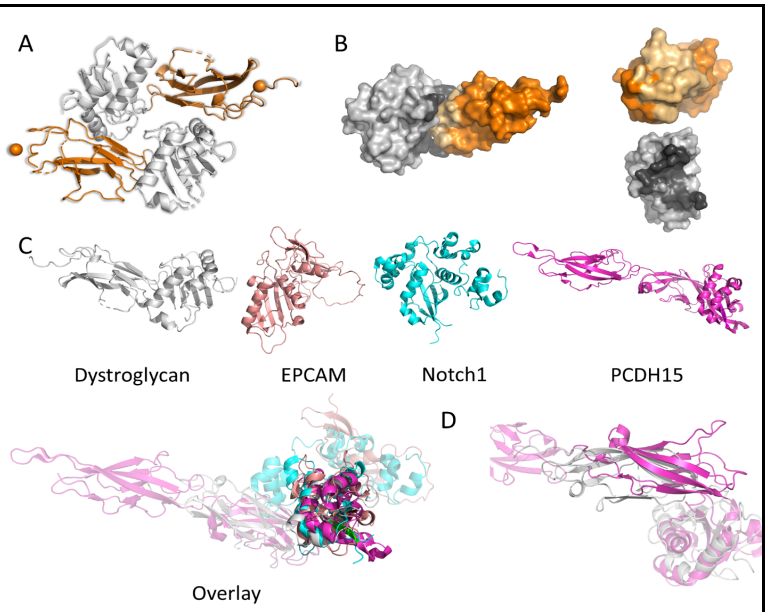

Figure 2. Dystroglycan structure highlights interaction between Ig-like and SEA-like domains and the similarity of dystroglycan with other SEA domain containing proteins. Cartoon representation of dystroglycan and other previously published proteins with SEA containing domains. **(A)** Crystal structure of dystroglycan proteolysis domain of asymmetric unit with Ig-like domain colored in orange and SEA-like domain in gray. **(B)** Surface representation of dystroglycan with Ig-like and SEA-like domain interactions colored differently to show interface **(C)** Dystroglycan (gray) structure compared to the published structures of other SEA domain proteins including EPCAM (salmon; PDB: 4MZV), Notch1 (cyan; PDB: 3I08), and PCDH15 (pink; PDB: 6BXZ). Bottom shows the superimposition of the SEA domains of these structures. **(D)** Overlay of the cadherin domains of dystroglycan (gray) and PCDH15 (pink), highlighting the difference in their interactions with the SEA domain.

processing (Fig 1C). However, the N-terminal truncations of the Ig-like domain exhibited an increasing ratio of unprocessed to processed dystroglycan, and

the constructs containing only some of the Ig-like domain ("540") or solely the SEA-like domain showed the lowest degree of α/β-processing. These data suggest that the Ig-like domain folds in concert with the SEA-like domain to support dystroglycan α/β-processing. We will term the portion of dystroglycan comprised of full-length Ig-like and SEA-like domains the dystroglycan "proteolytic switch domain." This is the same domain that is able to functionally substitute for Notch's NRR in facilitating proteolytic events required for Notch signaling. [26]

Structure of proteolytic switch domain reveals interaction between Ig-like and SEA-like domain

We next expressed this construct in E. Coli, purified it, and set up crystallization screens to solve an X-ray structure of the Ig-like + SEA domain region at 2.4 Å (Table 1). Two dystroglycan molecules comprise the asymmetric unit with density for residues G491-V722 in one structure and N494-P720 in the other (Figure 2A). The two molecules in the asymmetric unit interact in an antiparallel fashion and align with 0.673 RMSD when overlaid. The overall monomer structure shows the Ig-like and SEA-like domains interacting to form an L-shape. The N-terminal domain Ig-like domain shows a characteristic two layer Ig-domain β-sandwich and the C-terminal SEA-like domain is characterized by a **βαββαβ** sandwich as observed in other SEA domain structures (Figure 2A)[23,29,30]. The interface between the Ig-like and SEA-like domain is moderate, spanning over 550 $Å^2$ as calculated. Seven hydrogen bonds and 2 salt bridges form the strongest interaction between the two domains though a variety of hydrophobic interactions also exist (Figure 2B). Five amino acids (513Y, 562Y, 631K, 643S, and 647L) involved in hydrogen bond formation are highly conserved with other sites featuring conserved polarity. Importantly, these interactions directly stabilize one of the β-strands which is involved in positioning for autoproteolysis as well as the **α**-helix that sits atop the two β-strands involved in auto-proteolysis. This interaction helps to stabilize the torsional strain required for autoproteolysis in the SEA domain[20] This crystal structure provides further evidence that the Ig-like domain is important to stabilize the SEA-like domain in dystroglycan.

Structural Comparison to other juxtamembrane SEA-domain containing proteins

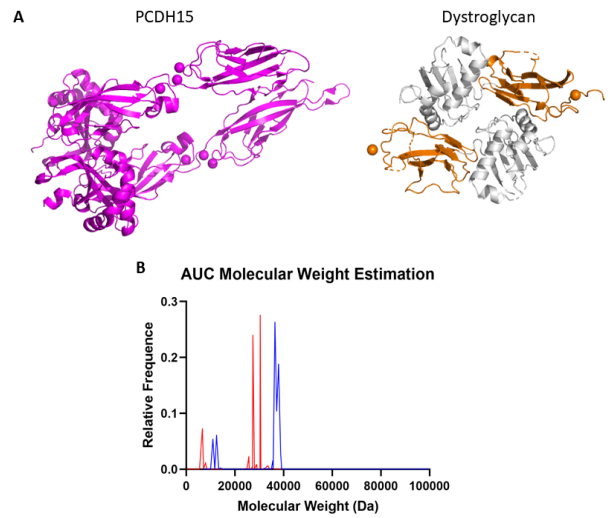

Figure 3. Determination of dystroglycan dimerization in solution. (A) Comparison of PCDH15 and dystroglycan dimerization. PCDH15 forms a parallel dimer whereas dystroglycan forms an anti-parallel dimer in crystal. **(B)** Molecular weight as determined by AUC for two different protein preparations.

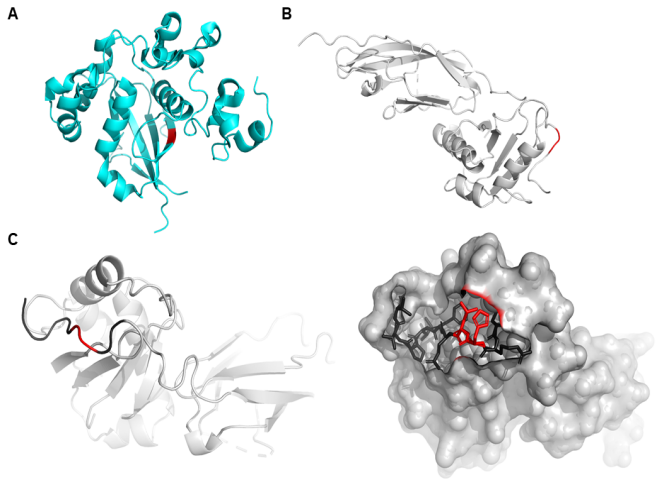

Figure 4. Dystroglycan MMP cleavage site is sterically regulated by association with the SEA domain. (A-B) proteolytic cleavage sites (red) of Notch (A) and Dystroglycan (B). **(C)** Cartoon (left) and surface (right representation of dystroglycan with MMP cleavage site (red) and surround amino acids (black).

Comparison of dystroglycan with other structures of proteins with SEA or SEA-like domains, including Notch1 (PDB: 3I08), Protocadherin-15 (PCDH15) (PDB: 6BXZ), EPCAM (PDB: 4MZV), Mucin1 (PDB: 2ACM), and Mucin16 (PDB: 1IVRZ) highlight how structurally conserved the domain is (Figure 2C). However, superimposition of the structures reveal that dystroglycan SEA domain is most structurally similar to that of Notch1 and EPCAM with RMSD ~3.0 $Å^2$ whereas RMSD with the other three is close to 5.5 $Å^2$. The overall architecture of PCDH15 and dystroglycan is strikingly similar including the N-terminal neighboring Ig-like (cadherin) domain, though the cadherin domain sits lower compared to dystroglycan (Figure

2D). The other structures all feature adjacent N-terminal domains that interact with the SEA-like domain, though their structure, nature of interaction, and extent of interaction all vary.

Observed dystroglycan dimer is crystallographic in contrast to Protocad-15 dimer

Recent PCDH15 crystal structures shows that the equivalent SEA-like and adjacent Ig-like domain form a crystallographic dimer, though PCDH15 forms parallel dimers while dystroglycan forms an antiparallel dimer (Figure 3A). Most of the strong bonds at the dimer interface are highly conserved between dystroglycan and PCDH15, though there are some differences. The study further showed that the PCDH15 domain is a dimer in solution, which is hypothesized to be important for protein function[30]. The structural similarities and similar crystallographic dimer led us to explore if dystroglycan is a dimer in solution.

To test for dimerization in solution, we used analytical ultracentrifugation (AUC) as it can not only determine if a dimer state exists but also quantify what percent of protein dimerizes as seen in the crystal structure. AUC gave a major sedimentation coefficient at 2.2 which corresponds to ~35 kDa, slightly higher than the molecular weight of a monomer (Figure 3B). Thus no evidence of dimer formation existed in AUC and we conclude that dystroglycan exists as a monomer in solution.

Dystroglycan MMP cleavage site is protected by interdomain interactions

The ADAM10/17 site of Notch maps to the middle of the terminal beta-strand of the SEA-like domain (Figure 4A). It is occluded from its protease by the adjacent N-terminal domain that drapes over the SEA-like domain and with the aid of a Leucine plug directly prevents access to the ADAM site. Though the chimeric Notch signaling assays where the dystroglycan domain studied here can functionally substitute for the

Notch proteolytic switch domain, suggesting a potentially similar mechanism of conformational protection of dystroglycan proteolysis, dystroglycan is cleaved by MMPs at a site His715-Leu716 that maps outside of the SEA-like domain and predicted to be unstructured. Moreover, the structure shows the adjacent Ig-like domain interacts with the SEA-like domain distal from the MMP sites (Figure 4B), suggesting different mode of protease site regulation than in Notch. A closer look at structure reveals the MMP site is located in a stretch of amino acids C-terminal to the SEA-like domain that folds back into a groove formed between the terminal beta strand and alpha helix (Figure 4C). Though the His-Leu residues comprising the MMP site are solvent exposed, access to MMP is likely prevented via tight anchoring of the peptide via a disulfide bond, multiple backbone hydrogen bonds and hydrophobic interactions that bury the phenylalanine c-terminal to the MMP site. Notably, the disulfide bond anchor is the site of the disease-associated mutation C669F found in patients with muscular dystrophy, suggesting that disruption of this motif might make dystroglycan susceptible to MMPs.

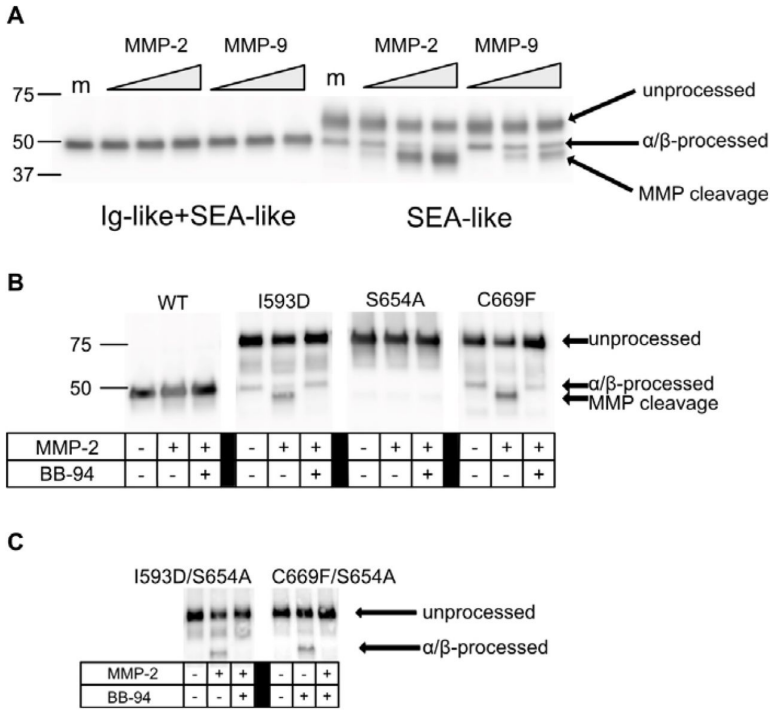

FIG 5. MMP cleavage of the dystroglycan proteolytic switch domain is conformationally regulated. Anti-HA western blot of immunoprecipitated dystroglycan proteolytic switch domains. **(A)** Comparing the proteolytic switch domain (cad-like and SEA-like domains) versus SEA-like domain only. Increasing concentrations of activated MMP-2 and MMP-9 were reacted in MMP buffer. **(B)** Proteolytic switch domains harboring disease mutations. **(C)** Combining disease mutations. Blots are representative of multiple independently performed experiments. Reactions in (B) and (C) contained 0.73 μg/mL activated MMP-2 and 40 μM BB-94 where noted.

MMP cleavage of the proteolytic switch domain is conformationally regulated

The structure suggests the MMP site in the proteolytic switch domain of dystroglycan should be resistant to exogenous MMPs. To probe the accessibility of dystroglycan's protease sites to exogenous MMPs, constructs consisting of either the proteolytic switch domain or the SEA-like domain alone lacking the adjacent Ig-like domain required for proper folding were secreted and immunoprecipitated in HEK-293T cells. The proteins were incubated with increasing amounts of recombinant activated MMP-2 or MMP-9 for 1 hour and blotted against the HA tag. Upon addition of exogenous MMPs, the proteolytic switch domain remained a single band while a faster-migrating species appears for the SEA-like domain alone construct (Fig 5A). This suggests that the MMP sites are exposed when the Ig-like domain is not present, likely due to incorrect folding of the domain, which is consistent with the α/β-processing truncation studies in Figure 1C. This result recapitulates what we observed

with the SNAPS assay, in which the dystroglycan proteolysis domain in the context of Notch was protected from MMP cleavage while the truncated domain was susceptible to cleavage [26].

Since protection of the MMP sites in the context of isolated proteolytic switch domains relies on the correct conformation of the domain, we next wanted to test whether mutations associated with muscle disease found in humans and model organisms alter the conformation of the domain and expose the protease sites. To date, three disease-related mutations within the dystroglycan proteolytic switch domain are associated with muscular dystrophy [21,31,32], but how they contribute to disease pathogenesis is unclear. I593D, a mutation located in the Ig-like domain, was identified in a forward screen for muscular dystrophy mutations in zebrafish [31]. The S654A mutation blocks α/β-processing and has been observed to cause a muscular dystrophy phenotype in mice [32]. This mutation, however, is not expected to alter the conformation of the Ig-like or SEA-like domains. Finally, the C669F mutation located near the C-terminus of the SEA-like domain, which disrupts a putative disulfide linkage, was discovered in patients with a severe form of muscular dystrophy [21]. The mechanism of how this mutation impacts dystroglycan function has not been elucidated.

The three disease constructs were individually expressed, immunoprecipitated, and blotted to assess both α/β-processing and MMP cleavage. In comparison to the wild type proteolytic switch domain, decreased α/β-processing is observed for all three mutations (Fig 5B). While expected for the S654A mutant in which the processing cleavage site is abolished, the incorrect α/β-processing of C669F and I593D suggests they disrupt the structure of the domain, much like the truncation constructs described in earlier experiments. Upon addition of exogenous MMP-2, the proteolytic switch domains harboring the I593D and C669F mutations both had the appearance of an MMP-cleaved species. The presence of this species was abrogated by the addition of batimastat (BB-94), a global metalloprotease inhibitor. In contrast, the S654A mutation did show any cleavage products resulting from MMP addition suggesting an intact proteolytic switch domain. Previous structural work of mucin SEA domains has shown that mutating the α/β-processing site as in S654A does not impact the overall fold of the SEA domain [28]. When the S654A mutation is combined with the other two MMP cleavage susceptible mutations, there is an expected complete loss of α/β-processing (Fig 5C). The addition of exogenous MMP-2 also results in MMP proteolysis. Therefore, we conclude that while all three disease mutations disrupt α/β-processing of dystroglycan's proteolytic switch domain, only the C669F and I593D mutations alter folding of the domain, and MMP cleavage is indeed conformationally regulated.

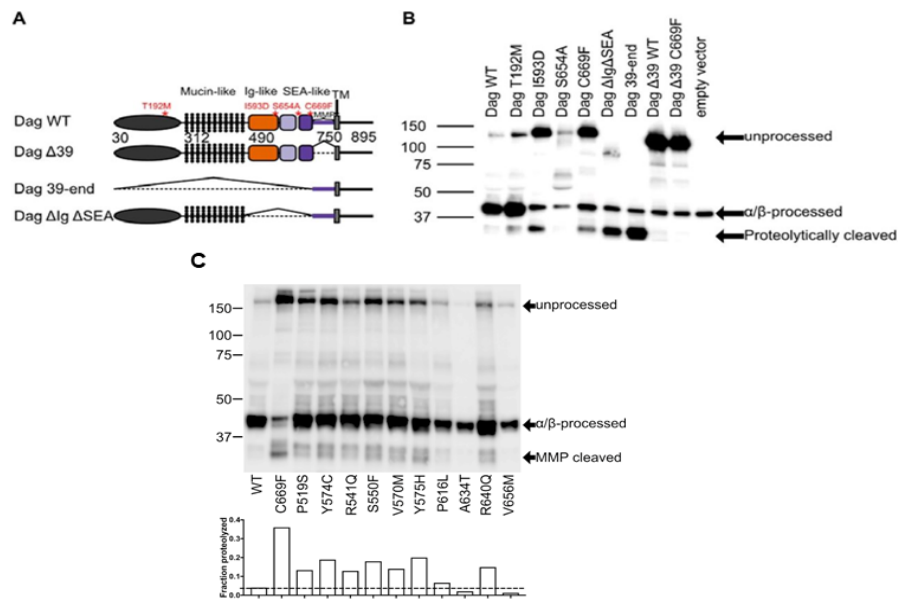

FIG 6. Presence of the β-dystroglycan proteolytic fragment depends on the structural integrity of the proteolytic switch domain. (A) Schematic of full-length dystroglycan constructs used in (B). Red asterisks denote location of muscle disease mutations. Dashed lines denote domain deletions. TM = transmembrane domain. **(B-C)** Anti-β-dystroglycan western blots of full-length dystroglycan constructs transiently transfected in Cos7 cells. Processed β-dystroglycan migrates at 42 kDa while MMP proteolytically-cleaved β-dystroglycan migrates at 31 kDa. **(C)** Cancer-related

Proteolysis in full-length receptor depends on the structural integrity of the proteolytic switch domain

We then explored whether the MMP resistance phenotype of the wild-type proteolytic switch domain and the proteolysis sensitive phenotype observed in truncation and muscular dystrophy disease mutants was recapitulated in the full-length receptors. Various full-length and truncated dystroglycan constructs were transiently transfected into Cos7 cells (Fig 6A). Cell lysates were blotted for β-dystroglycan using an antibody specific to the C-terminus of dystroglycan. Full-length wild-type dystroglycan undergoes normal α/β-processing and no appreciable MMP cleavage is observed, as expected (Fig 6B). In contrast, the I593D and C669F mutations both disrupt α/β-processing, and a new 31 kDa fragment appears corresponding to the size of β-dystroglycan cleaved by endogenous MMPs [15,27]. To confirm that this new band corresponds to a proteolytic fragment of the correct size, we expressed a construct termed "39-end" that begins 39 amino acids into the extracellular juxtamembrane domain, approximately corresponding to the expected MMP-cleaved species. This construct migrated similarly to the observed MMP cleaved fragment. To further test that the new species corresponds to proteolytic events near the C-terminus of the SEA-like domain and is not just an artifact of overexpression, we created a construct termed "Δ39" in which the putative MMP cleavage sites were deleted (removal of amino acids 711-749 directly before the transmembrane domain). When this deletion is present in the context of the C669F mutation, no MMP cleaved species is observed (Fig 6B). Moreover, in the context of a construct termed "ΔIgΔSEA" that removes all of the proteolytic switch domain except residues 711-749 in which constitutively exposes proteolysis sites, a large fraction of β-dystroglycan is further proteolyzed into the MMP cleaved fragment. This trend mirrors what was observed with immunoprecipitated proteolysis domains. The same pattern of disease mutants blocking α/β-processing and exposing MMP cleavage sites was also seen in other immortalized cell types, validating what

we observed in Cos7 cells. As a final test of how proteolysis is affected, we encoded dystroglycan with cancer associated mutations found in the Catalog of Somatic Mutations in Cancer (COSMIC). As observed with muscular dystrophy mutations, most cancer mutations increased MMP cleavage of full length dystroglycan (Figure 6C). Thus, dystroglycan proteolysis is dysregulated in both cancer and muscular dystrophy.

Discussion:

The adhesion receptor dystroglycan provides a critical mechanical link between the ECM and the actin cytoskeleton that can be disrupted by proteolytic cleavage by MMPs, and several lines of evidence suggest a role for abnormal dystroglycan proteolysis in disease pathogenesis. We demonstrated that dystroglycan proteolysis by MMPs is conformationally regulated by a juxtamembrane proteolytic switch domain, analogous to regulation of proteolysis in Notch receptors. We found that dystroglycan's SEA-like domain, which harbors both the α/β-processing site and the putative MMP cleavage sites, cooperates with its adjacent Ig-like domain to support the non-covalent complex formed from α/β-processing during receptor maturation. Autoproteolysis, which depends on correct folding of the domain, only occurs in the presence of its neighboring Ig-like domain. While MUC1 [33] and MUC16 [34] SEA domains and the SEA-like domain of the receptor tyrosine phosphatase IA-2 [35] exist in the absence of neighboring domains, the dystroglycan SEA-like domain appears to be more similar to the SEA-like domains of PCDH15[30], Notch [36–38] and EpCAM [29], which extensively interact with neighboring N-terminal domains. We find that the dystroglycan proteolytic switch domain in isolation is resistant to exogenous metalloproteases but becomes susceptible to proteolysis when the domain is truncated or encodes muscle disease-related mutations, much like in Notch receptors [39,40]. A similar trend is observed in full-length receptors; low levels of the MMP cleaved fragment are observed for the wild-type receptor, but levels are dramatically increased for constructs in which a majority of the proteolytic switch domain has been removed and for receptors encoding disease-related mutations. Moreover, the aberrant proteolysis in the C669F mutant is abrogated when the MMP sites are deleted from the construct.

The second conclusion of this work is that mutations in the dystroglycan proteolytic switch domain destabilize the domain, resulting in sensitization of the proteolytic switch domain to MMP cleavage and thus disruption of the critical ECM to actin tether. Impairment of the link the DGC provides is clearly important in muscular dystrophy and is an emerging concept in cancer. In addition to muscle disease related mutations, we also examined previously uncharacterized mutations mined from cancer data sets in COSMIC. The muscle disease related mutations disrupt α/β-processing, suggesting that they drastically alter the fold of the SEA-like domain and expose the normally occluded protease sites. Isolated proteolytic switch domains harboring these mutations are cleaved by exogenous MMPs, and a large fraction of the receptor is cleaved into the 31 kDa species in the context of the full-length receptor. On the other hand, the cancer mutations we investigated generally exhibit normal α/β-processing, suggesting that the structure of the SEA-like domain is not drastically altered. Further work is needed to explore the mechanism of proteolytic exposure in the cancer mutants, but the observed graded structural disruption is similar to observations of leukemia-derived mutations found in the NRR of Notch receptors [39]. Nevertheless, most mutations of both classes result in enhanced MMP proteolysis in the context of the full-length receptor. Future studies will be used to determine if this aberrant proteolysis is relevant to dystroglycan's critical roles in the cell by examining both cell migration and cell morphology phenotypes.

Our findings that MMP cleavage is normally controlled by the conformation of the proteolysis domain suggest that blocking dystroglycan proteolysis may also be a potential therapeutic avenue. Indeed, treatment of muscular dystrophy mouse models with protease inhibitors ameliorates symptoms [17–19]. Our study provides a rationale and platform for developing specific inhibitors of dystroglycan proteolysis. Since dystroglycan proteolysis is increased by the presence of disease mutations and in pathogenic states, developing an antibody that renders dystroglycan protease-resistant could be of great clinical interest. This strategy has been successfully employed in Notch [41,42] and with HER2 [43]. Antibodies to both of these targets have been utilized in our SNAPS assay to block proteolysis and the methodology could similarly be deployed with dystroglycan [26].

Acknowledgements

We would like to thank Jim Ervasti for helpful discussions. This study was supported by an NIH NIGMS R35GM119483 grant to WRG. ANH received salary support from AHA Fellowship 17PRE33330000 and is an ARCS Fellow. MJMA was supported by NIH MSTP grant T32 GM008244. WRG is a Pew Biomedical Scholar.

Author contributions
MJMA, ANH and WRG designed experiments. MJMA and ANH acquired data. All authors analyzed and interpreted data as well as wrote and revised the manuscript.

Materials and Methods:
Materials
Antibodies used in this study were: anti-FLAG (Sigma, F1804), anti-HA (Covance, MMS-101P), anti-β-dystroglycan (Leica, 43DAG1/8D5), anti-vinculin (Sigma, V9131). Restriction enzymes were purchased from New England Biolabs (NEB). Cloning was performed using In-Fusion HD Cloning Mix (Clontech). All electrophoresis supplies were purchased from Bio-Rad. MMP-2 and MMP-9 were purchased from R&D Systems. Point mutations within dystroglycan's proteolytic switch domain were created using Pfu Turbo Polymerase (Agilent) and DpnI (NEB).

Cloning
Human dystroglycan cDNA in CMV-6 plasmid was obtained from Origene (Cat # SC117393) and used as a template for all of the cloning. For expression of secreted proteolytic switch domains, cad-like+SEA-like domain (amino acids 490-749) and SEA-like domain (599-749) constructs were cloned into pcDNA3 vector containing N-terminal Notch signal sequence and Flag tag and C-terminal Fc and HA tags. For creating full-length dystroglycan constructs, dystroglycan and mentioned truncations were PCR amplified from cDNA and cloned into pcDNA3 digested with BamHI and HindIII. Disease relevant mutations were created using Quikchange II site-directed mutagenesis (Agilent). For x-ray crystallography and AUC, Ig-like and SEA-like domain (amino acids 490-749) were cloned into pTD68/His6-SUMO parent vector.

Mammalian cell culture
All of the cells used in this paper were cultured in Dulbecco's Modified Eagle Medium (DMEM, Corning) supplemented with 10% FBS, and 100 units/mL penicillin/streptomycin. Cells were maintained at 5% CO_2 at 37 ºC in a humidified incubator. Cos7 cells were a kind gift from Dr. Margaret Titus. U2-OS and HEK-293T cells were purchased from ATCC. All transfections were performed in Opti-MEM (Gibco) with Lipofectamine 3000 (Invitrogen) according to manufacturer instructions.

MMP activation
MMP-2 and MMP-9 were activated using p-aminophenylmercuric acetate (APMA) in MMP assay buffer (50 mM Tris pH 7.5, 10 mM $CaCl_2$, 150 mM NaCl, 0.05% Brij-35 (w/v)) according to manufacturer instructions. MMP-2 was activated for 1 hour at 37°C. MMP-9 was activated for 24 hours at 37°C according to manufacturer instructions.

MMP assay
10 cm plates of HEK-293T cells were transiently transfected with dystroglycan proteolytic switch domains. The conditioned media was collected 48 hours post-transfection. Conditioned media was bound to magnetic protein A beads and incubated for 1 hour at 4°C. The beads were washed thoroughly with wash buffer (50 mM Tris pH 7.5, 150 mM NaCl, 0.1% NP-40). The beads were split into separate tubes, MMP buffer was added (50 mM Tris, 10 mM $CaCl_2$, 150 mM NaCl, 0.05% Brij-35 (w/v), pH 7.5) and activated MMP and BB-94 added as appropriate. For the experiment using increasing amounts of MMP in Figure 2A: 0.07 µg/mL, 0.37 µg/mL, and 0.73 µg/mL of respective MMPs were added, and the tubes were incubated for 1 hour at 37°C. The remaining protein was eluted by incubating the beads with 2x SDS sample buffer with 100 mM DTT for 5 minutes at room temperature.

Western blot
Cell lysates from 24-well plates were collected 72 hours post-transfection using RIPA lysis buffer supplemented with protease inhibitor. Lysates were electrophoresed on a 4-20% SDS-PAGE gel in Tris/Glyine/SDS running buffer supplemented with 2 mM sodium thioglycolate. The protein was then transferred to a nitrocellulose membrane using a Genie Blotter (Idea Scientific) and blocked with 5% milk. The primary antibodies were diluted 1:1000 in 5% bovine serum albumin (BSA) with 0.2 % sodium azide in TBS-T and incubated overnight at 4 ºC. A goat-anti mouse HRP conjugated antibody (Invitrogen) (1:10000) was used as a secondary antibody. Western blots were imaged using chemiluminescent buffer (Perkin Elmer Western Lightning Plus ECL) and the

Amersham 600UV (GE) with staff support at the University of Minnesota Imaging Center. Band quantification was performed using ImageJ.

Homology model generation
The homology model for the proteolytic switch domain of dystroglycan was produced using SWISS-MODEL(?).

E. coli expression
Sequence confirmed pTD68 vector with dystroglycan (amino acids 490-749) were transformed into BL21(DE3) *E. coli* competent cells (Agilent). Cells cultured in 1L LB broth at 37°C and induced at OD600 of 0.6 with 0.5 mM IPTG (isopropyl-d-1-thio-galactopyranoside, Sigma Aldrich) and grown overnight at 18°C. Cell pellets were resuspended in 10 mL of lysis buffer (50 mM Tris pH=7.5, 250 mM NaCl, 1 mM EDTA) and pulse sonicated for three 30 second rounds with 2 minute incubation in ice in between. Lysate was centrifuged at 24k x G for 30 min and supernatant discarded. The pellet was resuspended in lysis buffer supplemented with 5M Urea and centrifuged again at 24k x G for 30 minutes. This supernatant was batch bound for 1 hour with 2 mL HisPure Ni-NTA agarose beads (ThermoFischer). After lysate cleared gravity columns, the beads were washed with 20 mL each of wash buffer 1 (50 mM Tris pH=7.5, 500 mM NaCl, 3 M urea) and wash buffer 2 (50 mM Tris pH=7.5, 200 mM NaCl, 100 mM imidazole, 3M urea). Elution was done with 5 mL 50 mM Tris pH=7.5, 200 mM NaCl, 250 mM imidazole, 1 M urea). The Ni column elution was mixed with an equal volume of refold buffer (50 mM Tris pH=8.5, 200 mM NaCl, 50 mM CaCl2, 2 mM cysteine, 0.5 mM cystine) and put in a 10 kDa snakeskin dialysis bag (ThermoFischer) and placed in 2L refold buffer and left at 4°C overnight. The protein was then moved into freshly made refold buffer and left at 4°C for 4 hours before a final switch in fresh refold buffer at 4°C overnight. It was then moved into dialysis buffer (50 mM Tris pH=7.5, 200 mM NaCl, 10 mM CaCl2) and left for 4 hours at 4°C. ULP-1 (5 U per 1 L of *E. coli* expression) was added into the dialysis bag which was placed in new dialysis buffer and placed at 4°C overnight. The contents was then batch bound with 2 mL HisPur Cobalt Resin (ThermoFisher) to removed His-ULP1 and His-SUMO. Flow through from the gravity column was collect, and the column was washed with 10 mL dialysis buffer (also collected). The flow through and wash were pooled and concentrated using spin concentrator (Amicon Ultra-15 CentrifugalFilter Unit, 3 kDa cut-off) in order to be further purified using ENrich SEC70 (Bio-Rad) size exclusion column (with filtered and degassed dialysis buffer). Samples were concentrated to 7 mg/mL.

Crystallization, data collection, and processing
We used Rigaku's CrystalMation system to perform a broad, oil-immersion, sitting drop screen of dystroglycan. Crystals were achieved using a 1:1 ratio of protein to a well solution of 0.1 M Tris, 0.2 M lithium sulfate, 20% PEG 4000, pH=8.3. Addition of any cryoprotectant caused lower resolution data collection so crystals were frozen directly into liquid nitrogen and shipped to Advanced Photon Source at the Argonne National Laboratory. All data was processed with the HKL suite.

The dystroglycan structure was solved with molecular replacement using a predicted structure from AlphaFold 2. Refinement was done using PHENIX software and visual correction using Coot . All statistical analyses and verification was done in PHENIX.

Analytical Ultracentrifugation (AUC)
Sedimentation velocity AUC experiments were performed in a Optima Analytical Ultracentrifuge (Beckman Coulter) following standard procedure (ref). Briefly, purified protein was diluted to absorbance of 0.6 at 280 nm and loaded into AUC cell assemblies with Ti centerpieces. These cells were loaded into the An-60 TI rotor and spun at 50k rpm at 20°C. Absorbance was measured at 280 nm. Data analysis was performed with the software UltraScan III following standard protocol. AUC experiments were done with two biological replicates which were run in duplicate to ensure data accuracy.

References

1. Barresi, R. & Campbell, K. P. Dystroglycan: from biosynthesis to pathogenesis of human disease. *J. Cell Sci.* **119**, 199–207 (2006).
2. Ervasti, J. M. & Campbell, K. P. Membrane organization of the dystrophin-glycoprotein complex. *Cell* **66**, 1121–1131 (1991).
3. Ibraghimov-Beskrovnaya, O. *et al.* Primary structure of dystrophin-associated glycoproteins linking dystrophin to the extracellular matrix. *Nature* **355**, 696–702 (1992).
4. Michele, D. E. & Campbell, K. P. Dystrophin-glycoprotein complex: post-translational processing and dystroglycan function. *J. Biol. Chem.* **278**, 15457–15460 (2003).
5. Ervasti, J. M. Dystrophin, its interactions with other proteins, and implications for muscular dystrophy. *Biochim. Biophys. Acta* **1772**, 108–117 (2007).
6. Petrof, B. J., Shrager, J. B., Stedman, H. H., Kelly, A. M. & Sweeney, H. L. Dystrophin protects the sarcolemma from stresses developed during muscle contraction. *Proc. Natl. Acad. Sci. U. S. A.* **90**, 3710–3714 (1993).
7. Agrawal, S. *et al.* Dystroglycan is selectively cleaved at the parenchymal basement membrane at sites of leukocyte extravasation in experimental autoimmune encephalomyelitis. *J. Exp. Med.* **203**, 1007–1019 (2006).
8. Williamson, R. A. *et al.* Dystroglycan is essential for early embryonic development : disruption of Reichert ' s membrane in Dag1-null mice. **6**, 831–841 (1997).
9. Cohn, R. D. *et al.* Disruption of DAG1 in differentiated skeletal muscle reveals a role for dystroglycan in muscle regeneration. *Cell* **110**, 639–648 (2002).
10. Moore, S. A. *et al.* Deletion of brain dystroglycan recapitulates aspects of congenital muscular dystrophy. *Nature* **418**, 422–425 (2002).
11. Jing, J. *et al.* Aberrant expression, processing and degradation of dystroglycan in squamous cell carcinomas. *Eur. J. Cancer* **40**, 2143–2151 (2004).
12. Losasso, C. *et al.* Anomalous dystroglycan in carcinoma cell lines. *FEBS Lett.* **484**, 194–198 (2000).
13. Singh, J. *et al.* Proteolytic enzymes and altered glycosylation modulate dystroglycan function in carcinoma cells. *Cancer Res.* **64**, 6152–6159 (2004).

14. Matsumura, K. *et al.* Proteolysis of beta-dystroglycan in muscular diseases. *Neuromuscul. Disord.* **15**, 336–341 (2005).
15. Yamada, H. *et al.* Processing of beta-dystroglycan by matrix metalloproteinase disrupts the link between the extracellular matrix and cell membrane via the dystroglycan complex. *Hum. Mol. Genet.* **10**, 1563–1569 (2001).
16. Nadarajah, V. D. *et al.* Serum matrix metalloproteinase-9 (MMP-9) as a biomarker for monitoring disease progression in Duchenne muscular dystrophy (DMD). *Neuromuscul. Disord.* **21**, 569–578 (2011).
17. Hindi, S. M., Shin, J., Ogura, Y., Li, H. & Kumar, A. Matrix metalloproteinase-9 inhibition improves proliferation and engraftment of myogenic cells in dystrophic muscle of mdx mice. *PLoS One* **8**, e72121 (2013).
18. Kumar, A., Bhatnagar, S. & Kumar, A. Matrix metalloproteinase inhibitor batimastat alleviates pathology and improves skeletal muscle function in dystrophin-deficient mdx mice. *Am. J. Pathol.* **177**, 248–260 (2010).
19. Li, H., Mittal, A., Makonchuk, D. Y., Bhatnagar, S. & Kumar, A. Matrix metalloproteinase-9 inhibition ameliorates pathogenesis and improves skeletal muscle regeneration in muscular dystrophy. *Hum. Mol. Genet.* **18**, 2584–2598 (2009).
20. Akhavan, A., Crivelli, S. N., Singh, M., Lingappa, V. R. & Muschler, J. L. SEA domain proteolysis determines the functional composition of dystroglycan. *FASEB J.* **22**, 612–621 (2008).
21. Geis, T. *et al.* Homozygous dystroglycan mutation associated with a novel muscle-eye-brain disease-like phenotype with multicystic leucodystrophy. *Neurogenetics* **14**, 205–213 (2013).
22. Pei, J. & Grishin, N. V. Expansion of divergent SEA domains in cell surface proteins and nucleoporin 54. *Protein Sci.* **26**, 617–630 (2017).
23. Gordon, W. R. *et al.* Structural basis for autoinhibition of Notch. *Nat. Struct. Mol. Biol.* **14**, 295–300 (2007).
24. Gordon, W. R. *et al.* Structure of the Notch1-negative regulatory region: implications for normal activation and pathogenic signaling in T-ALL. *Blood* **113**, 4381–4390 (2009).
25. Gordon, W. R. *et al.* Mechanical Allostery: Evidence for a Force Requirement in the Proteolytic Activation of Notch. *Dev. Cell* **33**, 729–736 (2015).
26. Hayward, A. N., Aird, E. J. & Gordon, W. R. A toolkit for studying cell surface shedding of diverse

transmembrane receptors. *Elife* **8**, (2019).

27. Bozzi, M. *et al.* Enzymatic processing of beta-dystroglycan recombinant ectodomain by MMP-9: identification of the main cleavage site. *IUBMB Life* **61**, 1143–1152 (2009).
28. Macao, B., Johansson, D. G. A., Hansson, G. C. & Härd, T. Autoproteolysis coupled to protein folding in the SEA domain of the membrane-bound MUC1 mucin. *Nat. Struct. Mol. Biol.* **13**, 71–76 (2006).
29. Pavšič, M., Gunčar, G., Djinović-Carugo, K. & Lenarčič, B. Crystal structure and its bearing towards an understanding of key biological functions of EpCAM. *Nat. Commun.* **5**, 4764 (2014).
30. De-la-Torre, P., Choudhary, D., Araya-Secchi, R., Narui, Y. & Sotomayor, M. A Mechanically Weak Extracellular Membrane-Adjacent Domain Induces Dimerization of Protocadherin-15. *Biophys. J.* **115**, 2368–2385 (2018).
31. Gupta, V. *et al.* The zebrafish dag1 mutant: a novel genetic model for dystroglycanopathies. *Hum. Mol. Genet.* **20**, 1712–1725 (2011).
32. Jayasinha, V. *et al.* Inhibition of dystroglycan cleavage causes muscular dystrophy in transgenic mice. *Neuromuscul. Disord.* **13**, 365–375 (2003).
33. Bork, P. & Patthy, L. The SEA module: a new extracellular domain associated with O-glycosylation. *Protein Sci.* **4**, 1421–1425 (1995).
34. Das, S. *et al.* Membrane proximal ectodomain cleavage of MUC16 occurs in the acidifying Golgi/post-Golgi compartments. *Sci. Rep.* **5**, 9759 (2015).
35. Primo, M. E. *et al.* Structure of the mature ectodomain of the human receptor-type protein-tyrosine phosphatase IA-2. *J. Biol. Chem.* **283**, 4674–4681 (2008).
36. Gordon, W. R. *et al.* Structural basis for autoinhibition of Notch. *Nat. Struct. Mol. Biol.* **14**, 295–300 (2007).
37. Xu, X. *et al.* Insights into Autoregulation of Notch3 from Structural and Functional Studies of Its Negative Regulatory Region. *Structure* **23**, 1227–1235 (2015).
38. Gordon, W. R. *et al.* Structure of the Notch1-negative regulatory region: implications for normal activation and pathogenic signaling in T-ALL. *Blood* **113**, 4381–4390 (2009).
39. Malecki, M. J. *et al.* Leukemia-associated mutations within the NOTCH1 heterodimerization domain fall into at least two distinct mechanistic classes. *Mol. Cell. Biol.* **26**, 4642–4651 (2006).
40. Sanchez-Irizarry, C. *et al.* Notch subunit heterodimerization and prevention of ligand-independent

proteolytic activation depend, respectively, on a novel domain and the LNR repeats. *Mol. Cell. Biol.* **24**, 9265–9273 (2004).

41. Li, K. *et al.* Modulation of Notch signaling by antibodies specific for the extracellular negative regulatory region of NOTCH3. *J. Biol. Chem.* **283**, 8046–8054 (2008).
42. Wu, Y. *et al.* Therapeutic antibody targeting of individual Notch receptors. *Nature* **464**, 1052–1057 (2010).
43. Molina, M. A. *et al.* Trastuzumab (Herceptin), a Humanized Anti-HER2 Receptor Monoclonal Antibody, Inhibits Basal and Activated HER2 Ectodomain Cleavage in Breast Cancer Cells1. *Cancer Res.* **61**, 4744–4749 (2001).